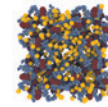


THIS WEEK

EDITORIALS

PUBLISHING Towards greater sharing of computer code **p.536**

WORLD VIEW Scientists must mobilize to track Ebola outbreak **p.537**



OXYGEN Sponge-like powder holds secret to gas storage **p.539**

Pillars of reform

The Chinese government's planned overhaul of its core research-funding system is vital if the country is to achieve its potential on the global scientific stage.

China has come to an unsettling conclusion: the system that it uses to invest in science and technology is broken. The nation's past efforts to become a great innovative state have produced clear signs of success: a flourishing space programme, a dominant global position in genome sequencing and some internationally prominent technology companies. But scratching beneath that surface reveals a creeping suspicion that China is not getting value for the money it diverts into research.

These misgivings seem to have reached crisis point. China today is full of new initiatives, reforms and an anti-corruption drive that together aim to set the nation on the right track. The impact on science is set to be monumental — if China follows it through.

The problems that lead to inefficiency in science are various. Research programmes overlap. Low-quality researchers, often selected more than a decade ago when they were the best that China could get, are now a dead weight in a system that has since managed to acquire much better. And worryingly, earlier this month, the government accused researchers at some of the country's most prestigious universities of misusing research funds (see *Nature* <http://doi.org/wpb;2014>). Shutting down such egregious abusers is a first step, and the ten-year prison terms handed down to two scientists who diverted funds certainly indicate that China is determined to make others think twice before they forge receipts.

The nation is also right to reform how the Chinese Academy of Sciences supports promising research projects. And perhaps most boldly — although the details are still sketchy — the finance and science ministries last week announced a joint effort to completely overhaul the way competitive funding is distributed.

In principle, this could drive improvements in the crucial area of how research and researchers are assessed. Critics have long argued that small grants in China are reviewed strictly, whereas large grants receive little or no review. The large grants, critics charge, are too often decided on the basis of personal connections, not scientific validity. The latest reform aims to stop the science ministry awarding these large grants and will instead distribute them through a new (albeit still vaguely defined) 'five-pillar system'.

The key to China's new system is how it will judge which research is worth pursuing. How can it promote the truly innovative and feasible while suppressing science that is either not original or unfeasible? This is not a new problem. Policy-makers and leaders of the science community worldwide grapple with these questions. The topic was again addressed at a meeting on research assessment and evaluation in Shanghai last week, co-sponsored by *Nature*.

Discussions there ranged from the evaluation of young researchers to the difficulty of balancing societal impact with research that drives innovation. For instance, Nobel-prizewinning biophysicist Kurt Wüthrich argued for focusing less on past achievements and failures, and more on "talented underachievers" — those with great

ideas who just might not have proved themselves yet.

Assessment of research is an especially urgent issue for China as it attempts to leave behind cronyism and an incestuous grant system. It would be too easy to fall back on hyper-quantitative assessment — a system that results in scientists running from one project to the next, trying to join as many as they can to maximize the number of papers and awards that they rack up.

The reform of the Chinese Academy of Sciences is meant to change that. An initiative in translational research, launched earlier this

"The impact on science is set to be monumental — if China follows it through."

month (see page 547), has put recruitment of the best researchers at the centre of its planning — and has arranged an impressive international committee to ensure that happens.

Will the reform of China's core competitive funding system work? It is a promising sign that the National Natural Science Foundation of China, the well-respected body that

distributes the smaller grants, will be one of the five pillars now tasked with broader research assessment.

But what about the other pillars? Will they be the old fragmented divisions of the science ministry with new names, or will they really be streamlined units that are determined to make every renminbi count? How the government acts on that point will largely determine the future of research in China.

In doing so, it will influence the rate at which China's overseas scientists are willing to head back home, and whether foreign scientists will respond to China's attempts to lure them. Such reforms could also help to resolve one of the big questions in modern science: if and when the impact of research in China will surpass that of the United States. ■

Call to action

Time to ramp up science's contribution to controlling the Ebola outbreak.

Science has so far taken a back seat as the Ebola outbreak has continued to spread. Research has deferred to the need to gear up the public-health response. But there is a growing sense that, unless science can somehow now change the game, the outbreak will be difficult to bring under control.

The Ebola virus has killed more than 4,800 people in six countries, and has affected people in another two, the latest being Mali. There are still not enough medical staff and treatment beds to handle the current caseload. The World Health Organization (WHO) projects that

as many as 10,000 new cases could be arising per week by December if the outbreak is not turned around.

Enter science. Speeding the development of treatments and vaccines is one area in which the international community is trying to move forward. On 22 October, the US Biomedical Advanced Research and Development Authority and the US Army awarded US\$17.1 million to Profectus BioSciences, a company based in Baltimore, Maryland, that is developing vaccines against Ebola based on vesicular stomatitis virus. It is the third candidate Ebola vaccine to have moved towards or into clinical trials this year. On 24 October, the WHO outlined plans to test the first two — one licensed to NewLink Genetics of Ames, Iowa, the other being developed by GlaxoSmithKline, headquartered in London. These two vaccines have already entered human safety trials and the WHO says that they could be tested in health-care workers and others in West Africa as early as December.

In the meantime, aid agencies such as Médecins Sans Frontières (also known as Doctors Without Borders) and researchers funded by the European Union will test candidate Ebola treatments, including experimental drugs, medicines already approved for other uses that could be made available 'off label', and purified plasma or blood from Ebola survivors.

Beyond treatments and vaccines, scientists have more fundamental questions, about both the Ebola virus behind the current outbreak and other viruses in the family to which it belongs, the filoviruses. This group includes Marburg virus, also capable of causing a lethal haemorrhagic fever, which killed a Ugandan health-care worker on 28 September. A third filovirus outbreak occurred this year in the Democratic Republic of the Congo, where an Ebola outbreak unrelated to that in West Africa has killed 49 people.

The emergence of three filovirus outbreaks this year and the increasing frequency and reach of such outbreaks — which have occurred every year except 2 in the past 21 years — should serve as the clearest warning possible: we urgently need to understand more about the pathology, distribution, epidemiology and clinical aspects of these viruses. A World View on page 537 argues that such science should help to steer the response; a News Feature on page 554 lays out the five most pressing questions about the filoviruses, and says why answering them might help to prevent a future outbreak or even help to bring this one under control.

For instance, new filoviruses have been discovered within the past five years, such as the Lloviu virus discovered in 2011 in bats in Spain. And scientists have learnt that these viruses have a much more widespread distribution than was suspected. The Reston virus, for example, an ebolavirus that does not seem to harm humans, has turned up in recent years in pigs in both the Philippines and China. Scientists suspect that there are more of these viruses to be found, in more places, and urgently want to understand why some are lethal to humans and others are not — and whether that could change.

“Whether it is Ebola virus or something completely different, there will be a next time.”

It is also not known which animals harbour Ebola virus in the wild, or how the first person infected in the West African outbreak last December contracted the disease. Understanding this is crucial if people are to avoid a possible reservoir in the future.

It has been difficult to answer these questions for many reasons, such as the (fortunate) relative rarity and unpredictability of human filovirus outbreaks. And laboratory studies require highly contained, specialized biosafety-level-4 (BSL-4) labs — of which there are too few around the world.

Thanks to a biodefence building boom over the past decade, there are now 13 such labs planned or operating in the United States. Canada, France, Australia, Germany, the United Kingdom, South Africa, Gabon and Russia are among the select nations that also have such facilities. But there are major research-funding nations, such as Japan, that do not have BSL-4 labs, or do not allow them to perform the highest-containment research because of worries that pathogens could escape and spark lethal local epidemics.

The current Ebola outbreak proves the fallacy of that decision. The world would not be in the position it is today, with the possibility of deploying an Ebola vaccine during the current outbreak, without the existence of both high-containment facilities and money for research on diseases that are, thankfully, rare in developing countries. More of both, in more places, can only hasten our understanding of Ebola and other diseases. Because one thing is clear: whether it is Ebola virus, another filovirus or something completely different, there will be a next time. ■

Code share

Papers in Nature journals should make computer code accessible where possible.

A theme in *Nature's* ongoing campaign for the replicability and reproducibility of our research papers is that key components of publications should be available to peers who wish to validate the techniques and results.

A core element of many papers is the computer code used by authors in models, simulations and data analysis. In an ideal world, this code would always be transportable and easily used by others. In such a world, our editorial policy would be to insist on sharing to allow free use, as we already do (as far as is practicable) with data and research materials. Unfortunately, such an ideal is not easy to attain owing to the amount of extra funding and effort it would require to render some major pieces of code shareable. Nevertheless, we at *Nature* and the *Nature* research journals want to encourage as much sharing as possible.

Climate modellers have made some strides in this regard. The journal *Geoscientific Model Development* has a good example of such a policy (see go.nature.com/jv8g1w), and an article in *Nature Geoscience* discusses some of the opportunities presented by code sharing, as well as

the obstacles (S. M. Easterbrook *Nature Geosci.* 7, 779–781; 2014).

As a leading example of transparency policies in other disciplines, the data journal *GigaScience* requires code used in its papers to be available, and hosts it in a way that allows others to analyse the data in publications. One point made by Easterbrook is that even if the code is shared, others might often make little or no use of it, but on some occasions the take-up will be large.

Nature and the *Nature* journals have decided that, given the diversity of practices in the disciplines we cover, we cannot insist on sharing computer code in all cases. But we can go further than we have in the past, by at least indicating when code is available. Accordingly, our policy now mandates that when code is central to reaching a paper's conclusions, we require a statement describing whether that code is available and setting out any restrictions on accessibility. Editors will insist on availability where they consider it appropriate: any practical issues preventing code sharing will be evaluated by the editors, who reserve the right to decline a paper if important code is unavailable. Moreover, we will provide a dedicated section in articles in which any information on computer code can be placed. And we will work with individual communities to put together best-practice guidelines and possibly more-detailed rules.

➔ NATURE.COM
To comment online,
click on Editorials at:
go.nature.com/xhunq

For full details, see our guide for authors at go.nature.com/o5ykhe. For an archive of our content and initiatives concerning reproducibility, see <http://www.nature.com/nature/focus/reproducibility>. ■



Developed nations must not fear sending Ebola help

The anxiety and stigma associated with Ebola are hampering Australia's willingness and ability to help with the control efforts in Africa, argues **Tim Inglis**.

The Ebola outbreak has pitted rational science against fear and superstition. We see this in Africa: the murder last month of eight people working to raise awareness of the disease close to the town of Nzerekore, in southeastern Guinea, is a tragic example. But we see it in the supposed scientific superpowers too.

In the United States, for example, the debate over how the nation should respond is being undermined by media hype and political panic over domestic cases; there are calls for total bans on flights from the outbreak region. And here in Australia, the government is reluctant to send people to help to control the spread of the disease until it can get guarantees from nations closer to the outbreak that those people would be able to receive treatment in the event that they became infected. Worse, health-care workers who voluntarily went to the affected region to help have been criticized on their return to Australia. These people — including a nurse who developed a fever after her return (but tested negative for Ebola) — have been turned on by politicians and the press for putting fellow Australians at risk, despite having followed expert guidance and quarantined themselves.

The threat posed by a few imported cases of Ebola is low: Australia has the procedures, resources and facilities here to minimize the risk of secondary transmission. And other developed countries are willing to accept that risk because they understand the crucial and urgent need to tackle the outbreak in Africa. Mathematical models predict that transmission will continue for many months, even if the rate of expansion levels out soon. The large number of infections clearly increases the probability of exports and that the virus's genome will continue to mutate. The real issue is that the threat to Australia, the United States and other developed countries will be much higher in six months. The best defence is to act now and in Africa.

I believe that Australia could handle isolated cases. Why do politicians and the public not seem to share that confidence? It is true that we do not understand the Ebola virus very well (see page 554), and when science is lacking, the default response to infectious disease is often fear. But before we are too critical, scientists must realize that some of the public and political response is driven by the mystique that we have created around Ebola and other viral haemorrhagic fever agents.

For example, when the outbreak began, several biomedical research centres in Australia promptly established procedures for screening possible cases — all negative so far. Eager to reassure and to show what a good job these public-health labs were doing, we entertained the media outside the high-security labs. The messages were supposed to be comforting, but the

images of staff entering secure facilities in full Tyvek protection suits was anything but. Such footage does little to challenge a popular view informed as much by disaster movies as by molecular biology. In fact, it could have inadvertently made the situation worse and entrenched a national preoccupation with border protection.

Research on Ebola lags behind the clinical need, but much has been learned about its virology, immunology and molecular epidemiology in the past nine months. Here, science can suggest a way to improve the international response. At present, lab resources — including those sent from abroad — tend to be located in disease hot spots. But if we want to restrict the outward spread of the virus, it makes little sense to busy these screening facilities — and new ones to be deployed — with samples that have a high probability of testing positive. We should focus support on rapid turnaround of tests for the field epidemiology teams working at the edge of the outbreak zone.

This would take diagnosis to where patients are most likely to become infected next, rather than sending those people or their blood samples to centres where the disease is concentrated. This could help to steer the countermeasure response to where it is likely to have the most impact. In view of the escalating case load and mortality figures, it is an experiment we should try as a matter of priority.

Crucial to such a response are mobile labs that can follow the disease control teams. My team has one. We have developed and tested a mobile molecular lab in other emerging infectious-disease settings, including epidemic influenza. It is a small technical step to transfer the standard

(fixed) lab assay for Ebola virus RNA to a portable thermocycler device.

Colleagues have volunteered and we are set to test the mobile lab in the tropical conditions of northern Western Australia in the coming weeks. We hope to be working in West Africa by the end of the year.

What will we bring home with us? Unlike clinical staff, we are unlikely to have to have much contact with actual patients. The blood samples we handle could be highly infectious, but I argue that researchers who understand the biology of the Ebola virus have a healthy respect for the occupational safety risks, and know how to mitigate those risks. Australia can therefore make an effective contribution to the control efforts with little risk to her own population.

Past experience has shown me that effective infection control needs a strong grounding in science. Scientists are needed at the front line. The risk is worth it. ■

Tim Inglis is a medical microbiologist at the University of Western Australia in Nedlands.
e-mail: tim.inglis@uwa.edu.au

**CRUCIAL TO THE
RESPONSE
ARE
MOBILE LABS
THAT CAN FOLLOW
THE DISEASE
CONTROL
TEAMS.**

➔ NATURE.COM
Discuss this article
online at:
go.nature.com/om9zen

RESEARCH HIGHLIGHTS

Selections from the scientific literature

PLANETARY SCIENCE

Mercury's ice is a recent arrival

Ice at Mercury's poles is a relatively new arrival — a finding that could help to resolve a debate about whether ice may have survived for billions of years on the planet closest to the Sun.

Using data from NASA's MESSENGER spacecraft, Nancy Chabot of the Johns Hopkins University Applied Physics Laboratory in Laurel, Maryland, and her colleagues studied how light scattered inside dimly lit polar craters. They found that in a northern crater called Prokofiev, highly reflective ice drapes over the underlying topography. This suggests that the ice has appeared on the surface relatively recently.

This ice was either delivered to Mercury, perhaps by comets, or churned up from below by impacts battering its surface, the authors conclude. *Geology* <http://doi.org/wjf> (2014)

MYCOLOGY

Teamwork helps yeast to infect

Immune responses that should combat a disease caused by yeast instead make the fungus grow, potentially worsening the infection.

Robin May at the University of Birmingham, UK, and his co-workers studied strains of *Cryptococcus gattii*, which can cause meningitis and other problems.

They found that oxygen-containing molecules produced by the host as a defence mechanism cause some of the fungal cells to become quiescent and resist immune responses. The remaining cells proliferate quickly, resulting in an overall increase in cell



GERRY ELLIS/MINDEN PICTURES/FLPA

EVOLUTION

Lizards adapt quickly to invaders

Lizards in Florida have rapidly evolved traits that make them better tree-climbers, probably in response to an invasive competitor.

Cuban brown anole lizards (*Anolis sagrei*) have spread over the past few decades across the southeastern United States, where they compete for territory and food with the native green anole (*Anolis carolinensis*; pictured). Yoel Stuart at the University of Texas, Austin, and his colleagues introduced small populations of the invader to three islands in a central Florida

lagoon and found that the native green anoles perched higher in trees than native lizards on nearby islands that had not been invaded. After just 20 generations of anoles, the team found that native lizards on invaded islands had evolved larger toepads, probably to help them cling to less-secure branches farther up in trees.

Competition between closely related species can drive rapid, observable evolutionary change, the authors say.

Science 346, 463–466 (2014)

numbers. The resistant cells contained energy-producing organelles called mitochondria that were tubular in shape.

The findings suggest that the fungal cells act as a team during infection, with non-dividing cells helping neighbouring ones to grow rapidly.

Nature Commun. 5, 5194 (2014)

ASTRONOMY

Mysterious signals may be from Earth

Radio pulses that look like they came from deep space could actually have earthly origins.

A team led by Pascal

Saint-Hilaire at the University of California, Berkeley, detected five short but intense radio bursts at the Bleien Radio Observatory in Switzerland. This is only the second location at which such pulses have been detected, and their origins are still unclear. Characteristics of the radio waves suggest that they were stretched after passing through vast amounts of plasma — usually indicating an origin outside of the Milky Way, such as exploding stars in other galaxies.

However, the pulses were detected only when the antenna was in a mode susceptible to ground

interference, and all but one pulse occurred in the late morning. This suggests that the signals could actually be coming from sources on Earth. *Astrophys. J.* 795, 19 (2014)

CANCER

Tumours linked to cellular rubbish

Discarded rubbish from tumours could trigger nearby healthy cells to become malignant.

Many cells shed exosomes: membrane-bound packages of proteins, DNA and RNA that are thought to be a

waste-management system. Raghu Kalluri at the University of Texas MD Anderson Cancer Center in Houston and his colleagues found that exosomes from cancer cells contain the building blocks for short RNA fragments that can shut off gene expression. Healthy cells that were exposed to cancer exosomes in culture caused tumours when the cells were later injected into mice, whereas cells exposed to normal exosomes did not.

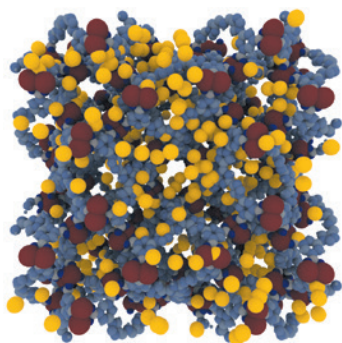
Samples of exosomes from people with breast cancer also caused tumours in 5 out of 11 mice treated. This finding could aid the discovery of markers for disease progression, or even represent a route to new cancer therapies. *Cancer Cell* <http://doi.org/wkg> (2014)

CHEMISTRY

Molecular sponges store oxygen

An oxygen cylinder could hold even more gas if it were filled with sponge-like powders, chemists report.

The powders are a type of metal-organic framework (MOF): sponge-like materials in which metal atoms are connected by organic groups, creating a porous network with many promising applications including gas storage. Omar Farha of Northwestern University in Evanston, Illinois, and his colleagues simulated oxygen adsorption (pictured, yellow spheres) on 10,000 MOFs and selected two to test. They experimentally showed that the MOFs could store and release oxygen over 50 cycles and outperformed



zeolites, another kind of chemical sponge.

The structures could allow soldiers or medical teams to carry oxygen using smaller, lighter containers that operate at lower pressures than cylinders, the team says.

Angew. Chem. Int. Edn <http://doi.org/f2vn85> (2014)

GLACIOLOGY

Channels hint at glacier hardness

Ancient channels preserved beneath the West Antarctic Ice Sheet suggest that part of the glacier prevailed during warm periods more than two million years ago.

By combining radio-echo soundings of the landscape underneath the glacier with satellite images of the ice surface, Kathryn Rose of the Bristol Glaciology Centre, UK, Martin Siegert at Imperial College London and their colleagues discovered a series of ancient wide, shallow channels. These suggest that, historically, there was a large flow of meltwater from the ice surface. The most recent period that was warm enough to generate such melting was during the Pliocene, 5.3 million to 2.6 million years ago.

Parts of this ice sheet, which some studies have predicted will destabilize in a warmer climate, may have existed at intervals during a period that was 2 °C warmer than now, the authors say.

Geology <http://doi.org/wkt> (2014)

BIOTECHNOLOGY

Paper-based gene tools

Functional biological circuits can be printed on paper, reports a team led by James Collins at Boston University in Massachusetts.

The team synthesized cell-free gene networks from off-the-shelf parts and freeze-dried them on to paper. When later rehydrated, the networks worked as programmable

SOCIAL SELECTION

Popular articles on social media

Rules for reproducibility win support

Nearly a decade after writing a scathing critique of biomedical research, 'Why Most Published Research Findings Are False', Stanford University scientist John Ioannidis has published a follow-up. The health-policy researcher suggests a blueprint for making scientific results more reliable, including increasing the statistical certainty of discoveries, giving more weight to negative results and changing how researchers earn kudos.

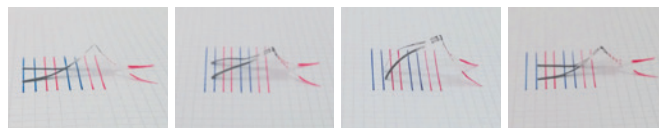
Many commenters chimed in with support for his paper, even if they did not believe that change could come easily. Simon Wheeler, a public-health nutritionist at the University of Cambridge, UK, endorsed Ioannidis's suggestions, tweeting that scientists should be "creating a culture where these are norms and expectations, not just lofty ideals". Mick Watson, a computational biologist at the University of Edinburgh's Roslin Institute, UK, tweeted, "I'm totally with John Ioannidis when he says the scientific reward system needs to change."

PLoS Med. 11, e1001747 (2014)



Based on data from altmetric.com. Altmetric is supported by Macmillan Science and Education, which owns Nature Publishing Group.

➔ NATURE.COM
For more on popular papers:
go.nature.com/kw7hck



in vitro diagnostics with a colorimetric output. The team demonstrated the technology by using the circuits to detect glucose and to distinguish between RNA fragments from two related virus species of the Ebolavirus genus, called Sudan and Zaire.

The paper technology should allow the easy and low-cost use of synthetic biology outside the lab, the team predicts.

Cell <http://doi.org/wkr> (2014)

MATERIALS

Sunshine drives graphene machines

Machines that move by bending in response to moisture can be made by exposing thin sheets of graphene oxide to sunlight.

A team led by Hong-Bo Sun at Jilin University in

Changchun, China, focused sunlight on one side of graphene oxide paper. The ultraviolet radiation induced a reaction that negatively charged the surface so that it repelled water more strongly than the layer below, causing the paper to curl in seconds on contact with moisture.

The team fashioned a claw from the paper that closes when approached by a sweaty finger, and a paper robot (pictured) that crawls when the humidity in its environment is raised and lowered. The authors say that the material could be used in devices including sensors and smart textiles.

Adv. Mater. <http://doi.org/wjj> (2014)

➔ NATURE.COM
For the latest research published by Nature visit:
www.nature.com/latestresearch

SEVEN DAYS

The news in brief

EVENTS

China's Moon shot

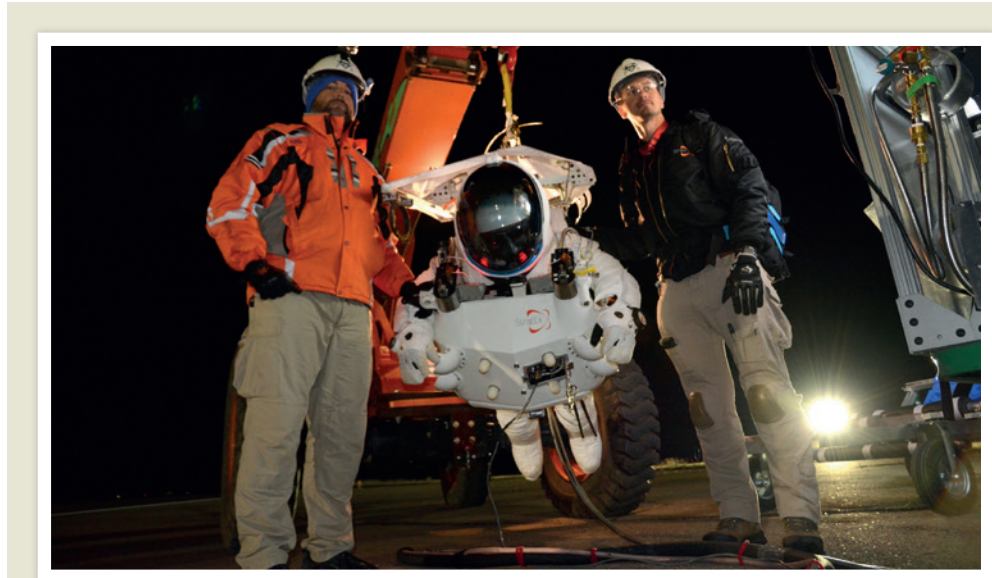
China launched an uncrewed lunar orbiter on 24 October, from the Xichang launch center in Sichuan province. The probe will slingshot around the Moon and is scheduled to land in northern China after an eight-day journey, becoming the nation's first round-trip lunar mission. The mission aims to demonstrate heat-shield and guidance technologies for a planned 2017 lander, which will be China's first to bring samples back to Earth.

Ebola transported

Health officials in Mali reported on 24 October that the nation's first confirmed person with Ebola — a two-year-old girl who had recently arrived from Guinea — had died. The girl began showing symptoms such as nosebleeds while still in Guinea, and started out for Mali with her grandmother on 19 October. Many opportunities for high-risk exposures occurred while the child travelled extensively by public transport, says the World Health Organization, which called the situation an emergency. An initial investigation has identified 43 close, unprotected contacts, including 10 health-care workers.

Hunt fraudsters

The US government has charged two South Africans with conspiracy to sell places on illegal rhinoceros hunts by misleading US hunters. According to an indictment released on 23 October, brothers Dawie and Janneman Groenewald travelled throughout the United States between 2005 and 2010, charging people between US\$3,500 and \$15,000 to hunt 'problem' rhinos in South Africa that they claimed



PARAGON SPACE DEVELOPMENT CORPORATION

Skydiving hits new heights

Alan Eustace, a senior vice-president at Google in Mountain View, California, broke the sound barrier and a world skydiving record on 24 October, when he jumped from a reported altitude of 41.4 kilometres over Roswell, New Mexico. Eustace began his roughly 2.5-hour ascent into the upper stratosphere around dawn, being dragged skyward by a helium-filled scientific balloon. A pressurized space

suit (pictured) designed by Paragon Space Development in Tuscon, Arizona, helped Eustace to breathe at high altitude and prevented fluids in his body from boiling. On his way down, Eustace reached an estimated peak speed of around 1,320 kilometres per hour. In October 2012, Austrian parachutist Felix Baumgartner set the previous record-high skydive, jumping from an altitude of 39 kilometres.

needed to be euthanized. The two were also charged with secretly trafficking rhino horns, and laundering their profits through bank accounts that they had set up for their business in Autauga County, Alabama.

FUNDING

Funding reform

China's government on 21 October announced the passage of a reform plan to overhaul the funding of research in the country. Under the plan, the main competitive government grant schemes will be eliminated by 2017, including two core programmes that supported China's science and technology

development in recent decades. Funding will be reorganized into five programmes managed by a new science and technology agency. The move is part of an effort to reduce wastefulness and to consolidate grant management, currently distributed across various government departments. See pages 535 and 547 and go.nature.com/pzzvf5 for more.

POLICY

Full text for free

More than half of all peer-reviewed research articles published between 2007 and 2012 are now free to download from the Internet, according to a 22 October report produced

for the European Commission. By comparison, last year's assessment found that papers from only one year (2011) had reached 50% free access. See go.nature.com/dbw76k for more.

Diversity grants

The US National Institutes of Health (NIH) has awarded US\$31 million in grants aimed at encouraging minorities and under-represented groups to pursue careers in biomedical fields, the agency announced on 22 October. The grants target institutions that receive relatively little NIH funding and enrol a high proportion of students from low-income backgrounds. The funds will support

ways to improve access to research opportunities for undergraduates, build a national mentoring network, and assess the results of the initiatives themselves to understand which strategies bring lasting change.

EU climate goals

European Union (EU) member states have agreed to reduce the region's greenhouse-gas emissions by at least 40% by 2030, relative to 1990 levels. The EU's new climate and energy policy framework, released on 24 October, includes a politically binding target of getting 27% of energy from renewable sources by 2030; and a goal of improving overall energy efficiency by 27%, to be reviewed in 2020 with the hope of increasing the latter target to 30%. See *Nature* <http://doi.org/wmb> (2014) for more.

Commission vote

The European Parliament approved a new European Commission on 22 October, which will begin a five-year term on 1 November. Portuguese economist Carlos Moedas will serve as commissioner for research, science and innovation, while Spain's Miguel Arias Cañete, a lawyer, will be commissioner for energy and climate change. See go.nature.com/t73tfi for more.

RESEARCH



Comet close-up

NASA's Mars Reconnaissance Orbiter snapped comet C/2013 A1 Siding Spring (pictured) as it buzzed past the red planet on 19 October. The image, released last week, is the first ever taken of the nucleus of a comet from the Oort Cloud, the reservoir of icy bodies that lies far beyond Pluto. The nucleus of Siding Spring turned out to be smaller than expected, measuring less than 500 metres across. Siding Spring passed just 138,000 kilometres from the Martian surface. See go.nature.com/psvju9 for more.

Ebola vaccines

On 24 October, the World Health Organization (WHO) announced plans to produce millions of doses of experimental Ebola vaccines by the end of 2015, with hundreds of thousands of doses available by June. The outbreak has infected more than 10,000 people and killed around 5,000, the WHO estimates. Phase I trials of two

candidate vaccines have begun, and the first phase II trials are set to start in Liberia in December and in Sierra Leone in January. See go.nature.com/pj71hi for more.

First light

The National Synchrotron Light Source II, a high-energy photon source at Brookhaven National Laboratory in Upton, New York, produced its first photons on 23 October — a major milestone towards achieving full research capabilities. Thirty beamlines are under development around the 792-metre electron-storage ring, where specialized magnets coax the electrons into producing intense beams of X-rays. These can be guided into experimental stations and used to probe matter. The first experiments are expected before the end of this year.

FACILITIES

Cancer centre

The University of Oxford, UK, announced plans on 23 October to establish the Precision Cancer Medicine Institute, a £110-million (US\$177-million) research facility funded by charitable and commercial partners and a £35-million government grant. The institute will test precision treatments such as proton-beam therapy for patients with early-stage

COMING UP

5–8 NOVEMBER

The annual meeting of the Society of Vertebrate Paleontology takes place in Berlin, where scientists will discuss vertebrate origins, evolution and functional morphology. go.nature.com/hmmzkm

6–9 NOVEMBER

At the International Conference of Crisis Mappers in New York City, aid workers, technologists and policy-makers will discuss best practices and recent innovations in humanitarian technology. go.nature.com/j19byw

cancers that currently have a poor prognosis. The centre is also expected to house a new research partnership with the non-profit Chan Soon-Shiong Institute of Molecular Medicine in Culver City, California, which is intended to target treatments more precisely using large amounts of genetic, molecular and clinical data.

AWARDS

Maddox prize

The John Maddox Prize for Standing up for Science was awarded on 27 October to physicist David Robert Grimes, a postdoc and journalist at the University of Oxford, UK, and US freelance journalist Emily Willingham. The award honours people who promote science in the public interest, often despite obstacles. Named after a former editor of *Nature*, the prize is jointly awarded by *Nature*, the Kohn Foundation in London and the UK campaign group Sense About Science. See go.nature.com/rbjzsg for more.

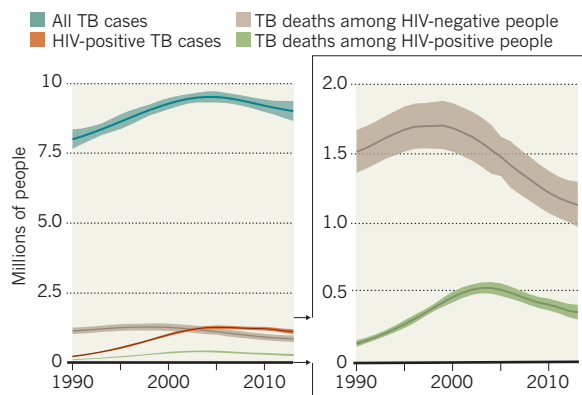
NATURE.COM
For daily news updates see:
www.nature.com/news

TREND WATCH

An estimated 9 million people developed tuberculosis (TB) last year, the World Health Organization (WHO) said in a 22 October annual update (see go.nature.com/rohvhp). Although rates of new cases and deaths are thought to be falling, this figure is 400,000 more than that estimated for the previous year, because improved measurements enable the WHO to be more confident in its figures. The agency said that annual funding for TB prevention needs to rise by some US\$2 billion, to \$8 billion.

TUBERCULOSIS TOLL REVISED UPWARDS

The World Health Organization estimates that 9 million people developed tuberculosis (TB) last year.



NEWS IN FOCUS

VIROLOGY Destruction of NIH smallpox sample on hold **p.544**

MARINE BIOLOGY Sea-floor destruction feeds China jellyfish blooms **p.545**

AGEING Pet dogs to be enrolled in trial of anti-ageing drug **p.546**



SCIENCE METRICS Top-100 cited papers thrust methods into limelight **p.550**

ARCTIC-IMAGES/CORBIS



Record amounts of gas have spewed from Iceland's Holuhraun plain.

GEOLOGY

Icelandic volcano stuns scientists

Sulphur-rich eruption defies preparations for an ashy blast.

BY ALEXANDRA WITZE IN REYKJAVIK

Icelandic sunrises and sunsets have been tinged blood red, of late. Above the maritime bustle of Reykjavik's harbour and the city's towering concrete Hallgrímskirkja church, volcanic pollution gives the skies an eerie glow.

For eight weeks, lava has been spurting out of a fissure in the ground radiating from the Bárðarbunga volcano, about 250 kilometres from Reykjavik. Sulphur dioxide has been spurting too — 35,000 tonnes of it a day, more than twice the amount spewing from all of Europe's smokestacks. The gas has spread

across the Icelandic countryside, causing people to wheeze and trapping some indoors.

The record-setting amount of pollution has surprised even volcanologists in the middle of a major project funded by the European Union to understand the island's fiery activity. They had been preparing for a repeat of the 2010 Eyjafjallajökull eruption, which led to a billowing ash plume that grounded planes across Europe. "Everybody was expecting a big ash cloud, and now we have something totally different," says Anja Schmidt, an atmospheric modeller at the University of Leeds, UK, who studies how volcanic gases spread.

The timing of the eruption was just about perfect for the project, called FUTUREVOLC. The initiative aims to use Iceland as a natural laboratory to understand how magma makes its way from deep in Earth's crust to the surface — to do so, its organizers have focused on four of Iceland's most active volcanoes, one of which is Bárðarbunga. The researchers used extra seismometers and global-positioning-system (GPS) stations to bolster the monitoring network maintained by the Icelandic Meteorological Office and the University of Iceland, both in Reykjavik. And they used these to measure the events leading up to and after the eruption with unprecedented detail.

THE LONG MARCH

Earthquakes began shaking Bárðarbunga on 16 August. For two weeks, researchers watched as seismic activity marched north and east, towards the edge of the ice cap that covers the volcano. GPS stations measured the ground flexing upward as huge amounts of magma shifted underground. "The seismic shows us the detail, and the GPS shows us the volume," says Kristín Vogfjörð, a seismologist at the Icelandic Meteorological Office and co-leader of FUTUREVOLC.

For 45 kilometres, the magma crept along, cooling and forming an underground sheet known as a dike. By 29 August, it had made it to the edge of the ice cap and begun erupting into a barren plain called Holuhraun.

Since then, the eruption has spewed at least half a cubic kilometre of lava, making it the largest lava-producing eruption in Iceland since 1947. But the quantity of gas is what has startled scientists most. Chunks of rock collected from the eruption show how gas-rich the lava is; the rocks are porous, filled with air pockets where the gas has leaked out.

With the right winds, the sulphur that Holuhraun produces can reach as far as the European continent, where Austria has recorded more sulphur in its air than any time since the industrial clean-up of the 1980s.

FUTUREVOLC has placed some basic gas-monitoring equipment at its volcanoes, including two types of spectrometer at the Holuhraun site. One measures concentrations of SO₂ by studying how it absorbs particular wavelengths of sunlight. The other scans for multiple gasses simultaneously.

Both types of instrument rely on daylight to make their measurements, and the ▶

► encroaching darkness of the high northern winter will limit their usefulness, says Sara Barsotti, an atmospheric physicist at the Icelandic Meteorological Office. (Even getting to the eruption site, which is remote, is getting harder as the winter sets in.) Researchers will try to continue making as many gas measurements on the ground as possible, and compare those to the rough estimates made by satellites, Barsotti says.

Icelandic officials are trying to work out how much of a hazard the sulphur is and whether they can predict its movement more accurately. The meteorological office has begun issuing forecasts of where the gas is likely to travel each day. Sulphur spikes as high as 21,000

micrograms per cubic metre were measured last weekend in the town of Höfn; the World Health Organization recommends no more than 500 micrograms per cubic metre for a 10-minute exposure.

No one has died in the eruption, and the plume is not high enough to penetrate the stratosphere and cause widespread climate perturbations. But the million tonnes of sulphur emitted so far are an unprecedented experiment in testing the effects of toxic-gas exposure, Barsotti says.

Lessons from Iceland may prove useful in understanding long-term gas exposure in other volcanic regions, such as Japan, Indonesia and Hawaii. In the early 2000s, residents

around the Miyake-jima volcano in Japan were evacuated when it began erupting with roughly the same level of sulphur emissions.

In Iceland, the last similar event was a fissure eruption known as the Krafla fires that began in 1975 and lasted on and off until 1984, says Freysteinn Sigmundsson, a volcanologist at the University of Iceland and co-leader of FUTUREVOLC. If the current eruption is tapping magma deep in the crust, as the lava's volume and chemistry suggest, then it, too, may continue for months or even years.

"This eruption comes at a good time for the project," Sigmundsson says, waving out his office window at Reykjavik's red skies. "We don't see an end in sight." ■

VIROLOGY

'Forgotten' NIH smallpox virus languishes on death row

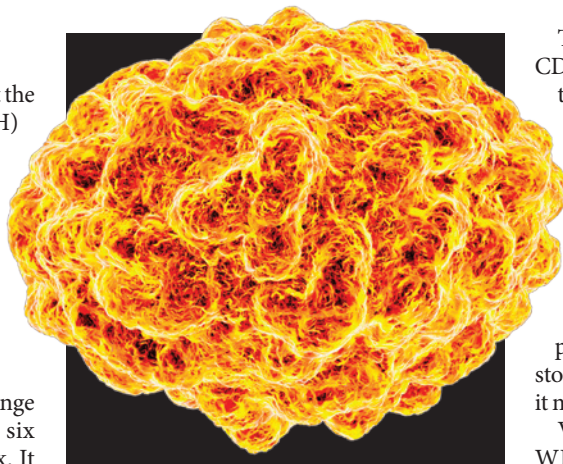
World Health Organization lacks resources to witness destruction of stocks.

BY SARA REARDON

Six vials of smallpox virus discovered at the US National Institutes of Health (NIH) in Bethesda, Maryland, in June are marked for death by autoclave — yet they linger in a high-security freezer. Chalk it up to bureaucracy: an international agreement requires that the World Health Organization (WHO) witnesses the destruction of the samples, but the agency is overwhelmed by the Ebola outbreak in West Africa.

The delay adds another twist to the strange tale of the vials, which sat forgotten for six decades in an unmarked cardboard box. It also mirrors the fierce debate over whether to destroy the two other known stocks of smallpox in the world. That decision is on hold while the WHO determines whether smallpox virus could be synthesized from scratch by those with nefarious intent, and whether virus stocks should therefore be preserved to help develop countermeasures against a terrorist attack.

The NIH smallpox was discovered in an agency 'cold room' used by the US Food and Drug Administration; soon after, the vials were shipped to the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia. That agency, and its Russian counterpart — VECTOR, in Koltsovo — are the only two laboratories authorized to work with smallpox under an international agreement that also calls for the eventual destruction of all stocks. (The



Research stocks of smallpox virus are limited.

virus was eradicated in the wild in 1979.)

The CDC promised to destroy the NIH samples immediately, with WHO officials present. But that has proved more difficult than anticipated. The WHO has dealt with just one other surprise smallpox discovery: viral DNA found in a South African laboratory in 2013. WHO officials were present when that DNA was destroyed in January this year — but no WHO employee is certified to enter the CDC's high-security smallpox lab. This means that a WHO official must fly to Atlanta to witness the destruction of the virus on closed-circuit television. Arranging the trip has been made more difficult by the Ebola crisis, says Alejandro Costa, head of the WHO team in Geneva, Switzerland, that monitors smallpox issues.

The authorized stocks of smallpox at the CDC and VECTOR are also in limbo. For more than 20 years, the US and Russian governments have fought efforts to set a deadline to destroy the samples. They argue that the virus should be preserved to test vaccines and antivirals that might be needed to respond to an accidental smallpox release or a terrorist attack.

In May, the WHO's policy-making body, the World Health Assembly, again put off a decision on the fate of the smallpox stocks. Unusually, the group did not say when it might revisit the issue.

Virologist D. A. Henderson, who led the WHO's programme to eradicate smallpox, argues that all stocks should be destroyed, in a kind of mutual disarmament. "Anyone found with isolates of smallpox virus after point X would be guilty of crimes against humanity," says Henderson, now at the Center for Health Security in Baltimore, Maryland.

At a meeting last week, members of the WHO Advisory Committee on Variola Virus Research agreed to commission a report assessing the risk that a bioterrorist could synthesize smallpox — a situation that might add weight to the argument for keeping the natural virus. Costa says that the WHO will postpone any decision on stock destruction until the report is released, in 2016 at the earliest.

The NIH smallpox is likely to be destroyed well before then: Costa estimates by January 2015. In the meantime, experts say that the virus is safe at the CDC. ■

MEHAU KULYK/SPL CORBIS

MARINE ECOLOGY

Coastal havoc boosts jellies

Five-year Chinese study suggests that human activity made gelatinous outbreaks worse.

BY JANE QIU IN QINGDAO

“It was a truly gelatinous world,” says marine ecologist Sun Song, recalling a jellyfish outbreak last year in the Yellow Sea off China’s eastern coast. “The slimy monsters were everywhere, their long tentacles fluttering ferociously in the rolling waves.” Such blooms have repeatedly choked Chinese waters in the past decade, posing substantial threats to tourism, fisheries and coastal facilities such as chemical plants and nuclear power stations.

Jellyfish blooms have probably struck the world’s seas periodically for hundreds of millions of years. But researchers suspect that human activity has made them worse, with knock-on effects on marine ecosystems and especially on fish, because jellyfish compete with fish for food and feed on their larvae. Now, one of the largest studies of jellyfish blooms so far — involving laboratory testing, ocean surveys and field experiments — has found direct evidence linking the proliferation of jellyfish to human disturbances along the coast, especially the disruption of sea-floor ecosystems.

The five-year, US\$4.5-million initiative started work in 2010 led by Sun, who is director of the Chinese Academy of Sciences’ Institute of Oceanology in the coastal city of Qingdao. Sun’s team, involving eight institutions and more than 100 researchers, presented its latest results earlier this month at the World Conference on Marine Biodiversity, also in Qingdao. China’s waters are not the only ones to have had jellyfish outbreaks in recent years, so the results “could help to tackle similar problems in other parts of the world” such as the Mediterranean or the Caspian Sea, says Mark Costello, a marine ecologist at the University of Auckland in New Zealand.

The survey shows that Chinese coasts are plagued primarily by three jellyfish species: *Aurelia aurita*, *Cyanea nozaki* and the most damaging one, the giant *Nemopilema nomurai*, which can grow to 2 metres in diameter and weigh as much as 200 kilograms.

Jellyfish are most recognizable in the short-lived ‘medusa’ stage of their life cycles, when they swim freely and reproduce sexually. But they also go through a stage in which they are bound to the sea floor as ‘polyps’. These can live for years and reproduce asexually, budding new medusae. Most research so far has focused on the medusa stage, but “the key to jellyfish outbreaks may lie in the polyps”, says Sun.

Polyps are notoriously difficult to find and



A polyp of *Nemopilema nomurai*, 1 millimetre long, will give rise to medusae more than 2 metres across.

collect: they are transparent and easy to break, and can be just micrometres long. “Numerous attempts by scuba diving, underwater robots and dredging up the sediments have led to nothing,” Sun says. So the team turned to the jellyfish in the oceanology institute’s aquarium. The researchers grew polyps on hard surfaces such as glass, plastic or stone in the laboratory and then lowered the slabs onto the seabed at half a dozen coastal locations, to see if specific environmental factors would spur a mini-outbreak.

As Sun reported at the conference, polyps of all three species tended to sprout medusae when sea-bottom temperatures were about 10–15 °C. The blooms were greatly enhanced by extreme temperature swings. Moreover, more medusae formed in waters that had fewer bottom dwellers, or ‘benthos’ — animals such as starfish, sea slugs, shrimp and fish — and greater environmental disturbance. Nutrient overloads from sewage and agricultural run-off, and construction of infrastructure such as ports and bridges, seemed to help the jellyfish to thrive.

The findings point to a link between the gelatinous scourges and coastal disturbances in China caused by overfishing, bottom trawling and oxygen depletion following algal blooms, as well as coastal development and increasingly frequent extreme-weather events, says Costello. Moreover, a decline in both the number and the diversity of benthos in the past 50 years may have also contributed, says Xu Kuidong, a marine ecologist at the Institute of Oceanology. Benthos that feed on polyps “are the key biological control of jellyfish populations”, he says.

The ocean currents that skirt China’s coast can affect benthos by changing temperatures and nutrient concentrations, so a broader approach will be necessary to unveil the true causes of jellyfish outbreaks, says Sun. In April, he and his colleagues launched a five-year, \$165-million project to study ocean currents. By measuring the motion of currents at depths of up to 6,000 metres, and learning how ecosystems respond, the researchers hope to piece together how local and regional factors combine to cause blooms. “This might allow us to forecast — and even prevent — a forthcoming outbreak,” says Sun.

And in May, an international collaboration launched the Jellyfish Database Initiative (JeDI), which will provide global, open-access data on jellyfish abundances and associated environmental factors. Such projects — as well as more-detailed studies of polyps — will attempt to improve understanding of outbreaks and the contribution of natural variations and human factors to changes in jelly populations around the world, says Cathy Lucas, a marine biologist at the University of Southampton, UK, who contributes to JeDI.

Sun says that increasing rates of jellyfish outbreaks in Chinese waters might be an indicator of worsening ecosystem health. Over the past 500 million years, the creatures have survived countless climatic and environmental upheavals — including all five known mass-extinction events, he points out. “When ecosystems deteriorate, they are likely to thrive when others fail.” ■

SUN SONG



TIMOTHY CLARY/AFP/GETTY

Large dogs like this Old English Sheepdog are promising test subjects for longevity studies.

large dogs that typically live for eight to ten years; they would start giving the drug to animals aged six to nine. A pilot trial would involve about 30 dogs, half of which would receive the drug, and would allow the researchers to dose the dogs for a short time and observe effects on heart function and some other health measures. The trial could be completed in as little as three years, but researchers will know long before that — perhaps in months — whether rapamycin improves cardiac function or other aspects of health.

Rapamycin acts on a protein that is involved in cell growth, but little is known about how it extends life. It might retard the ageing process itself or it might prevent age-related diseases. One hypothesis is that it works primarily by preventing the development of cancers.

Pet dogs should provide a more realistic test than lab mice of how the drug would work in humans. Pets experience some of the same environmental influences and get some of the same age-related diseases as their masters, says Kaerberlein. (He plans to enrol his own German shepherd dog when it is old enough.)

Other researchers say that Kaerberlein and Promislow's reasoning makes sense. "We're talking about whether aged pets will benefit, and that's a good model for a human population," says physiological geneticist David Harrison of the Jackson Laboratory in Bar Harbor, Maine, who has studied rapamycin in mice.

Kaerberlein and Promislow have collected about US\$200,000 in University of Washington institutional funds for the pilot. But they will need much more funding for a larger trial of several hundred dogs to test whether the drug, given over years, can extend lifespan — and to study the normal ageing process in thousands of animals to try and understand the mechanism of any life-extending effects.

Participants at the Seattle meeting are also considering whether pet owners could be asked to help fund a bigger trial and if so, whether researchers could promise donors that their dogs would be treated with rapamycin rather than a placebo without compromising the study. Usually, participants in a gold-standard clinical trial are randomly assigned to a treatment or control group.

"It's a kind of unique situation because they are people's pets," says Kaerberlein. "We haven't come to any decision on how that's going to be handled." ■

AGEING

Pet dogs set to test anti-ageing drug

Trials would study extension of lifespan in domestic setting.

BY ERIKA CHECK HAYDEN

Yeast, worms and mice: all have lived longer when treated with various chemical compounds in laboratory tests. But many promising leads have failed when tried in humans. This week, researchers are proposing a different approach to animal testing of life-extending drugs: trials in pet dogs. Their target is rapamycin, which is used clinically as part of an anti-rejection drug cocktail after kidney transplants and which has also been shown to extend the lives of mice by 13% in females and 9% in males (D. E. Harrison *et al. Nature* **460**, 392–395; 2009).

The compound's effect on lifespan has not yet been tested in people — human trials are expensive and it takes a long time to learn whether a drug can extend a human life. Furthermore, rapamycin is no longer patentable, so pharmaceutical companies are unwilling to invest effort in it. The drug can also cause some serious side

effects that might rule it out as a pharmaceutical fountain of youth. It has, for example, been linked to an increased risk of diabetes in people who have had kidney transplants (O. Johnston *et al. J. Am. Soc. Nephrol.* **19**, 1411–1418; 2008). But at low doses, researchers suspect that the drug will not be a problem for healthy dogs.

"Any compound has side effects of some sort; the question is what kind of side effects one can live with," says Randy Strong, a gerontologist at the Barshop Institute for Longevity and Aging Studies in San Antonio, Texas.

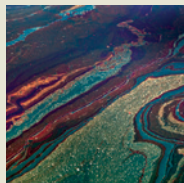
To answer that question, molecular biologists Matthew Kaerberlein and Daniel Promislow at the University of Washington in Seattle propose to give low doses of rapamycin to dogs in a study that would also test whether the drug can extend the animals' lives. The pair invited experts in canine health and the biology of ageing to a meeting in Seattle on 28–29 October to discuss how to structure such a trial.

The researchers hope to test rapamycin in

ARI COONROUGH/
BLOOMBERG VIA GETTY


**MORE
ONLINE**

VIDEO OF THE WEEK



Oil from Deep Horizon spill settled on more of the sea floor than expected
go.nature.com/urmycb

MORE NEWS

- 'Many worlds' interact to generate quantum mechanical effects
go.nature.com/qj31qf
- Ferns reach consensus to choose sex
go.nature.com/oxzilj
- Cancerous cells can 'infect' healthy ones
go.nature.com/33vtk5

NATURE PODCAST



Top 100 papers of all time; capturing asteroids; the first science writer; and a special on the 'hobbit'
nature.com/nature/podcast

BIOMEDICINE

China opens translational medicine centre in Shanghai

First of five linked institutes aims to capitalize on basic-research investments.

BY DAVID CYRANOSKI

Over the past decade, China has poured money into research, especially in the biomedical sciences. But as the nation's health-care costs have risen in the past few years, critics have argued that the investment has not paid off. A group of researchers and government officials now hopes to improve those returns with the official opening this month of the National Centre for Translational Medicine in Shanghai. The 1-billion-renminbi (US\$163-million) centre, slated to complete construction in 2017, is the first of five institutions meant to bridge the gap between basic research and clinical application by putting researchers, doctors and patients under one roof.

China's biologists have made impressive progress in fields such as genome sequencing and protein-structure analysis, but have produced little in terms of drugs or other medical products. "Bloggers and others are always complaining that China is just burning money," says Xiao-Fan Wang, a cancer researcher at Duke University in Durham, North Carolina, and one of the 21 people on the centre's international consulting committee. Some even ask why China, which has made progress in industry largely by copying other countries, has not succeeded by following the same strategy in the biomedical sciences, says Wang.

Wang says that doctors in China are overworked, often racing through whole days of back-to-back patient visits and procedures. And yet, because most Chinese hospitals are part of universities, the doctors must publish to get promoted. In this competitive environment, they often refuse to share data, but are rarely able to do thorough research themselves. "You can't expect them to do eight hours of surgery and then jump into a lab coat," says Wang. The Shanghai centre will change that, he says, by giving clinicians time to focus on research questions.

Saijuan Chen, the centre's director and a geneticist at Shanghai Jiaotong University, says that the centre began interviews this month to recruit some 50 principal investigators and 12 scientists to direct research in their disciplines. She says that the institution's focus will be on developing treatments for heart disease, stroke, metabolic diseases and cancer.

The centre's international consulting committee will help to recruit top talent in these



A major push is under way in China to develop new drugs and surgical procedures.

fields. Tak Mak, a committee member and an immunologist at the University of Toronto in Canada, says that the strategy shows a commitment to hiring on the basis of expertise rather than connections. It is "an effort to break out of the old system — which is probably the system in more places than you'd care to know", he says.

FINE BALANCE

One challenge will be attracting Chinese biomedical researchers who are working abroad. Wang says that China has been able to lure basic-research scientists back to the country with competitive pay. But clinicians in China receive a fraction of the salaries they can get in the United States.

Managing the centre may also be difficult, says committee member Sujuan Ba, chief operating officer of the US National Foundation for Cancer Research in Bethesda, Maryland. The centre's governing council comprises members from ten government agencies and institutions. "That shows the wide scope of support from China, but it is going to be a huge challenge for the leadership team to balance and to manage each council member's requisites," she says. "It is very important for the centre to steer away from bureaucratic red tape and stay focused on its mission and long-term vision of conducting high-impact translational research."

The 54,000-square-metre site will have 300 beds for patients and study volunteers.

It will also run a biobank that will collect hundreds of thousands of patient samples, and an 'omics' centre that will conduct high-throughput genome analyses and gather data on proteins and metabolic products.

Haematologist Zhu Chen of Shanghai Jiao Tong University, who is China's former health minister and chairman of the centre's scientific advisory board, hopes that the centre will emulate the spirit of St Jude Children's Research Hospital in Memphis, Tennessee, with its close relationship between clinicians and basic researchers. He also emphasizes the need to make therapeutic trials free: in China, participants are often charged for treatment. Zhu and Saijuan Chen, who are married, led one of the few successful translational research projects in China — treatment of a form of leukaemia using retinoic acid and arsenic trioxide.

The Shanghai centre will eventually have four sister institutions: a geriatrics centre at the People's Liberation Army general hospital in Beijing, a centre for rare and refractory diseases at the Peking Union Medical College in Beijing, a molecular-medicine research centre at the Fourth Military Medical University in Xian and a regenerative-medicine centre at the West China Hospital in Chengdu.

"These centres are at the historic moment to make a huge impact for China's drug development," says Ba. "We should be able to see the signs of success in five years, if not earlier." ■

Geneticists tap human knockouts

Sequenced genomes reveal mutations that disable single genes and can point to new drugs.

BY EWEN CALLAWAY

For decades, biologists have studied gene function by inactivating the gene in question in mice and other lab animals, and then observing how it affects the organism. Now researchers studying such gene ‘knockouts’ have another, ideal model at their disposal: humans.

The approach does not involve genetically engineering mutant people in the lab, as is done in mice. Instead, researchers scan the genomes of thousands or millions of people, looking for naturally occurring mutations that inactivate a particular gene. By observing how these mutations affect health, researchers hope to gain insight into basic biology and to unearth new disease treatments.

Geneticists discussed several such large-scale efforts during a packed session at the American Society of Human Genetics meeting in San Diego, California, last week. “So much of what we know is based on mice and rats, and not humans,” says Daniel MacArthur, a genomicist at Massachusetts General Hospital in Boston, whose team identified around 200,000 variations that naturally knocked out genes by trawling the protein-coding portion of the genome, or exome, in more than 90,000 people. “Now we can find people who actually have a particular gene inactivated or somehow modified, and that allows us to test hypotheses directly.”

On average, every person carries mutations that inactivate at least one copy of 200 or so genes and both copies of around 20 genes. However, knockout mutations in any particular gene are rare, so very large populations are needed to study their effects. These ‘loss of function’ mutations have long been implicated in certain debilitating diseases, such as cystic fibrosis. Most, however, seem to be harmless — and some are even beneficial to the persons carrying them. “These are people we’re not going to find in a clinic, but they’re still really informative in biology,” says MacArthur.

His group and others had been focusing on genome data, but they are now also starting to mine patient-health records to determine the — sometimes subtle — effects of the mutations. In a study of more than 36,000 Finnish people, published in July (E. T. Lim *et al.* *PLoS Genet.* **10**, e1004494; 2014), a team that included MacArthur discovered that people lacking a gene called *LPA* might be protected from heart disease, and that another knockout mutation, carried in one copy of a gene by up to 2.4% of Finns, may cause fetuses to miscarry if it is present in both copies.

Bing Yu of the University of Texas Health



Finland offers geneticists a rich seam of variation.

Science Center in Houston told the meeting how she and her collaborators had compared knockout mutations found in more than 1,300 people with measurements of around 300 molecules in their blood. The team found that mutations in one gene, called *SLCO1B1*, were linked to high levels of fatty acids, a known risk factor for heart failure. And a team from the Wellcome Trust Sanger Institute in Hinxton, UK, reported that 43 genes whose inactivation is lethal to mice were found to be inactivated in humans who are alive and apparently well.

POSTER CHILD

Following up on such insights will help researchers to unpick the functions of the thousands of human genes about which little or nothing is known, say MacArthur and others. It might even aid drug discovery by identifying genes or biological pathways that could protect against disease.

The poster child for human-knockout efforts is a new class of drugs that block a gene known as *PCSK9* (see *Nature* **496**, 152–155; 2013). The gene was discovered in French families with extremely high cholesterol levels in the early 2000s. But researchers soon found that people with rare mutations that inactivate one copy of *PCSK9* have low cholesterol and rarely develop heart disease. The first *PCSK9*-blocking drugs should hit pharmacies next year, with manufacturers jostling for a share of a market that could reach US\$25 billion in five years.

“I think there are hundreds more stories like *PCSK9* out there, maybe even thousands,”

in which a drug can mimic an advantageous loss-of-function mutation, says Eric Topol, director of the Scripps Translational Science Institute in La Jolla, California. Mark Gerstein, a bioinformatician at Yale University in New Haven, Connecticut, predicts that human knockouts will be especially useful for identifying drugs that treat diseases of ageing. “You could imagine there’s a gene that is beneficial to you as a 25-year-old, but the thing is not doing a good job for you when you’re 75.”

The human-knockout data presented last week will also make it easier to interpret the growing number of genomes being sequenced from people seeking treatment, says Nazneen Rahman, a medical geneticist at the Institute for Cancer Research in London. Rahman told the meeting about her team’s analysis of knockout mutations in 1,000 British people. “The take-home message is that these types of mutations occur much more commonly than people thought,” she says.

When a person’s genome is sequenced to identify the cause of a mysterious illness, any knockout mutations that turn up are obvious suspects. Having a complete list of such mutations and their health effects (or lack thereof) should help to identify the true causes of a disease, Rahman says.

To that end, MacArthur’s team last week released knockout and other data from some 63,000 people — and others are already making use of this trove. “One of the first things I did when they released the data was to look at all my favourite genes for very severe diseases, to see if there are people in those databases — and there are,” says John Belmont, a medical geneticist at Baylor College of Medicine in Houston. He found 11 people with mutations associated with Marfan syndrome, a disorder that affects connective tissue and that can cause heart problems if left untreated.

However, these people may be silent cases, they might somehow be protected against the disease or their genomes might have been sequenced incorrectly. People who carry disease-causing mutations but don’t get sick are especially intriguing, he says. “We should pay attention to them,” Belmont says. “They may hold a key for new treatments.” ■

CORRECTION

The News Feature ‘The ethics squad’ (*Nature* **514**, 418–420; 2014) misspelled Susan Kornetsky’s name.

THE TOP 100 PAPERS

Nature explores the most-cited research of all time.

BY RICHARD VAN NOORDEN,
BRENDAN MAHER AND REGINA NUZZO

The discovery of high-temperature superconductors, the determination of DNA's double-helix structure, the first observations that the expansion of the Universe is accelerating — all of these breakthroughs won Nobel prizes and international acclaim. Yet none of the papers that announced them comes anywhere close to ranking among the 100 most highly cited papers of all time.

Citations, in which one paper refers to earlier works, are the standard means by which authors acknowledge the source of their methods, ideas and findings, and are often used as a rough measure of a paper's importance. Fifty years ago, Eugene Garfield published the Science Citation Index (SCI), the first systematic effort to track citations in the scientific literature. To mark the anniversary, *Nature* asked Thomson Reuters, which now owns the SCI, to list the 100 most highly cited papers of all time. (See the full list at www.nature.com/top100.) The search covered all of Thomson Reuters' Web of Science, an online version of the SCI that also includes databases covering the social sciences, arts and humanities, conference proceedings and some books. It lists papers published from 1900 to the present day.

The exercise revealed some surprises, not least that it takes a staggering 12,119 citations to rank in the top 100 — and that many of the world's most famous papers do not make the cut. A few that do, such as the first observation¹

of carbon nanotubes (number 36) are indeed classic discoveries. But the vast majority describe experimental methods or software that have become essential in their fields.

The most cited work in history, for example, is a 1951 paper² describing an assay to determine the amount of protein in a solution. It has now gathered more than 305,000 citations — a recognition that always puzzled its lead author, the late US biochemist Oliver Lowry. "Although I really know it is not a great paper ... I secretly get a kick out of the response," he wrote in 1977.

The colossal size of the scholarly literature means that the top-100 papers are extreme outliers. Thomson Reuter's Web of Science holds some 58 million items. If that corpus were scaled to Mount Kilimanjaro, then the 100 most-cited papers would represent just 1 centimetre at the peak. Only 14,499 papers — roughly a metre and a half's worth — have more than 1,000 citations (see 'The paper mountain'). Meanwhile, the foothills comprise works that have been cited only once, if at all — a group that encompasses roughly half of the items.

Nobody fully understands what distinguishes the sliver at the top from papers that are merely very well known — but researchers' customs explain some of it. Paul Wouters, director of the Centre for Science and Technology Studies in Leiden, the Netherlands, says that many methods papers "become a standard reference that one cites in order to make clear

to other scientists what kind of work one is doing". Another common practice in science ensures that truly foundational discoveries — Einstein's special theory of relativity, for instance — get fewer citations than they might deserve: they are so important that they quickly enter the textbooks or are incorporated into the main text of papers as terms deemed so familiar that they do not need a citation.

Citation counts are riddled with other confounding factors. The volume of citations has increased, for example — yet older papers have had more time to accrue citations. Biologists tend to cite one another's work more frequently than, say, physicists. And not all fields produce the same number of publications. Modern bibliometricians therefore recoil from methods as crude as simply counting citations when they want to measure a paper's value: instead, they prefer to compare counts for papers of similar age, and in comparable fields.

Nor is Thomson Reuters' list the only ranking system available. Google Scholar compiled its own top-100 list for *Nature*. It is based on many more citations because the search engine culls references from a much greater (although poorly characterized) literature base, including from a large range of books. In that list, available at www.nature.com/top100, economics papers have more prominence. Google Scholar's list also features books, which Thomson Reuters did not analyse. But among the science papers, many of the same titles show up.

Yet even with all the caveats, the old-fashioned hall of fame still has value. If nothing else, it serves as a reminder of the nature of scientific knowledge. To make exciting advances, researchers rely on relatively unsung papers to describe experimental methods, databases and software.

Here *Nature* tours some of the key methods that tens of thousands of citations have hoisted to the top of science's Kilimanjaro — essential, but rarely thrust into the limelight.

BIOLOGICAL TECHNIQUES

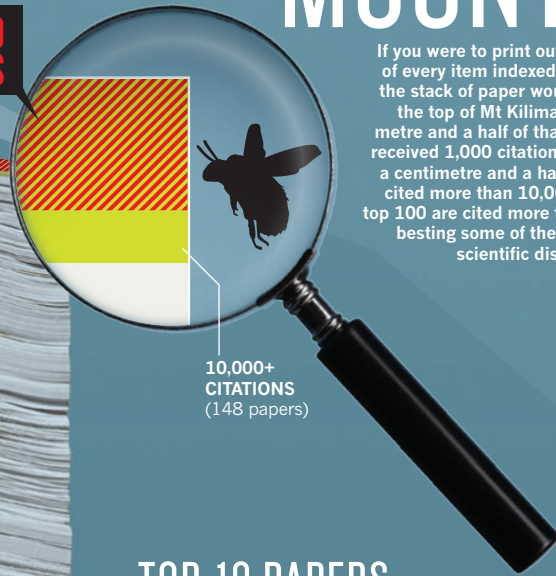
For decades, the top-100 list has been dominated by protein biochemistry. The 1951 paper² describing the Lowry method for quantifying protein remains practically unreachable at number 1, even though many biochemists say that it and the competing Bradford assay³ — described by paper number 3 on the list — are a tad outdated. In between, at number 2, is Laemmli buffer⁴, which is used in a different kind of protein analysis. The dominance of these techniques is attributable to the high volume of citations in cell and molecular biology, where they remain indispensable tools.

At least two of the biological techniques described by top-100 papers have resulted in Nobel prizes. Number 4 on the list describes the DNA-sequencing method⁵ that earned the late Frederick Sanger his share of the 1980 Nobel Prize in Chemistry. Number 63 describes polymerase chain reaction ▶

THE PAPER MOUNTAIN

If you were to print out just the first page of every item indexed in Web of Science, the stack of paper would reach almost to the top of Mt Kilimanjaro. Only the top metre and a half of that stack would have received 1,000 citations or more, and just a centimetre and a half would have been cited more than 10,000 times. All of the top 100 are cited more than 12,000 times, besting some of the most recognizable scientific discoveries in history.

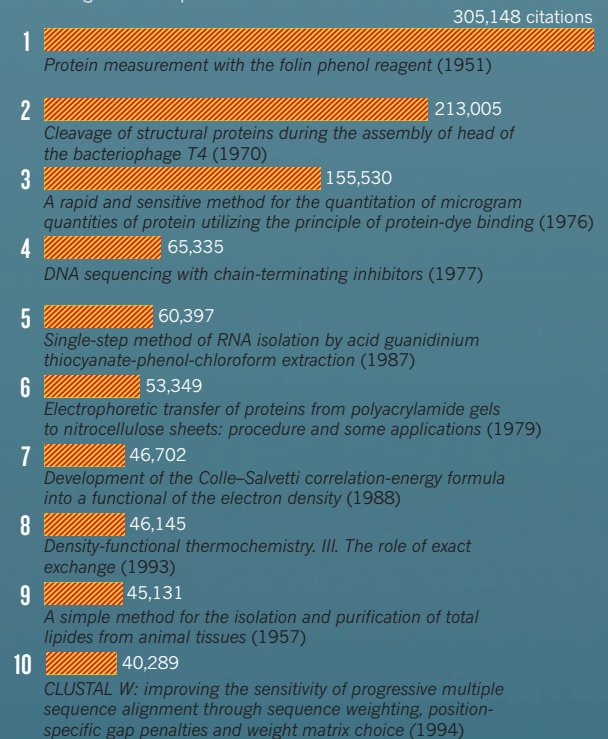
TOP-100 PAPERS



10,000+ CITATIONS
(148 papers)

TOP-10 PAPERS

Just 3 papers have received more than 100,000 citations, putting them well ahead of the rest. These runaway hits all cover biological lab techniques, which in general dominate the list of most-cited literature, including 7 of the top 10.



NATURE.COM

To explore the full list, and to see each paper's citation history over time, visit nature.com/top100

Height
6,000 m

5,500 m

5,000 m

4,500 m

4,000 m

3,500 m

3,000 m

2,500 m

2,000 m

1,500 m

1,000 m

500 m

0 m

100-999 CITATIONS (1,066,046 papers)

10-99 CITATIONS (18,280,005 items)

0 CITATIONS (25,332,701 items)

1,000-9,999 CITATIONS (14,351 items)

Light aircraft cruising altitude 3,000 m

Burj Khalifa 828 m

Eiffel Tower 301 m

Watson and Crick on structure of DNA (1953)
5,207 citations

Farman, Gardiner & Shanklin discover the ozone hole (1985)
1,871 citations

Hirsch proposes the h index to measure scientific productivity (2005)
1,797 citations

Data provided by Thomson Reuters/Web of Science. Individual paper citation figures extracted 7 October 2014. Distribution of citations in database: 19 September 2014

► (PCR)⁶, a method for copying segments of DNA that earned US biochemist Kary Mullis the prize in 1993. By helping scientists to explore and manipulate DNA, both methods have helped to drive a revolution in genetic research that continues to this day.

Other methods have received less public acclaim, but are not without their rewards. In the 1980s, the Italian cancer geneticist Nicoletta Sacchi linked up with Polish molecular biologist Piotr Chomczynski in the United States to publish⁷ a fast, inexpensive way to extract RNA from a biological sample. As it became wildly popular — currently, it is number 5 on the list — Chomczynski patented modifications on the technique and built a business out of selling the reagents. Now at the Roswell Park Cancer Institute in Buffalo, New York, Sacchi says that she received little in the way of monetary rewards, but takes satisfaction from seeing great discoveries built on her work. The technique played a part in the explosive growth in the study of short RNA molecules that do not code for protein, for example. “That is what I would consider, scientifically speaking, a great reward,” she says.

BIOINFORMATICS

The rapid expansion of genetic sequencing since Sanger’s contribution has helped to boost the ranking of papers describing ways to analyse the sequences. A prime example is BLAST (Basic Local Alignment Search Tool), which for two decades has been a household name for biologists wanting to work out what genes and proteins do. Users simply have to open the program in a web browser and plug in a DNA, RNA or protein sequence. Within seconds, they will be shown related sequences from thousands of organisms — along with information about the function of those sequences and even links to relevant literature. So popular is BLAST that versions^{8,9} of the program feature twice on the list, at spots 12 and 14.

But owing to the vagaries of citation habits, BLAST has been bumped down the list by Clustal, a complementary programme for aligning multiple sequences at once. Clustal allows researchers to describe the evolutionary relationships between sequences from different organisms, to find matches among seemingly unrelated sequences and to predict how a change at a specific point in a gene or protein might affect its function. A 1994 paper¹⁰ describing ClustalW, a user-friendly version of the software, is currently number 10 on the list. A 1997 paper¹¹ on a later version called ClustalX is number 28.

The team that developed ClustalW, at the European Molecular Biology Laboratory in Heidelberg, Germany, had created the program to work on a personal computer, rather than a mainframe. But the software was transferred when Julie Thompson, a computer scientist from the private sector, joined the lab in 1991. “It was a program written by biologists;

“ALTHOUGH I
REALLY KNOW
IT’S NOT A
GREAT PAPER,
I SECRETLY
GET A KICK
OUT OF THE
RESPONSE.”

I’m trying to find a nice way to say that,” says Thompson, who is now at the Institute of Genetics and Molecular and Cellular Biology in Strasbourg, France. Thompson rewrote the program to ready it for the volume and complexity of the genome data being generated at the time, while also making it easier to use.

The teams behind BLAST and Clustal are competitive about the ranking of their papers. It is a friendly sort of competition, however, says Des Higgins, a biologist at University College Dublin, and a member of the Clustal team. “BLAST was a game-changer, and they’ve earned every citation that they get.”

PHYLOGENETICS

Another field buoyed by the growth in genome sequencing is phylogenetics, the study of evolutionary relationships between species.

Number 20 on the list is a paper¹² that introduced the “neighbor-joining” method, a fast, efficient way of placing a large number of organisms into a phylogenetic tree according to some measure of evolutionary distance between them, such as genetic variation. It links related organisms together one pair at a time until a tree is resolved. Physical anthropologist Naruya Saitou helped to devise the technique when he joined Masatoshi Nei’s lab at the University of Texas in Houston in the 1980s to work on human evolution and molecular genetics, two fields that were starting to burst at the seams with information.

“We physical anthropologists were facing kind of the big data of that time,” says Saitou, now at Japan’s National Institute of Genetics in Mishima. The technique made it possible to devise trees from large data sets without eating up computer resources. (And, in a nice cross-fertilization within the top-100, Clustal’s algorithms use the same strategy.)

Number 41 on the list is a description¹³ of how to apply statistics to phylogenies. In 1984,

evolutionary biologist Joe Felsenstein of the University of Washington in Seattle adapted a statistical tool known as the bootstrap to infer the accuracy of different parts of an evolutionary tree. The bootstrap involves resampling data from a set many times over, then using the variation in the resulting estimates to determine the confidence for individual branches. Although the paper was slow to amass citations, it rapidly grew in popularity in the 1990s and 2000s as molecular biologists recognized the need to attach such intervals to their predictions.

Felsenstein says that the concept of the bootstrap¹⁴, devised in 1979 by Bradley Efron, a statistician at Stanford University in California, was much more fundamental than his work. But applying the method to a biological problem means it is cited by a much larger pool of researchers. His high citation count is also a consequence of how busy he was at the time, he says: he crammed everything into one paper rather than publishing multiple papers on the topic, which might have diluted the number of citations each one received. “I was unable to go off and write four more papers on the same thing,” he says. “I was too swamped to do that, not too principled.”

STATISTICS

Although the top-100 list has a rich seam of papers on statistics, says Stephen Stigler, a statistician at the University of Chicago in Illinois and an expert on the history of the field, “these papers are not at all those that have been most important to us statisticians”. Rather, they are the ones that have proved to be most useful to the vastly larger population of practising scientists.

Much of this crossover success stems from the ever-expanding stream of data coming out of biomedical labs. For example, the most frequently cited statistics paper (number 11) is a 1958 publication¹⁵ by US statisticians Edward Kaplan and Paul Meier that helps researchers to find survival patterns for a population, such as participants in clinical trials. That introduced what is now known as the Kaplan–Meier estimate. The second (number 24) was British statistician David Cox’s 1972 paper¹⁶ that expanded these survival analyses to include factors such as gender and age.

The Kaplan–Meier paper was a sleeper hit, receiving almost no citations until computing power boomed in the 1970s, making the methods accessible to non-specialists. Simplicity and ease of use also boosted the popularity of papers in this field. British statisticians Martin Bland and Douglas Altman made the list (number 29) with a technique¹⁷ — now known as the Bland–Altman plot — for visualizing how well two measurement methods agree. The same idea had been introduced by another statistician 14 years earlier, but Bland and Altman presented it in an accessible way that has won citations ever since.

The oldest and youngest papers in

the statistics group deal with the same problem — multiple comparisons of data — but from very different scientific milieux. US statistician David Duncan's 1955 paper¹⁸ (number 64) is useful when a few groups need to be compared. But at number 59, Israeli statisticians Yoav Benjamini and Yoşef Hochberg's 1995 paper¹⁹ on controlling the false-discovery rate is ideally suited for data coming from fields such as genomics or neuroscience imaging, in which comparisons number in the hundreds of thousands — a scale that Duncan could hardly have imagined. As Efron observes: "The story is one of the computer slowly, then not so slowly, making its influence felt on statistical theory as well as on practice."

DENSITY FUNCTIONAL THEORY

When theorists want to model a piece of matter — be it a drug molecule or a slab of metal — they often use software to calculate the behaviour of the material's electrons. From this knowledge flows an understanding of numerous other properties: a protein's reactivity, for instance, or how easily Earth's liquid iron outer core conducts heat.

Most of this software is built on density functional theory (DFT), easily the most heavily cited concept in the physical sciences. Twelve papers on the top-100 list relate to it, including 2 of the top 10. At its heart, DFT is an approximation that makes impossible mathematics easy, says Feliciano Giustino, a materials physicist at the University of Oxford, UK. To study electronic behaviour in a silicon crystal by taking account of how every electron and every nucleus interacts with every other electron and nucleus, a researcher would need to analyse one sextillion (10^{21}) terabytes of data, he says — far beyond the capacity of any conceivable computer. DFT reduces the data requirement to just a few hundred kilobytes, well within the capacity of a standard laptop.

Theoretical physicist Walter Kohn led the development of DFT half a century ago in papers^{20,21} that now rank as numbers 34 and 39. Kohn realized that he could calculate a system's properties, such as its lowest energy state, by assuming that each electron reacts to all the others not as individuals, but as a smeared-out average. In principle, the mathematics are straightforward: the system behaves like a continuous fluid with a density that varies from point to point. Hence the theory's name.

But a few decades passed before researchers found ways to implement the idea for real materials, says Giustino. Two^{22,23} top-100 papers are technical recipes on which the most popular DFT methods and software packages are built. One (number 8) is by Axel Becke, a theoretical chemist at Dalhousie University in Halifax, Canada, and the other (number 7) is by US-based theoretical chemists Chengteh Lee, Weitao Yang and Robert Parr. In 1992, computational chemist John Pople (who would share the 1998 Nobel prize with

Kohn) included a form of DFT in his popular Gaussian software package.

Software users probably cite the original theoretical papers even if they do not fully understand the theory, says Becke. "The theory, mathematics and computer software are specialized and are the concern of quantum physicists and chemists," he says. "But the applications are endless. At a fundamental level, DFT can be used to describe all of chemistry, biochemistry, biology, nanosystems and materials. Everything in our terrestrial world depends on the motions of electrons — therefore, DFT literally underlies everything."

CRYSTALLOGRAPHY

George Sheldrick, a chemist at the University of Göttingen in Germany, began to write software to help solve crystal structures in the 1970s. In those days, he says, "you couldn't get grant money for that kind of project. My job was to teach chemistry, and I wrote the programs as a hobby in my spare time." But over 40 years, his work gave rise to the regularly updated SHELX suite of computer programs, which has become one of the most popular tools for analysing the scattering patterns of X-rays that are shot through a crystal — thereby revealing the atomic structure.

The extent of that popularity became apparent after 2008, when Sheldrick published a review paper²⁴ about the history of the system, and noted that it might serve as a general literature citation whenever any of the SHELX programs were used. Readers followed his advice. In the past 6 years, that review paper has amassed almost 38,000 citations, catapulting it to number 13 and making it the highest-ranked paper published in the past two decades.

The top-100 list is scattered with other tools essential to crystallography and structural biology. These include papers describing the HKL suite²⁵ (number 23) for analysing X-ray diffraction data; the PROCHECK programs²⁶ (number 71) used to analyse whether a proposed protein structure seems geometrically normal or outlandish; and two programs^{27,28} used to sketch molecular structures (numbers 82 and 95). These tools are the "bricks and mortar" for determining crystal structures, says Philip Bourne, associate director for data science at the US National Institutes of Health in Bethesda, Maryland.

An unusual entry, appearing at number 22, is a 1976 paper²⁹ from Robert Shannon — a researcher at the giant chemical firm DuPont in Wilmington, Delaware, who compiled a comprehensive list of the radii of ions in a series of different materials. Robin Grimes, a materials scientist at Imperial College London, says that physicists, chemists and theorists still cite this paper when they look up values of ionic size, which often

correlate neatly with other properties of a substance. This has made it the highest formally-cited database of all time.

"We often cite these kinds of papers almost without thinking about it," says Paul Fossati, one of Grimes's research colleagues. The same could be said for many of the methods and databases in the top 100. The list reveals just how powerfully research has been affected by computation and the analysis of large data sets. But it also serves as a reminder that the position of any particular methods paper or database at the top of the citation charts is also down to luck and circumstance.

Still, there is one powerful lesson for researchers, notes Peter Moore, a chemist at Yale University in New Haven, Connecticut. "If citations are what you want," he says, "devising a method that makes it possible for people to do the experiments they want at all, or more easily, will get you a lot further than, say, discovering the secret of the Universe". ■

Richard Van Noorden is a reporter for Nature based in London, Brendan Maher is an editor for Nature based in New York, and Regina Nuzzo is a writer based in Washington DC.

1. Iijima, S. *Nature* **354**, 56–58 (1991).
2. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. Biol. Chem.* **193**, 265–275 (1951).
3. Bradford, M. M. *Anal. Biochem.* **72**, 248–254 (1976).
4. Laemmli, U. K. *Nature* **227**, 680–685 (1970).
5. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. Natl Acad. Sci. USA* **74**, 5463–5467 (1977).
6. Saiki, R. K. et al. *Science* **239**, 487–491 (1988).
7. Chomczynski, P. & Sacchi, N. *Anal. Biochem.* **162**, 156–159 (1987).
8. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. *J. Mol. Biol.* **215**, 403–410 (1990).
9. Altschul, S. F. et al. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
10. Thompson, J. D., Higgins, D. G. & Gibson, T. J. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
11. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. *Nucleic Acids Res.* **25**, 4876–4882 (1997).
12. Saitou, N. & Nei, M. *Mol. Biol. Evol.* **4**, 406–425 (1987).
13. Felsenstein, J. *Evolution* **39**, 783–791 (1985).
14. Efron, B. *Ann. Statist.* **7**, 1–26 (1979).
15. Kaplan, E. L. & Meier, P. J. *Am. Stat. Assoc.* **53**, 457–481 (1958).
16. Cox, D. R. *J. R. Stat. Soc. B* **34**, 187–220 (1972).
17. Bland, J. M. & Altman, D. G. *Lancet* **327**, 307–310 (1986).
18. Duncan, D. B. *Biometrics* **11**, 1–42 (1955).
19. Benjamini, Y. & Hochberg, Y. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
20. Kohn, W. & Sham, L. J. *Phys. Rev.* **140**, A1133 (1965).
21. Hohenberg, P. & Kohn, W. *Phys. Rev. B* **136**, B864 (1964).
22. Becke, A. D. *J. Chem. Phys.* **98**, 5648–5652 (1993).
23. Lee, C., Yang, W. & Parr, R. G. *Phys. Rev. B* **37**, 785–789 (1988).
24. Sheldrick, G. M. *Acta Crystallogr. A* **64**, 112–122 (2008).
25. Otwinowski, Z. & Minor, W. *Method. Enzymol. A* **276**, 307–326 (1997).
26. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
27. Kraulis, P. J. *J. Appl. Crystallogr.* **24**, 946–950 (1991).
28. Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr. A* **47**, 110–119 (1991).
29. Shannon, R. D. *Acta Crystallogr. A* **32**, 751–767 (1976).

NATURE.COM
For more analysis of
the top 100, see:
nature.com/top100



THE EBOLA QUESTIONS

Scientists know a lot about the virus that causes Ebola — but there are many puzzles that they have yet to solve.

BY ERIKA CHECK HAYDEN

To much of the world, the virus behind the devastating Ebola outbreak in Africa seems to have stormed out of nowhere. But Leslie Lobel thinks we should have seen it coming.

In 2012, Lobel and a team of researchers spent six months in Uganda studying the Ebola virus and related viruses. Over the course of their stay, these pathogens caused at least four separate outbreaks of disease in central Africa, affecting more than 100 people. To Lobel, a virologist at Ben-Gurion University of the Negev in Beer-Sheva, Israel, the outbreaks felt like the small tremors that can precede a major earthquake. “We all said, something is going on here; something big is going to happen,” he says.

Like Lobel, other scientists have predicted that these viruses would one day cause a major epidemic — and the current outbreak, which

has so far killed nearly 5,000 people, has proved them right. There are five species of closely related viruses that scientists refer to as ‘ebolaviruses’; the species behind the current outbreak, *Zaire ebolavirus*, is more generally known as ‘the Ebola virus.’ Along with Marburg virus and Lloviu virus, the ebolaviruses make up the filoviruses, a family that was unknown before the 1960s. All of the filoviruses share a common structure, and most of them cause life-threatening haemorrhagic fevers in humans.

Research on these once-ignored viruses took off after the 2001 anthrax attacks in the United States, which prompted officials to sink money into investigating lethal pathogens that might be used in bio-terror attacks and to build dedicated laboratories where they can be

An Ebola virus particle from the 2014 outbreak.

safely studied. Scientists have learned how these viruses work and have created the first experimental vaccines and therapies that might stop them. “The biodefence funding has been huge,” says microbiologist Thomas Geisbert of the University of Texas Medical Branch at Galveston, who has studied ebolaviruses for 26 years.

But the advances in knowledge about filoviruses have also exposed the gaps. Scientists suspect that more members of the filovirus family remain to be discovered and in other parts of the world. They are also working to understand which animals naturally harbour filoviruses and why human filovirus outbreaks seem to be rising in frequency: they have occurred in 19 of the past 21 years, and three times this year alone. Finding answers is difficult because outbreaks are unpredictable and laboratory work with filoviruses requires the highest security measures.

In the past few months, research has necessarily taken a back seat to efforts to control the Ebola outbreak — but as the epidemic escalates, science is coming to the fore. Researchers are recognizing that they might be able to stop this Ebola virus only if they understand its biology and how to control it. “We need a lot more information about the virology, the clinical presentation and the epidemiology of this virus,” says Michael Osterholm, a public-health scientist at the University of Minnesota’s Center for Infectious Disease Research and Policy in Minneapolis. “Nobody underestimates the difficulty of doing that research in these settings, but it is really important to get this information.” With this in mind, *Nature* asked leading researchers to discuss the most urgent scientific questions about the Ebola virus and other filoviruses — the questions that, if answered, might prevent another disastrous outbreak, or help to contain the current one. This is what they said.

WHERE DO FILOVIRUSES COME FROM?

In July 2007, a miner who had been prospecting for lead and gold in a Ugandan cave became infected with Marburg virus. Officials closed the cave, and a team of researchers led by the US Centers for Disease Control and Prevention (CDC) swept in to investigate. They hoped to answer a decades-old question: what animal is the natural host for filoviruses? The mystery had lingered since 1967, when Marburg virus — the first filovirus to be discovered — sickened European lab workers who had handled imported monkeys. The high lethality of filovirus infection in monkeys, humans and other apes suggested that primates were not the natural hosts: if a virus kills too many of its hosts, then it cannot propagate and dies out. There were clues that bats might be the ‘reservoir’ species. But to prove it, scientists needed to find an infected bat.

The researchers captured some 1,300 bats roosting in the cave and tested their blood for Marburg virus¹. They finally found what they were looking for: infectious Marburg viruses isolated from five Egyptian fruit bats, none of which showed symptoms of disease. The team also found more infected bats in a nearby cave that had been linked to a previous Marburg virus outbreak.

It is not entirely clear how the virus is transmitted from bats to people, although the most likely route is through contact with bodily fluids. Bats infected with Marburg virus in the lab shed the virus in their mouths, so wild bats might spread it by leaving traces on fruit that is later eaten by other animals². Knowing the host species for the other filoviruses is crucial. “Until we understand what that reservoir is, it is difficult to limit your encroachment on that species,” says virologist John Dye of the US Army Medical Research Institute of Infectious Diseases in Fort Detrick, Maryland.

Researchers now strongly suspect that bats are the natural host for ebolaviruses, too. In 1976, during one of the first known Ebola outbreaks, the six people who were initially infected worked in a factory room in Sudan that was home to roosting bats³. Researchers have since isolated antibodies to ebolaviruses from bats, as well as snippets of genetic material from the viruses. But proving that bats are the reservoir has been maddeningly difficult — no one has isolated an infectious ebolavirus from a wild bat, and it has been difficult to trace rare and sporadic outbreaks back to a source. The ebolavirus outbreaks have originated in

many locations, only sometimes among people or animals who have had contact with bats. “You’re pretty much looking at the entire tropical forest,” says Jonathan Towner, a molecular virologist at the CDC in Atlanta, Georgia, who trapped bats in Uganda for the Marburg investigations.

The current outbreak is thought to have begun in southeastern Guinea in December 2013, when a two-year-old boy died of a mysterious illness that quickly spread to family members and health-care workers. So far, the response to the outbreak has been much more focused on containing it than tracing its source. “The public-health crisis is unprecedented,” Towner says. “There’s no room for an ecology investigation at this point.”

HOW WIDESPREAD ARE FILOVIRUSES?

Filoviruses are not just found in bats and primates. That became clear in 2008, when Philippine officials called for help to investigate an outbreak of disease in pigs. When researchers arrived, they found that the pigs were infected with *Reston ebolavirus*⁴ — a species that was first discovered in monkeys imported to the United States from the Philippines in 1989. The pig discovery was a shock because until then, no ebolavirus had been known to naturally infect a farm animal. And it was not a freak event: *Reston ebolavirus* was also reported⁵ in pigs in China in 2012. However, the Reston virus seems to be relatively harmless to humans. People working on the Philippine pig farms harboured antibodies to it — a sign that the pigs had transmitted the virus to them — but no humans fell ill.

In 2011, scientists confirmed that pigs can also become infected with *Zaire ebolavirus*⁶. The concern now is that pigs could serve as a mixing vessel for filoviruses. Pigs could be simultaneously infected with several filoviruses, which might exchange genetic material to eventually produce new versions that are capable of sickening humans. “The relevant practical question is, do we need to worry about Reston? If it is truly not pathogenic in humans, is there a potential that it could change?” says Erica Ollmann Saphire, a structural biologist who studies filoviruses at the Scripps Research Institute in La Jolla, California.

“We all said, something is going on here; something big is going to happen.”

Scientists are probably just beginning to understand the different types of filoviruses and their geographic reach. The list of known filoviruses only recently expanded: the fifth ebolavirus (*Bundibugyo ebolavirus*) was discovered in Uganda in 2007, and the Lloviu virus was identified from dead bats in Spain in 2011. “We may find more members of the filovirus family elsewhere in the world,” says virologist Ayato Takada of Hokkaido University in Sapporo, Japan. The fact that the viruses are more common than was once thought suggests that they have been around for a very long time — perhaps for most of human history. And it could be that scientists are detecting only a small fraction of events in which they cross over from animals to humans. Researchers are now trying to work out how frequently they jump into people and, when they do, how often they cause disease. In 2010, one team reported that as many as 20% of people in some areas of Gabon carry antibodies to *Zaire ebolavirus* in their blood, indicating that they were exposed to the virus in the past without becoming ill⁷.

Lobel says that these data must be taken with a pinch of salt, and that the assays could actually be detecting immune responses to viruses similar to ebolaviruses. In his studies, people known to have survived definite Ebola virus infections had a different immune response from

KNOW YOUR ENEMY

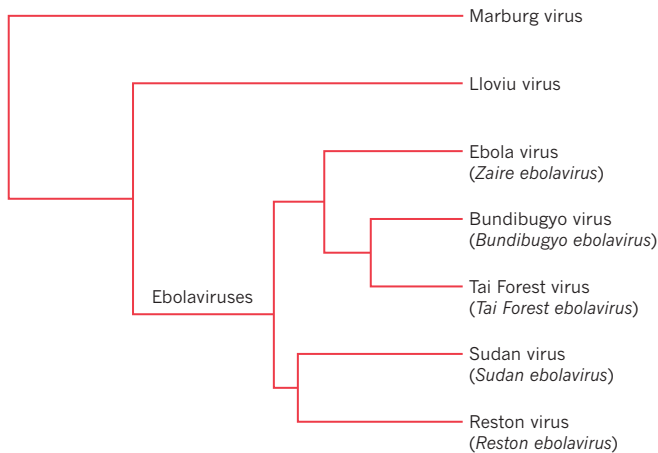
Regions at risk

By mapping risk factors, including the location of fruit bats that might harbour ebolaviruses, scientists estimated the areas in Africa most at risk of an outbreak.

- High risk
- Medium risk
- Low risk

Killers in the family

The filovirus family includes Marburg virus, Lloviu virus and five species of ebolavirus. Some family members have been traced to fruit bats, from which they may pass to humans.



Sierra Leone, Guinea and Liberia — the countries at the epicentre of the current outbreak — are at high risk because many people live in areas inhabited by bats that might carry ebolaviruses.

Before this year, all but one of the previous human Ebola outbreaks started in one of five central African countries: Republic of Congo, South Sudan, Democratic Republic of Congo, Uganda and Gabon.

The mapping exercise revealed 22 countries and a total of 22 million people at risk of an ebolavirus outbreak.

those who have never been infected. He and other researchers are now surveying immune responses among more filovirus survivors.

A pressing question in the present outbreak is how the currently circulating *Zaire ebolavirus* might be changing. Its rapid spread hints that there could be something different about this strain — possibly that it has become easier to pass from one person to another. “We have to start investigating whether the 2014 strain behaves like previous outbreak strains, or whether it could potentially be more transmissible,” says Kristian Andersen, a virologist at the Broad Institute in Cambridge, Massachusetts.

Andersen is careful to note that there are no data to suggest this yet. Although a handful of researchers have raised fears that the virus might mutate into an airborne form, most think it is unlikely that the virus could change so dramatically in such a short time. Genetic analyses have shown that this strain of *Zaire ebolavirus* has mutated hundreds of times since it diverged from an ancestral virus about ten years ago⁸ (see *Nature* <http://doi.org/vsd; 2014>), but no one yet knows whether any of these mutations have altered important properties of the virus. Instead, its unprecedented spread is thought to be due to the fact that it emerged in an area of Africa where people were unfamiliar with the virus and how to control it, allowing it to escape to urban centres. Researchers are now carrying out more-detailed studies of the virus to better understand its origins and characteristics.

ARE WE MAKING EBOLA OUR OWN WORST ENEMY?

In September, epidemiologists published an analysis in which they mapped the locations of all ebolavirus outbreaks in Africa along with the known ranges of three bat species that are candidate reservoirs for the viruses⁹. They also plotted changes in African populations and mobility — for instance, the proportion of each country that lives in rural

and urban areas. The team wanted to pinpoint the areas that might be at highest risk of future outbreaks.

Before this year, all but one of the human ebolavirus outbreaks could be traced back to central Africa, and the *Zaire ebolavirus* species had never been seen in West Africa. But “we shouldn’t have been massively surprised that it was there”, says epidemiologist Simon Hay of the University of Oxford, UK, who led the analysis. Even when he and his team did not include data from the current outbreak, they predicted that the three hardest-hit countries — Sierra Leone, Guinea and Liberia — would be at high risk for an ebolavirus outbreak because they have large numbers of people living in areas inhabited by the bats. In all, the analysis highlights 22 African countries where ebolavirus outbreaks are likely to start, putting 22 million people at risk (see ‘Know your enemy’).

The study also begins to explain why filovirus outbreaks seem to be growing more frequent, widespread and larger in size. Human populations in countries that are likely to harbour filoviruses have nearly tripled since the viruses were first discovered, and flight traffic has increased by one-third since 2005. The viruses are not coming to us; instead, we are encroaching on the viruses, as population growth and increasing travel put humans in contact with viral hosts, and then people unwittingly transport the viruses around the world. “There’s this perception that these outbreaks occur in completely isolated, remote parts of Africa where they burn out before they hit major population centres,” Hay says. “Obviously we’ve seen that this is not the case with the latest outbreak.” His team has released its data¹⁰ and

➔ **NATURE.COM**
For more on the spread of the 2014 Ebola outbreak, see: go.nature.com/m55ual

hopes that others will now use the information to look for more specific environmental factors — combinations of climate and geography that might pinpoint the precise places where future outbreaks are most likely to occur.

WHY IS EBOLA SO LETHAL?

Ebola virus is one of the most lethal viruses known. In the current outbreak, an estimated 60–70% of those infected have died, and in previous outbreaks the figure has reached almost 90%. (Only rabies, smallpox and a handful of other viruses are as fatal if left untreated.) The reason that Ebola virus and other filoviruses are so lethal is because they turn the body's own defences against it.

Normally when a virus invades the body, it triggers cells in the 'innate' arm of the immune system, which cause inflammation and other reactions to fight off the infection. The Ebola virus, however, infects and cripples innate immune cells, taking out this first line of defence. These dying cells also trigger a destructive flood of chemicals, called a cytokine storm, and cause the downstream death of cells that normally make protective antibodies.

Other highly pathogenic viruses also trigger cytokine storms, but filoviruses are thought to be particularly lethal because they affect a wide array of tissues. As well as the immune system, Ebola virus attacks the spleen and kidneys, where it destroys cells that help the body to regulate its fluid and chemical balance and that make proteins that help the blood to clot. At its worst, Ebola virus causes the liver, lungs and kidneys to shut down, other organs to fail and the blood vessels to leak fluid into surrounding tissues. Too often, this ends in death.

If scientists can understand how the immune systems of survivors were able to fight off the virus, they might be able to encourage this form of defence through a vaccine. Researchers have found that survivors of previous outbreaks managed to make antibodies to Ebola virus, avoid the cytokine storm and preserve their immune cells during the course of infection. But why they can do this, and others cannot, is a mystery. "The question in our mind is, how did they survive?" Lobel says.

The right treatment can raise the chance of survival. In the current outbreak, people infected with Ebola virus who are treated in developed countries have been more likely to survive than patients treated in Africa, because they receive much more intensive care. There is no targeted treatment for filoviruses, but doctors can closely monitor and correct blood chemistry and protein imbalances caused by organ failure and fluid loss — with intravenous drips or kidney dialysis, for example. "Right now, the Ebola virus has such a high lethality because of the level of care provided in the places where it is occurring," says Dye. "If health care gets better, the lethality rates will decrease."

Sadly, Osterholm says, even the most basic interventions that could help people infected with Ebola virus are not being provided right now in the hardest-hit areas. For instance, oral rehydration therapy is being widely used in place of intravenous fluid replacement because of the fear that health-care workers could be infected in the process of inserting an intravenous line. Osterholm says that it is urgent to ask how these choices are affecting patients: "In this setting, what treatments are working? Is what we're doing having any clinical impact?"

CAN THE VIRUS BE STOPPED?

Dozens of previous filovirus outbreaks have been halted using the same basic tools: isolating and treating patients, and tracing and monitoring their contacts. Public-health officials have also used this method to contain the spread of Ebola virus in Nigeria and Senegal. But across West Africa as a whole, the public-health response was completely inadequate at the start of the outbreak, which has allowed the virus to spread rapidly.

If, as some epidemiologists predict¹¹, the disease infects tens or even hundreds of thousands of people by January 2015, then it could become nearly impossible for tried-and-tested public containment measures to bring the epidemic under control. For one thing, it could be unfeasible to recruit and train the numbers of medical staff required. At this stage, some experts say that a new plan is needed.

Aid agencies and non-profit organizations are already trying new types of control. In Sierra Leone, officials are building isolation centres that can house patients away from their family and community, to

prevent the virus from spreading, but that have fewer qualified health-care workers than a standard treatment facility. It is a controversial move, because it risks sending the message that patients are being warehoused to die. But it acknowledges the harsh reality that existing clinics are full, patients are being turned away, and that this is fuelling the disease spread. "Everyone realizes these smaller-level community facilities are not ideal, but it is trying to do something rather than nothing to try to bring transmission down in communities," Hay says.

Another new approach in this epidemic has been the use of experimental therapies and vaccines developed specifically to target Ebola virus. Much attention has focused on ZMapp — a cocktail of antibodies that was first identified using mice vaccinated with Ebola virus proteins, and which has been given to several people in this outbreak. In August, researchers reported that ZMapp protected 18 monkeys from dying of Ebola virus — the first report of a highly successful therapy in animals already showing symptoms of the disease¹². Researchers are now organizing further tests of this and other medicines, as well as a handful of experimental vaccines.

"The question in our mind is, how did they survive?"

But even if these therapies are proven to work, the challenge is hardly over. The products that have progressed furthest in development are all focused on *Zaire ebolavirus*, the most lethal of the four species that sicken humans, but it is unlikely that they will be as effective against other filoviruses. Sapphire leads an international consortium that is systematically testing combinations of antibodies to find the ones that work best against different filovirus types. Researchers hope one day to produce therapies that work against multiple filoviruses and that could be used immediately when symptoms are detected, rather than spending precious time diagnosing exactly which virus is making someone ill.

All this work has taken on new urgency as predictions about the African outbreak have grown more dire. This year, the US National Institute of Allergy and Infectious Diseases, the Wellcome Trust, the European Union and various drug companies have channelled emergency funding into research on potential drugs and vaccines. And on 3 November, the US Institute of Medicine is convening a meeting in Washington DC to lay out an agenda for research on Ebola virus. That meeting takes place against growing worries that the virus might become endemic in parts of Africa, continuing to sicken and kill people for several years.

Researchers understand that the public-health response to this outbreak is a top priority, but they are also becoming more determined to understand this mysterious family of killer viruses. "What's happening in West Africa is a desperate situation," Sapphire says. "It makes people very willing to figure out what they need to do to solve problems." ■

Erika Check Hayden is a reporter for Nature in San Francisco, California.

1. Towner, J. S. *et al. PLoS Pathog.* **5**, e1000536 (2009).
2. Amman, B. R. *et al. J. Wildl. Dis.* (in the press).
3. Arata, A. A. & Johnson, B. in *Ebola Virus Haemorrhagic Fever* (ed. Pattyn, S. R.) 191–202 (Elsevier/North Holland Biomedical Press, 1978).
4. Barrette, R. W. *et al. Science* **325**, 204–206 (2009).
5. Pan, Y. *et al. Arch. Virol.* **159**, 1129–1132 (2014).
6. Kobinger, G. P. *et al. J. Infect. Dis.* **204**, 200–208 (2011).
7. Becquart, P. *et al. PLoS ONE* **5**, e9126 (2010).
8. Gire, S. K. *et al. Science* **345**, 1369–1372 (2014).
9. Pigott, D. M. *et al. eLife* <http://dx.doi.org/10.7554/eLife.04395> (2014)
10. Mylne, A. *et al. Sci. Data* **1**, 140042 (2014).
11. Meltzer, M. I. *et al. Morb. Mortal Wkly Rep.* **63**, (Suppl. 3) 1–14 (2014).
12. Qiu, X. *et al. Nature* **514**, 47–53 (2014).

COMMENT

METRICS Survey asks authors: is your most-cited work your best work? **p.561**

IMMUNOLOGY A history of autoimmune disease research and its pioneers **p.564**

MATHEMATICS John Napier book marks 400 years of logarithms **p.566**



CLIMATE 'Paris Protocol' should boost trade in green technology **p.567**

DETLEV VAN RAVENSWAAY/SPL



At least one roughly 10-metre-wide asteroid passes as close as the Moon each week (artist's impression).

Find asteroids to get to Mars

Asteroid retrieval is a distraction, says **Richard P. Binzel**. Better steps to interplanetary travel abound.

Interplanetary flight is the next giant leap for humans in space. Yet consensus on even the smallest steps forward has proven elusive. In June, a US National Research Council report¹ illuminated many options but offered no recommendations.

Return to the Moon? Head straight to Mars? Pluck a boulder off an asteroid and tug it to lunar orbit, just so that idle astronauts have somewhere to go and something to do? NASA must decide which path to follow before President Barack Obama's

budget announcement in January 2015.

Some options are better than others. The cost and complexity of human space exploration demands that each element be measured by its value towards the ultimate goal: Mars.

But NASA's stated next priority will not contribute to that aim. Its Asteroid Redirect Mission (ARM)² is a multibillion-dollar stunt to retrieve part of an asteroid and bring it close to Earth where astronauts can reach it. It will require an ancillary spacecraft deploying either a huge capture bag or a Rube Goldberg contraption resembling a giant arcade-game claw. Neither technology is useful for getting humans to Mars.

There is a better way. Thousands of shipping-container-sized and larger asteroids pass almost as close as the Moon each year. Many, such as September's near-miss asteroid 2014 RC, come closer. We need to find them far enough in advance, and abundant opportunities for crewed missions will open up.

This gateway for human space exploration requires three things: a thorough asteroid survey to find thousands of nearby bodies suitable for astronauts to visit; extending flight duration and distance capability to ever-increasing ranges out to Mars; and developing better robotic vehicles and tools to enable astronauts to explore an asteroid regardless of its size, shape or spin. The asteroid survey would also provide a prudent and overdue assessment of future impact hazards.

PASSERS BY

Asteroids orbit the Sun, most of them in the asteroid belt between Mars and Jupiter. Jupiter's gravity and other forces nudge a few onto paths that come within 40 million kilometres of Earth. The largest of these near-Earth asteroids (NEAs) is about 30 kilometres across, although most measure in metres. The smallest objects are most numerous and strike Earth frequently. Tonnes of their residual grains and pebbles rain down from space every day; a few metre-sized objects hit each year.

At about 20 metres across, the asteroid that lit the skies and shattered windows in Chelyabinsk, Russia, in 2013 demonstrated a threshold between merely delivering meteorites and posing a significant hazard. A Chelyabinsk-like airburst occurs somewhere on Earth on average every 50 years, usually over the oceans. A 10-kilometre-wide ▶

► impactor such as the one that hit Earth at the end of the Cretaceous period is, thankfully, a one-in-hundred-million year event³.

But asteroids can be friends, not foes. For nearly four decades⁴, NEAs have been recognized as human spaceflight destinations that are more accessible than the lunar surface. Owing to their minuscule gravity fields, a rendezvous with an asteroid merely entails sidling up and flying alongside one, with no need for a specialized landing craft.

Asteroids whose orbits pass between Earth and Mars offer a range of milestone destinations for testing distance and duration capabilities of human spaceflight. The first missions might last weeks and not go far; later excursions could last months and venture successively farther into interplanetary space, proving that Mars is within our grasp.

A 2009 presidential task force⁵ recognized NEAs as a low-cost and achievable option along such a 'flexible path' to Mars. Obama set course gently by mentioning the word 'asteroid' and the year '2025' in a speech to a receptive crowd of space workers in April 2010. Although barely a murmur compared with President John F. Kennedy's exhortation in May 1961 to go to the Moon, NASA planners took Obama's words as a directive.

Reality soon set in: getting humans to any known near-Earth asteroid by the mid-2020s was deemed beyond the range and budget of emerging flight systems. So the ARM scheme was hatched: instead of sending humans to an asteroid, let's bring a piece of one to within the anticipated reach of a 2025 crew. An uncrewed solar-electric propulsion vehicle would capture the booty and tug it back to a lunar orbit, where astronauts launched separately could explore it.

Mission accomplished? Not at all, in my

view. Hardware and operations to capture, contain and redirect an asteroid are dead-end elements with no value for long-duration crewed space travel¹. Delivering a supply module to lunar space would be a more sensible way to demonstrate solar-electric propulsion and benefit astronaut safety.

"Why retrieve an asteroid when we can wait for one to come near us?"

with a one-off costly stunt.

Equally specious in my opinion are arguments that ARM can deliver important new information on asteroid hazards or space resources. The ARM target would be scarcely one-quarter the size and one-hundredth the mass of the Chelyabinsk body — too small to survive atmospheric passage. And the idea that we might benefit any time soon by extracting water or rocket fuel out of an asteroid is fantastical, owing to its cost and complexity.

EXPLORATION GATEWAY

NASA needs a more pragmatic plan. Rededication to Mars as the long-term goal is clearly necessary (independent of any decision regarding a return to the Moon). Near-Earth asteroids are the most accessible interplanetary stepping stones to Mars. Why retrieve an asteroid when we can wait for one to come near us?

There are plenty to choose from (see 'Mission requirements'). The roughly one thousand objects of tens of metres and larger⁶ currently known to pass within a few times the distance to the Moon are just the tip of the

Conveying to the public that reaching Mars requires patient and diligent progression in capabilities is the honest alternative to distracting them

iceberg (see go.nature.com/9db89t). Nearly 10 million more such objects await discovery in orbits passing between Earth and Mars; one or more fly inside the Moon's orbit every week. So far, we have detected scarcely 0.1% of them because our current asteroid survey network is not up to the task.

The asteroids that are optimally located for exploration are also among those posing a potential impact hazard to Earth. 'Find them before they find us' has been the mantra of planetary scientists^{3,7} and national reports^{8,9} for decades. Yet neither the US Congress nor NASA has funded a large, dedicated telescopic survey. NASA spends only a minuscule fraction of its budget — \$US20 million per year — on searches using a haphazard array of adapted astronomical facilities.

Current surveys also fall short of US law. The George E. Brown, Jr. Near-Earth Object Survey Act of 2005 requires NASA to find 90% of hazardous asteroids 140 metres or larger by 2020, a goal that will not be met without significantly more investment. It is a grand challenge requiring a grand commitment from the space agency, the president and Congress: \$200 million per year. This budget line, comparable to NASA's New Frontiers programme that has sent uncrewed probes to the Solar System's planets, would allow the survey to be completed within a decade at a fraction of ARM's projected multibillion-dollar cost.

Once humans can reach one asteroid in its native orbit, the gateway is opened such that hundreds (if not thousands) more will be accessible, enabling a steady programme of exploration to be unrolled in the late 2020s and 2030s. Commercial development of space resources on select candidates could begin to be evaluated towards the middle of this century.

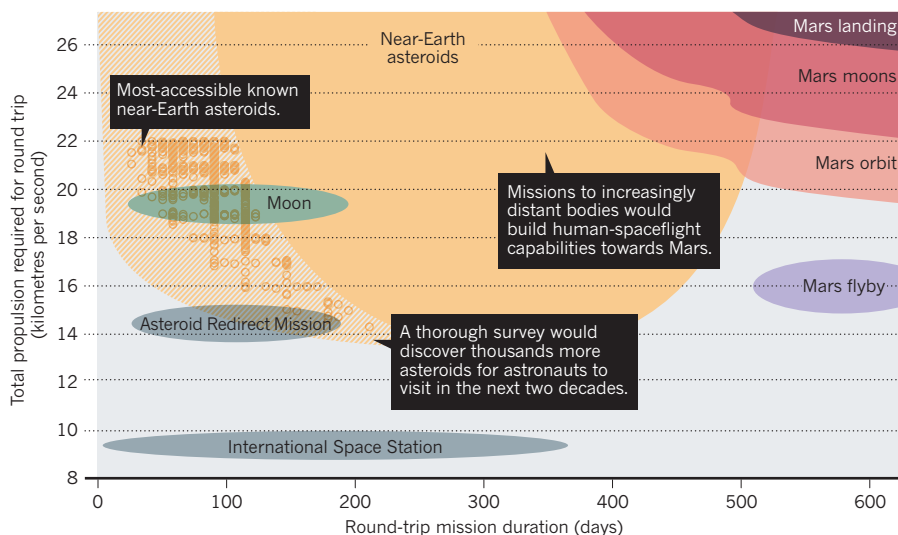
Robotic operated vehicles (ROVs) will need to be developed to explore asteroids, just as astronauts in orbit around or on the surface of Mars will need to command robotic workhorses. Most asteroids are irregularly shaped, boulder-strewn and tumble in their orbits. We have no reason to expect an accessible asteroid to have a slow and steady spin and safe, smooth surface hospitable to a space suit. ROVs such as those used for underwater exploration could be directed from the main crew module to carry out *in situ* investigations. With an array of arms or grappling devices to gain a toehold necessary for working in microgravity, and the capability to operate longer hours than possible for a spacewalk, exploration of these worlds could be achieved.

GRAND CHALLENGE

Over the next two months, Obama's 2015 budget will be shaped. NASA needs to make a clear choice about its priorities. It should abandon the ARM mission concept and

MISSION REQUIREMENTS

A mission to a near-Earth asteroid would require less propulsion and a shorter mission duration than a human mission to Mars. Less than 1% of the most-accessible asteroids are currently known (yellow circles), but a dedicated survey (filling in the yellow-hatched region) would reveal abundant asteroid stepping-stone opportunities as a gateway for human interplanetary exploration.



SOURCE: BRENT W. BARBEE/NASA GODDARD SPACEFLIGHT CENTER

make an asteroid survey its top priority to provide a basis for future crewed missions.

NASA's directorates for human exploration, space technology, science and others must pool resources to address the agency's declared grand challenge. Congress and the White House should fund a new series of missions that embrace the resulting synergy of exploration, technology and science for the benefit of humanity.

Such a 'Grand Challenge Mission' programme should support human spaceflight and humankind's future. It should be modelled after, and with a budget similar to, the New Frontiers programme of planetary probes each costing less than \$800 million and selected through competition. Advance definition of objectives would drive development by industry and academia towards the best ideas and optimize long-term planning. Competitive selection of proposals ensures the most cost-effective return on taxpayer investment.

Three asteroid-related concepts should be explored: an asteroid survey to find a series of human destinations on the path to Mars while fulfilling the requirements of the 2005 survey act; a competition to test robotic asteroid-deflection methods on which civilization's survival could one day depend; and another to test robotic methods of extracting water or mining other valuable resources from asteroids that might contribute towards sustaining human spacefaring decades from now.

NASA needs to get back on track towards achieving humankind's next giant leap in space. ■

Richard P. Binzel is a joint professor of planetary science and aeronautics and astronautics at the Massachusetts Institute of Technology in Cambridge, Massachusetts, USA.
e-mail: rpb@mit.edu

1. National Research Council *Pathways to Exploration: Rationales and Approaches for a U.S. Program of Human Space Exploration* (National Academies Press, 2014).
2. NASA Asteroid Initiative Opportunities Forum: *Update on Asteroid Redirect Mission* (NASA, 2014); available at go.nature.com/cu63ol.
3. Yeomans, D. K. *Near-Earth Objects: Finding Them Before They Find Us* (Princeton Univ. Press, 2013).
4. Shoemaker, E. M., Williams, J. G., Helin, E. F. & Wolfe, R. F. in *Asteroids* (ed. Gehrels, T.) 253–282 (Univ. Arizona Press, 1979).
5. US Human Spaceflight Plans Committee *Seeking A Human Spaceflight Program Worthy of a Great Nation* (NASA, 2009).
6. Barbee, B. et al. in *Proc. IAA Planet. Def. Conf. 2013 IAA-PDC13-04-13* (2013).
7. Chapman, C. R. & Morrison, D. *Nature* **367**, 33–40 (1994).
8. National Research Council *Defending Planet Earth: Near-Earth Object Surveys and Hazard Mitigation Strategies* (National Academies Press, 2010).
9. NASA Advisory Council *Report of the Ad-Hoc Task Force on Planetary Defense* (NASA, 2010).



Is your most cited work your best?

John P. A. Ioannidis and colleagues asked the most highly cited biomedical scientists to score their top-ten papers in six ways.

After half a century of citation indices, several intriguing questions remain. Are the most highly cited papers the most important ones? Does science make progress mostly through evolution or through revolution? Are these two processes mutually exclusive or complementary, and which do high citations most reflect? Are surprising findings difficult to publish?

Highly cited papers are nodes in the network of the dissemination and discussion of scientific information. But citation counts alone cannot reveal why a paper is considered so important as to attract repeated mention by other researchers. To contribute to these debates, we surveyed the most-cited authors of biomedical research for their views

on their own influential published work^{1–3}.

We got some intriguing feedback. The vast majority of this elite group felt that their most important paper was indeed one of their most-cited ones. Yet they described most of their chart-topping work as evolutionary, not revolutionary.

BEST OF THE BEST

We listed the 400 most-cited biomedical scientists in the period 1996–2011⁴. We selected each author's ten most-cited papers (adjusted for publication year) published in 2005–08, and asked them to score the papers (on a scale of 0 to 100) in six dimensions.

We restricted the period to 2005–08, because the perception of the importance ▶

► of a paper can change over time⁵. Old, highly cited papers might become stereotyped⁶: people unquestioningly treat them as canonical. Recent papers (those published within the past five years), have had insufficient time to accrue citations.

Overall, 123 scientists responded, scoring 1,214 papers between them. On average, investigators tended to give their blockbuster papers high scores for dimensions that reflect evolution: Continuous Progress, Broader Interest and Greater Synthesis (for definitions of these terms and extended data, see Supplementary Information; go.nature.com/76fqj2). They gave the papers lower scores on average for dimensions that reflect revolution: Disruptive Innovativeness and Surprise (see 'Self assessment'). They also indicated that blockbuster papers were easy to publish, with some exceptions.

Twenty scientists (16%) felt that their most important paper published in 2005–08 was not among their top ten most cited. However, most of these 20 papers were still heavily cited (on average in the top 3% published in the same year in terms of citations; seven were in the top 15 papers that the author published in 2005–08). Authors scored these papers higher for Disruptive Innovativeness (in nine cases) and Surprise (in five cases) than their ten most-cited papers.

Fifty-two papers were appraised by at least two authors. We evaluated co-author agreement by comparing the scores for each dimension to their median values for that author. We considered the paired responses as an example of an agreement if both authors scored a paper above their median, below their median or at their median. The expected proportion of agreement given random responses is 50%. The rate of agreement ranged from 74% to 86% for the six dimensions, but the sample size was limited.

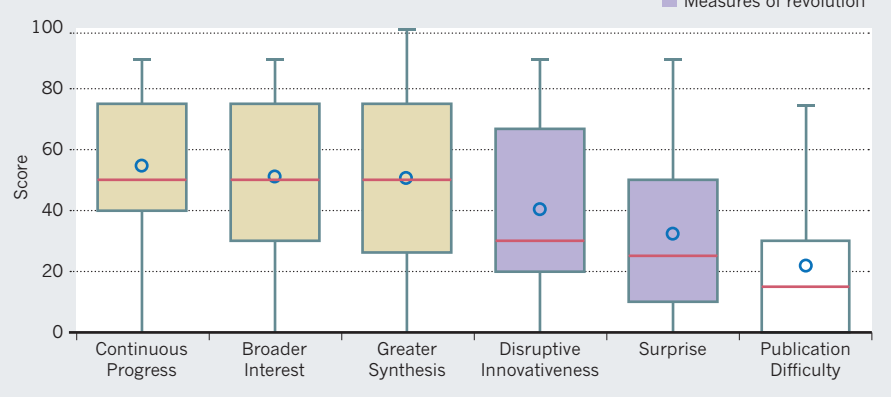
SURPRISE, SURPRISE

Predictably, the strongest correlations were between Disruptive Innovativeness and Surprise, and between Surprise and Publication Difficulty. We had expected that papers would be perceived as either evolutionary or revolutionary, not both. A skew towards evolution or revolution would have shown up as a strong negative correlation between Disruptive Innovativeness and Continuous Progress, but this relationship was not statistically significant. Instead, Broader Interest scores correlated with strong scores for both revolution and evolution.

Our survey, of course, has limitations. First, only just over 30% of the contacted authors replied. The non-responders might have given higher scores for revolution or evolution. Second, we assessed papers published only in a narrow time frame, in an effort to achieve uniformity. Third, authors may appraise their own work

SELF ASSESSMENT

One hundred and twenty-three of the most highly cited biomedical scientists scored their ten most-cited papers (1,214 between them) published in 2005–08 in six dimensions.



more positively than work by others.

A fourth caveat is that the sample of the most-cited scientists is biased in favour of individuals whose work has already been widely accepted. Those with paradigm-changing ideas that are not so accepted (and thus cited) would not be in this group. Finally, scientists whose citations are closer to the average may have scored these highly cited papers or their own middle-ranking papers quite differently.

We suspect that among more averagely cited work, evolution scores would have been even more prominent. True out-of-the-box innovation and major surprises are probably not common across the literature.

NOW WHAT?

This small survey is a salutary reminder both of how much information citations can convey, and of how much about the science of science we have yet to understand. For example, given that revolutionary papers are quite rare, how might we identify them early? Do innovative papers make connections between areas of knowledge that are not typically made⁷, or do they get cited early on by papers in remote fields?

In the future, with all the low-hanging fruit plucked, will we see (among the most-cited papers) a relative drop in the percentage of revolutionary papers and a corresponding increase in papers that provide a synthesis of the literature? How quickly does stereotyping of opinion about the importance of a paper happen, and how quickly and how much do such opinions change over time? What proportion of the most important papers across each of these six dimensions might be found among the output of the large majority of scientists with more-average citation profiles?

It would be particularly useful to know whether successful out-of-the-box ideas are generated and defended largely by the most influential scientists or by colleagues lower on the citation rankings. Would the opinion

of scientists who cited the top papers that we examined square with the opinions summarized here? Are there other dimensions in addition to the six that we examined that might capture the essence of important work? For example, some respondents pointed to the significance of translational potential and social impact for research, which might have been captured only in part under our Broader Interest dimension. And it would be interesting to know whether there are major differences in the evolution-versus-revolution pattern in the physical sciences.

One way to answer some of these questions would be to survey those who cite the highly cited papers or investigators with more-modest citation rankings. We must continue to hone indices other than citation-based metrics to complement appraisal of scientific accomplishment⁸. ■ **SEE NEWS FEATURE P.550**

John P. A. Ioannidis is co-director of the Meta-Research Innovation Center at Stanford (METRICS), Stanford University, California, USA. **Kevin W. Boyack** is at SciTech Strategies, Albuquerque, New Mexico, USA. **Henry Small** is at SciTech Strategies, Bala Cynwyd, Pennsylvania, USA. **Aaron A. Sorensen** is director of informatics at the Temple University School of Medicine, Philadelphia, Pennsylvania, USA. **Richard Klavans** is at SciTech Strategies, Berwyn, Pennsylvania, USA. e-mail: jioannid@stanford.edu

- Small, H. *Scientometrics* **60**, 305–316 (2004).
- Astin, H. S. in *The Outer Circle: Women in the Scientific Community* (eds Zuckerman, H., Cole, J. R. & Bruer, J. T.) 57–70 (W. W. Norton, 1991).
- Virgo, J. A. *Libr. Q.* **47**, 415–430 (1977).
- Boyack, K. W., Klavans, R., Sorensen, A. A. & Ioannidis, J. P. A. *Eur. J. Clin. Invest.* **43**, 1339–1365 (2013).
- Cozzens, S. E. *Knowl. Creat. Diffus. Util.* **9**, 511–529 (1988).
- Small, H. G. *Soc. Stud. Sci.* **8**, 327–340 (1978).
- Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. *Science* **342**, 468–472 (2013).
- Ioannidis, J. P. A. & Khoury, M. J. *J. Am. Med. Assoc.* **312**, 483–484 (2014).



ST BARTHOLOMEW'S HOSPITAL ARCH. & MUS./WELLCOME IMAGES

A nineteenth-century watercolour showing the effects of lupus, an autoimmune disease.

AUTOIMMUNITY

When the body betrays

Tilli Tansey surveys a magisterial, historically rich biography of autoimmunity.

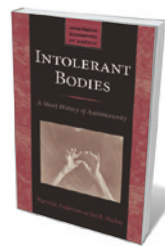
A terrifying incident opens *Intolerant Bodies*. In 1981, novelist Joseph Heller became paralysed within a few hours by the autoimmune disease Guillain-Barré syndrome. In this horrific physiological own goal echoing the catch-22 of his classic book, Heller's immune system turned on his peripheral nervous system.

It has been calculated that in any population, 5–10% of people — a large proportion of them women — will at some point have an autoimmune disease. *Intolerant Bodies*, by the medically trained historian Warwick Anderson and clinical immunologist Ian Mackay, draws on medical, scientific, historical and literary sources to highlight four such conditions: type 1 diabetes, rheumatoid arthritis, systemic lupus erythematosus and multiple sclerosis.

The term autoimmune first appeared in 1951, and autoimmunity in 1957, although neither was then — or now — universally accepted. Some immunologists continue to prefer 'autoallergy'; the late Robin Coombs, inventor of the first specific test for red-cell

antibodies, declared that he would "go down fighting" in his opposition to 'auto-immune'.

Mackay and Anderson focus less on the heyday of autoimmunity in the past few decades than on its 'prehistory' and emergence as a defined field of medical research and practice. Early-nineteenth-century studies of fevers led to the recognition that the body's own physiological responses could be harmful, and germ theory later that century spurred knowledge of the body's defence mechanisms. Medical luminaries such as Robert Koch and Joseph Lister struggled to understand basic mechanisms and isolate causative agents. In the 1880s, the Russian biologist Élie Metchnikoff argued that



Intolerant Bodies: A Short History of Autoimmunity
WARWICK ANDERSON
AND IAN R. MACKAY
Johns Hopkins
University Press: 2014.

circulating cells, the phagocytes, digested invading cells. In the first half of the twentieth century, in public-health labs and early pharmaceutical companies, the development of serotherapy — treating infectious diseases such as diphtheria by injecting serum raised in an immunized animal — stimulated much research on the immune system.

The authors emphasize how after the Second World War, an increase in government and private funding led to the expansion of research laboratories, including Henry Kunkel's at the Rockefeller Institute in New York City, Macfarlane Burnet's at the Walter and Eliza Hall Institute in Melbourne, Australia, and those at industrial concerns such as pharmaceutical giant Merck. Researchers developed tools, techniques and languages based on cellular and molecular approaches. In Britain, Peter Medawar unravelled the mechanisms of rejection after tissue transplantation. In New York City, Jules Freund demonstrated that an emulsion of paraffin oil and dead mycobacteria, the causative organisms of tuberculosis, enhanced and prolonged

immune reactions in animal models of disease. Burnet, who shared a Nobel prize with Medawar in 1960, recognized Freund's contribution as key to the experimental study of autoimmunity. The subject was now high on the agenda of many biomedical scientists. Work on subsets of lymphocytes, human-tissue antigens and the genetics of histocompatibility factors all contributed to the understanding of cellular mechanisms, and to devising diagnostic and therapeutic strategies for an array of disorders.

But *Intolerant Bodies* is more than a history of scientific concepts, clinical practices and experimental activities. It explores fundamental questions such as 'What is self?' — how and why the body's sophisticated defence mechanisms fail to recognize its cells and tissues and start to destroy them. The authors draw on immunological, philosophical, psychological and religious literature for this intriguing discussion. They mine thoughts from the likes of philosopher Jacques Derrida and anthropologist David Napier for their reflections on the definitions, boundaries and limitations of self.

Furthermore, Anderson and Mackay reveal an expert understanding of how to use 'lived experience' to bring a biography of disease to life. Personal accounts demonstrate how, as theories about the causes of inexplicable chronic and debilitating diseases abounded, the variety of treatments devised to alleviate or 'cure' them expanded. The nineteenth-century German poet Heinrich Heine was treated with leeches, sulphur baths and morphine, but these did little for an aggressive neurological complaint that the authors suggest was multiple sclerosis. And in 1957, US novelist Flannery O'Connor, who had lupus, wrote to a friend, "Every time something new is invented I get in on the ground floor with it. There have been fine improvements in the medicine". But steroids also destroyed her bones, a therapeutic stalemate that made her final years miserable. Particularly harrowing is the case of English diarist W. N. P. Barbellion (the pen name of Bruce Frederick Cummings), author of the 1919 *Journal of a Disappointed Man*. Barbellion, diagnosed with multiple sclerosis in his twenties, began to hoard laudanum and to keep a pistol close by. The disease killed him less than five years after diagnosis.

Unlike many biographies, *Intolerant Bodies* has no neat ending. With more than 80 conditions now considered to be autoimmune, diagnosis remains irregular and treatments unsatisfactory. Although more and more is understood about mechanisms, the causative stimuli are still unknown. The history of autoimmunity is far from over. ■

Tilli Tansey is professor of the history of modern medical sciences at Queen Mary, University of London.
e-mail: t.tansey@qmul.ac.uk

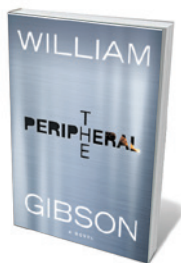
Books in brief



Making Space: How the Brain Knows Where Things Are

Jennifer M. Groh BELKNAP (2014)

The timing is spot-on for this study of the brain's navigational system, with a Nobel prize awarded to John O'Keefe, May-Britt Moser and Edvard Moser this month for their research in the field (see *Nature* **514**, 154–157; 2014). Neuroscientist Jennifer Groh deftly elucidates the mental computations that allow understanding of location and boundaries, interweaving well-judged snippets of history. The mechanisms, such as the brain's updates on eye movements, are fascinating — as is Groh's revelation that neurons can "do double duty" in tasks such as spatial navigation and memory.



The Peripheral: A Novel

William Gibson PUTNAM ADULT (2014)

With *Neuromancer* (Ace, 1984), which predicted the rise of the Internet, science-fiction writer William Gibson revealed his nose for near-future science and for seeing, as he puts it, a future already here but not "evenly distributed". He harnesses that ability anew for this techno-thriller. From plastic in the oceans to three-dimensional printing, today's concerns are beautifully repurposed as tomorrow's background. But for Gibson, technological change is just a different route to the same problems: his protagonists, who wield the humanoid-telepresence 'peripherals', are socially marginalized.



Note-by-Note Cooking: The Future of Food

Hervé This (transl. M. B. DeBevoise) COLUMBIA UNIVERSITY PRESS (2014)

Molecular cooking, that scientific wonderland of filter pumps and rotary evaporators, has evolved. The new revolution, proclaims physical chemist Hervé This, is "note-by-note cooking" — the orchestration of pure molecular compounds. Amid the dollops of theory are practical examples, such as chef Pierre Gagnaire's glucose "péligot" caramels (exemplars of contrasting consistency), and lists of compounds including the mushroom-scented 1-octen-3-ol. Recipes are scant, although Effervescence, a cocktail involving Syrah grape polyphenols and ethanol, sounds toast-worthy.



The Perfect Meal: The Multisensory Science of Food and Dining

Charles Spence and Betina Piqueras-Fiszman WILEY (2014)

Anyone who thoughtlessly masticates their way through a meal is missing something extraordinary. So might opine experimental psychologists Charles Spence and Betina Piqueras-Fiszman, who have mined neuroscience and behavioural economics for this study of the social, sensory and psychological factors that optimize the pleasures of the table. Their exhaustive analyses of everything from plate size to flavour incongruities are laced with details such as cutlery's evolution from Victorian marrow scoops and chocolate muddlers to today's textured spoons and "aromatic forks".



Bee Time: Lessons from the Hive

Mark L. Winston HARVARD UNIVERSITY PRESS (2014)

From the whirr of wings to the whiff of honey, the "full-body experience" of working in apiaries has, for biologist Mark Winston, sparked insights into humanity's relationship with nature. In this personal and scientific journey into the history we share with bees, he ranges over neonicotinoid pesticides and colony collapse, the control of African 'killer' bees and more. The charismatic social insects emerge as both icons of societal cohesion and symbols of nature's paradoxically mingled power and fragility. **Barbara Kiser**

Constructor of a marvellous canon

George Szpiro weighs up a life of John Napier, who gifted science with logarithms.

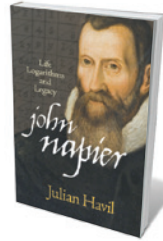
In Romanian playwright Eugène Ionesco's *The Lesson*, a professor asks a student to multiply 3,755,998,251 by 5,162,303,508. To his amazement, the answer comes quick as a shot: 19,390,002,844,219,164,508. The student declares that she had "learnt all possible results of all possible multiplications by heart".

If only it were that simple. (Indeed, the student got it wrong: the answer is 19,389,602,947,179,164,508.) Multiplication is computationally complex. The number of mathematical operations required to add two numbers grows linearly with the number of digits involved, but in multiplication using simple arithmetic the number of operations grows with the square of the number of digits. Computer algorithms now simplify this, but in centuries past multiplications were the bane of scientific research. The seventeenth-century astronomer Johannes Kepler, for instance, had to perform thousands of them to compute the positions of celestial objects.

His task would have been much more difficult had it not been for a discovery more than a decade before by mathematician John Napier: logarithms. Tables of these then-mysterious numbers meant that two figures could be multiplied by adding their logarithms. In *John Napier*, maths writer Julian Havil traces the life and work of the man who presented us with this gift 400 years ago.

Napier was born in Edinburgh in 1550, 14 years before Galileo. Little is known about his youth and education except that he travelled, presumably in Europe, and returned to Scotland a scholar at 21. He became a gentleman farmer, theologian and politician as well as a mathematician. A passionate intellectual, he often paced the countryside deep in thought in a dark cloak — his academic gear — which led neighbours to believe that he was in league with the devil. His groundbreaking book on logarithms, *Mirifici Logarithmorum Canonis Descriptio*, was published to little fanfare in 1614. It garnered less attention in his era than his 1593 blockbuster on the dangers of Catholicism, *A Plaine Discovery of the Whole Revelation of St. John*.

Descriptio consists of tables of logarithms and instructions on how to use them, in particular for multiplications and divisions in trigonometry, by adding and subtracting their entries. Within years of its publication, English mathematician William Oughtred



John Napier: Life, Logarithms, and Legacy
JULIAN HAVIL
Princeton University Press: 2014.

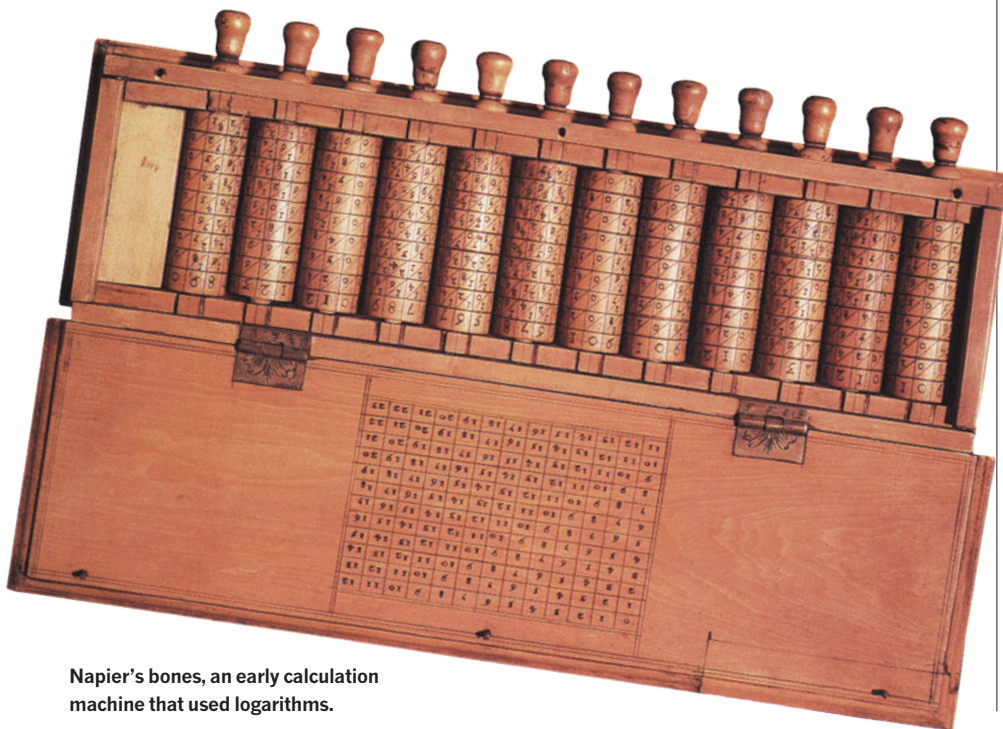
used Napier's logarithms to invent the slide rule, which remained indispensable for science, engineering and economics until the advent of the pocket calculator in the 1970s.

Napier published *Descriptio* without revealing how he had arrived at the logarithms. That secret was exposed only in *Mirifici Logarithmorum Canonis Constructio*, published after his death by his son Robert and friend Henry Briggs (to whom we owe the common logarithm to base ten). In *Constructio*, we find that Napier had discovered the numbers by considering two objects moving at different speeds along straight lines, and establishing a relationship between their locations at every moment. The most salient feature of logarithms is that they are needed to work out x in equations of the form $10^x = 1,000$. (In this case, $x = \log_{10}(1,000) = 3$.)

However, it took more than a century to clarify the relationship between logarithms and exponents (the power to which a number is raised). That was the work of Welsh mathematician William Jones, famous for being first to use π for the ratio of a circle's circumference to its diameter. Jones's contribution to Napier's work was unearthed only in 1771, long after his death, when Royal Society librarian John Robertson reported that the Welshman had formulated the law of logarithms, $\log(ab) = \log(a) + \log(b)$. This is exactly the relationship that Napier had sought to make calculations less tedious. Although developed to simplify computations, logarithms turned out to be of supreme importance for pure and applied mathematics, and for all branches of science, from physics and chemistry to psychology, information theory, computer science and economics.

Unfortunately, Havil, like Napier, keeps the reader in the dark about the essential nature of logarithms for most of the book. It would have been helpful to introduce logarithms in the way taught to most secondary students now — as functions that solve for exponents in simple equations — and move on from there. Furthermore, interminable computations, laid out in minute detail often over several pages, may discourage many readers; and the narrative suffers from the frequent use of long quotations in seventeenth-century English. Nevertheless, *John Napier* fills a gap concerning an important, and often ignored, chapter of mathematical history. ■

George Szpiro is New York correspondent for the Swiss newspaper *Neue Zürcher Zeitung*.
e-mail: georgeszpiro@gmail.com



Napier's bones, an early calculation machine that used logarithms.

Correspondence

Can GM marmoset use be justified?

We question the premise that genetically modified (GM) marmosets are essential to Japan's brain-mapping project for studying disorders such as schizophrenia and depression (*Nature* **514**, 151–152; 2014). It runs counter to efforts in fields such as toxicology to reduce or eliminate the use of non-human primates in research.

Our conclusions in the 2011 Review of Research Using Non-Human Primates (see go.nature.com/kgly27w), which included UK neuroscience research between 1997 and 2006, are relevant here. The review judged such research to be of high quality, yet few projects showed evidence of biomedical benefit — despite being completed many years before the review.

There was a tendency among researchers to make over-optimistic and unsubstantiated claims for their work and how it might improve the treatment of brain disorders, presumably to strengthen their case for funding. We concluded that “the size of the problem to which the science relates should not be accepted as sole justification for individual items of research” and that “health benefits should only be claimed when their potential is real”. In some instances, we judged that the case for using non-human primates — rather than other species, including humans — was not well made.

It could be argued that the unavailability of suitable GM animals at the time might account for the poor translation into health benefits. But the justification for using GM marmosets in research today needs to be much stronger, because of the substantial ethical and welfare issues at stake.

You quote US neuroscientist Terry Sejnowski, who proposes consideration of “the ethical issues that will inevitably arise up the road”. We contend that these should be considered

before the journey starts.
Patrick Bateson *Cambridge, UK.*
C. Ian Ragan *London, UK.*
ian.ragan1@btinternet.com

Sustainability: call to coordinate actions

The first step in implementing the United Nations' Sustainable Development Goals (SDGs; see M. Stafford-Smith *Nature* **513**, 281; 2014) will be to identify critical nodes at which issues such as water, energy and food need to be addressed simultaneously and that have the potential to set the agenda.

For example, schools could play a part in increasing water, food and energy security by improving students' knowledge of these systems as well as of flood and drought responses. These nodes could help to coordinate existing discussion structures into new national and local forums, and to overcome bureaucratic silos.

Contribution by governments to the SDGs will be voluntary, so efforts should focus on increasing the numbers of organizations, people and networks that can be linked to the SDG governance system. This would allow the system to address inequities in gender, geographic location and ethnicity, for example, at regional, national and local scales.

Effective monitoring will be pivotal to the success of the SDGs. If done globally through existing UN agencies, there is a risk that key resources would be diverted, whereas monitoring by individual countries might be inadequate or inconsistent. The UN High-level Political Forum could have an important role, but people outside the current sustainability discussions need to be heard as well. They too can help to craft context-specific goals and targets, maintain ambition and develop complex monitoring capacity.

Norichika Kanie *Tokyo Institute of Technology, Japan.*
Casey Stevens *Providence College, Rhode Island, USA.*
kanie@valdes.titech.ac.jp

Exempt green tech from trade rules

Next year in Paris, parties to the UN Framework Convention on Climate Change have the task of developing a global agreement to reduce carbon emissions — a successor to the failed Kyoto Protocol. They should strike a grand bargain with the World Trade Organization (WTO) to promote green industry.

For example, the new protocol could draw up an open-ended list of products and processes that would reduce emissions. This could then be used to exempt countries from the rules of fair trade for, say, five years, to allow them to build green industries.

Trade in green industries is crucial for transforming fossil-fuelled economies. Countries trying to promote a green economy should not have to experience trade disputes (see, for instance, go.nature.com/7joyyk).

There is a precedent for such exemptions in the WTO's General Agreement on Tariffs and Trade. The negotiation of a new 'Paris Protocol' provides a once-in-a-generation chance to create green industries and to cut emissions.
John A. Mathews *Macquarie University, Sydney, Australia.*
john.mathews@mgs.edu.au

Engage more cohort patients in research

Input from patients and the public is a requirement for funding by the UK National Institute of Health Research (NIHR), but is met by scepticism from some academics.

We analysed 70 birth cohort studies to evaluate the basis of this scepticism (P. J. Lucas *et al. BMC Med. Res. Methodol.* **13**, 56; 2013). Such studies rely on recruiting and retaining many people over many years — often a challenge. Retention was a factor in the June decision to pause the US National Children's Study, for example (see go.nature.com/mfmhhr).

These problems might

be overcome by consulting participants about research design and processes, rather than using them solely for data provision. But we found very few examples. Some scientists feared that involving participants could compromise data integrity; others lacked the right resources or were ignorant of the practice.

There is much to be gained from engaging participants in research, but persuading academics to do so will require better support and for other funders to follow the NIHR's lead.
Patricia Lucas, Maggie Leggett *University of Bristol, UK.*
Simon Denegri *NIHR, London.*
maggie.leggett@bristol.ac.uk

Translating public action into policy

Last month, millions took to the streets in climate rallies organized by some 1,500 organizations in 158 countries. Their message: that world leaders at the United Nations Climate Summit in New York must tackle the challenges of global warming head on. However, it is by no means clear that this will translate into firm political pledges (see *Nature* <http://doi.org/v3c>; 2014).

For instance, 80,000 people demonstrated ahead of the UN Climate Change Conference in Copenhagen in 2009. At the meeting, a political collusion between the United States and a group formed by Brazil, South Africa, India and China ended up sidelining the countries most in need of action against climate change (see A. P. Kythreotis *Prog. Hum. Geog.* **36**, 457–474; 2012).

When it comes to crucial issues of social justice and the economy, politicians should be guided by the people they are supposed to serve (see A. Petherick *Nature Clim. Change* **4**, 81–83; 2014).
Chrystal Mantyka-Pringle *University of Saskatchewan, Saskatoon, Canada.*
Andrew P. Kythreotis *Cardiff University, UK.*
c.mantyka-pringle@usask.ca

Ghost locus appears

The sequences of two sponge genomes provide evidence that the ParaHox developmental genes are older than previously thought. This has implications for animal taxonomy and for developmental and evolutionary biology. [SEE LETTER P.620](#)

JAMES O. MCINERNEY &
MARY J. O'CONNELL

An array of transcription factors controls development in animals¹. Transcription-factor-encoding genes belonging to one class — Antennapedia (ANTP) — are present throughout the animal kingdom and usually have a key role in development. The ANTP group includes the Hox, ParaHox and NK genes, all of which are paralogues, meaning that they have arisen in different animals from a shared ancestor as a result of gene-duplication events. However, the origins, evolution and, in particular, the timing of these duplication events have been unclear. On page 620 of this issue, Fortunato *et al.*² report the analysis of two new genomes from sponges of the class Calcarea (the calcisponges) that substantially clarifies the evolutionary history of the ANTP genes and provides insight into tissue structures that are shared across the animal kingdom.

A substantial puzzle has arisen concerning the repertoire of ParaHox and Hox genes in animals. The first sponge genome to be sequenced was that of the demosponge (class Demospongiae) *Amphimedon queenslandica*³. Although there seem to be no Hox or ParaHox genes in this genome, the evolutionary conservation of clusters of genes known to be neighbours of Hox and ParaHox genes in other organisms led to the proposal that Hox and ParaHox genes were present in the common ancestor of all animals but had been lost in sponges. The researchers called their idea the 'ghost locus' hypothesis³. However, this evidence, although intriguing, was indirect, because it was based on an inference of ancestral gene content.

Despite their simple morphologies, sponges from different classes are thought to be quite divergent and indeed it is not certain that they all form a monophyletic (natural) group^{4,5}. If they are monophyletic⁴, then their last common ancestor dates to approximately 600 million years ago; if they are a paraphyletic group, then the time at which the Calcarea and the Demospongiae diverged, for instance, might be as much as 800 million years ago⁵. Either way, the antiquity of their divergence means that the genomes of a demosponge and a

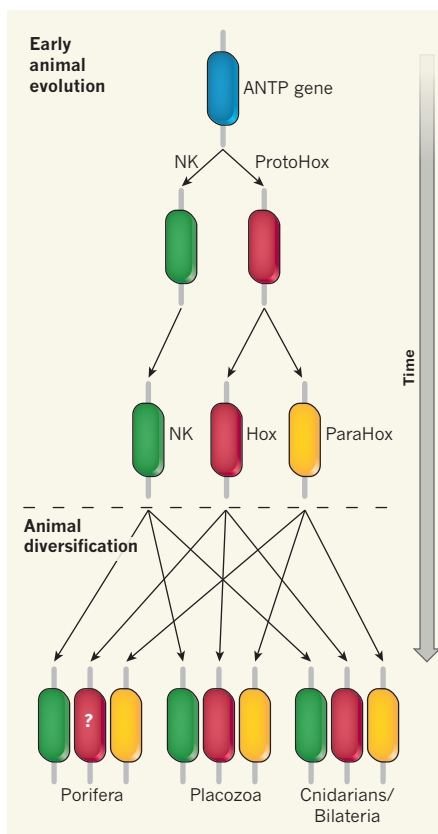


Figure 1 | Evolution of Hox, ParaHox and NK genes in animals. Fortunato and colleagues' comparison² of two calcisponge genomes (phylum Porifera) with genome data from other animals, including placozoans, cnidarians and bilaterians, suggests that an ancestral Antennapedia (ANTP) gene duplicated to give rise to the ancestor of NK genes and a 'ProtoHox' gene. The ProtoHox gene subsequently duplicated to give rise to the existing Hox and ParaHox gene families. Animal diversification occurred after these gene-duplication events, but gene losses have complicated the picture, explaining, for example, the seeming lack of Hox genes in calcisponges and of both Hox and ParaHox genes in demosponges.

calcisponge might be expected to differ from one another more than is usual for two animals from the same phylum, such that a single representative of the sponges can never be emblematic of the sponges — the phylum Porifera — as a whole. Sampling these divergent taxa is therefore crucial if we are

to understand the evolution of the genetic circuitry that effects animal development.

A 'superphylum' called the ParaHoxozoa⁶ has been defined on the basis of the presumed shared presence of Hox or ParaHox genes. This group includes the Bilateria (comprising all animals with bilateral symmetry), the Placozoa (simple, flat animals that live on the ocean floor) and the Cnidaria (radially symmetrical aquatic animals, such as jellyfish). The animals excluded from the ParaHoxozoa are the sponges and the Ctenophora (comb jellies), because of their presumed lack of these genes.

Fortunato *et al.* provide direct support for the ghost locus hypothesis, at least for the ParaHox genes. By combining data from their sponge genomes with existing data for phylogenetic, gene-neighbourhood and gene-expression analyses, they suggest not only that developmental-gene evolution has been very different in sponges compared with most other animals, but also that early animals did indeed contain a repertoire of ANTP genes that included the NK and ParaHox genes at least (Fig. 1).

The authors first constructed a phylogenetic tree of a large family of ANTP genes. Given the great phylogenetic breadth that is spanned by the tree, it is unsurprising that high statistical support for relationships is not achieved. Nonetheless, the tree topology does indicate the presence of NK and ParaHox genes in the genomes of the calcisponges.

Next, the researchers demonstrated that there is some conservation in the gene neighbourhoods between these putative ParaHox genes in sponges and their presumed equivalents in bilaterians. Conservation of gene neighbourhoods is a hallmark of the evolution of ANTP genes, but the authors find that the strong conservation observed in most animals is not found in sponges. This implies that conservation of ParaHox gene neighbourhoods is not a universal rule as was once thought.

Finally, the authors show that the expression patterns of the calcisponge ParaHox genes are remarkably similar to those of other animals. For example, the choanoderm cell layer in sponges has a similar pattern of ParaHox expression to the endoderm in bilaterian animals, and the pattern in the sponges'

cross cells is like that of sensory cells in other animals. This opens up the possibility that certain tissue types in sponges might have a common ancestry with other tissue types in bilaterian animals. However, it should be kept in mind that homology in developmental genes does not always equate to homology of organs⁷.

One implication of Fortunato and colleagues' findings is that, although the ParaHoxozoan superphylum may still be a useful taxonomic group, it can no longer be defined by the possession of ParaHox or Hox genes. In addition, the results suggest that the origin of the ParaHox genes is earlier than was previously thought and that the first animals probably had a ParaHox gene repertoire (Fig. 1). This idea leads to the intriguing question of whether the common ancestor of all animals was in fact more developmentally complex than present-day sponges, cnidarians and placozoans, and

that these groups have lost complexity, rather than that complexity has been gained in other animal lineages.

Our existing sample of animal genomes has heavily influenced our current ideas about ANTP gene evolution. There are two other classes of sponge, the Homoscleromorpha⁸ and the Hexactinellida⁹, and a draft genome sequence for the homoscleromorph *Oscarella carmela* is available¹⁰. Investigating its ANTP gene repertoire should shed further light on the evolution of ParaHox and Hox genes. More generally, the sequencing of additional genomes from organisms at the sparsely sampled base of the animal phylogenetic tree will be key to refining our understanding of the changing complexity of animals over the course of evolution. ■

James O. McNerney is in the Department

of Biology, National University of Ireland, Maynooth, County Kildare, Ireland.

Mary J. O'Connell is in the School of Biotechnology, Dublin City University, Glasnevin, Dublin 9, Ireland.
e-mails: james.o.mcinerney@nuim.ie;
mary.oconnell@dcu.ie

1. Gehring, W. J. *Trends Biochem. Sci.* **17**, 277–280 (1992).
2. Fortunato, S. A. V. *et al. Nature* **514**, 620–623 (2014).
3. Mendivil Ramos, O., Barker, D. & Ferrier, D. E. *Curr. Biol.* **22**, 1951–1956 (2012).
4. Moroz, L. L. *et al. Nature* **510**, 109–114 (2014).
5. Peterson, K. J., Cotton, J. A., Gehling, J. G. & Pisani, D. *Phil. Trans. R. Soc. B* **363**, 1435–1443 (2008).
6. Ryan, J. F. *et al. EvoDevo* **1**, 9 (2010).
7. Scotland, R. W. *BioEssays* **32**, 438–449 (2010).
8. Gazave, E. *et al. PLoS ONE* **5**, e14290 (2010).
9. Dohrmann, M., Janussen, D., Reitner, J., Collins, A. G. & Worheide, G. *Syst. Biol.* **57**, 388–405 (2008).
10. Hemmrich, G. & Bosch, T. C. *BioEssays* **30**, 1010–1018 (2008).

ASTROPHYSICS

Secret ingredient exposed

Astronomers have suspected for some time that magnetic fields are a key ingredient in the accretion of material that surrounds young stars. New observations have just begun to reveal these fields in action. [SEE LETTER P.597](#)

CHRISTOPHER M. JOHNS-KRULL

Most stars are born surrounded by disks of gas and dust, and it is in these disks that planets form. However, the gas and dust that makes up these disks (Fig. 1) does not all go into forming planets. Indeed, much of the disk material slowly falls, or accretes, onto the newly formed star, setting its final mass. For many years, astronomers have studied such accretion disks and

measured the rate at which the material accretes onto young stars, but exactly why such accretion occurs so efficiently has remained elusive. Over the past several years, researchers have started to zero in on a solution in theoretical studies for the evolution of an accretion disk that involves the action of magnetic fields. Unfortunately, until just this year there were no observations of the required magnetic fields in the disks. On page 597 of this issue, Stephens *et al.*¹ report how they have now clearly

detected magnetic fields in one such disk.

The key to understanding how an accretion disk works is to figure out what causes some of the material to spiral inward and eventually merge onto the star. Viscosity in the disk is one possible mechanism. At each point in the disk, the material slightly inside this position moves around the star just a bit faster than the material just outside the given point. This difference in speed results in viscosity, a type of friction, that slightly slows down the material on the inside, causing it to fall somewhat closer to the star. This viscosity, acting at all radii in the disk, leads to the accretion of some of the disk material onto the central star. The greater the viscosity, the more rapid the accretion of material. The trouble is, although astronomers can estimate the viscosity in disks on the basis of the temperature and density of the material there, the result fails by many orders of magnitude to produce the rate of accretion that is observed^{2,3}.

Magnetic fields, through a process known as the magnetorotational instability⁴, have been proposed as a possible means by which disks accrete material. If the material in the



Figure 1 | Accretion disks around newly formed stars. These visible-light images show disks of gas and dust (dark rings) around four infant stars (central bright regions) in the Orion Nebula, a star-forming molecular cloud about 400 parsecs away from Earth. The dust in the disks, which are surrounded by hot gas from the nebula, makes them look dark at visible wavelengths. Magnetic fields threading these disks, now observed directly by Stephens *et al.*¹ in the disk of the young star HL Tau, are probably the main element causing much of the observed disk material to accrete onto the stars. Each square region shown is about 260 billion kilometres across.

disk is threaded through by magnetic fields, then different regions of the disk will be connected to each other by these fields. Although the temperature of the disk is relatively low compared with that of the star, the material is warm enough for a few ions and free electrons to form. The radiation from the star can also cause ion formation. Ions are forced to move with magnetic fields — essentially, they are tied to the field lines. Two ions at slightly different radii in the disk, tied to the same field line, will exert forces on each other that are transmitted through the magnetic field. Effectively, the magnetic fields act as springs holding the ions together. The ion closer to the star should be moving faster in its orbit than the ion slightly farther out, but the field tying them together causes the inner particle to slow down a little and the outer particle to speed up, in exactly the same way as the viscosity described above. However, this magnetorotational-instability mechanism is much better at causing accretion than standard viscosity, and so seems to be a promising solution to the problem of how accretion disks work. But the question has been: are there magnetic fields in the disks around young stars?

Yes, according to Stephens and colleagues' study. The authors detected the magnetic fields in the disk of a young star by using the Combined Array for Research in Millimeter-wave Astronomy (CARMA) to measure the polarization of the light (the preferred direction in which the light's electric field oscillates) that is emitted by the dust in the disk. Dust particles in interstellar space and in the disks of young stars are not perfect spheres — they are somewhat oblong. They also spin, and usually have a small amount of electric charge, like the ions that are present in the gas. The result of their spinning oblong shape and electric charge is that the dust grains generally align their long axes perpendicular to the magnetic field⁵. This then means that they will emit polarized light.

It is more likely that the light emitted by a dust grain will have its electric field aligned with the long axis of the grain. If the dust grains all have random alignment relative to each other, the resulting total emission of radiation will have no preferred orientation and no polarization. By contrast, if most or all of the dust grains are aligned with each other owing to a magnetic field, the total emission will be polarized. This is exactly what Stephens *et al.* saw in the disk surrounding the young star HL Tau. Furthermore, the authors were able to spatially map the disk around HL Tau and measure the polarization at several locations in the disk, making these observations unusually powerful for testing our understanding of accretion disks. Similar observations were also made in 2014 of an even younger star⁶; however, the orientation of its disk and potential confusion from the surrounding environment did not allow as clear a study of the magnetic-field structure as that obtained by

Stephens *et al.* in the disk of HL Tau.

This result is exciting because it strongly suggests that magnetorotational instability may be the long-sought answer to just how accretion disks work. However, this is far from the last word on this astrophysical puzzle. The theoretical models of this phenomenon generally predict that the magnetic fields in the disk should be wrapped around the star, basically following the motion of the disk material as it orbits the star. Unfortunately, this is not exactly what Stephens and colleagues observed. Their polarization measurements of HL Tau's disk indicate that the magnetic fields are all more or less pointing in the same direction instead of wrapping around the star. This is a mystery with no good explanation at this time.

So what does the future hold for accretion-disk physics? A new, more sensitive radio-telescope array than the one used by Stephens and colleagues, the Atacama Large Millimeter/submillimeter Array (ALMA) in Chile, has begun operations. This facility will provide

detailed views of the structure and magnetic fields of disks around young stars. We can hope that this will produce observational results that will also spur new theoretical investigations into the role of magnetic fields in accretion disks. Our understanding of this phenomenon is not complete yet, but it has just taken a dramatic step forward. ■

Christopher M. Johns-Krull is in the Department of Physics and Astronomy, Rice University, Houston, Texas 77005, USA. e-mail: cmj@rice.edu

1. Stephens, I. W. *et al.* *Nature* **514**, 597–599 (2014).
2. Valenti, J. A., Basri, G. & Johns, C. M. *Astron. J.* **106**, 2024–2050 (1993).
3. Gullbring, E., Hartmann, L., Briceño, C. & Calvet, N. *Astrophys. J.* **492**, 323–341 (1998).
4. Balbus, S. A. & Hawley, J. F. *Rev. Mod. Phys.* **70**, 1–53 (1998).
5. Lazarian, A. J. J. *Quant. Spectrosc. Radiat. Transf.* **106**, 225–256 (2007).
6. Rao, R., Girart, J. M., Lai, S.-P. & Marrone, D. P. *Astrophys. J.* **780**, L6 (2014).

This article was published online on 22 October 2014.

STRUCTURAL BIOLOGY

Enzyme–chromatin complex visualized

The structure of an enzyme that is bound to a nucleosome — a protein complex around which DNA is wrapped — reveals how contacts between the two orient the enzyme so that it can modify a specific amino-acid residue. [SEE ARTICLE P.591](#)

JÜRIG MÜLLER & CHRISTOPH W. MÜLLER

In the nucleus, genomic DNA is organized in a highly compact protein–DNA complex called chromatin. The basic building block of chromatin is the nucleosome, an eight-subunit core comprised of two copies of each of the histone proteins H2A, H2B, H3 and H4, around which DNA is wrapped in two helical turns. Many nucleosome-remodelling and histone-modifying enzymes modulate chromatin structure and thereby affect the transcription, replication and repair of genomic DNA. However, despite the importance of these enzymes, we have only a limited understanding of how they interact physically with nucleosomes. On page 591 of this issue, McGinty *et al.*¹ describe the crystal structure of the H2A-modifying module of an enzyme called Polycomb repressive complex 1, bound to a nucleosome. For the first time, the structure reveals how a histone-modifying enzyme recognizes its nucleosomal substrate.

Polycomb repressive complex 1 (PRC1) acts with other protein complexes of the Polycomb group to repress the transcription of many of

the genes that control developmental processes in animals. PRC1 can compact chromatin and inhibit nucleosome remodelling through a non-enzymatic mechanism². But the complex also has 'E3' ubiquitin-ligase enzyme activity, which links a single ubiquitin molecule to an amino-acid residue, lysine (Lys) 119, in H2A³ — a process called monoubiquitylation. The function of this protein modification in transcriptional repression is not well understood. However, studies published this year^{4,5} found that Lys 119 ubiquitylation creates a binding site for one form of the histone-methyltransferase enzyme complex PRC2, and thereby promotes the addition of three methyl groups to H3 at Lys 27, a modification that is crucial for the transcriptional repression of Polycomb target genes.

Six forms of mammalian PRC1 have been identified. All contain two proteins that have RING fingers — domains that bind zinc and are often found in proteins involved in ubiquitylation. One protein is a subunit of E3 ubiquitin ligase, either Ring1B or the closely related protein Ring1A, and the other is either the Polycomb-group RING finger (PCGF)

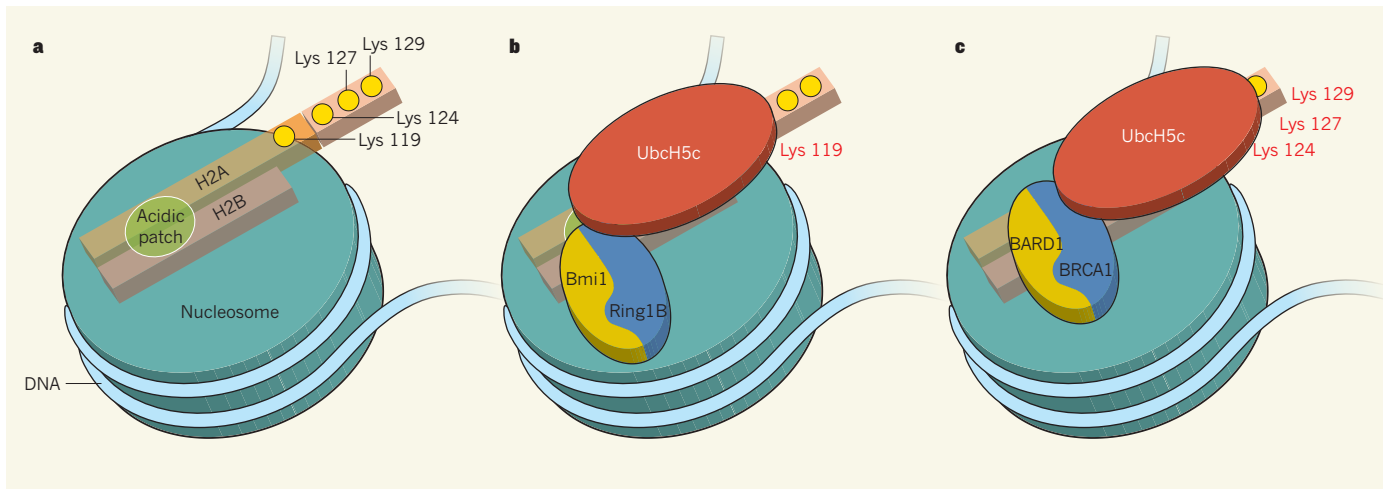


Figure 1 | Enzyme–nucleosome contacts dictate substrate specificity.

a, This schematic diagram shows the disc face of the nucleosome and the acidic patch formed by histone proteins H2A and H2B. The H2A lysine (Lys) amino-acid residues Lys 119, Lys 124, Lys 127 and Lys 129 are indicated by yellow circles. **b**, McGinty *et al.*¹ report that the Ring1B–Bmi1

ubiquitin-ligase enzyme, which is part of the Polycomb repressive complex PRC1, positions another enzyme, UbcH5c, just above Lys 119, enabling monoubiquitylation of H2A at this residue. **c**, A hypothetical model of how the BRCA1–BARD1 E3 ligase positions the UbcH5c enzyme above Lys 124, Lys 127 and Lys 129.

protein Bmi1 or one of five other PCGF-related proteins⁶. *In vitro* studies^{7–9} have established that Ring1B alone shows only poor E3 activity on nucleosomes, but that a Ring1B–Bmi1 complex functions as an E3 ligase, and acts with the ‘E2’ enzyme UbcH5c to monoubiquitylate nucleosomal H2A at Lys 119 (E2 enzymes provide the ubiquitin molecule for ubiquitylation).

Building on previous structural studies^{7–10}, McGinty and colleagues obtained crystals of the human Ring1B–Bmi1–UbcH5c complex bound to a nucleosome. They found that the complex binds to the flat sides of the nucleosome disc (Fig. 1), forming a crescent shape. Half of the crescent, the Ring1B–Bmi1 RING-finger dimer, binds the central histone surface, and the other half, UbcH5c, is positioned directly above the carboxy-terminal end of H2A, which contains the target of the complex, Lys 119. The researchers’ structure does not contain the ubiquitin molecule that is transferred to H2A, but the location of the active-site cysteine amino acid of UbcH5c suggests that only Lys 119 or the adjacent residue, Lys 118, would be close enough to become ubiquitylated.

Although UbcH5c is in contact with the nucleosomal DNA, it barely touches the histones in the nucleosome. Rather, the precise orientation of UbcH5c is controlled by the position of the Ring1B–Bmi1 dimer on the nucleosome surface. Here, a recurring theme of nucleosome recognition comes into play. McGinty and colleagues report that Ring1B uses a positively charged arginine amino-acid residue as an anchor to contact an ‘acidic patch’ formed by H2A and H2B. The acidic patch is implicated in nucleosome–nucleosome interactions¹⁰, and has also been identified as the crucial interaction site in all other nucleosome–protein complexes that have been structurally determined so far. As the authors note,

LANA peptide and RCC1, CENP-C and Sir3 proteins all use an arginine residue to make contact with this acidic patch (see Fig. 4 of the Extended Data for the paper¹).

In the case of the Ring1B–Bmi1 complex, McGinty *et al.* found that other positively charged Ring1B residues also interact with the acidic patch, extending the Ring1B–nucleosome interaction surface. Nevertheless, a mutational analysis performed by the authors shows that the arginine anchor is primarily responsible for binding. Mutating this residue resulted in a 50-fold decrease in binding affinity and H2A monoubiquitylation, whereas mutating adjacent positively charged residues had less-dramatic effects. The Bmi1 subunit seems to make a lesser contribution to nucleosome binding affinity, although contact with the amino-terminal helix of H3 is crucial for binding, and presumably also for fixing the orientation of the complex on the nucleosome surface.

What predictions can be made on the basis of the authors’ structure? The BRCA1–BARD1 RING-finger dimer is an E3 ligase that is related to Ring1B–Bmi1, but BRCA1–BARD1 ubiquitylates Lys 124, Lys 127 or Lys 129, which are located nearer to the C terminus of H2A than residues modified by Ring1B–Bmi1 (ref. 11). McGinty and colleagues found that, like Ring1B–Bmi1, the BRCA1–BARD1 dimer seems to act with UbcH5c, and that the nucleosome-binding region of Ring1B and the corresponding region in BRCA1 are evolutionarily conserved — a confirmation of findings from another study¹¹. Moreover, they show that mutation of the acidic patch eliminates BRCA1-mediated H2A monoubiquitylation in nucleosomes. These observations make it likely that BRCA1–BARD1 binds the acidic patch in a similar fashion to Ring1B–Bmi1. However, the overall geometry

probably differs, such that only H2A residues Lys 124, Lys 127 or Lys 129 can become modified (Fig. 1).

Only some of the six different Ring1B–PCGF complexes seem to be able to monoubiquitylate H2A at Lys 119 *in vivo*⁵. This raises the possibility that one or more subunits of the other PRC1-type complexes occupy different interfaces on the E3 ligase or on the nucleosome in a way that precludes either the binding of an E2 enzyme or the transfer of ubiquitin to H2A. Undoubtedly, McGinty and colleagues’ study will serve as the foundation for future structural investigations that will resolve how other E3 ligases or larger PRC1 assemblies interact with nucleosomes. ■

Jürg Müller is at the Max Planck Institute of Biochemistry, Laboratory of Chromatin Biology, Martinsried 82152, Germany.

Christoph W. Müller is at the European Molecular Biology Laboratory, Structural and Computational Biology Unit, Heidelberg 69117, Germany. e-mails: muellerj@biochem.mpg.de; cmueller@embl.de

1. McGinty, R. K., Henrici, R. C. & Tan, S. *Nature* **514**, 591–596 (2014).
2. Simon, J. A. & Kingston, R. E. *Mol. Cell* **49**, 808–824 (2013).
3. Wang, H. *et al.* *Nature* **431**, 873–878 (2004).
4. Kalb, R. *et al.* *Nature Struct. Mol. Biol.* **21**, 569–571 (2014).
5. Blackledge, N. P. *et al.* *Cell* **157**, 1445–1459 (2014).
6. Gao, Z. *et al.* *Mol. Cell* **45**, 344–356 (2012).
7. Buchwald, G. G. *et al.* *EMBO J.* **25**, 2465–2474 (2006).
8. Li, Z. *et al.* *J. Biol. Chem.* **281**, 20643–20649 (2006).
9. Bentley, M. L. *et al.* *EMBO J.* **30**, 3285–3297 (2011).
10. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. *Nature* **389**, 251–260 (1997).
11. Kalb, R., Mallery, D. L., Larkin, C., Huang, J. T. J. & Hiom, K. *Cell Rep.* **8**, 999–1005 (2014).

Radicals promote magnetic gel assembly

Engineering complex tissues requires high-throughput, three-dimensional patterning of materials and cells. A method to assemble small gel components using magnetic forces from encapsulated free radicals could be just the ticket.

CHRISTOPHER B. RODELL
& JASON A. BURDICK

The ability to control the micrometre-scale assembly of water-swollen polymer networks known as hydrogels has great potential in many disciplines. It might, for example, advance the development of robots from soft materials, of lab-on-a-chip diagnostic devices and of tissue-engineering constructs. Writing in *Nature Communications*, Tasoglu *et al.*¹ describe just such an ability. They have used paramagnetic forces (the attraction of certain materials to an external magnetic field) to guide the two- and three-dimensional assembly of micrometre-sized hydrogel subunits called microgels. Like the blocks in the game Tetris, these subunits can be turned and manipulated to assemble a desired structure. Because the technique can simultaneously accommodate several types and compositions of material in the presence of cells, it is highly appealing for tissue-engineering applications.

Tissue engineering depends on a precise interplay between cells and their surroundings, and on the hierarchical organization of three-dimensional (3D) tissues. Typical behaviours of cells, such as proliferation, differentiation and migration, are controlled in part through spatially and temporally distributed signals from the local surrounding matrix (the cell niche), as well as by soluble signals that diffuse through it.

Researchers have used a variety of fabrication and patterning approaches to introduce these signals into synthetic materials in an attempt to understand and recapitulate cell–matrix interactions on multiple size scales.

Scaffolds for tissue engineering are produced by either encapsulating cells within, or seeding them into, materials. Historically, such scaffolds had uniform composition and did not replicate the complexity of, for example, developing tissues or the wound-healing response. Many researchers have therefore expanded their methods to incorporate top-down design — the production of a uniform material that is subsequently patterned. In one such approach (photopatterning), spatially controlled exposure of materials to light generates changes in biomolecule presentation², mechanical cues³, or both⁴, to control cell behaviour through interactions with the matrix.

Although top-down approaches provide some control over uniform cell distributions in a single material, many tissues are highly non-uniform in matrix composition and cell distribution. Bottom-up construction methods, in which tissue constructs are assembled from smaller components, may thus be better suited for replicating biological complexity. 3D printing, in which the direct deposition of material creates precise 3D structures, embodies this strategy. Recent advances in technology have allowed 3D printing of

tissues through the deposition of cellular aggregates or cell-laden materials⁵. However, these processes still rely on repeated layer-by-layer deposition — a method⁶ reported in 1986 — and thus require long fabrication times that currently limit construct size.

Assembly of preformed microgel components may offer an alternative method, because this would allow rapid assembly and therefore the potential formation of tissue constructs at clinically relevant scales. So far, the assembly of such components has been achieved largely through passive thermodynamic processes⁷ or direct serial manipulation⁸. To improve efficiency, attention has recently turned to methods that drive assembly through external forces. In their approach, Tasoglu and colleagues used a chemical component that can be easily incorporated into nearly any hydrogel through diffusion (Fig. 1), and which contains a stable free radical (an unpaired electron). When exposed to an external magnetic field, the encapsulated radicals exert a paramagnetic force, which enables hydrogel components to be easily and rapidly manipulated into a desired location.

The authors show that the magnetization process can be temporally controlled and is compatible with cell viability and proliferation. Moreover, the radical-containing hydrogels are magnetically responsive for up to 24 hours after encapsulation of the radicals. However, the materials can be quickly rendered magnetically inactive by treatment with the antioxidant vitamin E, which acts as a radical scavenger. The researchers observed that vitamin E treatment tended to improve the viability and proliferation of encapsulated cells when compared with untreated gels.

To demonstrate the general applicability of their method, Tasoglu *et al.* have used it to form constructs containing many material types, densities and porosities. They also show that, following the assembly of hydrogel components into a desired configuration, another hydrogel may be added as a precursor

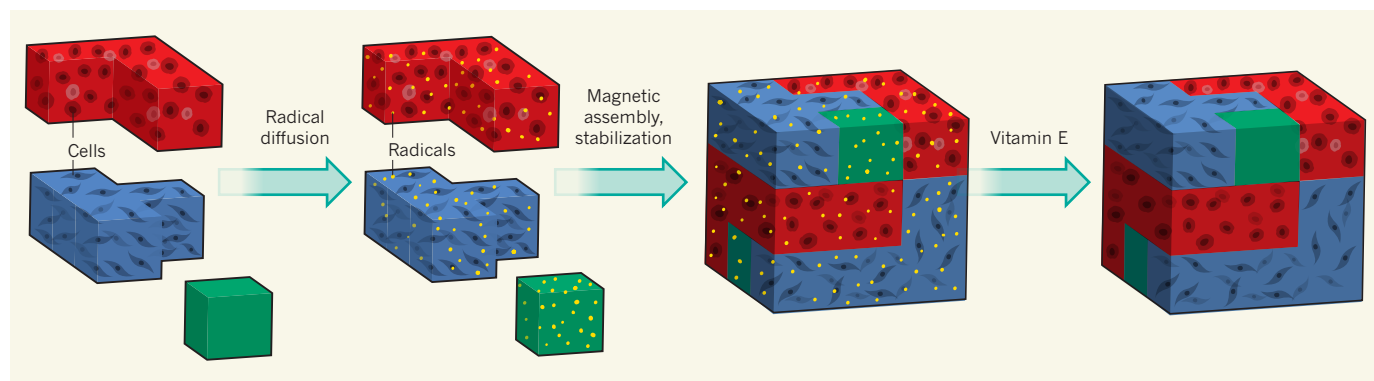


Figure 1 | Radical magnetic microgel assembly. Tasoglu *et al.*¹ describe a method for assembling micrometre-sized hydrogels (water-swollen polymer networks) into larger constructs with defined compositions and structures. The microgels are formed with or without encapsulated cells, and then a stable radical compound (yellow circles) is allowed to diffuse into them. The radicals generate

forces in the presence of an external magnetic field, driving microgel motion and assembly. The resulting constructs are stabilized by the addition of a hydrogel precursor that polymerizes on exposure to ultraviolet light. Finally, the radicals are quenched (converted to non-radical products) by the addition of vitamin E, to prevent them from damaging the encapsulated cells.

solution that is solidified by ultraviolet-light-initiated polymerization — a process that stabilizes the assembled construct. Although not demonstrated, the method should also be able to control the assembly of components that provide different stiffnesses, biochemical cues or other biologically relevant features. Such generality is important for tissue-engineering applications.

Although tissues are inherently 3D, most microgel-assembly methods have been limited to manipulation in two dimensions. By contrast, the authors demonstrate that they can levitate many microgel components, and simultaneously drive the assembly of truly 3D structures of up to a few millimetres across by arranging external magnets in appropriate orientations. Massively parallel assembly in three dimensions will be needed to achieve the formation of larger tissue constructs.

In principle, the new method addresses several major challenges in tissue engineering.

However, practical applications will be realized only when guided assembly of constructs can occur on the tissue scale. Incorporating molecular recognition between hydrogel subunits⁹, to automate the stabilization of 3D assemblies, might help to achieve this. Moreover, it will probably be necessary to pattern multiple cell types and to introduce a means of perfusing thicker constructs to provide oxygen and nutrients for long-term cell viability.

Nevertheless, Tasoglu and colleagues' method is likely to stimulate the growing interest in guided micro-assembly. The use of external driving forces, such as magnetism, allows previously inaccessible levels of parallel assembly and might therefore propel this bottom-up approach to clinical use. Before then, however, the method will probably have a more direct impact on the formation of smaller assemblies, such as lab-on-a-chip devices for diagnostic applications and organized co-culture systems for studies of cell–cell interactions. ■

Christopher B. Rodell and Jason A. Burdick are in the Department of Bioengineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia 19104, USA.
e-mail: burdick2@seas.upenn.edu

1. Tasoglu, S., Yu, C. H., Guven, G. S., Vural, T. & Demirci, U. *Nature Commun.* **5**, 4702; <http://dx.doi.org/10.1038/ncomms5702> (2014).
2. DeForest, C. A., Polizzotti, B. D. & Anseth, K. S. *Nature Mater.* **8**, 659–664 (2009).
3. Khetan, S., Katz, J. S. & Burdick, J. A. *Soft Matter* **5**, 1601–1606 (2009).
4. Gramlich, W. M., Kim, I. L. & Burdick, J. A. *Biomaterials* **34**, 9803–9811 (2013).
5. Murphy, S. V. & Atala, A. *Nature Biotechnol.* **32**, 773–785 (2014).
6. Hull, C. W. US patent 4575330 A (1986).
7. Du, Y., Lo, E., Samsher, A. & Khademhosseini, A. *Proc. Natl Acad. Sci. USA* **105**, 9522–9527 (2008).
8. Tasoglu, S., Diller, E., Guven, S., Sitti, M. & Demirci, U. *Nature Commun.* **5**, 3124; <http://dx.doi.org/10.1038/ncomms4124> (2014).
9. Harada, A., Kobayashi, R., Takashima, Y., Hashidzume, A. & Tamaguchi, H. *Nature Chem.* **3**, 34–37 (2011).

CARDIAC BIOLOGY

Cell plasticity helps hearts to repair

Fibroblast cells are known as key players in the repair of damaged heart structures. New findings show that injury also induces fibroblasts to become endothelial cells, helping to mend damaged blood vessels. SEE ARTICLE P.585

TORU MIYAKE & RAGHU KALLURI

Heart attacks caused by a blockage in the coronary artery induce severe injury to cardiac muscle cells, leading to cell dysfunction and death. The damage elicits repair and regenerative responses that provoke the removal of dying cells and cell debris, recruit immune cells and initiate the formation of new blood vessels to recover blood supply. Fibroblast cells play a central part in this repair response. In this issue, Ubil *et al.*¹ (page 585) describe how the plasticity of cardiac fibroblasts contributes to this process, by showing that fibroblasts can, in response to the activity of the transcription factor p53, convert into the endothelial cells that line the interior surface of blood vessels.

The heart consists of myocytes (muscle cells) and non-myocytes, which include cardiac fibroblasts and endothelial cells. The fibroblasts produce growth factors and extracellular matrix (ECM) proteins to maintain proper cardiac architecture, contraction and function². They also interact with endothelial cells and myocytes to aid angiogenesis (blood-vessel formation) and maintain physiological homeostasis³. Following heart damage,

cardiac fibroblasts are activated to produce ECM proteins and soluble factors to compensate for structural defects, contain the spread of damage, reinforce cardiac stiffness and prevent cardiac rupture. These activities, collectively referred to as fibrosis, aid in remodelling the heart musculature. Controlled fibrosis is crucial for restoring cardiac function after injury. However, excessive fibrosis is considered a pathological process that can lead to adverse effects, including reduced cardiac stiffness (diastolic dysfunction) and irregular electrical connectivity (arrhythmia).

Although endothelial cells in blood vessels are typically thought of as terminally differentiated cells, they can take on the characteristics of mesenchymal cells^{4,5}, which are generally mobile cells surrounded by interstitial ECM proteins. During this endothelial-to-mesenchymal transition (EndMT), the endothelial cells lose the tight junctions that hold neighbouring cells together, and gain the ability to move, produce ECM proteins and contribute to excessive fibrosis, while also depleting functional capillaries and the endocardium tissue layer. EndMT is induced in cardiac endothelial cells by signalling pathways that depend on the growth factor

TGF- β 1 and that can be reversed by the activity of the protein BMP7 (ref. 4).

Now, Ubil and colleagues show that, following acute cardiac injury, fibroblasts (which belong to the mesenchymal cell lineage) can undergo a reverse conversion — from mesenchymal to endothelial cells (MEndT) — and become components of blood vessels (Fig. 1). To study this plasticity, they used mice in which cells that gain or lose expression of cell-type-specific markers can be tracked by fluorescence, a technique called genetic fate mapping.

The authors induced ischaemia-reperfusion injury by blocking the coronary artery and then restoring blood flow in the hearts of these mice. Three days later, they found that 35% of fibroblasts in the injury zone expressed the endothelial marker VECAD and were located in the interior of the vessel. Of these fibroblast-derived endothelial cells, 41% took up acetylated low-density lipoprotein, which is suggestive of endothelial-cell functionality. Most of the cells undergoing this transition expressed the fibroblast markers Col1 α 2 or FSP1, whereas very few expressed α SMA, a marker shared by a subset of cardiac fibroblasts (myofibroblasts) and the mesenchymal cells generated through EndMT (Fig. 1). This finding highlights a functional heterogeneity of recruited fibroblasts in injured cardiac tissue with respect to their plasticity. Understanding this heterogeneity will require future studies using fate mapping of myofibroblasts⁶.

Ubil and co-workers also found that the fibroblast-derived endothelial cells in the mice express increased levels of p53, a transcription factor known for its multiple functions, including regulation of the cell cycle, apoptotic cell death and DNA repair. To investigate the involvement of the p53 signalling pathway

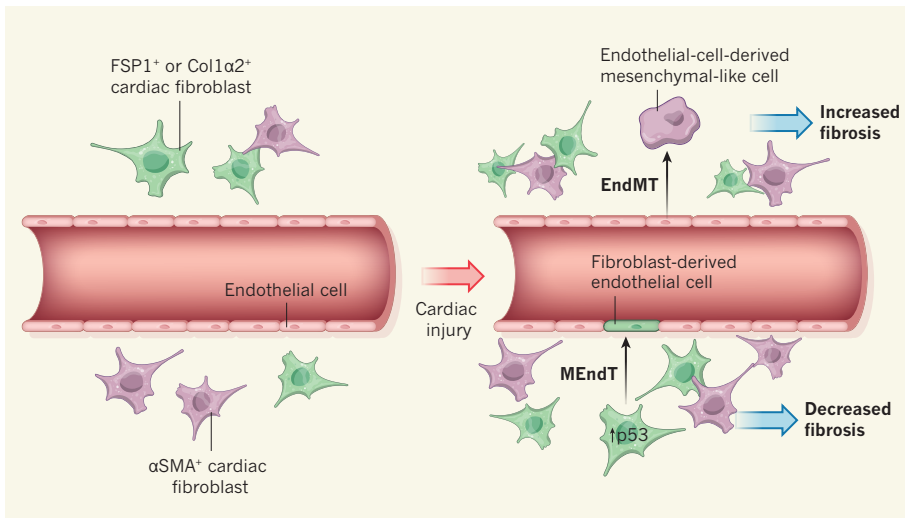


Figure 1 | Cardiac-cell conversions. The heart's response to injury involves the proliferation and activation of cardiac fibroblasts (a type of mesenchymal cell). Although this fibrosis is essential for repair, an excessive response can lead to cardiac dysfunction, and evidence is building that this balance is regulated by transitions between cell types. Previous studies^{4,5} have shown that endothelial cells lining blood vessels can convert to mesenchymal-like cells in a process called the endothelial-to-mesenchymal transition (EndMT). These cells, which express α SMA, a marker shared by a subset of cardiac fibroblasts (myofibroblasts), contribute to increased fibrosis. Now, Ubil *et al.*¹ show that some cardiac fibroblasts (distinguished by the markers FSP1 or Col1 α 2) can undergo a reverse, mesenchymal-to-endothelial transition (MEndT). This conversion is induced by activity of the transcription factor p53, and leads to reduced fibrosis and increased blood-vessel formation.

in the MEndT program, the authors studied cardiac fibroblasts cultured *in vitro* under serum deprivation — a stress condition that induces upregulation of p53. The cells formed tubular structures reminiscent of endothelial-cell architecture and expressed endothelial markers, including VECAD and transcription factors such as HoxA9 and HoxD3. However, the fibroblasts did not generate these tubules without serum deprivation, even when p53 was artificially overexpressed, suggesting that p53 expression alone is not sufficient to induce MEndT and that other signalling pathways are involved in launching this program.

The authors went on to show that treating mice with the small molecule RITA, which enhances p53 signalling, for three days after cardiac injury increased the number of fibroblast-derived endothelial cells. The treatment also accelerated angiogenesis and decreased cardiac fibrosis, leading to an improvement in cardiac function. These *in vitro* and *in vivo* findings suggest that p53 expression in fibroblast-derived endothelial cells has a key role in the recovery of cardiac function following injury.

Administration of other p53-activator molecules has previously been shown to impair angiogenesis, increase apoptotic cell death and cause dysfunctional muscle contraction in a mouse model of a condition called cardiac hypertrophy⁷. Conversely, Ubil *et al.* show no increase in myocyte p53 expression or the number of apoptotic cells after RITA treatment compared with control mice. These contrasting results could be explained by the

different cardiac-injury models and drugs used, which will need to be considered in investigations of the potential clinical application

IMMUNOLOGY

Starve a fever, feed the microbiota

A study finds that the cells lining the gut are modified in response to systemic infection, increasing the host's tolerance to infection in a manner that is dependent on the microorganisms that inhabit the gut. [SEE LETTER P.638](#)

SETH RAKOFF-NAHOUM & LAURIE E. COMSTOCK

Mammals have a mutualistic relationship with the consortium of microorganisms that inhabit their intestines, collectively known as the gut microbiota. The microbiota profits from an environment that is rich in dietary and host-derived nutrients, and provides its host with numerous benefits, ranging from immune and digestive capabilities to resistance to pathogen colonization. On page 638 of this issue, Pickard *et al.*¹ describe another facet of this mutualism. They show that in a model of systemic infection, the host modifies the surface molecules of the epithelial cells lining its intestine, which increases host fitness in a microbiota-dependent manner.

of p53-inducing drugs for targeting cardiac fibrosis.

Nevertheless, Ubil and colleagues' study provides insight into the cardiac-repair process and highlights potential new therapeutic strategies. It also adds to the debate of whether the term 'terminally differentiated' in adult tissue might be too confining when cellular plasticity is rampant and seemingly functional in situations of disease or injury, just as in embryonic development. ■

Toru Miyake and Raghu Kalluri are in the Department of Cancer Biology, Metastasis Research Center, University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. R.K. is also in the Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, and in the Department of Bioengineering, Rice University, Houston.

e-mail: rkalluri@mdanderson.org

1. Ubil, E. *et al.* *Nature* **514**, 585–590 (2014).
2. Souders, C. A., Bowers, S. L. & Baudino, T. A. *Circ. Res.* **105**, 1164–1176 (2009).
3. Kakkar, R. & Lee, R. T. *Circ. Res.* **106**, 47–57 (2010).
4. Zeisberg, E. M. *et al.* *Nature Med.* **13**, 952–961 (2007).
5. Von Gise, A. & Pu, W. T. *Circ. Res.* **110**, 1628–1645 (2012).
6. LeBleu, V. S. *et al.* *Nature Med.* **19**, 1047–1053 (2013).
7. Sano, M. *et al.* *Nature* **446**, 444–448 (2007).

This article was published online on 15 October 2014.

showed that systemic LPS administration leads to rapid and sustained fucosylation of glycoproteins throughout the small intestine. These fucosylated glycoproteins are released into the intestinal lumen where the fucose residues are liberated and consumed by the resident bacteria in the colon.

Pickard *et al.* found that systemic administration of LPS induces release of the cell-signalling molecule interleukin-23 (IL-23) from dendritic cells of the innate immune system, which in turn activates innate lymphoid cells to release IL-22. This induces expression of a specific fucosyltransferase gene, *Fut2*, in epithelial cells of the small intestine, resulting in addition of fucose to their surface molecules (Fig. 1).

Next, the authors demonstrated that mice lacking *Fut2* recovered weight more slowly after LPS-induced anorexia than did control mice, indicating that under these conditions, fucosylation is beneficial to the host. Furthermore, this benefit is dependent on the presence of the microbiota, because germ-free or microbiota-depleted mice also showed impaired weight gain in LPS-induced anorexic conditions, despite the fact that their epithelial cells were fucosylated.

Finally, to investigate whether fucosylation improves the host's fitness during infection, Pickard and co-workers infected *Fut2*-deficient or control mice with the intestinal bacterial pathogen *Citrobacter rodentium*. They showed that inducible fucosylation in response

to LPS is crucial for limiting both proliferation of colonic cells (hyperplasia), which is indicative of tissue damage, and weight loss during infection. This benefit is probably not due to increased resistance to infection, because the total levels of *C. rodentium* in the small intestine were similar in mice with or without *Fut2*. Because the damage to the host is independent of pathogen burden, it is probable that this microbiota-dependent effect improves host fitness through disease tolerance^{2,3}.

Recently, Goto *et al.*⁴ identified a molecular pathway involved in microbiota-triggered fucosylation on the small intestinal epithelium in normal, 'steady-state' conditions, which shares similarities with the systemic-LPS-inducible pathway described by Pickard and colleagues. Both studies provide evidence that fucosylation is dependent on a subset of IL-22-producing innate lymphoid cells. Similar to Pickard *et al.*, Goto and colleagues show that fucosylation is important for host defence against an intestinal bacterial pathogen, *Salmonella typhimurium*. However, in this second study, fucosylation inhibited bacterial invasion of intestinal tissues and, therefore, host protection seems to be due to pathogen resistance rather than disease tolerance.

An apparent discrepancy in Pickard and colleagues' study is that they did not detect fucose on small intestinal epithelial cells in mice that had not been treated systemically with LPS, whereas Goto and colleagues'

and other studies^{5,6} showed steady-state fucosylation in the distal small intestine of mice. This may be accounted for by differences in the intestinal microbiota of the mice used in each study. Pickard and co-workers used mice from the Jackson Laboratory whose microbiota does not contain segmented filamentous bacteria⁷, which Goto *et al.* identify as potent inducers of fucosylation. It will be interesting to determine the factors derived from the microbiota (and, in particular, from the segmented filamentous bacteria) that mediate intestinal fucosylation, which may act independently of Toll-like receptors⁷. It is likely that the steady-state fucosylation of the distal small intestine observed by Goto and colleagues would be augmented if mice were injected with LPS, because the fucosylation reported in the current study¹ occurred not only in this distal region, but throughout the small intestine.

Perhaps the most exciting aspect of Pickard and colleagues' work is what it adds to our understanding of beneficial interactions between the host and its microbiota. Although there is a growing list of host factors, cells and conditions that shape the microbiota, little is known about whether, under certain conditions, the host can select for a beneficial microbiota that increases the host's own fitness. This study suggests a mechanism by which such selection may occur. An outstanding question is how the microbiota contributes to disease tolerance.

Furthermore, the benefit to the host of fucosylation of intestinal epithelial cells shown by the authors sheds light on potential trade-offs with the known negative consequences of fucosylation. These include the use of fucosylated receptors by viruses⁸, and the use of liberated fucose by bacterial pathogens⁹. The complex interactions between the host, its microbiota and pathogens, especially regarding fucose, will continue to be a stimulating area for investigation. ■

Seth Rakoff-Nahoum and Laurie E. Comstock are in the Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115, USA. S. R.-N. is also in the Department of Medicine, Boston Children's Hospital.
e-mail: lcomstock@rics.bwh.harvard.edu

- Pickard, J. M. *et al.* *Nature* **514**, 638–641 (2014).
- Ayres, J. S. & Schneider, D. S. *Annu. Rev. Immunol.* **30**, 271–294 (2012).
- Medzhitov, R., Schneider, D. S. & Soares, M. P. *Science* **335**, 936–941 (2012).
- Goto, Y. *et al.* *Science* <http://dx.doi.org/10.1126/science.1254009> (2014).
- Urnesaki, Y., Tohyama, K. & Mutai, M. *J. Biochem.* **90**, 559–561 (1981).
- Bry, L., Falk, P. G., Midtvedt, T. & Gordon, J. I. *Science* **273**, 1380–1383 (1996).
- Ivanov, I. I. *et al.* *Cell* **139**, 485–498 (2009).
- Lindesmith, L. *et al.* *Nature Med.* **9**, 548–553 (2003).
- Ng, K. M. *et al.* *Nature* **502**, 96–99 (2013).

This article was published online on 1 October 2014.

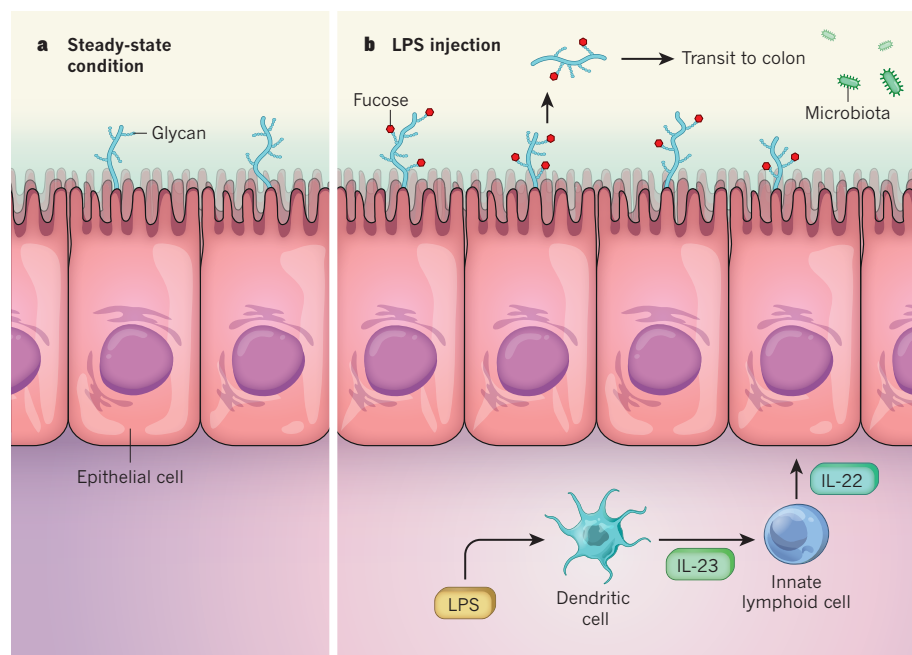


Figure 1 | Making a meal of disease. **a**, In steady-state conditions, fucosylation (the addition of fucose groups to the glycan chains of molecules secreted to the cell surface) is not induced in the small intestine. **b**, Pickard *et al.*¹ injected mice with lipopolysaccharide (LPS), which mimics systemic infection, including diminished food consumption. They report that this leads to the release of the cell-signalling molecule interleukin-23 (IL-23) from dendritic cells of the immune system. IL-23 drives IL-22 production from innate lymphoid cells, inducing expression of the gene encoding the fucosyltransferase enzyme *Fut2* (not shown) in the epithelial cells of the small intestine. *Fut2* catalyses fucosylation, and fucose-containing molecules are then released into the gut lumen where they transit to the colon. Here, fucose is liberated and consumed by members of the microbiota. This process correlates with an increased tolerance to infection in the host.

Three new Jurassic euharamiyidan species reinforce early divergence of mammals

Shundong Bi^{1,2}, Yuanqing Wang¹, Jian Guan³, Xia Sheng⁴ & Jin Meng⁵

The phylogeny of Allotheria, including Multituberculata and Haramiyida, remains unsolved and has generated contentious views on the origin and earliest evolution of mammals. Here we report three new species of a new clade, Euharamiyida, based on six well-preserved fossils from the Jurassic period of China. These fossils reveal many craniodental and postcranial features of euharamiyidans and clarify several ambiguous structures that are currently the topic of debate. Our phylogenetic analyses recognize Euharamiyida as the sister group of Multituberculata, and place Allotheria within the Mammalia. The phylogeny suggests that allotherian mammals evolved from a Late Triassic (approximately 208 million years ago) *Haramiyavia*-like ancestor and diversified into euharamiyidans and multituberculates with a cosmopolitan distribution, implying homologous acquisition of many craniodental and postcranial features in the two groups. Our findings also favour a Late Triassic origin of mammals in Laurasia and two independent detachment events of the middle ear bones during mammalian evolution.

Mammalia Linnaeus, 1758

Allotheria Marsh, 1880

Euharamiyida (new clade)

Shenshou lui gen. et sp. nov. Bi, Wang, Guan, Sheng and Meng

Eymology. *Shen*, from pinyin of the Chinese word, meaning deity, divinity or cleaver; *shou*, from pinyin of the Chinese word for creature, animal or beast; specific name after Lu Jianhua, the collector of the holotype. **Holotype.** A nearly complete skeleton from an adult individual (LDN HMF2001, Lande Museum of Natural History, Tangshan, Hebei Province, China) (Figs 1a and 2d, and Extended Data Figs 1a and 2). Three specimens are referred to as the paratypes (see Supplementary Information, section A).

Paratypes. Three specimens are referred to as the paratypes (see section A of Supplementary Information).

Locality and horizon. The Tiaojishan Formation, Daxishan site of Linglongta, Jianchang County, Liaoning Province, China; the locality was dated as approximately 160 million years ago (within the Oxfordian) (see Supplementary Information, section C, for age constraint).

Diagnosis. Medium-sized euharamiyidan with an estimated body mass of 300 g (Supplementary Information, section D). Dental formula $I^1-C^0-P^2-M^2/I_1-C_0-P_1-M_2$ (I, incisor; C, canine; P, premolar; M, molar; superscript, upper teeth; subscript, lower teeth). The only pair of upper incisors (I^2) are in contact so that a facet is present on the medial side of each tooth; I^2 with two cusps. The mesial upper premolar (P^3) small (not basined) and the ultimate premolar (P^4) not significantly larger than upper molars. Upper molars with two main cusps in the buccal row, separated by a low ridge; three cusps in the lingual one, of which the penultimate cusp (B2) is the largest. P_4 sub-molariform with a main mesiolingual cusp (a1) and a long basined heel with two rows of cusps; M_1 with three cusps in each cusp row, M_2 with four cusps in the lingual and three in the labial row (Fig. 2d; see Supplementary Information, section A, for differential comparisons). The terminology we use to designate cusps in allotherian teeth is as follows: for upper teeth, the buccal (labial or lateral) cusp row is A, lingual row is B, cusps are numbered from the

distal end; lower teeth, the lingual cusp row is a, the buccal row is b (lower case), numbering starts from the mesial (anterior) end.

Eleutherodontidae Kermack *et al.*, 1998

Xianshou gen. nov. Wang, Meng, Bi, Guan and Sheng

Eymology. *Xian*, from pinyin of the Chinese word meaning celestial being or immortal.

Locality and horizon. Same as *Shenshou*.

Diagnosis. Dental formula: $I^2-C^0-P^2-M^2/I_1-C_0-P_1-M_2$; upper and lower molars ovoid in outline, with a shallow central basin. Differ from *Sineleutherus* in having three well-separated cusps of I^2 , non-molariform P_4 with a hypertrophic mesiolingual cusp (a1) and a weakly basined heel. Differ from *Eleutherodon* in the ovoid upper molars by absence of third row of cusps, and lack of cuspules and transverse ridges in the central basin. Differ from *Shenshou* and *Arboroharamiya* in having an extra-small I^1 , ovoid upper and lower molars, more distally positioned distobuccal cusp (A1) on P^4 and M^1 ; a hypertrophic mesiolingual cusp (a1) on lower molars (Fig. 2e, f; see Supplementary Information, section A, for differential comparisons, and Extended Data Figs 5 and 6 for additional figures).

Xianshou linglong sp. nov. Wang, Meng, Bi, Guan and Sheng

Eymology. *linglong*, from pinyin of the Chinese word, meaning 'exquisite', and also after the town name Linglongta where the specimen came from.

Holotype. A skeleton preserved on a split slab of laminated siltstone (IVPP V16707A-B, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing, China; Figs 1b and 2e, and Extended Data Figs 1b and 5).

Diagnosis. Estimated body mass of 83 g. I^2 with three cusps; I^3 extremely small and budlike. Upper molars with sharp cusps and stronger and sharper ridges (flutings) than those in *Shenshou* and *X. songae*; M^1 with two main cusps at the mesial and distal end and connected by a low and sharp ridge in each cusp row; an enlarged and more distally extended distobuccal cusp (A1) on P^4 and M^1 ; P_4 enlarged with a hypertrophic mesiolingual cusp (a1) and a small talonid heel; cusp a1 of lower molar procumbent, projecting mesially beyond the crown.

¹Key Laboratory of Vertebrate Evolution and Human Origins of Chinese Academy of Sciences, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, Beijing 100044, China. ²Department of Biology, Indiana University of Pennsylvania, Indiana, Pennsylvania 15705, USA. ³Beijing Natural History Museum, 126 Tianqiao Street, Dongcheng District, Beijing 100050, China. ⁴Paleontological Museum of Liaoning, Shenyang Normal University, Shenyang, Liaoning 110034, China. ⁵Division of Paleontology, American Museum of Natural History, Central Park West at 79th Street, New York, New York 10024, USA.

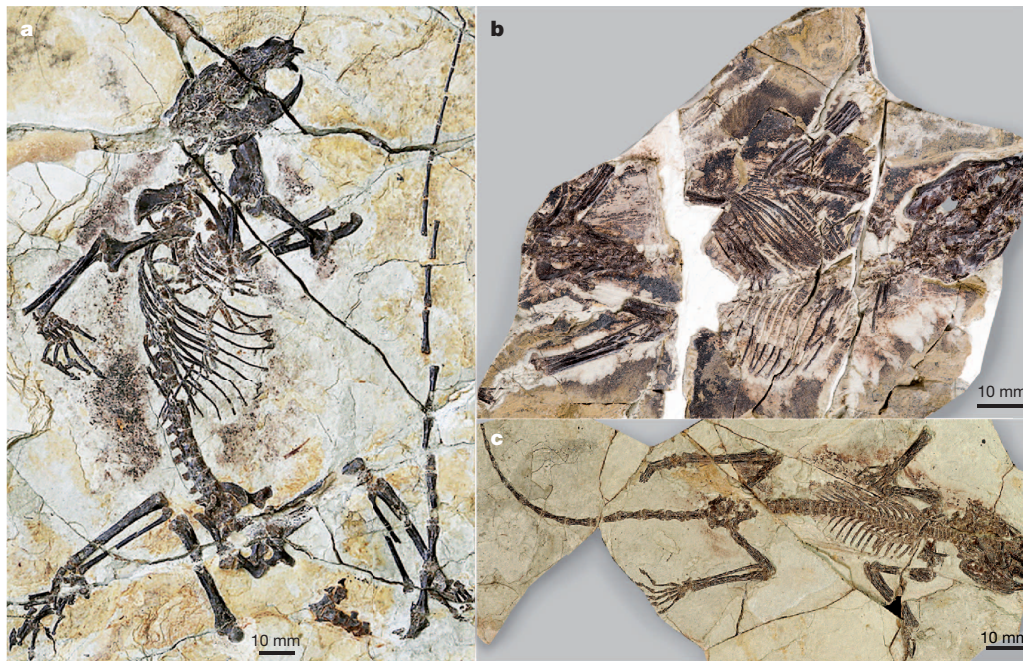


Figure 1 | The holotypes of three euharamiyidan species. **a**, Holotype (LDNHMF2001) of *Shenshou lui*. **b**, Holotype (IVPP V16707A) of *Xianshou linglong*. **c**, Holotype specimen (BMNH-PM003253) of *Xianshou songae*. See Extended Data Fig. 1 for interpretations of the skeletal features.

Xianshou songae sp. nov. Meng, Guan, Wang, Bi and Sheng
Etymology. The specific name is after Rufeng Song, the collector of the holotype specimen.

Holotype. A skeleton preserved partial skull, mandible, and most of the postcranial skeleton (BMNH-PM003253, Beijing Natural History Museum, China) (Figs 1c and 2f, and Extended Data Fig. 6).

Diagnosis. A small euharamiyidan with an estimated body mass of 40 g. Differ from *X. linglong* by having significantly smaller body size and in having P^4 more transversely oval, cusp A1 on P^4 and M^1 proportionally smaller and less distally extended. The lingual row of M^1 bears three cusps, of which the middle (B2) is the largest; lower molars anteroposteriorly short with a vertical hypertrophic mesiolingual cusp (a1) and low buccal cusps.

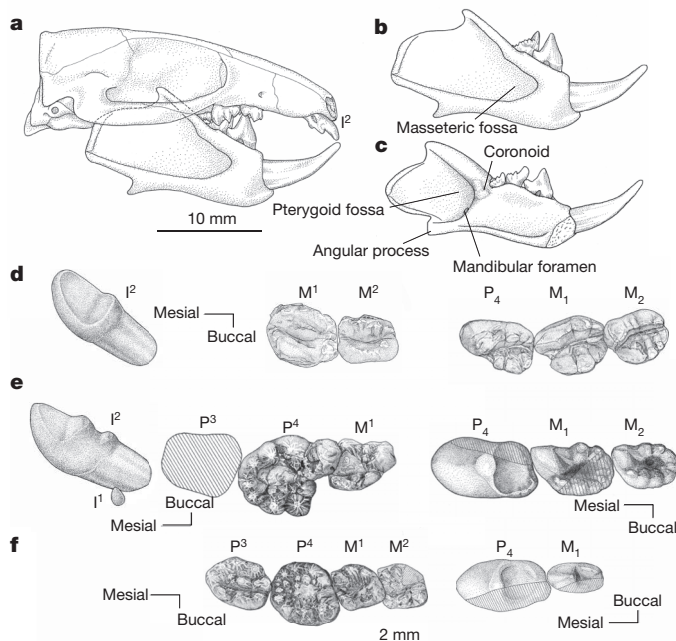


Figure 2 | Teeth, skull and mandibles of euharamiyidans. **a**, Reconstruction of the skull and jaws of *Xianshou linglong* (the small I^1 is on the medial side of the large I^2). **b**, The labial view of the lower jaw of *Xianshou linglong*. **c**, Lingual view of the lower jaw of *Xianshou linglong*. Part of the ascending ramus and coronoid bone have not been preserved and are reconstructed based on *Shenshou lui*. **d**, Right I^2 , $M^{1,2}$, and left P_4 - M_2 of *Shenshou lui* in occlusal view. **e**, Left I^1 , I^2 , P^4 , M^1 and P_4 - M_2 of *Xianshou linglong* in occlusal view. Dashed lines represent the alveolus of P^3 . **f**, Right P^3 - M^2 , and P_4 and M_1 of *Xianshou songae* in occlusal view. See text for description of the skull and mandible. Dental formula are $I^{1?}-C^0-P^2-M^2/I_1-C_0-P_2-M_2$ (the upper incisor is not preserved in this species). *Shenshou* and *Xianshou* have a multicusped I^2 and a greatly enlarged I_2 that is fully covered with enamel. There is no upper or lower canine. There are one lower and two upper premolars, and two upper and lower molars. Small enamel ridges (flutings) vary in size and density on cheek tooth cusps and basins. Upper premolars are broadly basined with main cusps located peripherally. Lower premolar has an enlarged a1, which is hypertrophic in *Xianshou* and *Arboroharamiya*¹, but there is no serration. Cusp A1 (distobuccal) of the upper premolar and molars is the largest and extends distally. Cusp a1 (mesiolingual) of the lower molars is the largest cusp, extending mesially. The lingual cusps of lower molars bear wear on their lingual sides, indicating that they bit into the central valley of the upper molars in mastication. Molars have a fusiform (spindle-shaped) basin that is closed mesially by a ridge in the upper molars and distally in the lower molars. The ridge can be erased owing to wear so that the central valley becomes confluent in *Shenshou*. Scale bars: **a-c**, 10 mm; **d-f**, 2 mm. Cusp terminology follows ref. 12. See photographs of teeth in Extended Data Figs 2, 4b, c, 5, 6 and 7a-c.

Craniodental features

The general skull shape of the new euharamiyidans is therian-like with a broad basicranial region and a tapered rostrum in dorsal and ventral views (Extended Data Figs 3 and 4) but multituberculate-like in lateral view (Fig. 2a, b). A small septomaxilla is probably present (Extended Data Fig. 2a). The zygomatic arch is slender but fully developed, with the anterior root lateral to P^4 . As reflected by the skull shape, the cranial cavity is more inflated than those of eutriconodontans. The glenoid fossa is anterolateral to the bulging promontorium of the petrosal and antero-posteriorly oriented, and lacks the postglenoid process (Extended Data Figs 4b and 5b). The dentary is similar to that of *Arboroharamiya*¹ and multituberculates in having a distinct diastema between the incisor and premolar, a small coronoid process, a masseteric fossa extending anteriorly to the level of P_4 and a low-positioned mandibular condyle that is

orientated ventrodorsally. A vestigial coronoid is present on the medial side of the jaw (Extended Data Fig. 2b), posteroventral to M_2 , similar to that of *Arboroharamiya* and the Jurassic multituberculate *Kuehneodon*². A reduced coronoid, usually indicated by a scar on the dentary, is also present in some more advanced mammals, such as *Amphitherium*³ and *Zhangheotherium*⁴. This is in contrast to sizable coronoids that cover the anterior portion of the postdentary unit in primitive mammaliaforms, such as *Morganucodon*⁵ and *Haldanodon*⁶. Unlike multituberculates but similar to *Arboroharamiya*, the dentary has a small angular process that inflects medially.

Dentitions are preserved *in situ* in the six specimens (Figs 1 and 2, and Extended Data Figs 1–7), confirming the tooth identification and orientation of *Arboroharamiya*¹ as well as reinterpretation¹ of some previously known euharamiyidans^{7–10}. The cusp arrangement and morphology from *Haramiyavia*, *Thomasia*, *Arboroharamiya* and the new species reported here show consistently larger row A cusps (buccal) than row B cusps in upper molars, and larger row a cusps (lingual) than row b cusps in lower molars. In the lower cheek teeth, the mesiolingual cusp (a1) is the largest, whereas the distobuccal one (A1) is the largest in the upper teeth. The lingual lower cusp row bites into the central valley of the upper molars as indicated by wear facets in the upper and lower teeth (Extended Data Fig. 7a–c). These dental features and occlusal patterns should also be applicable to those known only from isolated teeth, such as *Eleutherodon* and *Sineleutherus*. It is also noteworthy that similar tooth cusp morphology and occlusal patterns are present in the upper and lower second molars of some Jurassic multituberculates¹¹, in which

the mesiolingual cusps are the tallest among cusps in the lower molar. The molars with two rows of multiple cusps and their wear patterns, coupled with the morphologies of the lower jaw and the glenoid fossa, convincingly indicate palinal jaw movement during mastication¹².

Postcranial skeleton

Although the vertebral column varies considerably in extant mammals, particularly those from the southern continents¹³, the general or primitive cervical and thoracolumbar counts for therian mammals are conservative¹⁴. The axial skeleton of the new euharamiyidans is therian-like in possessing 7 cervical and 19 thoracolumbar vertebrae (13 thoracic and 6 lumbar; Fig. 3). In contrast to the docodontan *Castorocauda*¹⁵, the eutriconodont *Liaconodon* and *Yanoconodon*^{16,17}, and the symmetrodont *Akidolestes*¹⁸, which all have lumbar ribs, euharamiyidans (including *Arboroharamiya*) have true lumbar vertebrae that lack ribs but have obliquely oriented facets of the prezygapophysis, elaboration of the transverse process, and the laminapophysis (a mammalian neomorph that split into the anterior metapophysis and posterior anapophysis)¹⁹ (Fig. 3b). These features define a distinct thoracolumbar boundary (Fig. 3a) and, along with fixation of cervical vertebrae²⁰ and well-developed ribcage, are indicative of development of the diaphragm, a unique feature that allows mammals to progressively increase ventilation in adapting to fast movement^{19,20}.

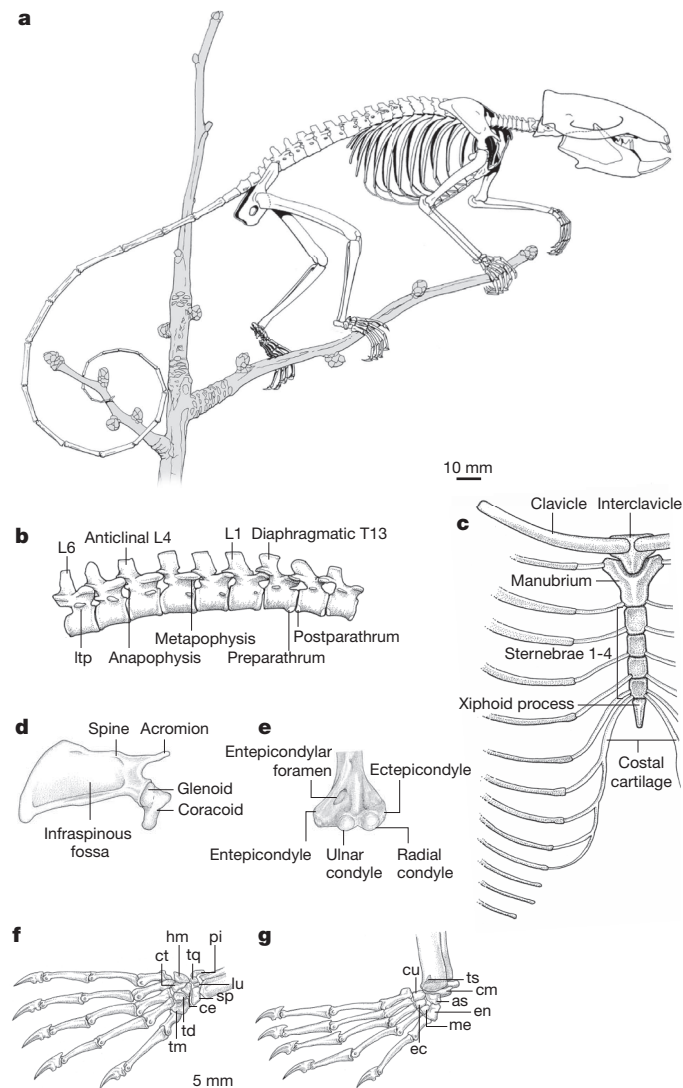


Figure 3 | Skeletal features of euharamiyidans. **a**, Reconstruction of the skeleton of *Shenshou lui*, based in part on *Xianshou linglong* and *Xianshou songae*. **b**, The vertebral column of *Shenshou lui*. **c**, The anterior section of the rib cage in *Shenshou lui*. **d**, The right scapular of *Shenshou lui*. **e**, The distal end of the humerus in *Shenshou lui*. **f**, Reconstruction of the hand of *Shenshou lui*, based on LDNHMF2001 and WGMV-001. **g**, Reconstruction of the foot of euharamiyidans, based on *Shenshou lui* (JZT-D061) and *Xianshou songae*. The cervical ribs, present in monotremes and many Mesozoic mammals^{4,17,20,24}, are absent. Similar to therians, such as *Eomaia*³² and *Sinodelphys*²⁷, the first ten thoracic ribs are connected to the ossified manubrium and sterna via costal cartilages and the last three ribs float freely. The capitulum of each rib articulates between two thoracic centra at a ventral position. The last thoracic rib (13th) is identified as the diaphragmatic vertebra because the articular facets of its zygapophyses with the first lumbar are more vertical than horizontal, as in many extant mammals⁴⁹. The anticline vertebrae is located at lumbar four. The tail consists of 18–22 caudal vertebrae, with the proximal ones bearing expanded transverse processes as in extant prehensile therians⁵⁰. The scapular blade is roughly rectangular and has a large and deep infraspinous fossa but lacks a supraspinous fossa. The clavicle is strap-shaped and slightly curved anteriorly. Its rough proximal end indicates a flexible contact with the depression on the clover-leaf-shaped interclavicle. The distal end of the humerus has a bulbous radial condyle, a relatively smaller spherical ulna condyle, and a prominent entepicondyle. The femur has a hemispherical head with an extensive articular surface but a short neck, the greater trochanter is triangular and does not extend above the head. The distal condyles are small with a shallow patellar groove. The ankles are comparable to those of multituberculates, such as *Kryptobaatar*²⁴, in that the calcaneus has a short and mediolaterally compressed tuber calcanei and are partly superposed by the astragalus. The cuboid is positioned obliquely and has little contact with the calcaneus so that metatarsal V is separated from the calcaneus. The entocuneiform is elongated, and its joint with metatarsal I is offset anteriorly from the joint of the intermediate cuneiform and metatarsal II, similar to that of *Asioryctes* and *Eomaia*³². Metapodials and proximal phalanges possess a well-developed palmar groove for the digital flexor muscle tendon. Metatarsals I and V are short and robust compared to others. The distal ends of proximal and intermediate phalanges are well trochleated. Terminal phalanges are compressed laterally and curved strongly with a sharp tip; each has a massive digital flexor tubercle on the ventral side and a small dorsally extended ridge. as, astragalus; ce, centrale; cm, calcaneum; ct, capitate; cu, cuboid; ec, ectocuneiform; en, entocuneiform; hm, hamate; L, lumbar; ltp, lumbar transverse process; lu, lunate; me, mesocuneiform; T, thoracic; td, trapezoid; tm, trapezium; sp, scaphoid; pi, pisiform; tq, triquetrum; ts, tarsal spur. The 10 mm scale applies to panel **a** and the 5 mm scale bar applies to all other postcranial elements. See photographs of skeletons and postcranial structures in Extended Data Figs 1, 3, 4a, 5a and 7d–g.

The scapula (Fig. 3d and Extended Data Fig. 7g) is primitively similar to that of *Megazostrodon*, *Haldanodon* and monotremes^{21,22} in that it lacks the suprascapular fossa, but differs from those taxa in that it has a reduced coracoid that is fused to the scapula and forms a small part of the glenoid fossa that faces ventrally. The clavicle (Fig. 3c) is similar to that of multituberculates²³. As in multituberculates the humerus has a slight shaft torsion (approximately 15°)²³, in contrast to a strong proximodistal torsion in cynodonts and other early mammals^{4,24}. The distal end of the humerus is highly comparable to that of multituberculates^{23,24} (Fig. 3e and Extended Data Fig. 7e). In the wrist, the hamate is hypertrophied and the scaphoid and triquetrum are enlarged, as in some extant arboreal marsupials^{25,26} and the scansorial *Sinodelphys*²⁷. Metacarpal V is offset from the hamate and the pisiform is sizable, similar to those of *Zhangheotherium*⁴ (Fig. 3g and Extended Data Fig. 8a).

The epipubic bone, a plesiomorphic feature common in several groups of Mesozoic mammals, including eutherians²⁸, is absent. The pelvis is similar to that of therians in being shallow and differs from the deep pelvis of multituberculates and monotremes in having a reduced pubis and an ischium with a straight dorsal margin and an ischial tuberosity that is slender and less extended dorsoposteriorly²⁴. The ilium is shorter than that of multituberculates. The femur has a short neck, similar to eutriconodonts^{16,29,30}, symmetrodonts^{4,18,31} and more primitive forms^{21,22}, but different from multituberculates, in which the femoral head has a long neck and the greater trochanter projects beyond the head²⁴. There is no parafibula, differing from *Eomaia* and *Jeholodens*^{30,32}. The proximal end of the tibia is roughly symmetrical, contrasting to that of multituberculates³³. The bony extratarsal spur is distinctive and displays different relative sizes in different species reported here. The spur or its os calcaris exists in monotremes³⁴ and is common in Mesozoic mammals and their close relatives^{15,35}. As in *Arboroharamiya* and some extant arboreal didelphids and cheirogaleid primates¹, the manus and pes are characterized by relatively short metapodials and long phalanges (Fig. 3f, g and Extended Data Fig. 8). The limb features are collectively indicative of scansorial and/or arboreal adaptation^{1,26,32}.

Phylogenetic relationship

A major unsolved problem in mammalian taxonomy and phylogeny concerns Allotheria (Multituberculata and Haramiyida)^{9,12,33,36–39}, which affects how we view the early evolution of mammals. Most recent studies present contrasting hypotheses that either place allotherians in mammals¹, indicating an explosive model for the origin of mammals in the Late Triassic (approximately 208 million years ago) and a long-fuse model for the origin of therians, or separate them from multituberculates and place them outside mammals⁴⁰, suggesting an explosive model for the origin of the Mammalia in the Middle Jurassic epoch. In light of the new data reported here, we are able to revise characters used in previous studies^{1,11,40}, discuss existing problems relating to *Megaconus* (Supplementary Information, section E), the postdentary trough in ‘haramiyidans’ (now demonstrated to be a paraphyletic group) (Supplementary Information, section F) and *Hadrocodium* (Supplementary Information, section G) and conduct phylogenetic analyses including the new species (Supplementary Information, sections H–J). The result shown in Fig. 4 (see also Extended Data Fig. 9) is consistent with most recent phylogenetic analyses^{1,11,41,42}, which suggest that the ‘haramiyidans’ are related to multituberculates to form Allotheria within the Mammalia. This supports a Late Triassic origin of mammals¹. This hypothesis gains support from some Late Triassic ‘symmetrodontans’, such as *Kuehneotherium* and *Woutersia*, that are considered taxonomically to be members of the trechnotherian mammals^{33,43,44}. Our analyses also recognize a new clade, here named Euharamiyida, that pairs with Multituberculata; this is consistent with the view that previously discovered ‘haramiyidans’ seem to form a paraphyletic group, from which multituberculates were derived^{12,36,39}. Primitive species traditionally placed in ‘haramiyidans’, such as *Haramiyavia* and *Thomasia*, form the stem members of allotherians. This resultant topology remains the same whether *Megaconus* was included (Extended Data Fig. 10a) or both *Megaconus* and *Hadrocodium* were

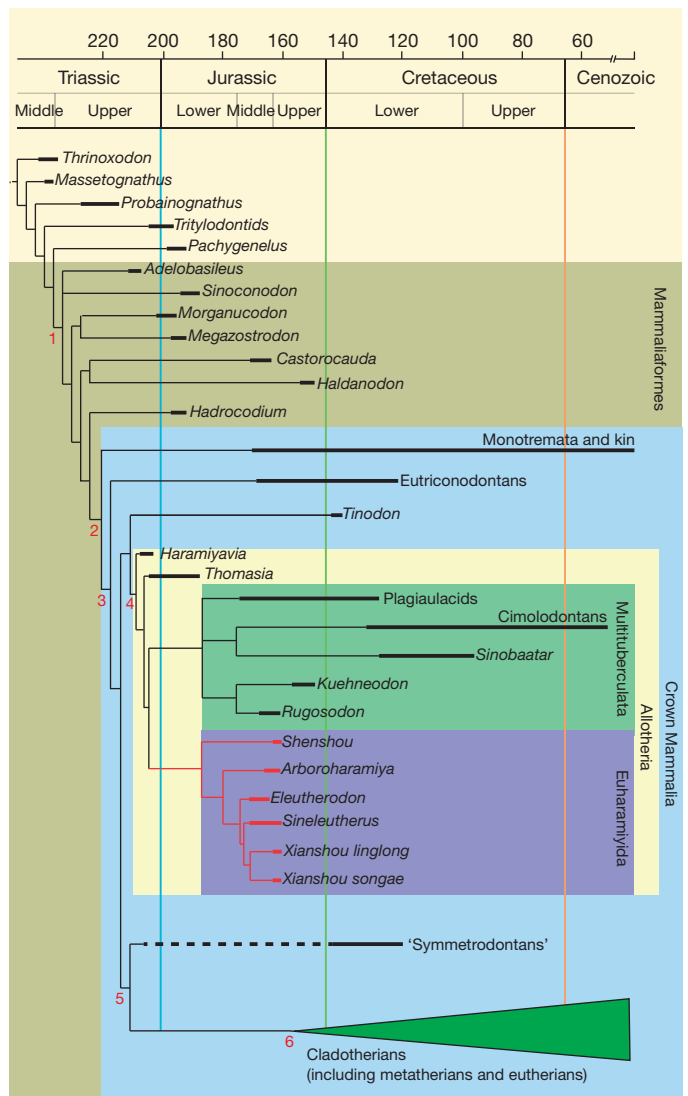


Figure 4 | Phylogeny of mammals with focus on Allotheria. This simplified cladogram is based on the consensus tree (Extended Data Fig. 9) computed from a data matrix with 113 taxa and 495 characters (see Supplementary Information). The dashed line of ‘symmetrodontans’ indicates occurrences of taxa that existed in the Late Triassic, such as *Kuehneotherium* and *Woutersia*^{33,43}, but are not included in our phylogenetic analyses because of their fragmentary preservation. Tree nodes represent the following clades: (1) Mammaliaformes; (2) Mammalia; (3) unnamed clade consisting of Eutriconodonta, Allotheria and Trechnotheria³³; (4) Allotheria; (5) Trechnotheria; and (6) Cladotheria. Mammaliaformes, Crown Mammalia, Allotheria, Multituberculata and Euharamiyida are also marked as nested colour blocks from the most to the least inclusive group.

excluded (Extended Data Fig. 10b) (see discussions in Supplementary Information, sections E and G).

Character evolution of early mammals

With the discoveries of the new euharamiyidans, it becomes increasingly evident that the cranial and postcranial features of euharamiyidans and multituberculates are similar to each other and to other mammals. However, the fundamental obstacle in interpreting their mammalian affinity remains the fact that the tooth pattern consists of two main rows of multiple cusps that are capable of longitudinal (palinal) chewing function in allotherians^{12,39,45}. If allotherians were placed outside mammals, it is equally difficult to derive the allotherian tooth pattern from other mammaliaformes, such as tritylodontids. Our phylogenetic analyses (Fig. 4) suggest that the primitive allotherian tooth pattern, as represented by *Haramiyavia*, was probably derived by developing an extra cusp row,

or rows, from a triconodont-like tooth pattern or even from a tooth pattern with an initially reversed triangular cusp arrangement. It was noted that in lateral view the teeth of *Haramiyavia* are more similar to those of *Sinoconodon* and *Morganucodon* than to those of multituberculates⁴⁶. We note, however, that the tooth pattern of *Haramiyavia* is also similar to that of *Woutersia*, which co-existed with *Theroteinus*, another Late Triassic haramiyid^{36,43}. The tooth morphology and occlusion of euharamiyidans indicate that, if the allotherian tooth pattern was derived from a triconodont tooth pattern, the secondary cusp row has to be added on the buccal side in the lower teeth. In the conventional view, however, development of extra cusps on the lingual cingula is common, but buccal cingula are rare on lower molars⁴⁵, although exceptions exist, such as *Hallautherium*⁴⁷. Nonetheless, the orientation of an isolated tooth in early mammals is not always certain, as demonstrated in the case of eleutherodontids¹ (this study). There is no convincing evidence to rule out the possibility that additional cusps could be added on the buccal side of the tooth in early mammals. Better material with teeth *in situ* from each taxon of interest, such as *Woutersia*, is needed to test this hypothesis.

Interpretations of character evolution in early mammals depends on their phylogeny. If 'haramiyids' were separated from multituberculates and placed outside mammals, while multituberculates fell within mammals⁴⁰, then numerous similar craniodental and postcranial features, particularly the molar pattern with two cusp rows and bilateral occlusion, must have evolved independently in 'haramiyidans' and multituberculates during different periods of time. In addition, detachment of the postdentary bones from the dentary would have evolved at least four times independently in 'haramiyidans', multituberculates, monotremes and therians. However, our phylogeny (Fig. 4 and Extended Data Figs 9 and 10) indicates that Euharamiyida and Multituberculata were probably derived from a *Haramiyavia*-like common ancestor at a minimum oldest age (according to current fossil records; future finds may reveal an earlier ancestor) in the Late Triassic and diversified thereafter during the Jurassic epoch, with known euharamiyidans adapting to a scansorial and/or arboreal lifestyle which may explain their rare fossil record. In contrast to interpreting numerous parallelisms in 'haramiyidans' and multituberculates⁴⁰, our hypothesis favours homologous acquisition of many similar craniodental and postcranial features in euharamiyidans and multituberculates, such as reduction of teeth, enlargement of the lower incisors, possessing only two molars in each side of the upper and lower jaws, and a palinal chewing motion. Moreover, euharamiyidans are similar to multituberculates in lacking the postdentary trough and Meckelian groove, indicating the presence of the definitive mammalian middle ear³. If the reinterpretation is correct—that the dentary of *Haramiyavia* has only the Meckelian groove (see Supplementary Information, section F)—then the clade containing Eutricondonta, Allotheria and Trechnotheria³³ (Fig. 4 and Extended Data Fig. 9) would have evolved from a common ancestor that had a transitional mammalian middle ear¹⁶. This clade and the geological and geographic occurrences of its earliest known members are consistent with accumulating evidence from Gondwana landmasses that shows a cosmopolitan distribution of members in the clade⁴⁸ and suggest a Laurasian origin of mammals. Finally, by reinterpreting *Hadrocodium* as having postdentary bones (see Supplementary Information, section G), our phylogeny suggests that detachment of the postdentary bones evolved twice independently during the early evolution of mammals, once in the clade leading to monotremes and once towards the clade containing Eutricondonta, Allotheria and Trechnotheria.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 April; accepted 25 July 2014.

Published online 10 September 2014.

- Zheng, X., Bi, S., Wang, X. & Meng, J. A new arboreal haramiyid shows the diversity of crown mammals in the Jurassic period. *Nature* **500**, 199–202 (2013).

- Hahn, G. Neue Schädel-Reste von Multituberculaten (Mamm.) aus dem Malm Portugals. *Geol. Palaeontol.* **11**, 161–186 (1977).
- Allin, E. F. & Hopson, J. A. in *The Evolutionary Biology of Hearing* (eds Webster, D. B., Fay, R. R. & Tavolga, W. N.) Ch. 28, 587–614 (Springer New York, 1992).
- Hu, Y., Wang, Y., Luo, Z. & Li, C. A new symmetrodont mammal from China and its implications for mammalian evolution. *Nature* **390**, 137–142 (1997).
- Kermack, K., Mussett, F. & Rigney, H. The lower jaw of *Morganucodon*. *Zool. J. Linn. Soc.* **53**, 87–175 (1973).
- Lillegraven, J. A. & Krusat, G. Cranio-mandibular anatomy of *Haldanodon expectatus* (Docodontia; Mammalia) from the Late Jurassic of Portugal and its implications to the evolution of mammalian characters. *Contrib. Geol. Univ. Wyo.* **28**, 39–138 (1991).
- Averianov, A., Lopatin, A. & Krasnolutskii, S. The First Haramiyid (Mammalia, Allotheria) from the Jurassic of Russia. *Dokl. Biol. Sci.* **437**, 103–106 (2011).
- Kermack, K. A., Kermack, D. M., Lees, P. M. & Mills, J. R. E. New multituberculate-like teeth from the Middle Jurassic of England. *Acta Palaeontol. Pol.* **43**, 581–606 (1998).
- Butler, P. M. & Hooker, J. J. New teeth of allotherian mammals from the English Bathonian, including the earliest multituberculates. *Acta Palaeontol. Pol.* **50**, 185–207 (2005).
- Martin, T., Averianov, A. O. & Pfretzschner, H. U. Mammals from the Late Jurassic Qigu Formation in the southern Junggar Basin, Xinjiang, Northwest China. *Palaeobiodiversity and Palaeoenvironments* **90**, 295–319 (2010).
- Yuan, C.-X., Ji, Q., Meng, Q.-J., Tabrum, A. R. & Luo, Z.-X. Earliest evolution of multituberculate mammals revealed by a new Jurassic fossil. *Science* **341**, 779–783 (2013).
- Butler, P. M. Review of the early allotherian mammals. *Acta Palaeontol. Pol.* **45**, 317–342 (2000).
- Asher, R., Lin, K., Kardjilov, N. & Hautier, L. Variability and constraint in the mammalian vertebral column. *J. Evol. Biol.* **24**, 1080–1090 (2011).
- Narita, Y. & Kuratani, S. Evolution of the vertebral formulae in mammals: a perspective on developmental constraints. *J. Exp. Zool. B* **304**, 91–106 (2005).
- Ji, Q., Luo, Z.-X., Yuan, C.-X. & Tabrum, A. R. A swimming mammaliaform from the Middle Jurassic and ecomorphological diversification of early mammals. *Science* **311**, 1123–1127 (2006).
- Meng, J., Wang, Y.-Q. & Li, C.-K. Transitional mammalian middle ear from a new Cretaceous Jehol eutriconodont. *Nature* **472**, 181–185 (2011).
- Luo, Z.-X., Chen, P.-J., Li, G. & Chen, M. A new eutriconodont mammal and evolutionary development in early mammals. *Nature* **446**, 288–293 (2007).
- Li, G. & Luo, Z.-X. A Cretaceous symmetrodont therian with some monotreme-like postcranial features. *Nature* **439**, 195–200 (2006).
- Filler, A. G. Homeotic evolution in the mammalia: diversification of therian axial seriation and the morphogenetic basis of human origins. *PLoS ONE* **2**, e1019 (2007).
- Buchholtz, E. A. *et al.* Fixed cervical count and the origin of the mammalian diaphragm. *Evol. Dev.* **14**, 399–411 (2012).
- Jenkins, F. & Parrington, F. R. The postcranial skeletons of the Triassic mammals *Eozostrodon*, *Megazostrodon* and *Erythrotherium*. *Phil. Trans. R. Soc. Lond. B* **273**, 387–431 (1976).
- Martin, T. Postcranial anatomy of *Haldanodon expectatus* (Mammalia, Docodontia) from the Late Jurassic (Kimmeridgian) of Portugal and its bearing for mammalian evolution. *Zool. J. Linn. Soc.* **145**, 219–248 (2005).
- Sereno, P. C. & McKenna, M. C. Cretaceous multituberculate skeleton and the early evolution of the mammalian shoulder girdle. *Nature* **377**, 144–147 (1995).
- Kielan-Jaworowska, Z. & Gambaryan, P. P. *Postcranial Anatomy and Habits of Asian Multituberculate Mammals*. (Scandinavian Univ. Press, 1994).
- Szalay, F. S. *Evolutionary History of the Marsupials and an Analysis of Osteological Characters*. (Cambridge Univ. Press, 1994).
- Argot, C. Functional-adaptive anatomy of the forelimb in the didelphidae, and the paleobiology of the paleocene marsupials *Mayulestes ferox* and *Pucadelphys andinus*. *J. Morphol.* **247**, 51–79 (2001).
- Luo, Z.-X., Ji, Q., Wible, J. R. & Yuan, C.-X. An Early Cretaceous tribosphenic mammal and metatherian evolution. *Science* **302**, 1934–1940 (2003).
- Novacek, M. J. *et al.* Epipubic bones in eutherian mammals from the Late Cretaceous of Mongolia. *Nature* **389**, 483–486 (1997).
- Jenkins, F. A. & Schaff, C. R. The Early Cretaceous mammal *Gobiconodon* (Mammalia, Triconodonta) from the Cloverly Formation in Montana. *J. Vertebr. Paleontol.* **8**, 1–24 (1988).
- Ji, Q., Luo, Z. & Ji, S.-A. A Chinese triconodont mammal and mosaic evolution of the mammalian skeleton. *Nature* **398**, 573–574 (1999).
- Rougier, G. W., Ji, Q. & Novacek, M. J. A new symmetrodont mammal with fur impressions from the Mesozoic of China. *Acta Geol. Sin. (Engl.)* **77**, 7–14 (2003).
- Ji, Q. *et al.* The earliest known eutherian mammal. *Nature* **416**, 816–822 (2002).
- Kielan-Jaworowska, Z., Cifelli, R. & Luo, Z. X. *Mammals from the Age of Dinosaurs: Origins, Evolution, and Structure*. (Columbia Univ. Press, 2004).
- Griffiths, M. *The Biology of the Monotremes*. (Academic Press, New York, 1978).
- Hurum, J. H., Kielan-Jaworowska, Z. & Luo, Z. Were mammals originally venomous? *Acta Palaeontol. Pol.* **51**, 1–11 (2006).
- Hahn, G., Sigogneau-Russell, D. & Wouters, G. New data on Theroteinidae: their relations with Paulchoffatiidae and Haramiyidae. *Geol. Palaeontol.* **23**, 205–215 (1989).
- Sigogneau-Russell, D. Haramiyidae (Mammalia, Allotheria) en provenance du Trias supérieur de Lorraine (France). *Palaeontogr. Abt. A* **206**, 137–198 (1989).

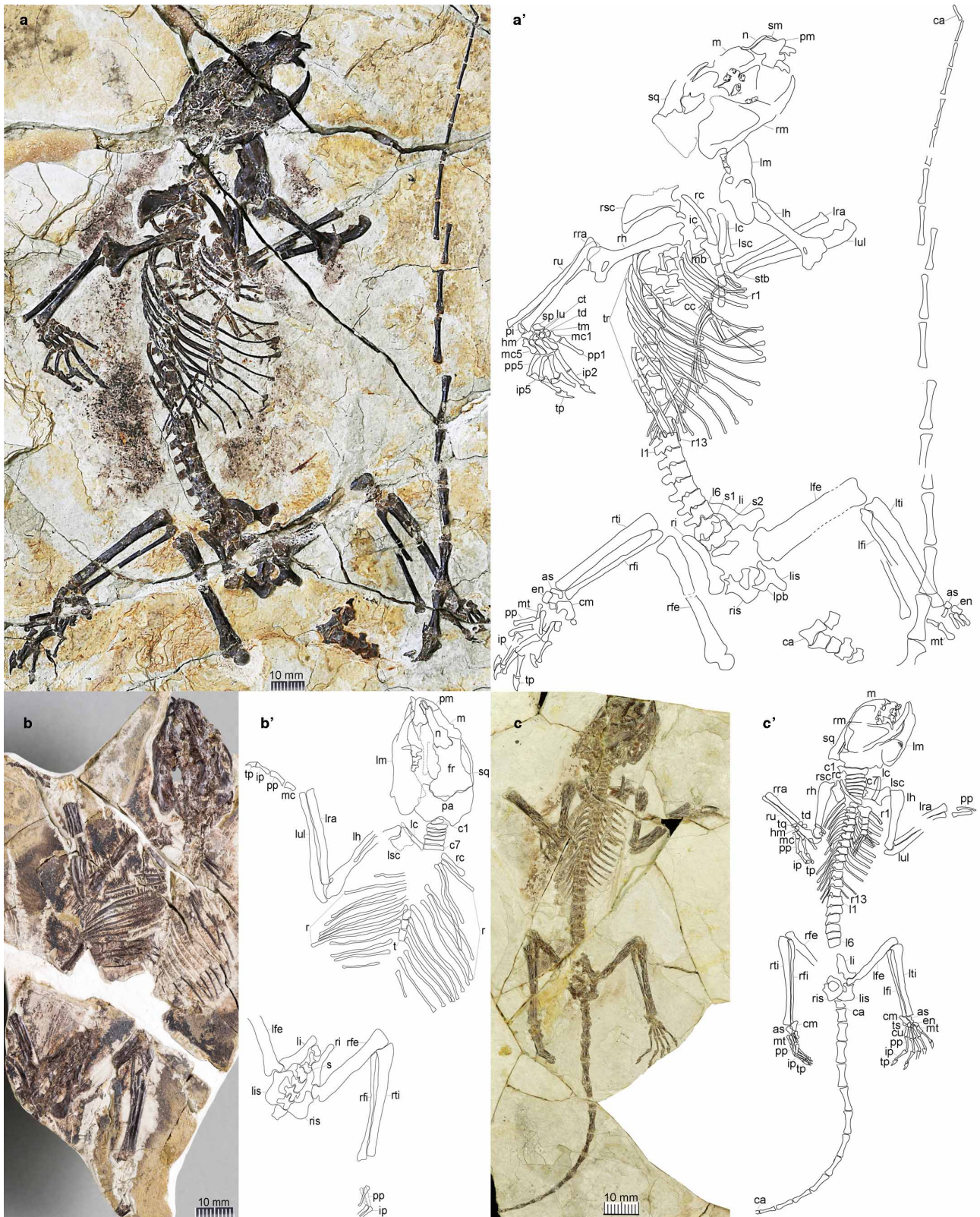
38. Jenkins, F. A., Gatesy, S. M., Shubin, N. H. & Amaral, W. W. Haramiyids and Triassic mammalian evolution. *Nature* **385**, 715–718 (1997).
39. Hahn, G. & Hahn, R. Evolutionary tendencies and systematic arrangement in the Haramiyida (Mammalia). *Geol. Palaeontol.* **40**, 173–193 (2006).
40. Zhou, C.-F., Wu, S., Martin, T. & Luo, Z.-X. A Jurassic mammaliaform and the earliest mammalian evolutionary adaptations. *Nature* **500**, 163–167 (2013).
41. Luo, Z.-X. & Wible, J. R. A Late Jurassic digging mammal and early mammalian diversification. *Science* **308**, 103–107 (2005).
42. Rowe, T., Rich, T. H., Vickers-Rich, P., Springer, M. & Woodburne, M. O. The oldest platypus and its bearing on divergence timing of the platypus and echidna clades. *Proc. Natl Acad. Sci. USA* **105**, 1238–1242 (2008).
43. Sigogneau-Russell, D. & Hahn, R. Reassessment of the Late Triassic symmetrodont mammal *Woutersia*. *Acta Palaeontol. Pol.* **40**, 245–260 (1995).
44. Kermack, D. M., Kermack, K. & Mussett, F. The Welsh pantothere *Kuehneotherium praecursoris*. *Zool. J. Linn. Soc.* **47**, 407–423 (1968).
45. Butler, P. & MacIntyre, G. Review of the British Haramiyidae (? Mammalia, Allotheria), their molar occlusion and relationships. *Phil. Trans. R. Soc. Lond. B* **345**, 433–458 (1994).
46. Kielan-Jaworowska, Z. *In Pursuit of Early Mammals*. 272 (Indiana Univ. Press, 2013).
47. Clemens, W. Rhaeto-Liassic mammals from Switzerland and West Germany. *Zitteliana* **5**, 51–92 (1980).
48. Krause, D. W. Gondwanatheria and ?Multituberculata (Mammalia) from the Late Cretaceous of Madagascar. *Can. J. Earth Sci. Rev. Can. Sci. Terre* **50**, 324–340 (2013).
49. Argot, C. Functional-adaptive anatomy of the axial skeleton of some extant marsupials and the paleobiology of the paleocene marsupials *mayulestes ferox* and *pucadelphys andinus*. *J. Morphol.* **255**, 279–300 (2003).
50. Schmitt, D., Rose, M. D., Turnquist, J. E. & Lemelin, P. Role of the prehensile tail during ateline locomotion: experimental and osteological evidence. *Am. J. Phys. Anthropol.* **126**, 435–446 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank H.-J. Li and Z.-J. Gao for using the specimens housed at the Jizantang Paleontological Museum; Z.-Y. Sun and D.-Y. Sun for using the specimen housed at the Museum of Wuyishan Mountain; J.-C. Lü for assistance in specimen collecting; S.-H. Xie for specimen preparation; M. A. Klingler and F.-X. Wu for specimen drawing and photography; W.-D. Zhang for SEM imaging; Z.-H. Zhou, X. Xu, F.-C. Zhang and X.-L. Wang for discussion on stratigraphy and faunal compositions; and A. Weil and G. W. Rougier for instructive comments. The study was supported by the National Basic Research Program of China (973 program, 2012CB821906), the Strategic Priority Research Program of Chinese Academy of Sciences (XDB03020501), the National Science Foundation of China (41128002) and the Hundred Talents Programs of the Chinese Academy of Sciences.

Author Contributions S.B., J.M. and Y.W. designed the study, performed the comparative and analytical work and wrote the paper. J.G. and X.S. collected data and contributed to the writing and discussion.

Author Information The Life Science Identifiers (LSIDs) for this publication have been deposited at <http://zoobank.org/> and include: urn:lsid:zoobank.org:pub:766EBC08-EF77-41E5-AC59-9F69E9F59BAA (for this publication), urn:lsid:zoobank.org:act:32B0742D-2DD3-47F9-A35B-E1FB284E9EA5 (*Shenshou*), urn:lsid:zoobank.org:act:BA8BF69D-0F21-4DE7-9EA5-BA6082973940 (*Shenshou lui*), urn:lsid:zoobank.org:act:4C10F9F1-A920-450B-8E5F-C88F3BD0920C (*Xianshou*), urn:lsid:zoobank.org:act:3DB1D738-70E9-4966-93FE-3FE5269F9C91 (*Xianshou linglong*), urn:lsid:zoobank.org:act:1F7EEE49-5E88-4941-AE8B-B2C7C31F9788 (*Xianshou songae*). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M. (jmeng@amnh.org) or Y.W. (wangyuanqing@ivpp.ac.cn).



Extended Data Figure 1 | The holotypes of three euhamiyidan species.

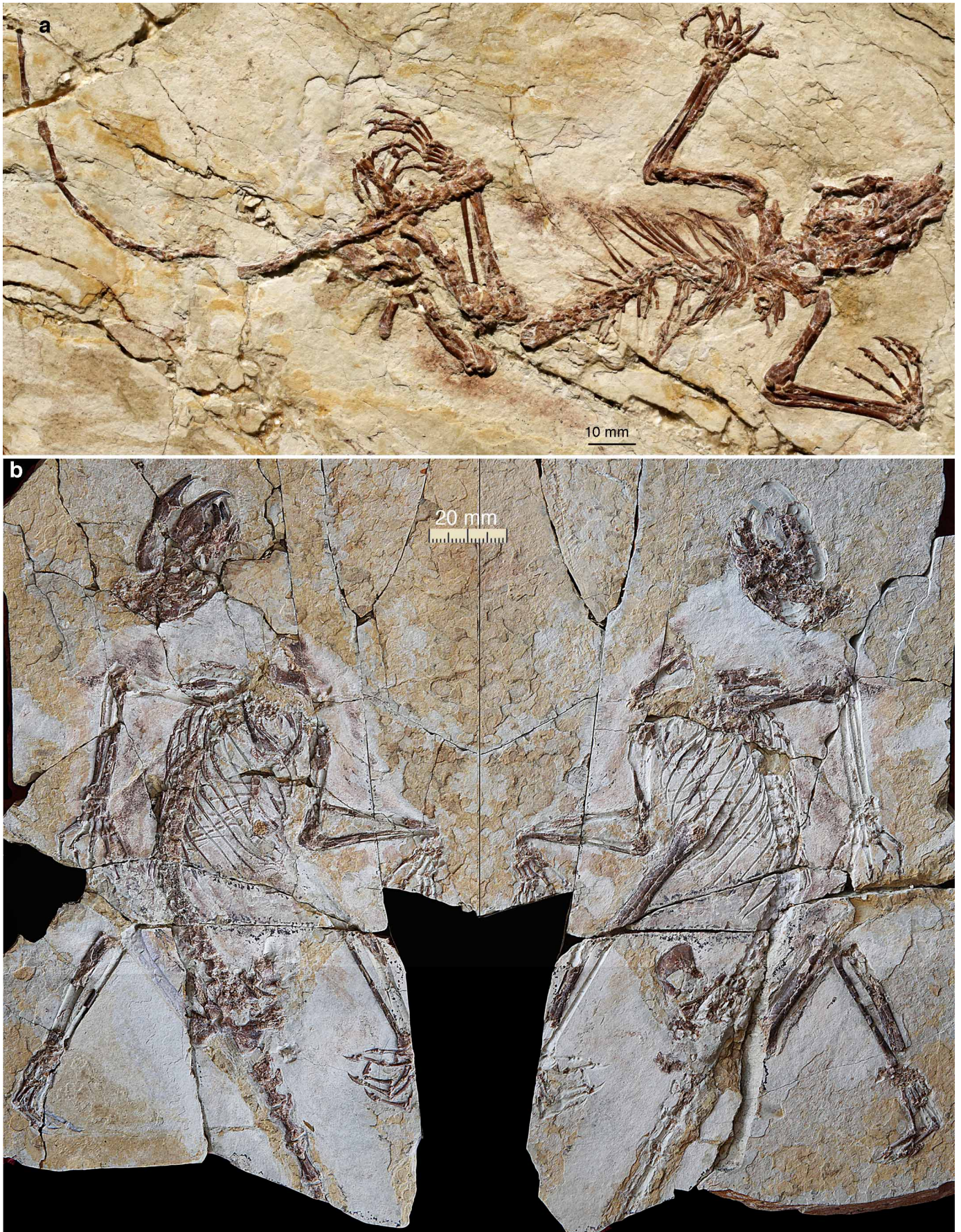
a, Holotype specimen (LDNHMF2001) and line drawing of *Shenshou lui*.
b, Holotype specimen (IVPP V16707A) and line drawing of *Xianshou linglong*.
c, Holotype specimen (BMNH-PM003253) and line drawing of *Xianshou songae*. as, astragalus; c, cervical vertebrae; ca, caudal vertebrae; cc, costal cartilage; cm, calcaneum; ct, capitata; cu, cuboid; en, entocuneiform; fr, frontal; hm, hamate; ic, interclavicle; ip, intermediate phalanges; l, lumbar vertebrae; lc, left clavicle; lfe, left femur; lfi, left fibula; lh, left humerus; li, left ilium; lis, left ischium; lm, left mandible; lra, left radius; lsc, left scapular; lpb, left pubis; lti, left

tibia; lu, lunata; lul, left ulna; m, maxilla; mb, manubrium; mc, metacarpal; mt, metatarsal; n, nasal; pa, parietal; pi, pisiform; pm, premaxilla; pp, proximal phalanges; r, rib; rc, right clavicle; rfe, right femur; rfi, right fibula; rh, right humerus; ri, right ilium; ris, right ischium; rm, right mandible; rra, right radius; rsc, right scapular; rti, right tibia; ru, right ulna; s, sacral vertebrae; sm, septomaxilla; sp, scaphoid; sq, squamosal; stb, sternbra; t, thoracic vertebrae; td, trapezoid; tm, trapezium; tp, terminal phalanges; tq, triquetrum; tr, thoracic ribs; ts, tarsal spur.

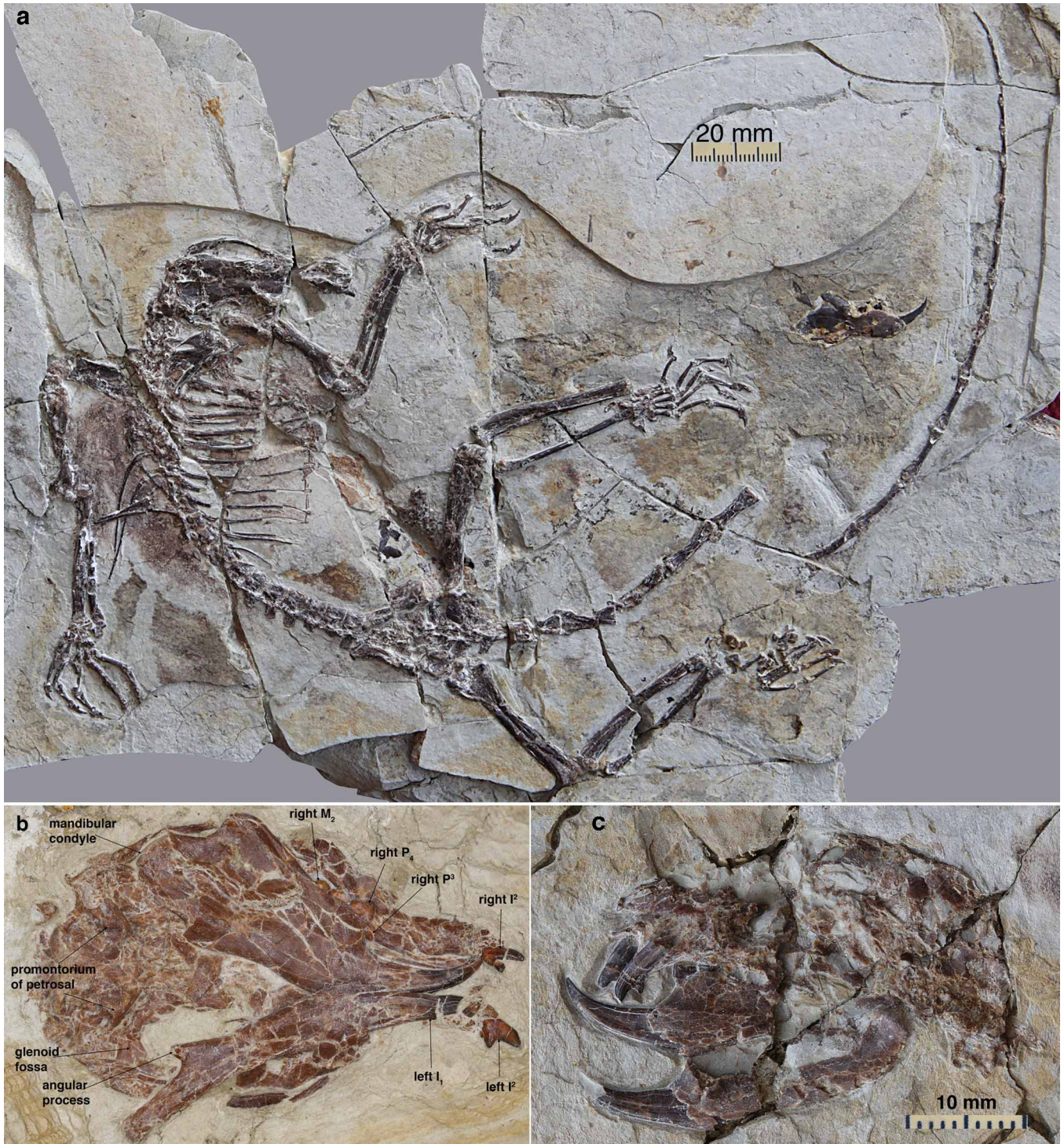


Extended Data Figure 2 | Close-up views of craniodental structure of *Shenshou lui* (LDNHMF2001). **a**, The skull and mandibles of *Shenshou lui* (LDNHMF2001). ap, angular process; m, maxilla; maf, masseteric fossa; n, nasal; pal, palatine; pm, premaxilla; ptf, pterygoid fossa; sm, septomaxilla; sq, squamosal. **b**, The lingual surface of the mandible of *Shenshou lui* (LDNHMF2001). The close-up view shows the vestige of the coronoid bone at the ventrodistal side of the ultimate lower molar and the wear facets on the lingual sides of the lingual cusps of $M_{1,2}$. **c**, The rostrum of the skull of *Shenshou*

lui (LDNHMF2001). The figure is a close-up view of the rostrum region in **a**. It shows the relationship of the alveoli for $P^{3,4}$ with the molars and the wear facets on the medial side of left I^2 . **d**, Photographs and stippling drawings of the dentition of *Shenshou lui* (LDNHMF2001). The dentitions were coated with ammonium chloride to enhance contrast. The holotype of *Shenshou* is an old individual, judging from the deeper wear on its teeth. The mesial end of the upper molar basin and the distal end of the lower molar basin appear to be open, not closed by ridge or cusps, but this may be partly due to the tooth wear.

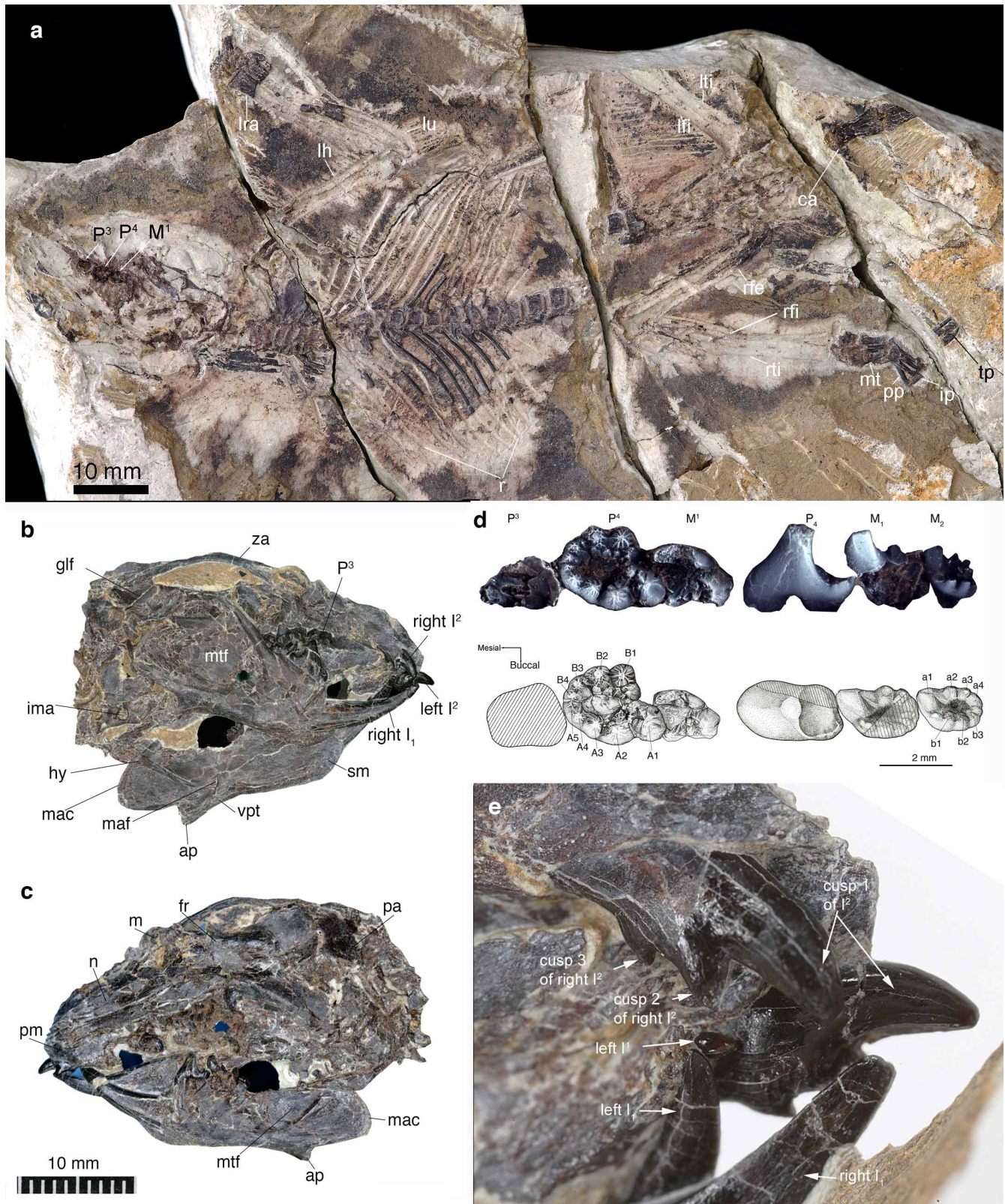


Extended Data Figure 3 | Paratypes 1 and 2 of *Shenshou lui*. a, Dorsal view of the skeleton of *Shenshou lui* (paratype 1, WGMV-001). b, The split skeleton of *Shenshou lui* (paratype 2, JZT-CK005A, B).



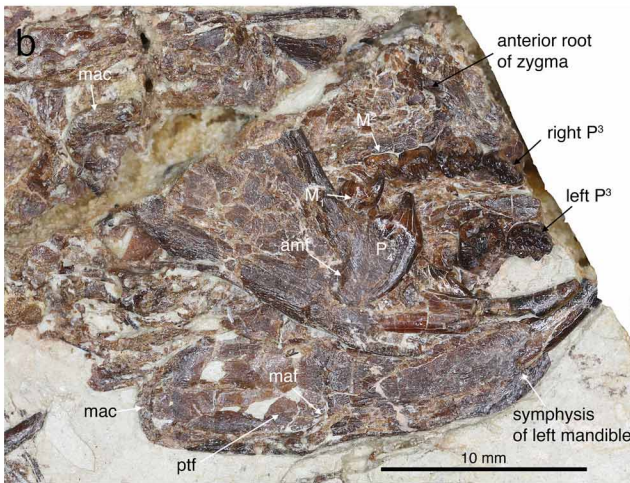
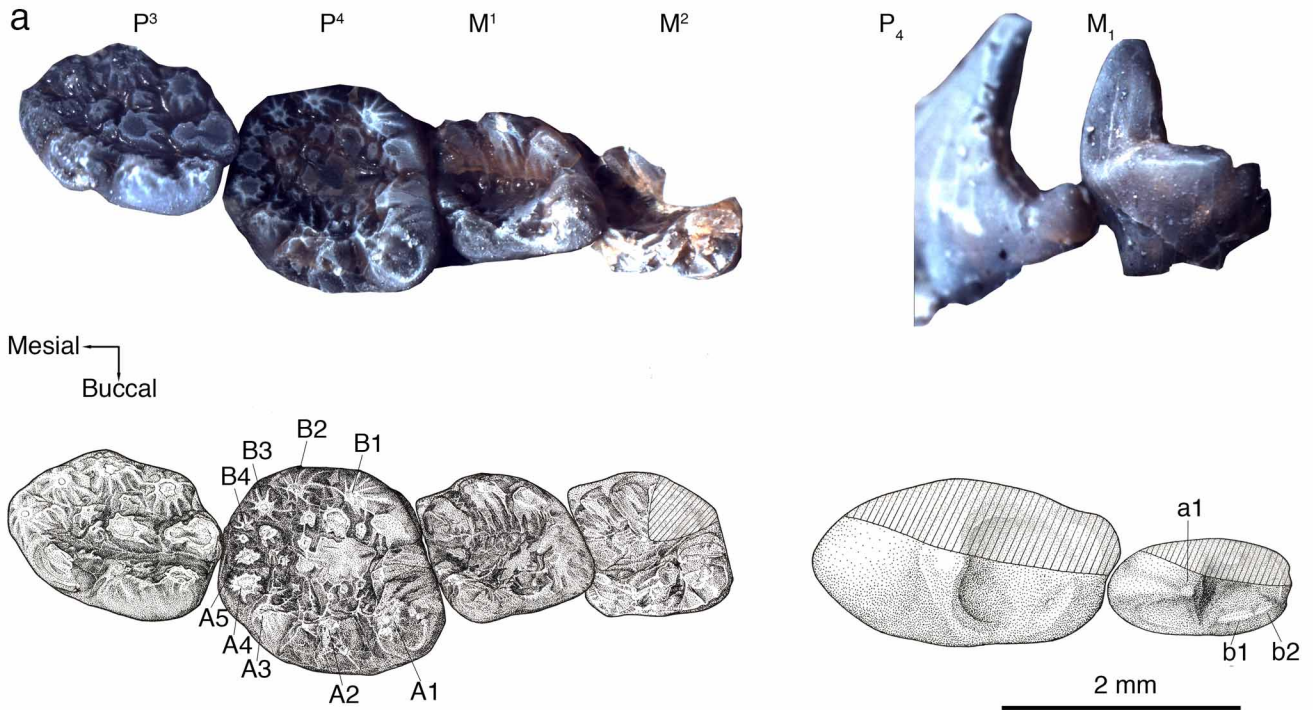
Extended Data Figure 4 | Paratype 3 and skulls of paratypes 1 and 2 of *Shenshou lui*. a, Dorsal view of the skeleton of *Shenshou lui* (paratype 3, JZT-D061). b, Ventral view of the skull of *Shenshou lui* (paratype 1, WGMV-001).

Note the bulging promontorium of the petrosal, the inflected angular process and small right P^3 . c, Lateral view of the skull and mandible of *Shenshou lui* (paratype 2, JZT-CK005).



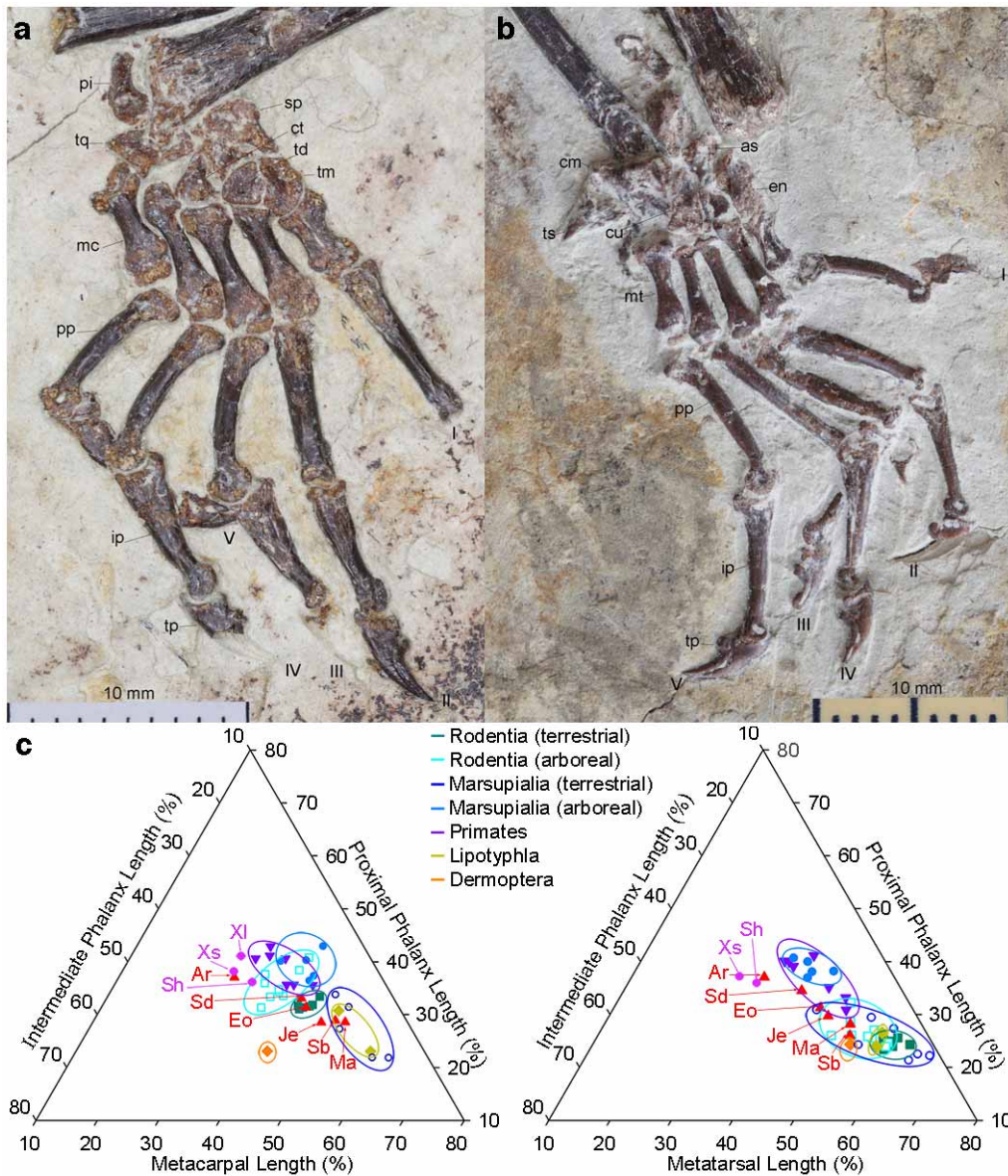
Extended Data Figure 5 | Holotype and close-up views of craniodental features of *Xianshou linglong* (IVPP V16707B). **a**, Counterpart of the holotype specimen of *Xianshou linglong* (IVPP V16707B). **b**, **c**, Skull (two sides) and mandibles of *Xianshou linglong* (IVPP V16707A). **b**, the skull (ventral view) and left mandible (medial view) and right mandible (labial view). **c**, the skull (dorsal view) and left mandible (later view); ap, angular process; fr, frontal;

glf, glenoid fossa; hy, hyoid; ima, internal acoustic meatus; m, maxilla; mac, mandibular condyle; maf, mandibular foramen; mtf, masseteric fossa; n, nasal; pa, parietal; pm, premaxilla; sm, symphysis of mandible (left); vpt, ventral pterygoid fossa; za, zygomatic arch. **d**, Photographs and stippling drawings of the dentition of *Xianshou linglong* (IVPP V16707). P^3 is broken; its crown was left in the matrix so that its outline can be reconstructed in the drawing. The lingual side of P^4 and labial side of M_1 which are exposed or broken are line-hatched. The dentitions were coated with ammonium chloride to enhance contrast. **e**, Incisors of *Xianshou linglong* (IVPP V16707A).



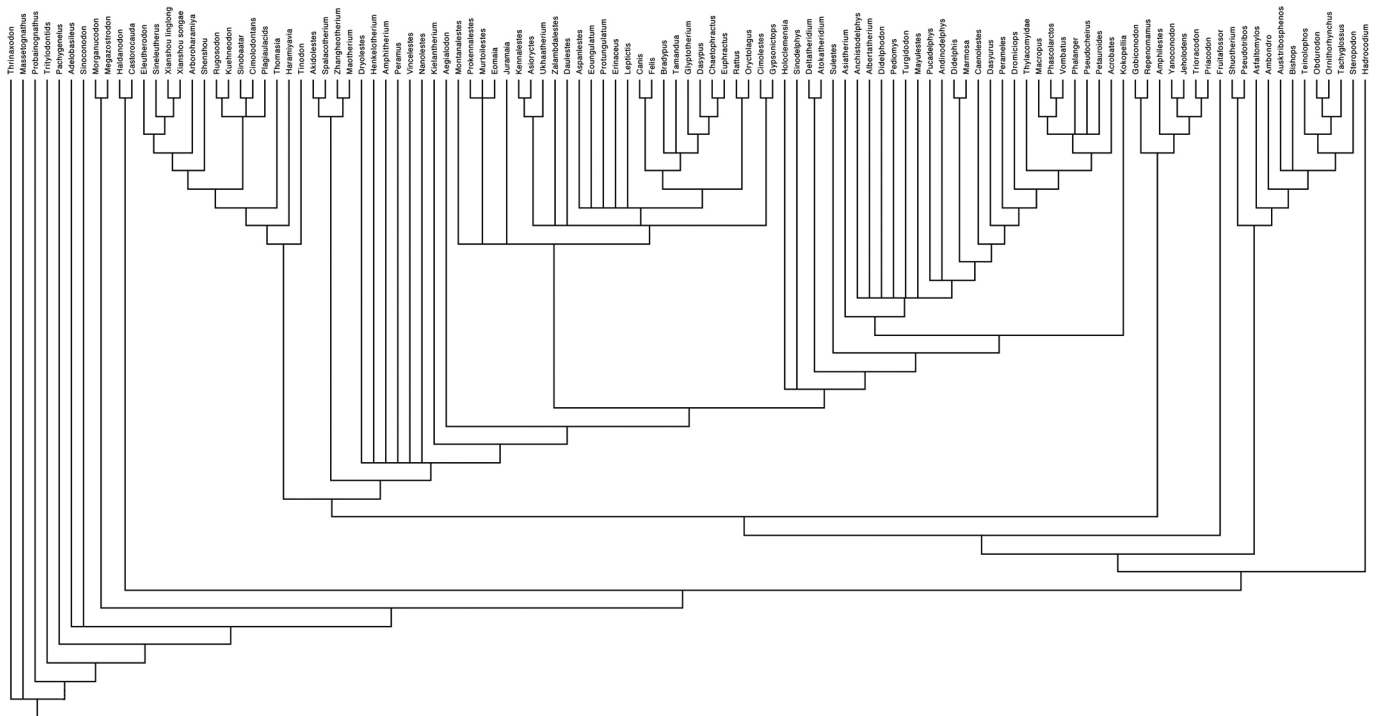
Extended Data Figure 6 | Close-up views of craniodental features of *Xianshou songae* (BMNHC-PM003253). a, Photographs and stippling drawing of the dentition of *Xianshou songae* (BMNHC-PM003253). The distolingual corner of M^2 is not exposed, indicated by hatched lines. P_4 and M_1 are also shown flipped horizontally and the lingual sides of P_4 and M_1 are not exposed, indicated by hatched lines. The dentition was coated with ammonium

chloride to enhance contrast. b, The skull and mandibles of *Xianshou songae* (BMNHC-PM003253). Note that the right upper and lower dentition are preserved in occlusion, showing that lingual cusp A_1 of M_1 bites into the valley of upper M_2 . amf, anterior border of the masseteric fossa; mac, mandibular condyle; maf, mandibular foramen; ptf, pterygoid fossa. c, Teeth *in situ* of *Xianshou songae* (BMNHC-PM003253).



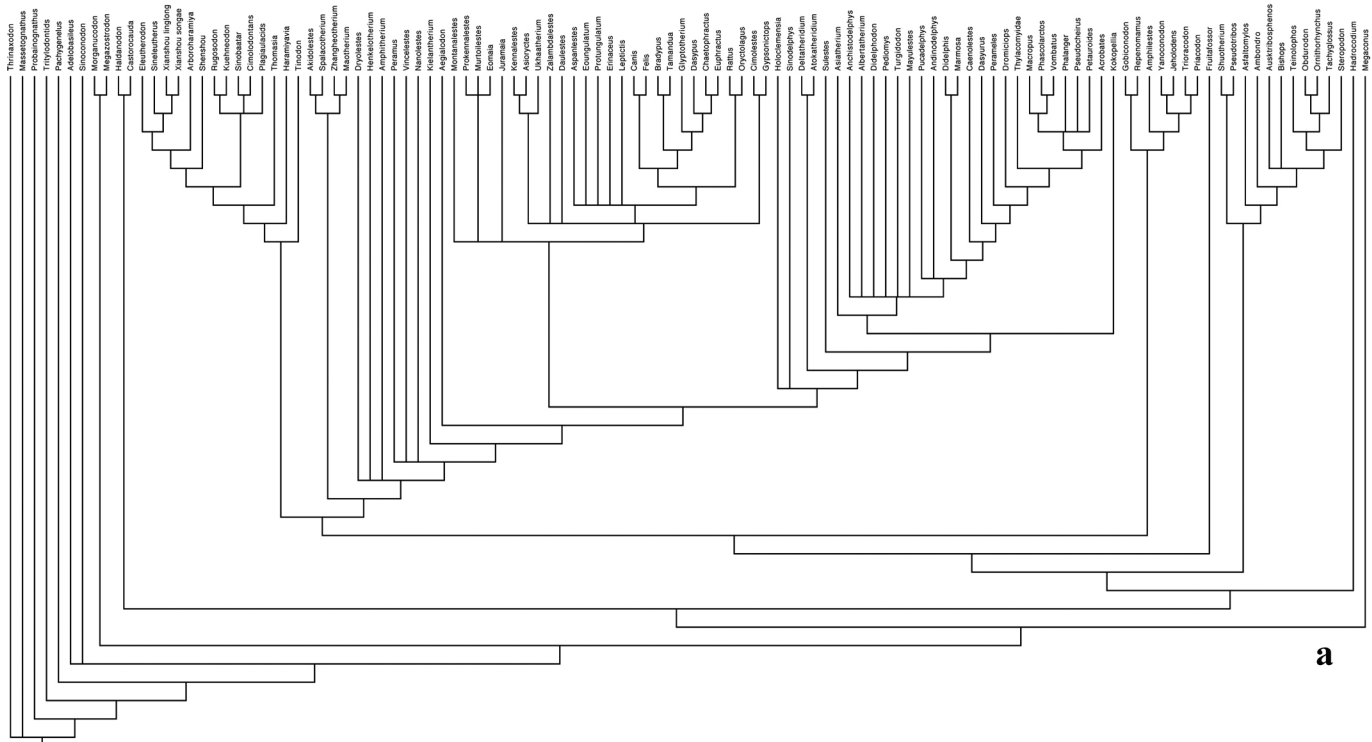
Extended Data Figure 8 | Euharamiyidan manus and pes, and their ternary plots. a. The manus of *Shenshou lui* (LDNHMF2001). **b.** The pes of *Shenshou lui* (JZT-D061). as, astragalus; cm, calcaneum; ct, capitate; cu, cuboid; en, entocuneiform; ip, intermediate phalanges; mc, metacarpal; mt, metatarsal; pi, pisiform; pp, proximal phalanges; sp, scaphoid; td, trapezoid; tm, trapezium; tp, terminal phalanges; tq, triquetrum; ts, tarsal spur. **c.** Ternary diagrams showing intrinsic manual and pedal ray III proportions. Ternary plots showing relative metapodial, proximal and intermediate phalangeal lengths for the third digit ray of the hand and foot. The lengths of the third metapodial, proximal

phalanx and intermediate phalanx are shown on their respective axes as a percentage of the combined length of the three segments. Compared to both fossil and extant taxa, euharamiyidans have the intrinsic manual and pedal ray proportions typical of arboreal species in which the proximal and intermediate phalanges are long relative to the metapodials. Ar, *Arboroharamiya jenkinsi*; Eo, *Eomaia scansoria*; Je, *Jeholodens jenkinsi*; Ma, *Maothierium sinensis*; Sb, *Sinobaatar lingyuanensis*; Sd, *Sinodelphys szalayi*; Sh, *Shenshou*; XI, *Xianshou linglong*; Xs, *Xianshou songae*.

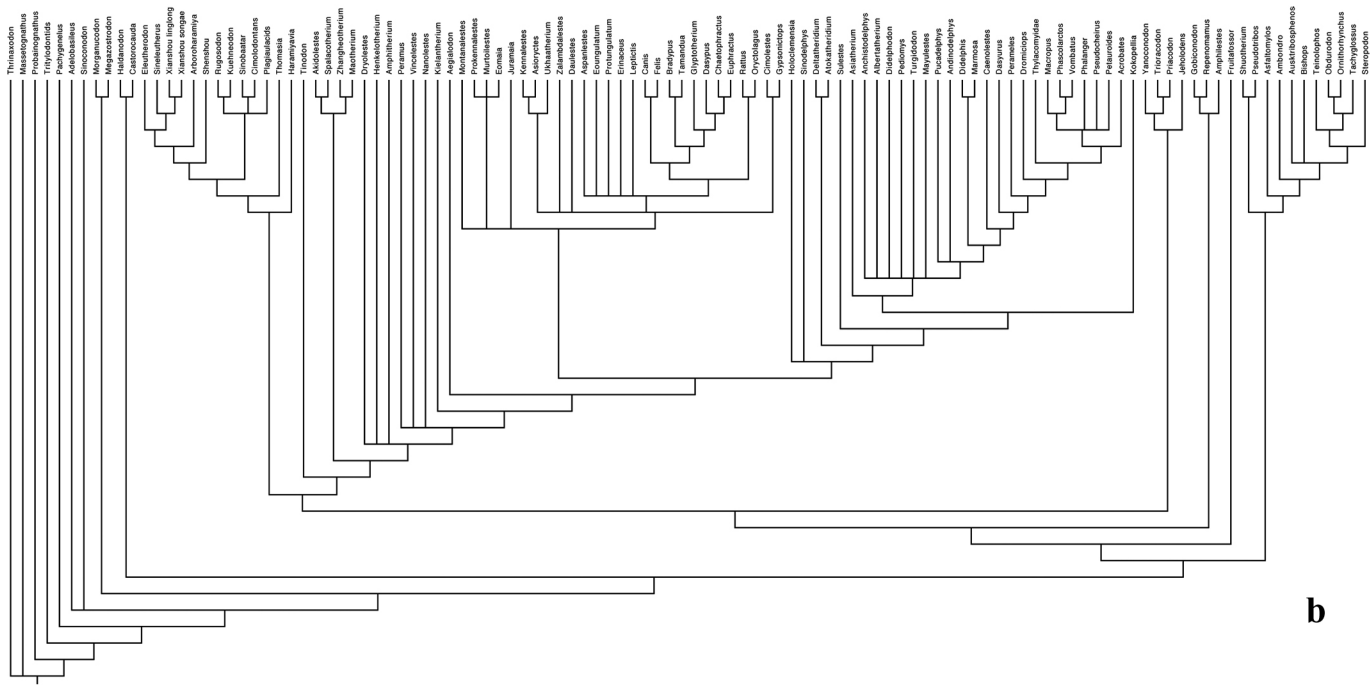


Extended Data Figure 9 | The strict consensus tree of 198 equally most parsimonious trees. Tree length = 2,509; consistency index (CI) = 0.3304; homoplasy index (HI) = 0.6696; retention index (RI) = 0.7901. The simplified version of this consensus tree is presented in Fig. 4. Note that *Tinodon* was placed as the sister taxon of Allotheria in this cladogram, which is inconsistent with the taxonomic position of the genus³³. This tree topology and that in Extended Data Fig. 10b show the phylogenetic instability of *Tinodon*, as in other studies^{11,16,17,40}, that may be attributed to a significant amount of missing

data. However, the current phylogenetic position of *Tinodon* may reflect the possibility that the allotherian tooth pattern, as represented by *Haramiyavia*, was derived by developing an extra cusp row from a triconodont-like tooth pattern or even from a tooth pattern with an initially reversed triangular cusp pattern or even from a tooth pattern with an initially reversed triangular cusp pattern or even from a tooth pattern with an initially reversed triangular cusp pattern, as discussed in the main text. Better material is needed to secure the phylogenetic position of *Tinodon* and to test our hypothesis. See Supplementary Information, section H, for methods as well as tree 1 and related data.



a



b

Extended Data Figure 10 | Strict consensus trees with *Megaconus* and without both *Megaconus* and *Hadrocodium*. **a**, The strict consensus tree of 133 equally most parsimonious trees (including *Megaconus*). Tree length = 2,528; CI = 0.3283; HI = 0.6717; RI = 0.7910. Note that *Megaconus* is separated from all other mammals and placed outside the crown mammals. This suggests that the placement of *Megaconus* in all other mammals⁴⁰ needs to be tested when new material becomes available. See Supplementary Information, section

E, for discussion. **b**, The strict consensus tree of 101 equally most parsimonious trees without *Megaconus* and *Hadrocodium*. Tree length = 2,448; CI = 0.3386; HI = 0.6614; RI = 0.7952. See Supplementary Information, section G, for discussion. Note that *Tinodon* is clustered with other ‘symmetrodonts’ plus the clade leading to therians, and that eutriconodonts are split into two groups. See Supplementary Information, sections E and G, for discussion.

Mesenchymal–endothelial transition contributes to cardiac neovascularization

Eric Ubil^{1*}, Jinzhu Duan^{2,3,4,5,6,7*}, Indulekha C. L. Pillai^{2,3,4,5,6,7}, Manuel Rosa–Garrido^{2,3,8,9}, Yong Wu⁸, Francesca Bargiacchi¹⁰, Yan Lu^{2,3,4,5,6,7}, Seta Stanbouly^{2,3,4,5,6,7}, Jie Huang^{2,3,4,5,6,7}, Mauricio Rojas¹¹, Thomas M. Vondriska^{2,3,7,8,9}, Enrico Stefani^{3,8,9} & Arjun Deb^{2,3,4,5,6,7}

Endothelial cells contribute to a subset of cardiac fibroblasts by undergoing endothelial–to–mesenchymal transition, but whether cardiac fibroblasts can adopt an endothelial cell fate and directly contribute to neovascularization after cardiac injury is not known. Here, using genetic fate map techniques, we demonstrate that cardiac fibroblasts rapidly adopt an endothelial–cell–like phenotype after acute ischaemic cardiac injury. Fibroblast–derived endothelial cells exhibit anatomical and functional characteristics of native endothelial cells. We show that the transcription factor p53 regulates such a switch in cardiac fibroblast fate. Loss of p53 in cardiac fibroblasts severely decreases the formation of fibroblast–derived endothelial cells, reduces post–infarct vascular density and worsens cardiac function. Conversely, stimulation of the p53 pathway in cardiac fibroblasts augments mesenchymal–to–endothelial transition, enhances vascularity and improves cardiac function. These observations demonstrate that mesenchymal–to–endothelial transition contributes to neovascularization of the injured heart and represents a potential therapeutic target for enhancing cardiac repair.

The mammalian heart after acute injury heals primarily by fibrosis. Cardiac fibroblasts proliferate at the site of injury¹ and fibroblast proliferation is accompanied by recruitment of endothelial cells. Endothelial cells contribute to neovascularization of the injury region² and promote repair³. A close interaction between fibroblasts and endothelial cells is thought to regulate wound healing⁴. A subset of endothelial cells, by undergoing endothelial–mesenchymal transition, generates fibroblasts in the injury region⁵ and cardiac fibroblasts express pro-angiogenic molecules that in turn promote angiogenesis^{6,7}. However, cardiac fibroblasts are thought to be terminally differentiated cells^{8,9} and whether they have the ability to adopt an endothelial phenotype and directly contribute to neovascularization after cardiac injury is not known. Here we demonstrate that cardiac fibroblasts undergo mesenchymal–endothelial transition to generate *de novo* endothelial cells in the injured heart and show that mesenchymal–endothelial transition (or MEndoT) can be augmented to enhance cardiac repair.

Cardiac fibroblasts adopt endothelial–cell–like fates

We used a genetic fate map strategy to label cardiac fibroblasts by crossing transgenic mice harbouring a tamoxifen-inducible Cre recombinase driven by fibroblast-specific regulatory sequence of the $\alpha 2$ (type 1) collagen gene (Col1a2–CreERT)^{10–12} with the lineage reporter strain (Rosa26R^{tdTomato})¹³ to create Col1a2–CreERT:R26R^{tdTomato} progeny mice. In these mice, administration of tamoxifen results in activation of Cre recombinase, and cells expressing Col1a2 at the time of tamoxifen administration are irreversibly labelled by tdTomato fluorescence. We administered tamoxifen for 10 days to adult Col1a2–CreERT:R26R^{tdTomato} mice. Five days after cessation of tamoxifen administration, we observed that approximately 55% of all non-myocyte cells exhibited tdTomato fluorescence and greater than 96% and 99% of tdTomato fluorescent

cells expressed the cardiac fibroblast markers domain discoidin receptor 2 (DDR2) and vimentin (Extended Data Fig. 1a–c). Immunofluorescent staining showed that $87 \pm 9\%$ and $99 \pm 0.5\%$ (mean \pm s.e.m.) of tdTomato-labelled cells expressed DDR2 and vimentin, respectively, supporting flow cytometry data (Extended Data Fig. 1d, e). tdTomato cells did not express endothelial markers VECAD and CD31 ($99.9 \pm 0.06\%$ and $99.8 \pm 0.02\%$ negative, respectively; mean \pm s.e.m.) (Extended Data Fig. 1f, g), and did not express the cardiac progenitor marker c-Kit nor markers of smooth muscle, macrophages and lymphatics (Extended Data Fig. 1h–k). Cardiac myocytes did not express Cre recombinase as previously shown¹⁰. Taken together these data strongly suggest that cells exhibiting tdTomato fluorescence in hearts of Col1a2–CreERT:R26R^{tdTomato} mice are cardiac fibroblasts and do not express canonical markers of other cardiovascular cell types.

We subjected Col1a2–CreERT:R26R^{tdTomato} mice to ischaemia-reperfusion cardiac injury 5 days after cessation of tamoxifen injection. By day 3 post-injury, $35 \pm 3\%$ (mean \pm s.e.m.) of labelled cardiac fibroblasts in the region of injury expressed the endothelial-specific marker VECAD, while in sham-injured animals only rare labelled cells expressed VECAD ($<0.3\%$) (Fig. 1a–c). Approximately $24 \pm 4\%$, $44 \pm 4\%$ and $35 \pm 3\%$ (mean \pm s.e.m.) of labelled cardiac fibroblasts also expressed other endothelial markers such as endothelial nitric oxide synthase (eNOS) and the endothelial tight junctional proteins claudin 5 (ref. 14) and occludin¹⁴, respectively (Fig. 1a–c). Mesenchymal–endothelial transition was most pronounced in the injury border zone and significantly decreased in regions remote from the infarct (Fig. 1c). The fraction of cardiac fibroblasts expressing VECAD increased between 1 and 3 days post-injury and remained similar at 3, 7 and 14 days (Fig. 1d). The fraction of tdTomato-positive cells expressing VECAD in sham-injured animals at 3, 7 and 14 days was $0.3 \pm 0.1\%$, $1.4 \pm 1.4\%$ and $0.6 \pm 0.4\%$ (mean \pm s.e.m.),

¹Department of Cell Biology & Physiology, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599, USA. ²Division of Cardiology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. ³Cardiovascular Research Laboratory, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. ⁴Eli and Edythe Broad Institute of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, California 90095, USA. ⁵Department of Molecular, Cell and Developmental Biology, College of Letters and Sciences, University of California, Los Angeles, California 90095, USA. ⁶Jonsson Comprehensive Cancer Center, University of California, Los Angeles, California 90095, USA. ⁷Molecular Biology Institute, University of California, Los Angeles, California 90095, USA. ⁸Division of Molecular Medicine, Department of Anesthesiology, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. ⁹Department of Physiology, David Geffen School of Medicine, University of California, Los Angeles, California 90095, USA. ¹⁰Department of Pharmacology, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599, USA. ¹¹Department of Medicine, McAllister Heart Institute, University of North Carolina, Chapel Hill, North Carolina 27599, USA.

*These authors contributed equally to this work.

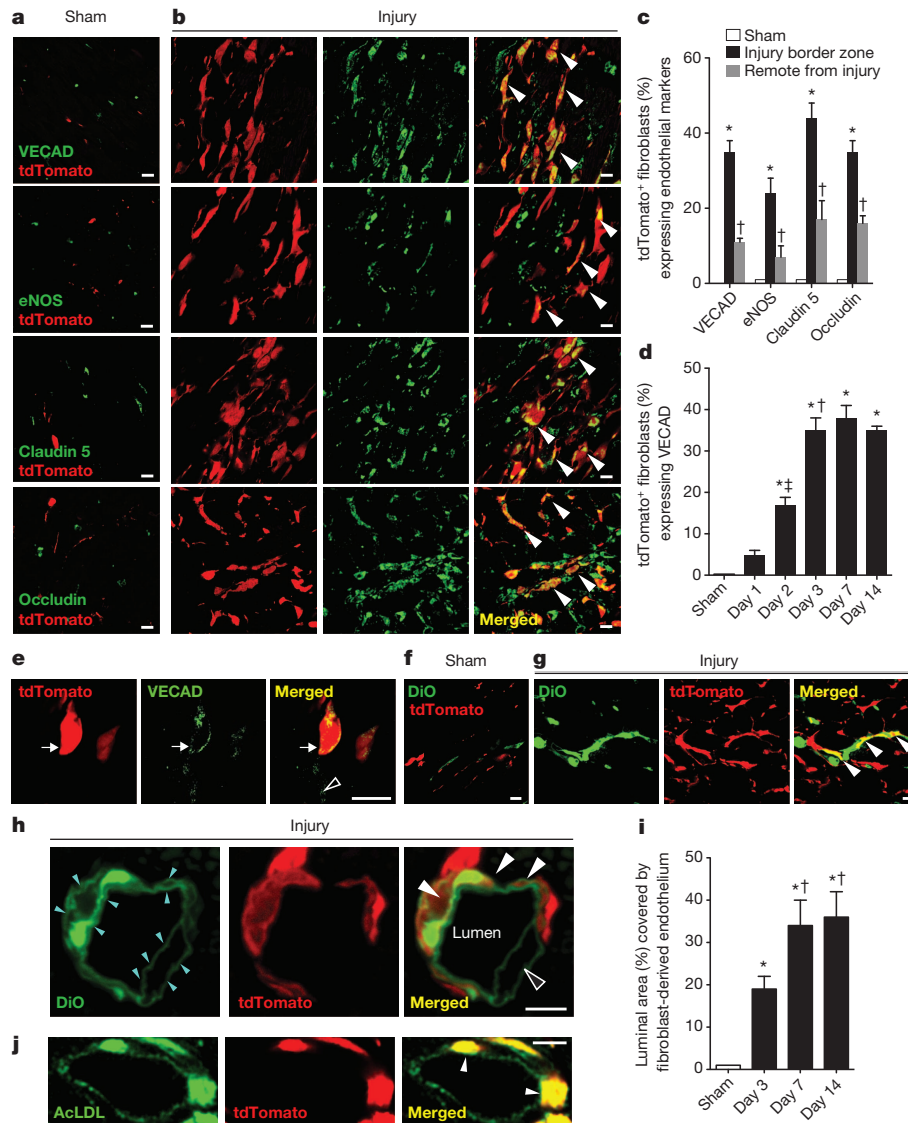


Figure 1 | Cardiac fibroblasts adopt endothelial cell fates after cardiac injury. **a, b,** Hearts from Col1a2-CreERT:R26R^{tdTomato} mice immunostained for endothelial markers (arrowheads). **c,** tdTomato⁺ fibroblasts (%) expressing endothelial markers (**P* < 0.05 versus sham, †*P* < 0.05 versus injury border zone). **d,** Temporal expression of VECAD (**P* < 0.005 versus sham, †*P* < 0.05 versus day 1 and 2, ‡*P* < 0.05 versus day 1). **e,** STED microscopy demonstrating tdTomato⁺VECAD⁺ cell (arrow; open arrowhead shows tdTomato⁻ endothelial cell). **f–h,** DiO-stained capillary in longitudinal section

(**f, g;** arrowheads show tdTomato⁺DiO⁺ cells) or cross-section (**h;** cyan arrowheads show DiO-stained inner and outer endothelial cell membranes; white arrowheads show tdTomato⁺ endothelial cell; unfilled arrowhead shows tdTomato⁻ endothelial cell). Scale bar: 5 μm. **i,** Luminal surface area occupied by fibroblast-derived endothelium (**P* < 0.005 versus sham, †*P* < 0.05 versus day 3). **j,** acLDL uptake by tdTomato⁺ endothelium (arrowheads) (*n* = 3 animals per group per time point). All graphs show mean ± s.e.m.; scale bar, 10 μm unless stated in the legend.

P > 0.05, one way ANOVA), demonstrating no temporal difference in the fraction of tdTomato-labelled cells expressing VECAD after sham injury.

As fibroblasts lie in close apposition to endothelial cells and pericytes, we performed super-resolution microscopy to validate our observations with confocal microscopy. Stimulated emission depletion (STED) microscopy is a form of super-resolution microscopy that provides an average lateral resolution of 30–40 nm compared to confocal microscopy, which provides an average lateral resolution of 250 nm (ref. 15). Using STED we observed that tdTomato cells after injury express VECAD and can be distinguished from closely apposed endothelial cells not expressing tdTomato label (Fig. 1e). STED microscopy also demonstrated that tdTomato cells did not express pericyte markers NG2 or CD146 (Extended Data Fig. 1, m).

We next investigated whether fibroblast-derived endothelial cells incorporate into capillaries in the infarct border zone. We perfused Col1a2-CreERT:R26R^{tdTomato} mice with a fluorescent lipophilic dye

DiO that labels endothelial cell membranes and has been used to identify the vasculature of solid organs including the heart¹⁶. In both longitudinal (Fig. 1f, g) and transverse sections of capillaries (Fig. 1h) in the injury border zone, we observed that endothelial cells bearing the fibroblast label lined the lumen of the blood vessel, but the contribution was minimal in sham-injured animals (Fig. 1f). Luminal DiO⁺ tdTomato⁺ cells expressed VECAD (Extended Data Fig. 2a), thus confirming the endothelial phenotype. Luminal surface area occupied by fibroblast-derived endothelial cells peaked at 7 days post-injury and remained similar at 14 days (Fig. 1i). Endothelial cells are known to take up acetylated LDL (acLDL)¹⁷. We systemically injected Col1a2-CreERT:R26R^{tdTomato} mice 3 days after cardiac injury with fluorescent-labelled acLDL and observed that 41 ± 8% (mean ± s.e.m.) of fibroblast-derived endothelial cells took up acLDL, demonstrating their functional similarity to native endothelial cells (Fig. 1j).

A subset of cardiac fibroblasts after cardiac injury become myofibroblasts and express α-smooth muscle actin (αSMA)^{18,19}. We stained for

α SMA and found that 3 days after injury, $11 \pm 3\%$ (mean \pm s.e.m.) of labelled cardiac fibroblasts in the injured region expressed α SMA but very few labelled cells co-expressed α SMA and VECAD, suggesting that cardiac fibroblasts undergoing mesenchymal–endothelial transition are distinct from myofibroblasts (Extended Data Fig. 2b–f).

We next determined whether substantial recombination in endothelial cells (that is, leaky promoter elements) could potentially confound our findings. Stochastic induction of Cre can occur after cardiac injury²⁰. We injected Col1a2-CreERT:R26R^{tdTomato} mice with vehicle (corn oil) instead of tamoxifen and observed rare labelled cardiac fibroblasts ($<0.07\%$ at 3 and 7 days) after ischaemic injury, demonstrating that injury alone did not lead to significant Cre activation (Extended Data Fig. 3a–c).

Endothelial cells are known to undergo endothelial–mesenchymal transition and adopt a fibroblast phenotype after cardiac injury⁵. We investigated whether endothelial cells undergoing endothelial–mesenchymal transition would exhibit tdTomato fluorescence and be mistakenly included in our analysis. We performed immunofluorescent staining for Col1 and observed that approximately $6 \pm 1\%$ (mean \pm s.e.m.) of endothelial cells after injury expressed Col1 but none of these cells exhibited tdTomato fluorescence (Extended Data Fig. 3d). Conversely, tdTomato-labelled cells in the injury region that expressed VECAD did not stain for Col1 (Extended Data Fig. 3e). Next, we crossed Col1a2-CreERT:R26R^{tdTomato} mice with Col1-GFP transgenic mice, which have GFP expression directly driven by the *Col1a1* promoter. GFP expression thus serves as a useful real-time reporter of Col1 expression^{21,22}. Progeny mice (Col1a2-CreERT:R26R^{tdTomato}:Col1-GFP) were administered tamoxifen and subjected to ischaemic myocardial injury. At 3 days post-injury, tdTomato-labelled cells that expressed VECAD did not exhibit GFP fluorescence (Extended Data Fig. 4a). Approximately $3 \pm 1\%$ (mean \pm s.e.m.) of VECAD-positive cells were GFP positive but with the exception of rare cells, VECAD⁺GFP⁺ cells did not exhibit tdTomato fluorescence (Extended Data Fig. 4b).

Immunofluorescent staining for Cre protein in Col1a2-CreERT:R26R^{tdTomato} mice hearts after injury demonstrated that tdTomato⁺ VECAD⁺ cells did not express Cre (Extended Data Fig. 4c–i), consistent with the earlier observation that these cells do not express type 1 collagen. In the absence of tamoxifen, tdTomato cells expressing Col1a2 would be expected to have Cre in their cytoplasm. In hearts of Col1a2-CreERT:R26R^{tdTomato}:Col1-GFP mice, tdTomato-labelled cells that expressed type 1 collagen (GFP positive) exhibited cytoplasmic but not nuclear Cre staining (Extended Data Fig. 4j–m). Taken together, these observations demonstrate the fidelity of the Cre driver and argue against a ‘leaky’ Cre as the basis for our findings.

Next we investigated whether Col1a2-expressing endothelial progenitors in the heart or bone marrow could have been labelled during administration of tamoxifen and subsequently generated tdTomato⁺ endothelial cells. We analysed the heart and bone marrow but observed that less than 0.5% of tdTomato cells in the heart (Extended Data Fig. 5a, b) or bone marrow (Extended Data Fig. 5c–f) expressed endothelial progenitor markers (CD133, CD34 and Flk1)²³. Bone-marrow-derived mesenchymal stem cells did not exhibit tdTomato fluorescence, with the exception of rare cells (Extended Data Fig. 5g). Collectively, these observations suggest that substantial labelling of endothelial progenitors did not occur, making them an unlikely source of tdTomato⁺ endothelial cells.

Finally, we subjected FSP1-Cre:R26R^{tdTomato} mice to ischaemic cardiac injury to confirm mesenchymal–endothelial transition with another Cre driver. The FSP1-Cre driver has been used to track cardiac fibroblast after *in vivo* reprogramming^{8,9}. By day 3 post injury, $31 \pm 4\%$ and $23 \pm 5\%$ of tdTomato-labelled cells in the injury region expressed the endothelial markers isolectin and VECAD, respectively, compared to $2 \pm 1\%$ and $4 \pm 1\%$ in sham-injured animals (Extended Data Fig. 6a–c), confirming mesenchymal–endothelial transition observations made using the Col1a2-CreERT driver.

p53 mediates MEndoT *ex vivo*

We next investigated the mechanisms regulating mesenchymal–endothelial transition and hypothesized that cellular stress after cardiac injury has a role in mesenchymal–endothelial transition. p53 is an important cellular stress-response gene²⁴, modulates reprogramming²⁵ and regulates epithelial–mesenchymal transition²⁶. We observed that 37% of tdTomato-labelled cardiac fibroblasts expressed p53 at 3 days post injury. In contrast, rare labelled cells expressed p53 in the sham-injured heart (Fig. 2a, b). p53 expression in labelled fibroblasts peaked at 7 days after cardiac injury (Fig. 2c). By day 7 after injury, approximately $91 \pm 7\%$ (mean \pm s.e.m.) of tdTomato-labelled cells expressing p53 co-expressed VECAD (Fig. 2d), demonstrating a strong association between p53 and VECAD expression in tdTomato-labelled cells.

To determine whether p53 has a regulatory role in mesenchymal–endothelial transition, we first established an *ex vivo* model. Cells subjected to types of cellular stress *ex vivo*, such as serum deprivation, upregulate p53 levels^{27,28}. We subjected cardiac fibroblasts to serum starvation and observed increased p53 levels by western blotting (Extended Data Fig. 7a, b). We next seeded labelled cardiac fibroblasts (99% purity by flow cytometry) on matrigel (a mixture of basement membrane proteins that facilitates capillary tube formation) and subjected them to serum starvation. In contrast to control fibroblasts (Fig. 2e, g), serum-starved fibroblasts formed capillary-tube-like structures, expressed VECAD (Fig. 2f) and took up acLDL (Fig. 2h), consistent with adoption of an endothelial-cell-like phenotype.

Using gain- and loss-of-function approaches, we determined whether p53 affects serum-starvation-induced mesenchymal–endothelial transition (Fig. 2i–o). Serum deprivation induced cardiac fibroblasts to form tubes on matrigel (Fig. 2i, j), but addition of pifithrin- α , an inhibitor of p53 (ref. 29), decreased tube formation by 67% (Fig. 2k, n). Pifithrin- α did not have an effect on serum-fed cardiac fibroblasts (Extended Data Fig. 7c, d). Next, we generated mice specifically deficient in fibroblast p53 by crossing Col1a2-CreERT:R26R^{tdTomato} mice with mice having both p53 alleles floxed³⁰. Progeny mice (Col1a2-CreERT:R26R^{tdTomato}:p53^{fl/fl}) were injected with tamoxifen at 8 weeks to delete p53 and label Col1a2-expressing p53-deficient cells (tdTomato⁺). Cardiac fibroblasts were subsequently harvested and subjected to serum starvation. The absence of p53 decreased capillary tube formation on matrigel by 95% (Fig. 2l, n). We subsequently investigated whether activation of p53 signalling enhanced serum-deprivation-induced mesenchymal–endothelial transition. Cardiac fibroblasts subjected to serum starvation were treated with the small molecule RITA (reactivation of p53 and induction of tumour apoptosis) that inhibits ubiquitin-mediated p53 degradation and enhances p53 signalling³¹. RITA enhanced tube formation by 50% (Fig. 2m, n). RITA also increased tube formation in serum-fed cardiac fibroblasts but to a lesser degree (Extended Data Fig. 7c, d). After serum deprivation, quantitative real-time polymerase chain reaction (qPCR) demonstrated 4–20-fold increased expression of endothelial-specific genes but treatment with pifithrin- α or the genetic deletion of p53 in cardiac fibroblasts significantly blunted induction of endothelial gene expression (Fig. 2o). Conversely, treatment of serum-starved cardiac fibroblasts with RITA significantly increased expression of endothelial-specific genes by 3–7-fold (Fig. 2o). Interestingly, overexpressing p53 in wild-type cardiac fibroblasts grown under serum-fed conditions did not induce expression of VECAD or other endothelial-specific genes, suggesting that enhancing p53 signalling in the absence of serum starvation is not sufficient to induce mesenchymal–endothelial transition.

We next performed chromatin immunoprecipitation (ChIP) for p53 on serum-starved cardiac fibroblasts and observed that p53 directly binds to promoter regions of endothelial-specific genes and the endothelial transcription factors *Hoxa9* (ref. 32) and *Hoxd3* (ref. 33) that are known to have an important role in endothelial differentiation (Fig. 2p). Gene expression of *Hoxa9* and *Hoxd3* was also increased in cardiac fibroblasts by 6.53 ± 1.74 and 7.18 ± 1.16 -fold, respectively, after serum starvation (mean \pm s.e.m., $P < 0.05$ compared to cardiac fibroblasts in 10% serum). Although an indirect effect of p53 cannot be excluded,

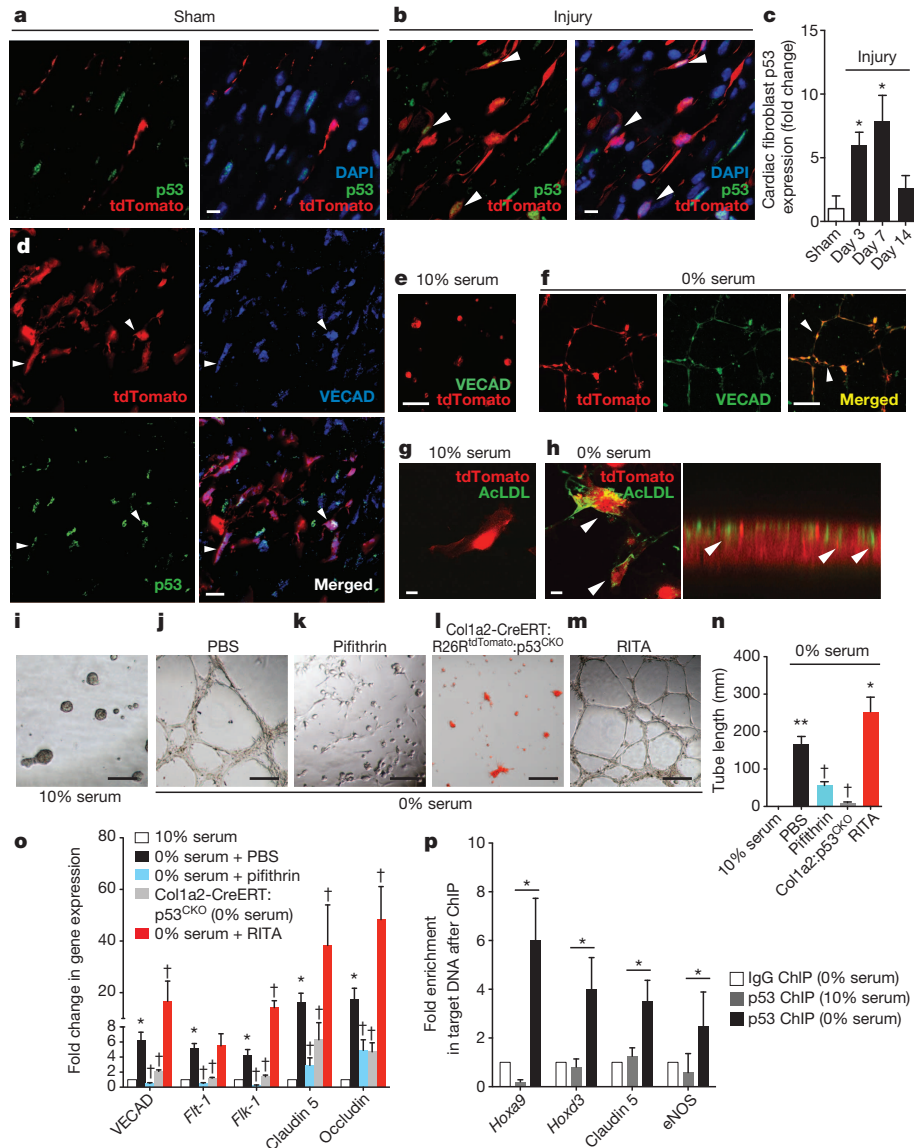


Figure 2 | Cardiac fibroblasts upregulate p53 after injury and p53 mediates mesenchymal-endothelial transition *ex vivo*. **a, b**, p53 immunostaining in injured hearts (arrowheads show tdTomato⁺p53⁺ cells). **c**, Temporal p53 expression in labelled fibroblasts (**P* < 0.05 versus sham, *n* = 3 animals per time point). **d**, Co-expression of p53, VECAD and tdTomato (arrowheads). **e–h**, tdTomato⁺VECAD⁺ tubes (**e, f**) (scale bar, 250 μm) and acLDL uptake (**g, h**) after serum starvation (arrowheads, *n* = 4) (scale bar, 20 μm). Right panel of **h** shows confocal image (xz plane) showing acLDL internalization.

these observations suggest that p53 at least in this model initiates an endothelial gene expression program by directly inducing transcription of endothelial-specific genes.

Using this *ex vivo* model, we next asked whether mesenchymal-endothelial transition was reversible. Transforming growth factor-β (TGF-β) enhances endothelial-mesenchymal transition⁵ and TGF-β added to cardiac fibroblasts at the onset of serum starvation prevented tube formation (Extended Data Fig. 8a, b) and induction of VECAD expression (0.98 ± 0.03-fold change in VECAD with TGF-β, mean ± s.e.m., *P* > 0.05, *n* = 3). Moreover, when TGF-β was added to serum-starved cardiac fibroblasts after they had already formed tubes, it led to 99% regression of tube formation (Extended Data Fig. 8c, d, g). A similar effect was observed with adding back serum (Extended Fig. 8e–g). VECAD expression also decreased by 56.4 ± 2.4% (mean ± s.e.m., *P* < 0.05, *n* = 3). Addition of pifithrin-α to serum-starved cardiac fibroblasts that

i–m, Tube formation of cardiac fibroblasts in 10% serum (**i**) or 0% serum with PBS (**j**), 100 μM pifithrin-α (**k**), 0.1 μM RITA (**m**), or p53 deletion (**l**) (bright-field and fluorescence overlay). Scale bar, 250 μm. **n**, Quantification of tube length (***P* < 0.005 versus 10% serum, †*P* < 0.005 and **P* < 0.05 versus starved cells, *n* = 3). **o**, Endothelial gene expression in cardiac fibroblasts (**P* < 0.005 versus 10% serum, †*P* < 0.05 versus PBS, *n* = 8). **p**, ChIP with p53 (**P* < 0.05). All graphs show mean ± s.e.m.; scale bar, 10 μm unless mentioned.

had already formed tubes demonstrated significant disruption of formed tubes compared to PBS-treated controls (Extended Fig. 8h–n). These observations suggest that p53 is required for maintaining the endothelial phenotype of the fibroblast-derived endothelial-like cell. Notably, the fraction of fibroblast-derived endothelial cells was maintained at 14 days after cardiac injury *in vivo* (Figs 1d and 2c) despite declining p53 levels suggestive of other factors stabilizing the endothelial phenotype.

p53 mediates MendoT in injured heart

To determine whether p53 mediates mesenchymal-endothelial transition *in vivo*, we subjected Col1a2-CreERT:R26R^{tdTomato};p53^{fl/fl} mice to ischaemic cardiac injury 5 days after cessation of tamoxifen. Mice with intact p53 exhibited 6 ± 0.64-fold increase (mean ± s.e.m.) in cardiac fibroblast p53 levels 3 days after cardiac injury, but in mice with fibroblast-specific p53 deletion (referred to as Col1a2-CreERT:R26R^{tdTomato};p53^{CKO})

(conditional knock out) the mean p53 expression in border zone cardiac fibroblasts failed to increase significantly (Fig. 3a–c). The degree of mesenchymal–endothelial transition at 3 days post-injury in Col1a2-CreERT:R26R^{tdTomato}:p53^{CKO} mice decreased by 57% (Fig. 3d–f) and was associated with a decrease in capillary density in the injury region (Fig. 3g). Increase in vessel density in the injured heart is a critical post-ischæmic repair mechanism. Diminished neovascularization can lead to rapid decline in cardiac function after cardiac injury and is associated with worsened scarring³⁴. Echocardiography on hearts of mice 7 days after cardiac injury showed significant worsening of cardiac function in p53^{CKO} mice (Fig. 3h, i). Masson's trichrome staining demonstrated a greater amount of collagen deposition in the hearts of Col1a2-CreERT:R26R^{tdTomato}:p53^{CKO} mice (Fig. 3j, k). The degree of inflammatory infiltrate measured by the number of CD68-expressing macrophages was also significantly higher in the Col1a2-CreERT:R26R^{tdTomato}:p53^{CKO} mice (Extended Data Fig. 9a, b, d). These observations demonstrate that p53 is necessary for mesenchymal–endothelial transition to occur after ischaemic cardiac injury and that disruption of mesenchymal–endothelial transition is associated with diminished post-injury vascularity and cardiac function.

We next looked at mechanisms inducing p53 expression in cardiac fibroblasts. Carboxy-terminal phosphorylation of histone H2A family member X (γ H2AX) is a DNA damage response pathway upstream of p53, is activated by hypoxic endothelial cells and shown to have a critical role in endothelial cell proliferation and hypoxic neovascularization³⁵. We observed that 90% of tdTomato cells that upregulated p53 co-expressed γ H2AX, suggesting that the γ H2AX pathway is strongly associated with p53 expression in fibroblasts undergoing mesenchymal–endothelial transition (Extended Data Fig. 10a–c).

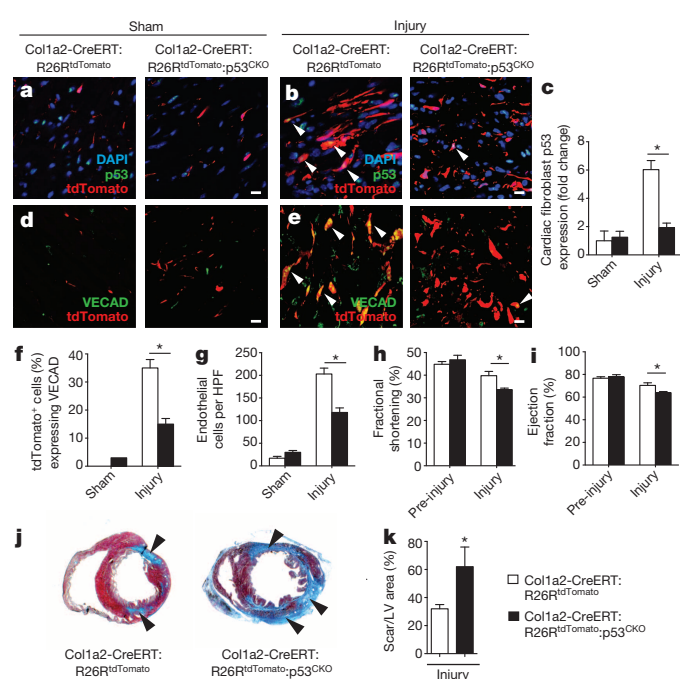


Figure 3 | Mesenchymal–endothelial transition after cardiac injury is p53 dependent. a–c, p53 immunostaining in Col1a2-CreERT:R26R^{tdTomato} and Col1a2-CreERT:R26R^{tdTomato}:p53^{CKO} hearts (a, b; arrowheads) and quantification of p53 expression in labelled fibroblasts (c; **P* < 0.005). d–f, VECAD immunostaining in labelled fibroblasts (d, e; arrowheads) and labelled fibroblasts (%) expressing VECAD (f; **P* < 0.005). g, Number of endothelial cells per high-power field (HPF) (**P* < 0.005, *n* = 5 mice). h, i, Cardiac function before and 7 days after injury (**P* < 0.05, *n* = 9 mice). j, Masson's trichrome staining 14 days after injury. k, Quantification of fibrosis (**P* < 0.05) (LV, left ventricle, *n* = 4 mice unless mentioned). All graphs show mean ± s.e.m.; scale bar, 10 μ m.

p53 activation enhances MEndoT after cardiac injury

We investigated whether stimulation of p53 signalling after cardiac injury enhances mesenchymal–endothelial transition after acute ischaemic cardiac injury. Col1a2-CreERT:R26R^{tdTomato} mice injected with RITA daily for 3 days after cardiac injury exhibited significantly increased p53 expression in labelled fibroblasts (13 ± 0.8-fold in RITA-injected animals versus 6-fold increase in PBS-injected animals) (Fig. 4a–c). RITA significantly enhanced the degree of mesenchymal–endothelial transition (Fig. 4d–f) and enhanced mesenchymal–endothelial transition was associated with a 26% increase in endothelial cells in the injury border zone (Fig. 4g). Increased neovascularization can lead to decreased collagen deposition and rapid improvement in post-injury cardiac function³. The area of collagen deposition 7 days after cardiac injury decreased by 57% in RITA-injected animals (Fig. 4h, i). Echocardiography demonstrated significantly improved cardiac function in RITA-injected mice compared to PBS-injected control mice (Fig. 4j, k).

Consistent with greater vascularity and better cardiac function, the inflammatory infiltrate was substantially reduced in RITA-injected animals (Extended Fig. 9a, c, d). As RITA increases p53 levels, we also looked at deleterious effects of increased p53 levels, such as apoptosis. We subjected Col1a2-CreERT:R26R^{tdTomato} mice to cardiac injury and consistent with published reports, did not observe increased p53 expression in myocytes after injury (Extended Data Fig. 9e, f)³⁶. In RITA-treated mice 3 days after injury there was not a significant increase in p53-expressing TUNEL-stained nuclei compared to PBS-injected controls (Extended Data Fig. 9g).

To demonstrate that increased mesenchymal–endothelial transition after administration of RITA was p53-dependent in cardiac fibroblasts, we injected RITA into Col1a2-CreERT:R26R^{tdTomato}:p53^{CKO} mice and

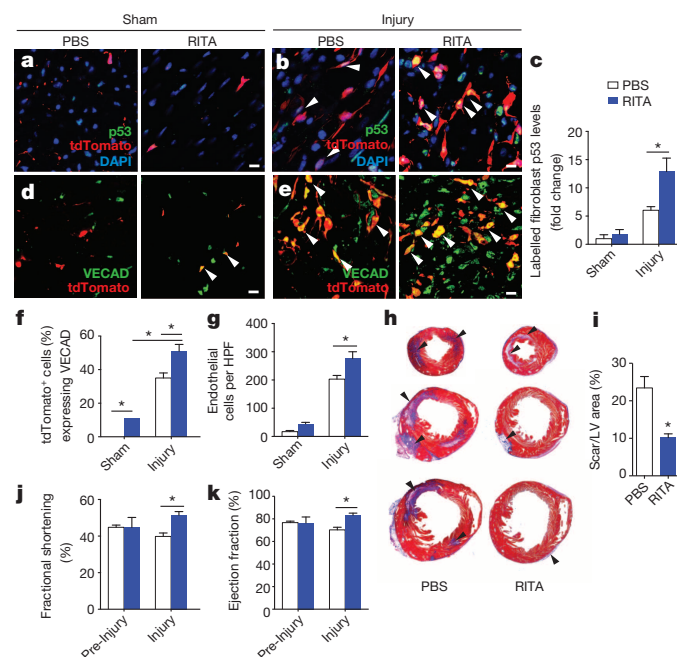


Figure 4 | RITA enhances mesenchymal–endothelial transition after cardiac injury. Hearts from PBS-treated (white bars) or RITA-treated (blue bars) Col1a2-CreERT:R26R^{tdTomato} mice. a–c, p53 immunostaining in labelled fibroblasts (a, b; arrowheads) and quantification of p53 expression (c; **P* < 0.005). d, e, VECAD immunostaining in labelled fibroblasts (arrowheads). f, Labelled fibroblasts (%) expressing VECAD (**P* < 0.05). g, Endothelial cells per HPF (**P* < 0.05). h, Masson's trichrome staining (blue, arrowheads) near apex (top row), between mid-ventricle and apex (middle row) and mid-ventricle (bottom row) (*n* = 14 mice). i, Quantification of fibrosis (**P* < 0.001, *n* = 14 mice). j, k, Cardiac function before and 7 days after injury (**P* < 0.05, *n* = 8 mice). *n* = 4 mice unless mentioned; all graphs show mean ± s.e.m.; scale bar, 10 μ m.

observed that RITA failed to significantly enhance mesenchymal–endothelial transition. The degree of mesenchymal–endothelial transition was not statistically different from that observed in PBS-treated Col1a2-CreERT:R26R^{tdTomato}.p53^{CKO} mice (Extended Data Fig. 9h, i). These observations suggest that RITA predominantly enhances mesenchymal–endothelial transition by activating p53 in cardiac fibroblasts.

Discussion

Our report suggests that cardiac fibroblasts possess a degree of native cellular plasticity that enables them to adopt endothelial-cell-like fates after cardiac injury. Recruitment and proliferation of endothelial cells in the ischaemic heart is a critical cardiac repair response. Our data point to the fibroblast as a novel and robust source of endothelial cell generation in the injured heart. Mesenchymal–endothelial transition (or MEndoT), a hitherto unreported biological phenomenon, appears to have an important physiological role in cardiac repair, as disruption of mesenchymal–endothelial transition worsened post-infarct vascularity and cardiac function. Teleologically, mesenchymal–endothelial transition provides the heart with an efficient strategy to increase neovascularization rapidly in the injured region. The use of a small molecule to augment mesenchymal–endothelial transition and increase neovascularization after injury suggests that mesenchymal–endothelial transition may represent a therapeutic target for enhancing vascularity and repair of ischaemic tissues.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 June 2013; accepted 4 September 2014.

Published online 15 October 2014.

- Christia, P. *et al.* Systematic characterization of myocardial inflammation, repair, and remodeling in a mouse model of reperfused myocardial infarction. *J. Histochem. Cytochem.* **61**, 555–570 (2013).
- Cohn, J. N., Ferrari, R. & Sharpe, N. Cardiac remodeling—concepts and clinical implications: a consensus paper from an international forum on cardiac remodeling. Behalf of an International Forum on Cardiac Remodeling. *J. Am. Coll. Cardiol.* **35**, 569–582 (2000).
- Kocher, A. A. *et al.* Neovascularization of ischemic myocardium by human bone-marrow-derived angioblasts prevents cardiomyocyte apoptosis, reduces remodeling and improves cardiac function. *Nature Med.* **7**, 430–436 (2001).
- Souders, C. A., Bowers, S. L. & Baudino, T. A. Cardiac fibroblast: the renaissance cell. *Circ. Res.* **105**, 1164–1176 (2009).
- Zeisberg, E. M. *et al.* Endothelial-to-mesenchymal transition contributes to cardiac fibrosis. *Nature Med.* **13**, 952–961 (2007).
- Chintalgattu, V., Nair, D. M. & Katwa, L. C. Cardiac myofibroblasts: a novel source of vascular endothelial growth factor (VEGF) and its receptors Flt-1 and KDR. *J. Mol. Cell. Cardiol.* **35**, 277–286 (2003).
- Zhao, L. & Eghbali-Webb, M. Release of pro- and anti-angiogenic factors by human cardiac fibroblasts: effects on DNA synthesis and protection under hypoxia in human endothelial cells. *Biochim. Biophys. Acta* **1538**, 273–282 (2001).
- Qian, L. *et al.* *In vivo* reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature* **485**, 593–598 (2012).
- Song, K. *et al.* Heart repair by reprogramming non-myocytes with cardiac transcription factors. *Nature* **485**, 599–604 (2012).
- Duan, J. *et al.* Wnt1/β-catenin injury response activates the epicardium and cardiac fibroblasts to promote cardiac repair. *EMBO J.* **31**, 429–442 (2012).
- Zheng, B., Zhang, Z., Black, C. M., de Crombrugge, B. & Denton, C. Ligand-dependent genetic recombination in fibroblasts: a potentially powerful technique for investigating gene function in fibrosis. *Am. J. Pathol.* **160**, 1609–1617 (2002).
- Liu, S., Thompson, K. & Leask, A. CCN2 expression by fibroblasts is not required for cutaneous tissue repair. *Wound Repair Regen.* **22**, 119–124 (2014).
- Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13**, 133–140 (2010).
- Dejana, E. Endothelial cell-cell junctions: happy together. *Nature Rev. Mol. Cell Biol.* **5**, 261–270 (2004).
- Singh, H. *et al.* Visualization and quantification of cardiac mitochondrial protein clusters with STED microscopy. *Mitochondrion* **12**, 230–236 (2012).
- Li, Y. *et al.* Direct labeling and visualization of blood vessels with lipophilic carbocyanine dye Dil. *Nature Protocols* **3**, 1703–1708 (2008).

- Urbich, C. & Dimmeler, S. Endothelial progenitor cells: characterization and role in vascular biology. *Circ. Res.* **95**, 343–353 (2004).
- Tomasek, J. J., Gabbiani, G., Hinz, B., Chaponnier, C. & Brown, R. A. Myofibroblasts and mechano-regulation of connective tissue remodelling. *Nature Rev. Mol. Cell Biol.* **3**, 349–363 (2002).
- Weber, K. T., Sun, Y., Bhattacharya, S. K., Ahokas, R. A. & Gerling, I. C. Myofibroblast-mediated mechanisms of pathological remodelling of the heart. *Nature Rev. Cardiol.* **10**, 15–26 (2013).
- Zhou, B. *et al.* Adult mouse epicardium modulates myocardial injury by secreting paracrine factors. *J. Clin. Invest.* **121**, 1894–1904 (2011).
- Acharya, A. *et al.* The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. *Development* **139**, 2139–2149 (2012).
- Lin, S. L., Kisseleva, T., Brenner, D. A. & Duffield, J. S. Pericytes and perivascular fibroblasts are the primary source of collagen-producing cells in obstructive fibrosis of the kidney. *Am. J. Pathol.* **173**, 1617–1627 (2008).
- Melo, L. G. *et al.* Molecular and cell-based therapies for protection, rescue, and repair of ischemic myocardium: reasons for cautious optimism. *Circulation* **109**, 2386–2393 (2004).
- Sharpless, N. E. & DePinho, R. A. p53: good cop/bad cop. *Cell* **110**, 9–12 (2002).
- Hong, H. *et al.* Suppression of induced pluripotent stem cell generation by the p53-p21 pathway. *Nature* **460**, 1132–1135 (2009).
- Chang, C. J. *et al.* p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nature Cell Biol.* **13**, 317–323 (2011).
- Hasan, N. M., Adams, G. E. & Joiner, M. C. Effect of serum starvation on expression and phosphorylation of PKC-α and p53 in V79 cells: implications for cell death. *Int. J. Cancer* **80**, 400–405 (1999).
- Shi, Y. *et al.* Starvation-induced activation of ATM/Chk2/p53 signaling sensitizes cancer cells to cisplatin. *BMC Cancer* **12**, 571 (2012).
- Komarov, P. G. *et al.* A chemical inhibitor of p53 that protects mice from the side effects of cancer therapy. *Science* **285**, 1733–1737 (1999).
- Marino, S., Vooijs, M., van Der Gulden, H., Jonkers, J. & Berns, A. Induction of medulloblastomas in p53-null mutant mice by somatic inactivation of Rb in the external granular layer cells of the cerebellum. *Genes Dev.* **14**, 994–1004 (2000).
- Issaeva, N. *et al.* Small molecule RITA binds to p53, blocks p53-HDM-2 interaction and activates p53 function in tumors. *Nature Med.* **10**, 1321–1328 (2004).
- Bruhl, T. *et al.* Homeobox A9 transcriptionally regulates the EphB4 receptor to modulate endothelial cell migration and tube formation. *Circ. Res.* **94**, 743–751 (2004).
- Boudreau, N., Andrews, C., Srebrow, A., Ravanpay, A. & Cheresh, D. A. Induction of the angiogenic phenotype by Hox D3. *J. Cell Biol.* **139**, 257–264 (1997).
- Fazel, S. *et al.* Cardioprotective c-kit⁺ cells are from the bone marrow and regulate the myocardial balance of angiogenic cytokines. *J. Clin. Invest.* **116**, 1865–1877 (2006).
- Economopoulou, M. *et al.* Histone H2AX is integral to hypoxia-driven neovascularization. *Nature Med.* **15**, 553–558 (2009).
- Bishopric, N. H. *et al.* Hypoxia-activated apoptosis of cardiac myocytes requires reoxygenation or a pH shift and is independent of p53. *J. Clin. Invest.* **104**, 239–252 (1999).

Acknowledgements We thank the Michael Hooker Confocal Microscopy and the Histology Research Core facilities at UNC, Chapel Hill. We thank the confocal and advanced light microscopy facilities in the Division of Molecular Medicine and California Nanosystems Institute at UCLA, Histology facilities at Brain Research Institute & Pathology and Laboratory Medicine, UCLA, Broad Stem Cell Research Center UCLA flow cytometry laboratories and the Vector core facility at UCLA for technical assistance. We thank K. Caron, V. Bauch, J. Taylor, C. Otey at UNC, Chapel Hill and J. Lusia at UCLA for scientific suggestions and discussion. This work was supported by grants from the National Institutes of Health (NIH R01HL102190) to A. Deb. Super-resolution studies and ChIP studies were supported by NIH grant HL088640 to E.S. and NIH grant HL105699 to T.V. E. Ubil was supported in part by a pre-doctoral grant from the American Heart Association. This work was initiated when A.D. was at the University of North Carolina, Chapel Hill and completed at the University of California, Los Angeles.

Author Contributions E.U. and J.D. performed experiments, obtained and analysed data and did statistical analysis. I.C.L.P., Y.W. and E.S. performed super-resolution microscopy experiments. I.C.L.P., S.S. and J.H. performed flow cytometry and *in vitro* experiments. M.R.-G. and T.M.V. performed ChIP experiments. F.B. did western blotting. M.R. and Y.L. performed animal surgeries and echocardiograms. A.D. conceptualized the project, performed statistical analysis, supervised the work and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.D. (adeb@mednet.ucla.edu).

METHODS

Animal care and use. All animal studies were approved by the Institutional Animal Care and Use Committee at the University of North Carolina, Chapel Hill and the University of California, Los Angeles.

Generation of transgenic and conditional knockout mice. Collagen1a2-CreERT:R26R^{tdTomato} mouse lines were obtained by crossing collagen1a2-CreERT mice with lineage reporter R26R^{tdTomato} mice. To obtain collagen1a2-CreERT:R26R^{tdTomato}:p53^{CKO} mice, collagen1a2-CreERT:R26R^{tdTomato} mice were crossed to p53^{fl/fl} mice and backcrossed to generate collagen1a2-CreERT:R26R^{tdTomato}:p53^{fl/fl} mice. Col1a2-CreERT:R26R^{tdTomato} mice were crossed with Col1-GFP mice to create progeny Col1a2-CreERT:R26R^{tdTomato}:Col1-GFP mice. FSP1-Cre mice were crossed with the R26R^{tdTomato} mice to create FSP1-Cre:R26R^{tdTomato} mice. Tamoxifen (1 mg) (Sigma) was injected intraperitoneally for 10 days to induce Cre-mediated recombination in Col1a2-CreERT mice. Five days following cessation of tamoxifen animals were subjected to ischaemic cardiac injury. All mice were on a C57BL/6 background.

Murine cardiac injury model. Mice (both male and female), 8–10 weeks old, were randomly allocated to sham or ischaemia-reperfusion cardiac injury. Investigators performing surgeries and cardiac function studies were blinded to mouse genotype or treatment. Mice were initially anaesthetized with 3% isoflurane, maintained at 2% isoflurane, and intubated using a Harvard Rodent Volume-Cycled ventilator. Myocardial injury was induced by 30 min ligation of the left anterior descending (LAD) coronary artery followed by reperfusion. Sham injury was performed in the same manner, with a ligature passed under the LAD, but the LAD was not ligated. Group size was estimated based on previously observed mortality rates following surgery. No animals were excluded from analysis unless they died during the surgical procedure. At 1, 2, 3, 7 or 14 days after injury, the mice were anaesthetized with pentobarbital (45 mg kg⁻¹) and the left ventricle was perfused with 5 ml PBS followed by 2 ml 4% PFA. For RITA-treated mice, beginning 24 h after injury RITA (Millipore) was administered intraperitoneally at 0.3 mg kg⁻¹ once daily for 3 days or until harvest. For acLDL staining, 1 µg fluorescent labelled acLDL (Invitrogen) per gram body weight was injected via a catheter inserted in the jugular vein and hearts harvested as described 4 h later. For DiO staining, DiO was prepared as described¹⁶ and 2 ml of 120 µg ml⁻¹ DiO perfused before PFA perfusion¹⁶.

After perfusion, the heart was further fixed for 1 h in 4% PFA at 4 °C and after 1 h the PFA was replaced with fresh 4% PFA and incubated at 4 °C for another 3 h. At the end of 4 h, the hearts were sucrose embedded overnight in a 30% sucrose solution (MP Biochemicals) and then frozen in OCT (Tissue-Tek). 7 µm sections were prepared in a longitudinal or transverse orientation.

Echocardiography. Echocardiography was performed in conscious mice. The hair over the anterior chest was removed by a depilation cream and warmed Aquasonic gel applied over the thorax. Mice were held firmly by hand for the duration of the procedure (approximately 5–10 min) and conditioned daily for the procedure starting 3 days before the procedure. The probe was positioned over the chest in a parasternal position. Parasternal long axis B-mode and M-mode images are recorded. Measurements and analysis were then performed as described¹⁰. The echocardiographer was blinded to the genotype and treatment of the animal being examined.

Immunohistochemistry and histology, confocal imaging and quantitation, super-resolution microscopy. Immunofluorescent staining on frozen sections (7 µm) was performed using primary antibodies to VECAD (catalogue number ab33168, Abcam), eNOS (ab66127, Abcam), claudin 5 (ab53765, Abcam), occludin (ab31721, Abcam), γH2AX (ab2893, Abcam), p53 (ab31333 and ab26, Abcam), Col1 (ab6308, Abcam), podoplanin (ab11936, Abcam), α-SMA (ab5694, Abcam), CD68 (ab125212, Abcam), CD146 (ab75769, Abcam), Cre (BIOT-106L & PRB-106P, Covance), troponin (SC-8121, Santa Cruz), NG2 (AB5320, EMD Millipore) and associated APC or fluorescein conjugated secondary antibodies (Abcam, Millipore, Invitrogen) as per manufacturer instructions. Labelled sections were imaged using a Leica SP2 AOBIS Upright Laser Scanning Confocal Microscope (Leica Microsystems). Five independent images for each area (that is, injury border zone, remote from injury, sham) were obtained and used for quantitative analysis. Co-localization analysis of confocal images was performed using Image J software (NIH). Mesenchymal-endothelial transition percentages were derived by counting the number of dually labelled cells and dividing by the number of tdTomato-positive cells. Quantification of vascular area derived from fibroblasts was performed using the JACoP Image J plugin³⁷. p53 expression levels in tdTomato-positive cells were determined using the JACoP Image J plugin and expressed as a Manders coefficient normalized to p53 levels in sham injury³⁸. In each case, five independent images from each area were analysed from sections prepared from each mouse. For super-resolution microscopy, a custom-made Stimulation Emission Depletion Super-resolution microscope was used. This microscope has a confocal channel for tdTomato and an ATTO647 fluorophore super-resolution channel that was used for VECAD, NG2 or CD146 visualization. Masson's trichrome staining was performed on heart sections as described¹⁰.

Fibroblast isolation and culture. Cardiac fibroblasts were isolated from the explanted hearts of euthanized uninjured mice. The hearts were explanted and washed three times with 1× HBSS (Gibco). The heart was minced into approximately 1-mm² sized pieces and digested using 10 ml of a 0.1% trypsin solution (Gibco) with 50 U ml⁻¹ collagenase II (Worthington)³⁹. Five sequential digestions were performed at 37 °C, the cells collected and passed through a 40 µm strainer and plated in IMDM, 1× penicillin/streptomycin, 10% FBS for 1 h at 37 °C. After 1 h the medium was changed to F12K 1× penicillin/streptomycin, 10% FBS (Gibco), 10 ng ml⁻¹ leukaemia inhibitory factor (LIF) (Millipore) and 10 ng ml⁻¹ basic fibroblast growth factor (bFGF) (Millipore). Cells were maintained under these conditions until they became confluent in 7–10 days.

Matrigel tube formation and LDL uptake assay. 6 × 10⁴ fibroblasts per cm² were cultured overnight at 37 °C, 5% CO₂ on Growth Factor Reduced Matrigel Basement Membrane Matrix (BD Biosciences) coated wells of Nunc Lab-Tek II CC2 chamber slides (Thermo). Serum starved or unstarved cells were cultured in IMDM, 1× penicillin/streptomycin or IMDM, 1× penicillin/streptomycin, 10% fetal bovine serum (Gibco), respectively. Pifithrin-α 100 µM (P4359, Sigma) or RITA 0.1 µM (506149, EMD Chemicals) were added to the cells cultured with the above culture medium. Acetylated LDL (Invitrogen) uptake was performed as described⁴⁰.

Flow cytometry. Flow cytometric analysis for cell surface markers were done using antibodies CD31-APC (17-0311, eBioscience), CD34-FITC (11-0341, eBioscience), CD133-FITC (11-1331, eBioscience), Flk1-APC (17-5821, eBioscience), VECAD-APC (17-1441, eBioscience), CD45-APC (103111, Biolegend), c-Kit APC (561074, BD Bioscience), DDR2 (SC-7555, Santa Cruz) and vimentin (Ab1620, Millipore). For vimentin and DDR2, a secondary APC conjugated anti-goat antibody was used (SC3860, Santa Cruz). Cultured cardiac fibroblasts were dissociated using accutase (Innovative Cell Technologies) and immunostained in FACS buffer (0.1% BSA PBS) at 1 × 10⁶ per ml for 20–30 min at 4 °C, followed by washing twice with FACS buffer and subsequently analysed in Beckman-Coulter (Dako) CyAn ADP. Data obtained was analysed and represented using Flowjo software.

Quantitative real-time PCR. RNA was isolated from cardiac fibroblasts cultured in the presence or absence of serum (IMDM, 1× penicillin/streptomycin, with or without 10% FBS), or in medium without serum and containing either 100 µM pifithrin-α or 0.1 µM RITA, grown for 48 h at 37 °C, 5% CO₂. RNA isolation and reverse transcription was performed using the SV Total RNA Isolation Kit and the Reverse Transcription System (Promega). qPCR was performed using the SensiMix SYBR and Fluorescein Kit (Quanta) on an iQ5 thermal cycler (BioRad).

Western blot. Protein was harvested from cardiac fibroblasts cultured in the presence or absence of serum (IMDM, 1× penicillin/streptomycin, with or without 10% FBS) for 24, 48 and 72 h at 37 °C, 5% CO₂.

Concentration normalized protein was prepared with SDS loading buffer and run on a 12× Mini-Protein TGX gel (BioRad) at 300V for 25 min. The protein was transferred from the gel to a nitrocellulose membrane using the Trans-Blot Turbo System (BioRad). After blocking with TBST +3% cold fish gelatin, the membrane was probed using primary antibodies to p53 (ab31333, Abcam) and α-tubulin (T6199, Sigma), washed with TBST, and labelled using fluorescently conjugated secondary antibodies (LI-COR Biosciences). After washing with TBST, the membrane was visualized on an Odyssey scanner (LI-COR Biosciences). Densitometry analysis was performed using the Gel Analyzer unit of ImageJ software (NIH).

Chromatin immunoprecipitation (ChIP). Cardiac fibroblasts (30 × 10⁶) were fixed in 1% formaldehyde, lysed in lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS, protease inhibitor cocktail Set I CALBIOCHEM) and sonicated using a EpiShear Multi-Sample Sonicator (Active Motif), leading to fragments between 300 and 1,000 bp. ChIP was performed using a commercially available ChIP-IT High Sensitivity Kit (Active Motif) according to manufacturer's instructions. DNA-bound protein was immunoprecipitated using an anti-p53 antibody (ab31333, Abcam) and anti IgG (SC-2025, Santa Cruz) as a negative control. The DNA recovered was analysed by quantitative real time-PCR using different primers sets that amplified the promoter region of *Hoxa9*, *Hoxd3*, *CLDN5*, *Nos3* and negative control genes: *Hoxa9* (1) 5'-TAAACTGCTCAGGCCATGCT-3' sense, 5'-CAGCCTGGACCCACTGAAAG-3' antisense; *Hoxd3* (1) 5'-TGCCTCAGTGTATTTCTCCCC-3' sense, 5'-ATGGATAACAGTGCCCGGTG-3' antisense; claudin 5 5'-CTCCGGAAGCCAACCTGGAG-3' sense, 5'-GGACCCAGTGTGCTAAACC-3' antisense; eNOS 5'-GGACAGCCAGAAATGGGA-3' sense, 5'-ACTGCTGTCGGTCTCTTTGT-3' antisense.

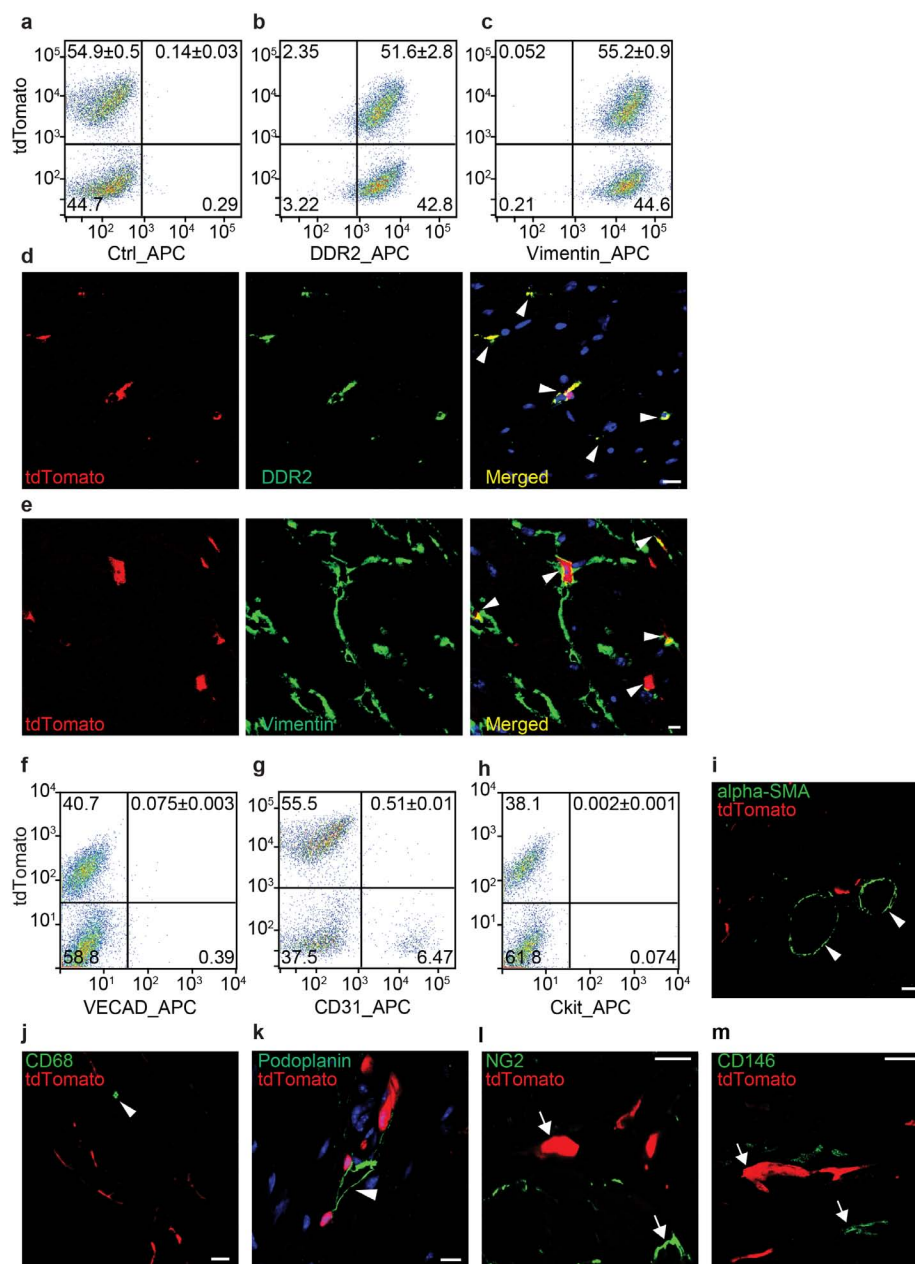
Primers for p53 binding were designed using the SABiosciences' proprietary database (DECOD, DECipherment Of DNA Elements), available from <http://www.sabiosciences.com/chippcr.php>.

PCR was performed with equal amounts of specific antibody immunoprecipitated sample, control (IgG) and input. Values were normalized to input measurements and enrichment was calculated using the comparative Δ-ΔCt (ΔΔCt)

method. Data shown correspond to one representative assay (that is, three PCRs) from a total of three independent assays each run with different sets of treated cells.

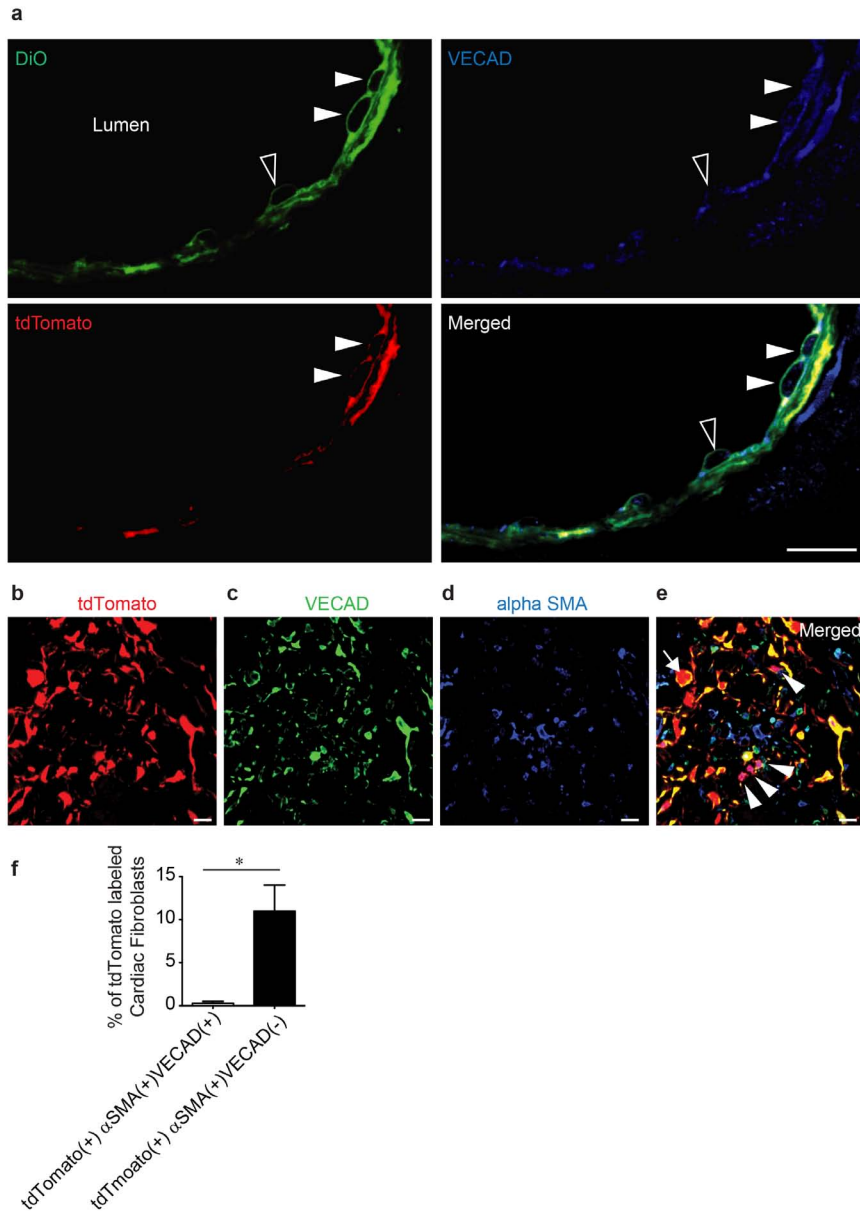
Statistical analysis. Statistical analysis was performed using GraphPad software (Prizm) using Student's *t*-test (two-tailed), one-way or two-way ANOVA with Bonferroni post-test analysis as appropriate. Welch's correction was used if variances between groups were significantly different. A *P* value of <0.05 was considered statistically significant. Graphs present the mean value \pm standard error of the mean (s.e.m.).

37. Bolte, S. & Cordelières, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224**, 213–232 (2006).
38. McDonald, J. H. & Dunn, K. W. Statistical tests for measures of colocalization in biological microscopy. *J. Microsc.* **252**, 295–302 (2013).
39. Agocha, A. E. & Eghbali-Webb, M. A simple method for preparation of cultured cardiac fibroblasts from adult human ventricular tissue. *Mol. Cell. Biochem.* **172**, 195–198 (1997).
40. Nagelkerke, J. F., Barto, K. P. & van Berkel, T. J. In vivo and in vitro uptake and degradation of acetylated low density lipoprotein by rat liver endothelial, Kupffer, and parenchymal cells. *J. Biol. Chem.* **258**, 12221–12227 (1983).



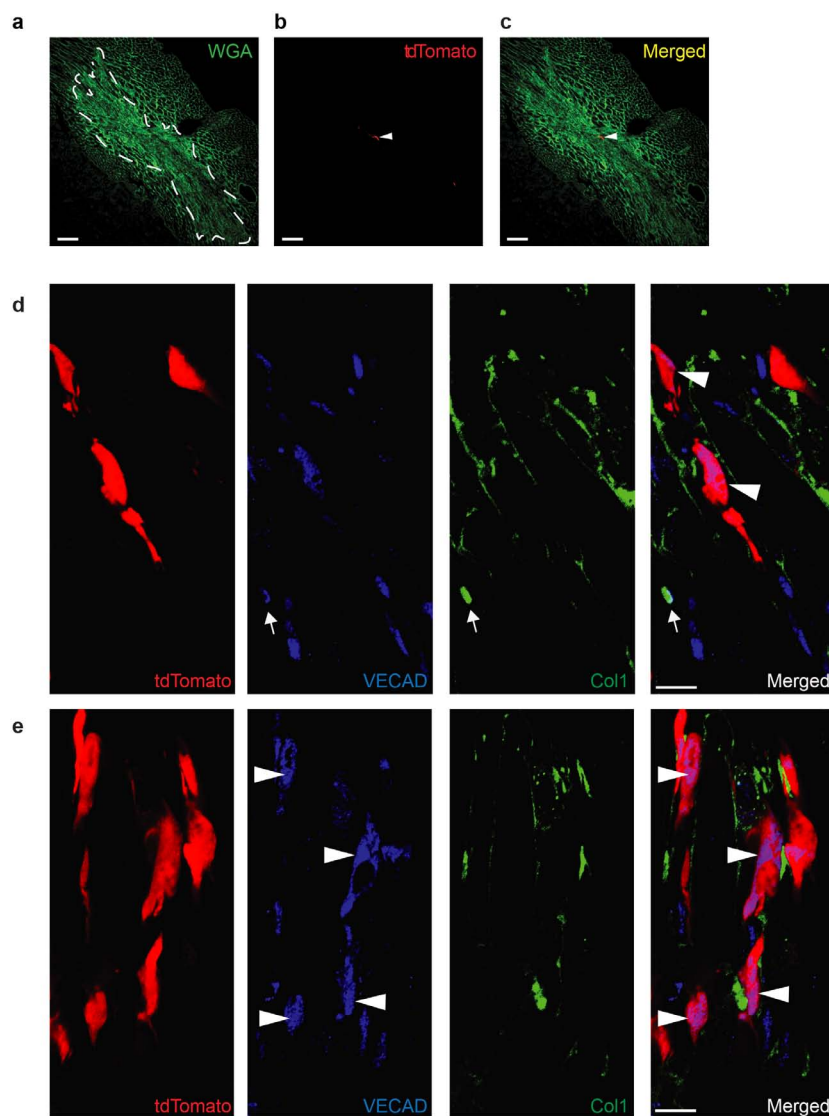
Extended Data Figure 1 | Phenotypic characterization of tdTomato-labelled cells in hearts of *Col1a2-CreERT:R26R^{tdTomato}* mice. **a–c**, Hearts were digested to obtain a non-myocyte population and subjected to flow cytometry to determine expression of DDR2 (**b**) and vimentin (**c**). Panel **a** serves as a control with no primary antibody added (number in each quadrant represents fraction of total population of cells). **d, e**, Immunofluorescent staining on sham-injured hearts of *Col1a2-CreERT:R26R^{tdTomato}* mice to determine expression of DDR2 (**d**) (green) and vimentin (**e**) (green) in tdTomato-labelled cells (red) with merged image (right; arrowheads point to tdTomato-labelled cells staining for DDR2 or vimentin). Scale bar, 10 μ m. **f–h**, Non-myocyte cells from the heart were subjected to flow cytometry to determine expression of VECAD (**f**), CD31 (**g**) or c-Kit (**h**) (number

in each quadrant represents fraction of total population of cells). **i–k**, Immunofluorescent staining for α -smooth muscle actin (**i**), CD68 (**j**) or podoplanin (**k**) was performed on frozen heart sections prepared from 8-week-old sham-injured *Col1a2-CreERT:R26R^{tdTomato}* mice after tamoxifen injections and co-localization analysis performed to determine the number of labelled cardiac fibroblasts expressing α -smooth muscle actin (99.4% negative for α -SMA, 1,500 cells examined), CD68 (100% negative for CD68, 1,000 cells examined) or podoplanin (arrowheads). Scale bar: **i–k**, 10 μ m. **l, m**, STED super-resolution microscopy demonstrating tdTomato-labelled cells (arrows) not expressing NG2 (**l**) (green, arrows, STED channel) or CD146 (**m**) (green, arrows, STED channel). Scale bar, 10 μ m.



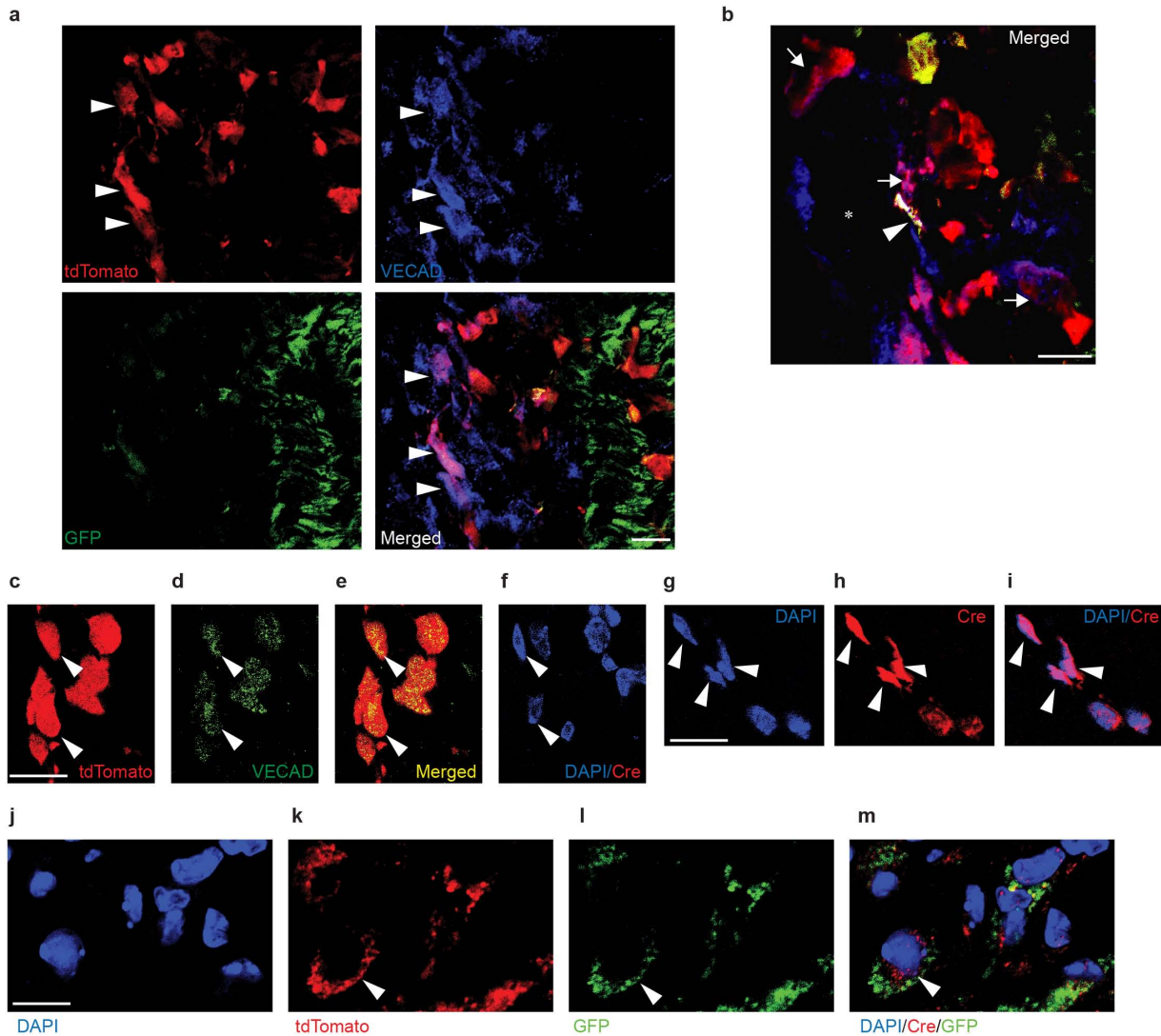
Extended Data Figure 2 | Expression of VECAD in fibroblast-derived endothelial cells that take up DiO and expression of α -smooth muscle actin and VECAD in tdTomato-labelled cells. **a**, Col1a2-CreERT:R26R^{tdTomato} mice were subjected to ischaemic injury, injected with DiO 7 days after cardiac injury and then harvested. High magnification of a wall of a blood vessel in the injured region demonstrates luminal cells staining for DiO (green, arrowheads), VECAD (blue, arrowheads), tdTomato (red, arrowheads) and merged image demonstrating co-localization of all three fluorophores (filled arrowheads); open arrowhead points to VECAD⁺DiO⁺ cell that does not bear the tdTomato label. Scale bar, 10 μ m. **b–f**, Expression of α -smooth muscle actin and VECAD in tdTomato-labelled cells. **b–e**, Tamoxifen-injected

Col1a2-CreERT:R26R^{tdTomato} mice underwent ischaemia reperfusion injury and hearts were harvested 3 days after injury and immunofluorescent staining performed for α -smooth muscle actin and VECAD. Section shows tdTomato-labelled cardiac fibroblasts (**b**), VECAD (**c**), α -smooth muscle actin expressing cells (**d**) and merged image (**e**) showing co-localization of fluorophores. Arrowheads show labelled cardiac fibroblasts expressing α -smooth muscle actin and arrow shows a labelled cardiac fibroblast expressing VECAD but not smooth muscle actin. Scale bar: **b–e**, 10 μ m. **f**, Fraction of labelled cardiac fibroblasts that are α -smooth muscle actin⁺ and VECAD⁺ or VECAD⁻ (mean \pm s.e.m., * $P < 0.05$, $n = 3$).



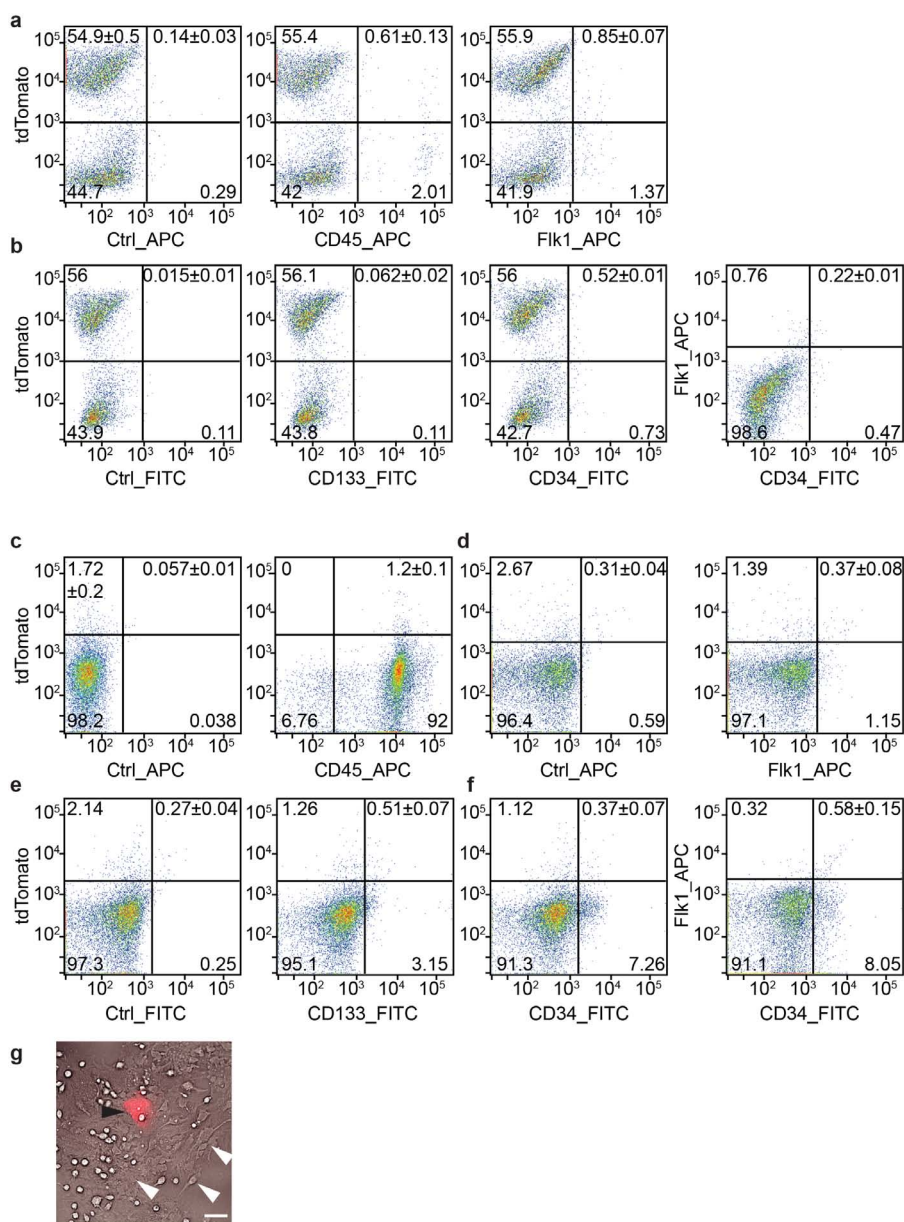
Extended Data Figure 3 | tdTomato expression in hearts of vehicle-injected Col1a2-CreERT:R26R^{tdTomato} mice and immunostaining for Col1 and VECAD in tamoxifen-injected Col1a2-CreERT:R26R^{tdTomato} mice after cardiac injury. a–c, Oil-injected Col1a2-CreERT:R26R^{tdTomato} mice underwent ischaemia reperfusion injury and 3 days after injury, hearts were harvested and sectioned. Area of injury (white lines) was stained with Alexa488-labelled wheat germ agglutinin (a) (WGA, stains cell membranes), rare tdTomato-expressing cells in same field (b) (arrowhead) and merged image (c) showing the presence of rare labelled cells in the injury region (arrowhead) (28 labelled cells out of 38,000 cells counted (0.07%), $n = 3$ mice). Scale bar: 100 μm . d, e, Col1a2-Cre:R26R^{tdTomato} mice (injected with tamoxifen

as described in text) were subjected to ischaemic cardiac injury and hearts harvested 3 days post injury and stained for Col1 and VECAD. d, Region of injury demonstrating a VECAD-expressing cell (blue, arrow) staining positive for Col1 (green) but negative for the tdTomato label (merged arrow). Arrowheads in merged panel show tdTomato-labelled cells expressing VECAD but not Col1. Scale bar, 10 μm . e, tdTomato-labelled cells expressing VECAD (arrowheads) that do not stain for Col1 (green) with merged image showing tdTomato⁺VECAD⁺ cells not staining with the Col1 antibody. Out of 225 cells counted ($n = 3$ mice), we did not observe a single tdTomato⁺VECAD⁺ cell to stain for Col1. Conversely, not a single VECAD⁺Col1⁺ cell exhibited tdTomato fluorescence. Scale bar, 10 μm .



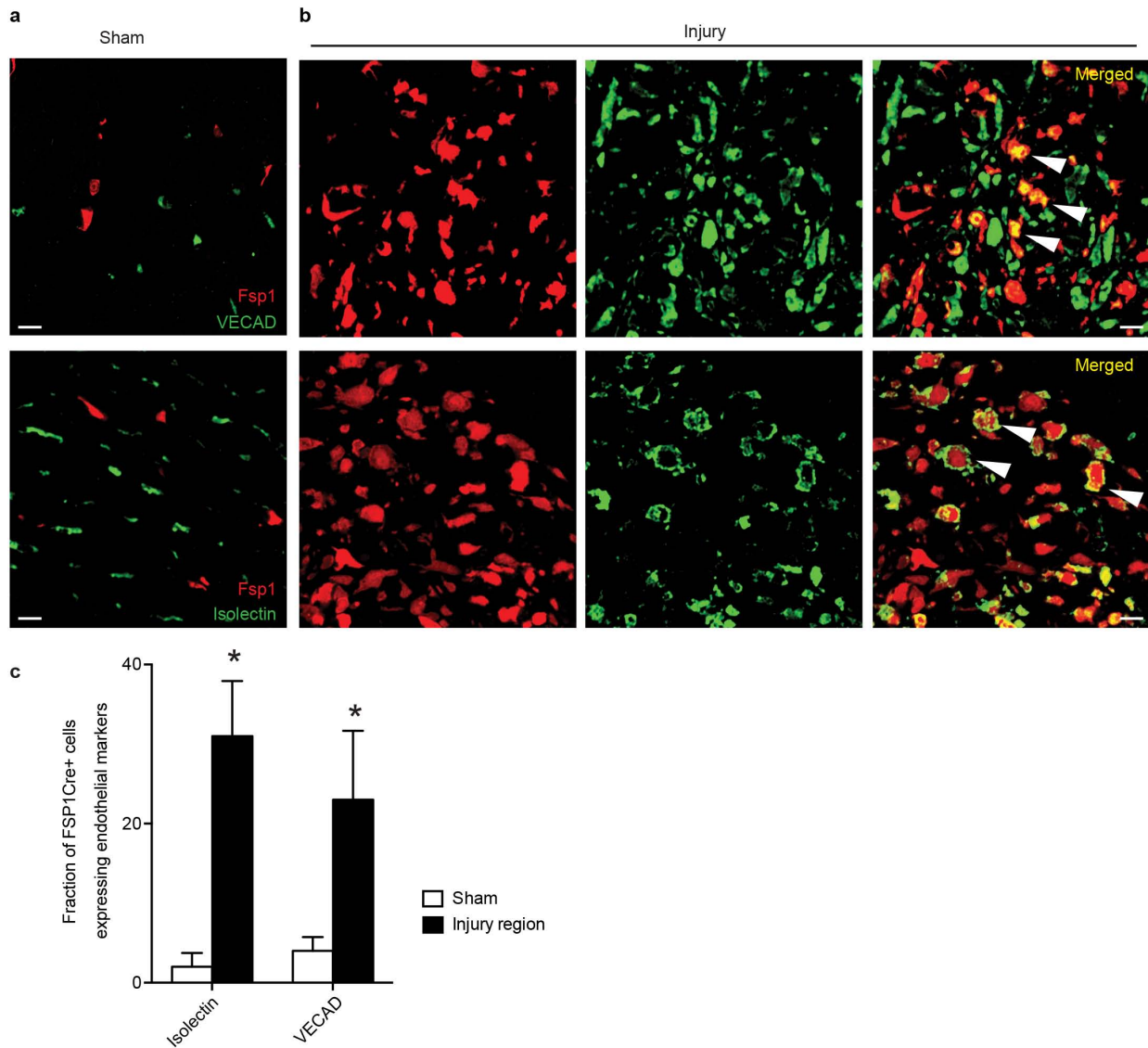
Extended Data Figure 4 | VECAD and Cre immunostaining in heart sections of *Col1a2-Cre:R26R^{tdTomato}:Col1-GFP* mice 3 days after injury. **a, b,** *Col1a2-Cre:R26R^{tdTomato}:Col1-GFP* mice were subjected to ischaemic cardiac injury and hearts harvested at 3 days post injury and stained for VECAD. **a,** tdTomato-labelled cells (red, arrowheads) expressing VECAD (blue, arrowheads) but not GFP (green), with merged image demonstrating co-localization of tdTomato and VECAD but not GFP (arrowheads). Scale bar, 10 μ m. **b,** High magnification of a blood vessel (asterisk) outlined by VECAD staining (blue) demonstrating rare cell that co-localizes all three fluorophores (tdTomato⁺VECAD⁺GFP⁺, white, arrowhead). Arrows point to tdTomato-positive cells expressing VECAD but not GFP. Scale bar, 10 μ m. **c-f,** Immunostaining for Cre protein on hearts of *Col1a2-Cre:R26R^{tdTomato}:Col1-GFP* mice harvested 3 days following ischaemic injury to detect Cre expression in the nucleus of tdTomato-labelled cells expressing

VECAD. tdTomato-labelled cells (**c**) in area of injury (arrowheads) expressing VECAD (**d**) (arrowheads), merged image showing co-localization of fluorophores (**e**) (arrowheads) and nuclei stained for DAPI and Cre (**f**) demonstrating absence of any detectable nuclear Cre protein. Scale bar, 10 μ m. **g-i,** Positive control demonstrating section of heart of Wt-1Cre transgenic mouse heart 3 days after injury with region of injury stained for DAPI (**g**), Cre (**h**) (red) and merged image (**i**) demonstrating numerous cells in the injury region expressing nuclear Cre (arrowheads). Scale bar, 10 μ m. **j-m,** Cre immunostaining in tdTomato-labelled cells expressing GFP. Area in region of injury with nuclei stained for DAPI (**j**), tdTomato expression (**k**) (arrowhead), GFP expression (arrowhead) (**l**) and merged image (DAPI, Cre and GFP) (**m**) showing Cre staining (red channel) localized to the cytoplasm of GFP-expressing cell (arrowhead). Scale bar, 10 μ m.



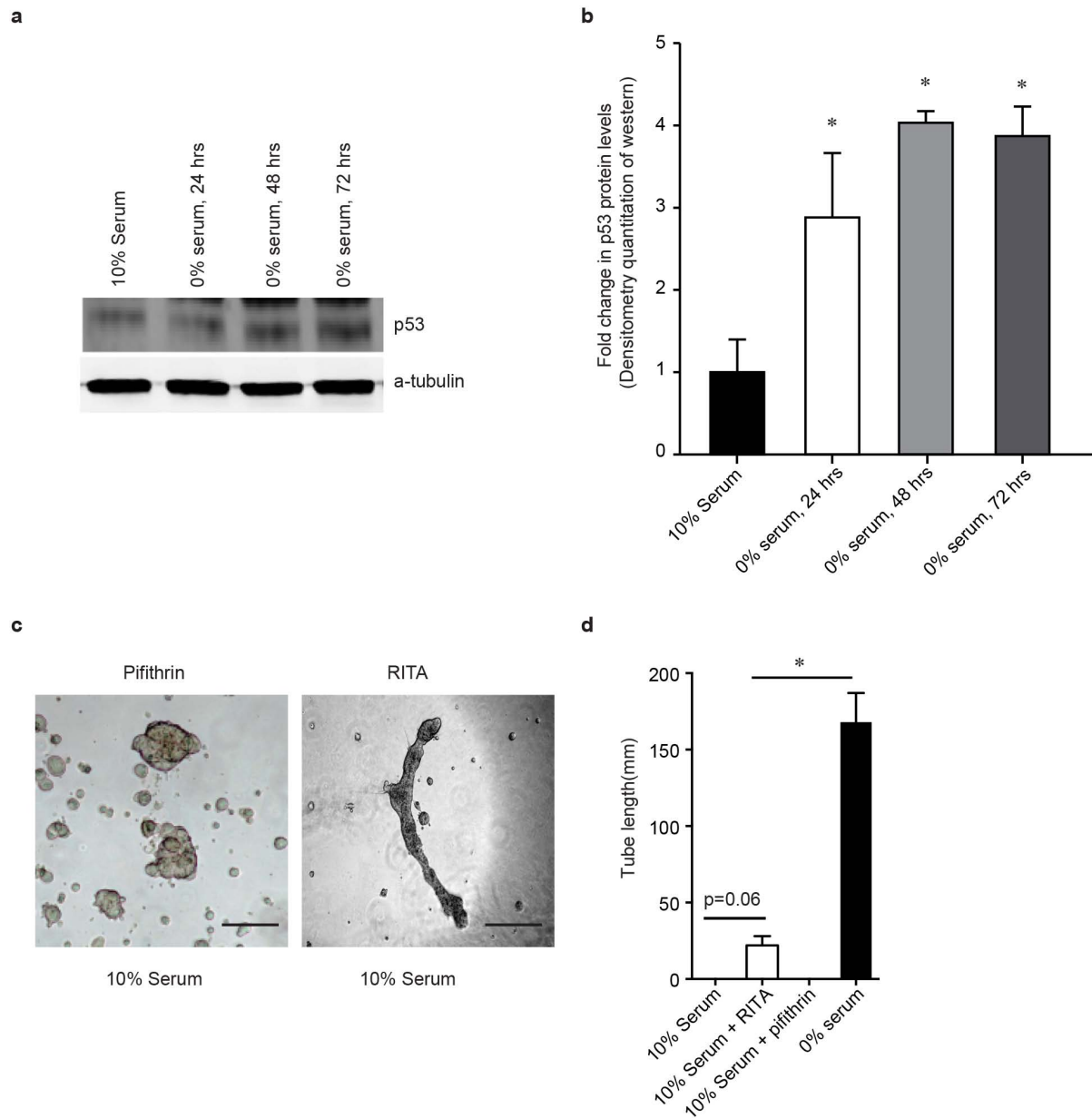
Extended Data Figure 5 | Flow cytometry for endothelial progenitor markers on non-myocyte cells harvested from uninjured hearts of *Col1a2-CreERT:R26R^{tdTomato}* mice and bone marrow cells isolated from the same animal. a, b, Hearts of mice were digested, myocytes discarded and the entire non-myocyte population without any further selection was subjected to flow cytometry. Expression of tdTomato and CD45 and Flk-1 (APC fluorophore) (a) and CD133, CD34 and combined expression of CD34 and Flk-1 in

non-myocyte population (b). c–f, Bone marrow cells were isolated from *Col1a2-CreERT:R26R^{tdTomato}* mice and without further culture subjected to flow cytometry. Expression of tdTomato and CD45 (c), Flk-1 (d), CD133 (e), and CD34 and combined expression of Flk1 and CD34 (f) in bone marrow cells. g, Expression of tdTomato in bone-marrow-derived mesenchymal stem-cell colonies (black arrowhead points to rare tdTomato-positive cell; white arrowheads point to mesenchymal stem cells). Scale bar, 50 μm .



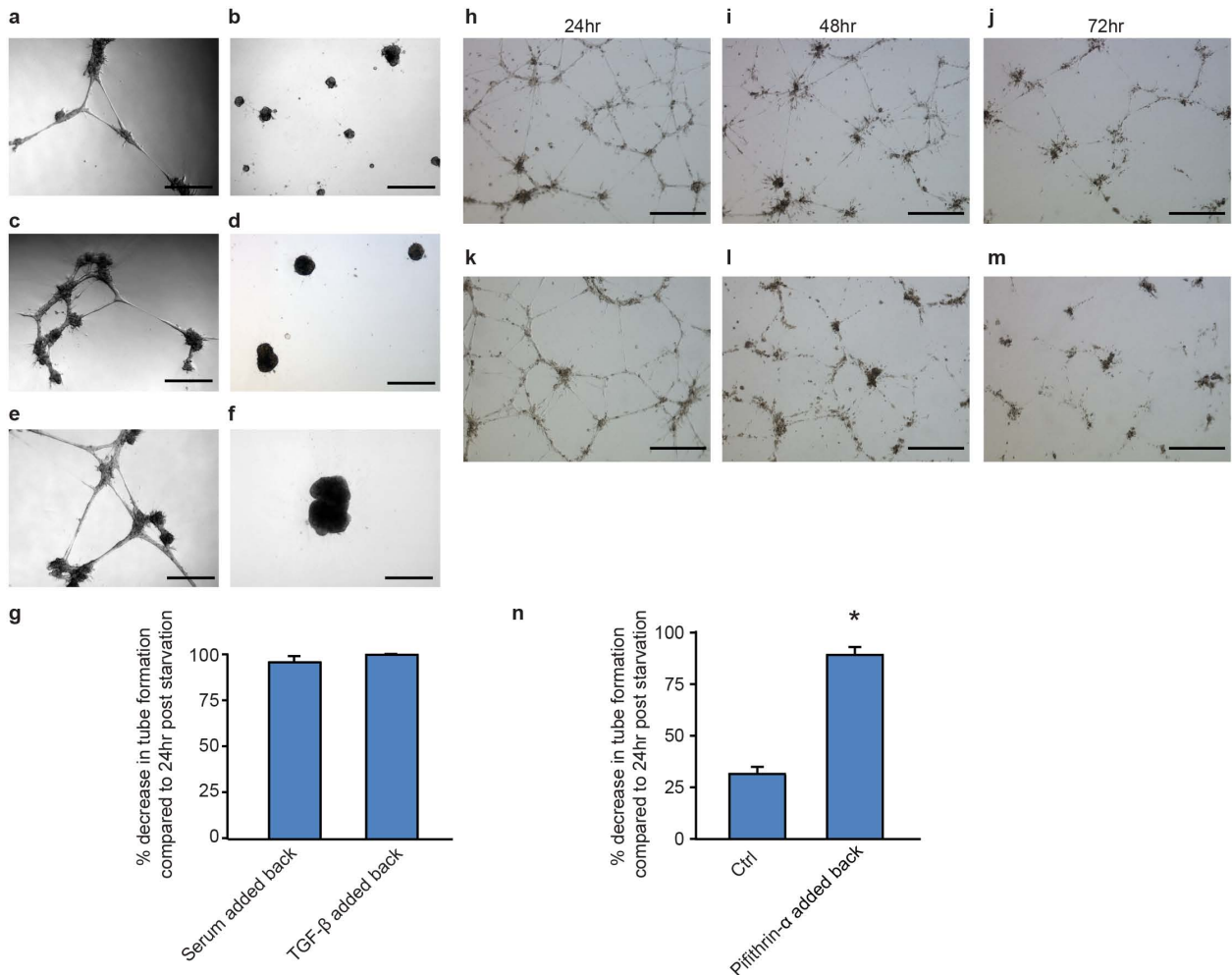
Extended Data Figure 6 | Mesenchymal–endothelial transition in FSP1-Cre:R26R^{tdTomato} mice. **a, b,** FSP1-Cre:R26R^{tdTomato} mice were subjected to sham (**a**) or ischaemia-reperfusion (**b**) cardiac injury. Hearts were harvested 3 days after injury and stained for endothelial marker VECAD or isolectin.

Injury region demonstrated tdTomato-labelled cells expressing VECAD or isolectin (arrowheads, $n = 4$). Scale bar, 10 μm . **c,** Quantification of labelled fibroblasts that express VECAD or isolectin in sham-injured animals and in the injury border zone (mean \pm s.e.m., $*P < 0.01$, $n = 4$).



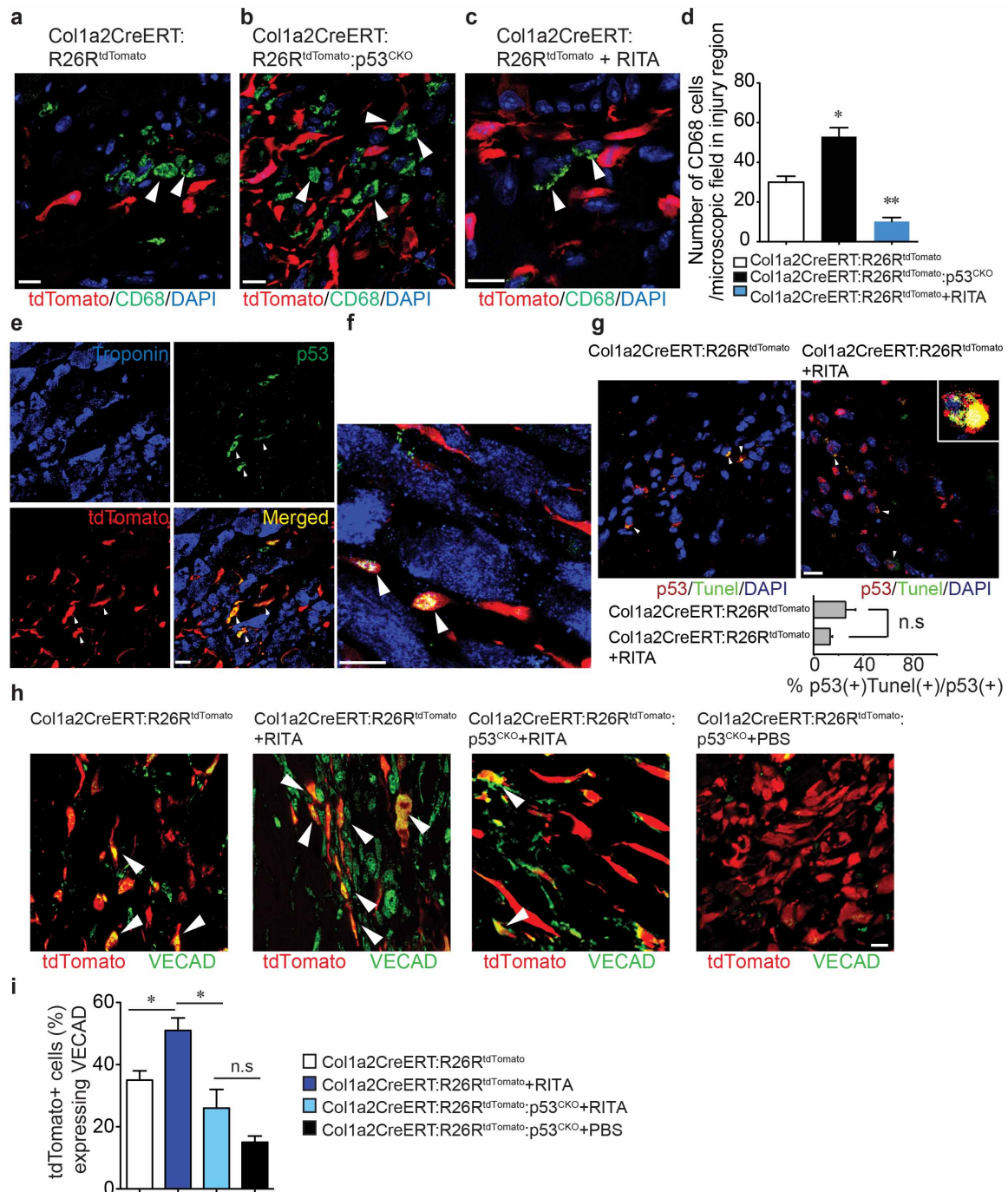
Extended Data Figure 7 | Effect of serum starvation on p53 levels and effects of pifithrin- α and RITA on tube formation of serum-fed cardiac fibroblasts. **a**, Western blot for p53 in cardiac fibroblasts subjected to serum starvation for 24, 48 and 72 h (representative sample from $n = 3$). **b**, Densitometric

quantification of western blot (mean \pm s.e.m., $*P < 0.05$ compared to cells in 10% serum). **c**, Effect on tube formation after adding pifithrin- α or RITA to cardiac fibroblasts grown in 10% serum. Scale bar, 250 μ m. **d**, Quantification of tube formation (mean \pm s.e.m., $*P < 0.05$, $n = 3$).



Extended Data Figure 8 | Effect of adding TGF- β to serum-starved cardiac fibroblasts, or adding TGF- β , serum or pifithrin- α on tubes that have already formed. a, b, Tube formation of cardiac fibroblasts subjected to serum starvation in the absence (a) or presence (b) of TGF- β . (TGF- β was added at the onset of serum starvation.) Scale bar, 250 μ m. **c–f,** Effect of adding TGF- β or serum to tubes that had already formed. **c, e,** Twenty-four hours after serum starvation (after tubes had already formed), PBS was added to tubes shown in **c** and **e** and photographs taken after another 24 h. **d, f,** After tubes had already formed (24 h of serum starvation), TGF- β (**d**) or serum (**f**) was added and photographs taken after another 24 h (note clumping of cells and regression of tubes in **d** and **f**). Scale bar: **c–f**, 250 μ m. **g,** Effect of adding TGF- β or serum to

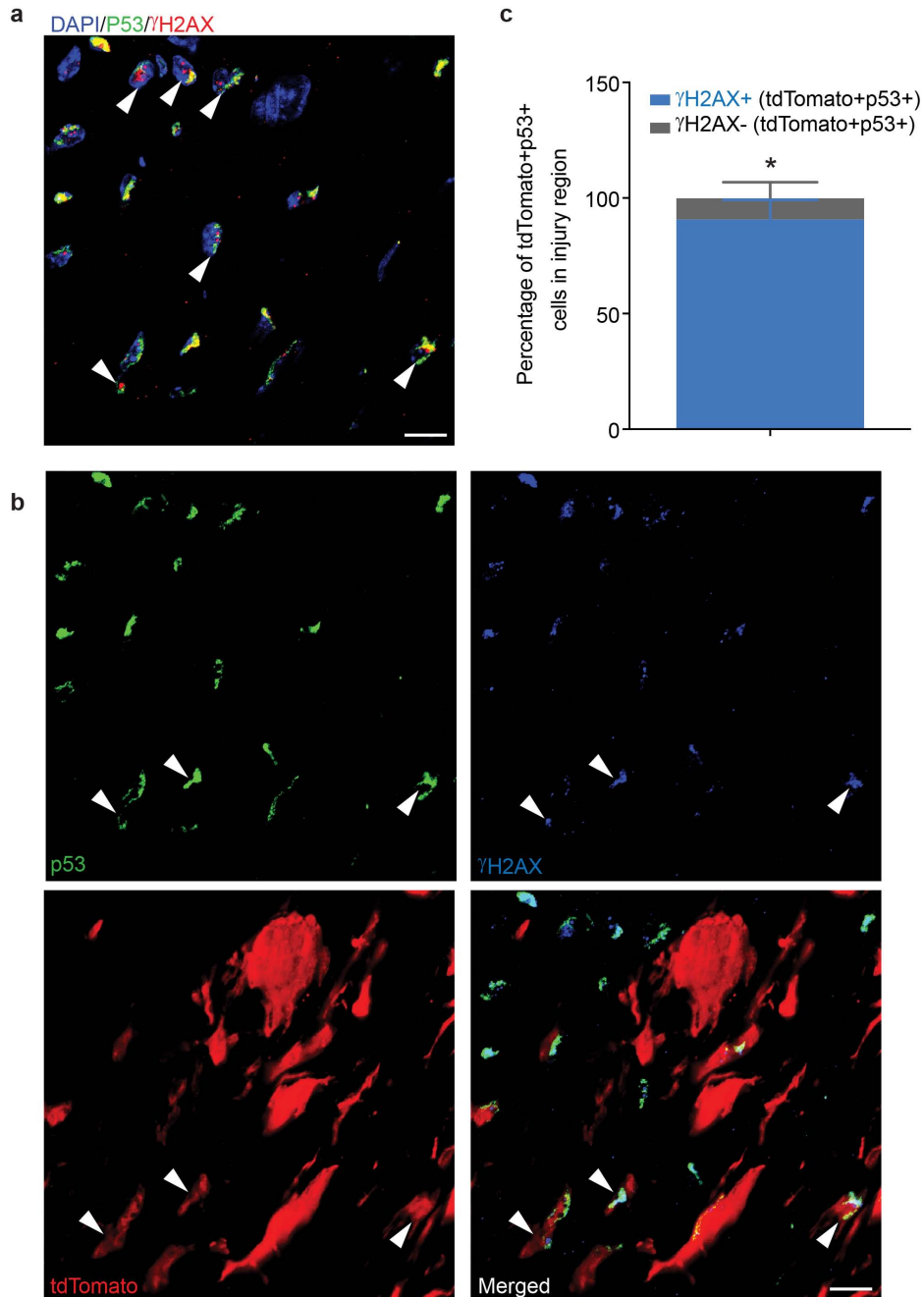
tubes that had already formed, expressed as a percentage decrease in tube length. **h–n,** Effect of adding pifithrin- α to serum-starved cardiac fibroblasts that had already formed tubes. **h, k,** Tube formation in cardiac fibroblasts serum starved for 24 h. Scale bar, 250 μ m. **i, j,** PBS was then added to cardiac fibroblasts shown in **h** and photographs were taken after another 24 h (**i**) or 48 h (**j**) of serum starvation. Scale bar, 250 μ m. **l, m,** Pifithrin- α was added to cardiac fibroblasts shown in **k** and photographs were taken after another 24 h (**l**) or 48 h (**m**) of serum starvation in the presence of pifithrin- α . Scale bar, 250 μ m. **n,** Tube length in **j** and **m** was expressed as a per cent change from their respective control (**h** and **k**) (mean \pm s.e.m., * P < 0.05 compared to control, n = 3).



Extended Data Figure 9 | RITA decreases inflammatory infiltrate after cardiac injury, does not increase apoptosis in myocytes and does not enhance mesenchymal-endothelial transition in Col1a2-CreERT:R26R^{tdTomato};p53^{CKO} mice. a–d, Sections of hearts harvested at 3 days after cardiac injury were stained for the monocyte/macrophage marker CD68 (green, arrowheads) in Col1a2-CreERT:R26R^{tdTomato} (a), Col1a2-CreERT:R26R^{tdTomato};p53^{CKO} (b) and RITA-injected Col1a2-CreERT:R26R^{tdTomato} (c) mice. Scale bar, 10 μ m. d, Quantification of the number of CD68 cells per high-power field in the injury region (mean \pm s.e.m., * P < 0.05 versus Col1a2-CreERT:R26R^{tdTomato}, ** P < 0.05 versus Col1a2-CreERT:R26R^{tdTomato}, n = 3).

e–g, p53 expression in myocytes after cardiac injury and effect of RITA on apoptosis in injury region. e, f, Col1a2-CreERT:R26R^{tdTomato} mice were subjected to ischaemic cardiac injury, hearts harvested at 3 days and sections stained for p53 and the cardiomyocyte marker troponin. e, p53 (green, arrowheads) staining is observed in tdTomato-expressing cells (red, arrowheads) but not in cardiomyocytes (blue); merged image shows arrowheads pointing to tdTomato-labelled cells expressing p53. Scale bar, 10 μ m. f, Higher magnification in injury region demonstrating tdTomato cells (arrowheads) expressing p53 (merged, yellow, arrowheads); p53 staining is not visible in cardiomyocytes (blue). Scale bar, 10 μ m. g, TUNEL staining and quantification to determine p53⁺ apoptotic cells after RITA injection (arrowheads point to p53⁺TUNEL⁺ cells in Col1a2-CreERT:R26R^{tdTomato} mice (left panel) and RITA-injected Col1a2-CreERT:R26R^{tdTomato} mice (right panel); inset shows p53⁺TUNEL⁺ cell in higher magnification (data shown as mean \pm s.e.m., ns, not significant, n = 3). Hearts in both cases were examined 3 days after injury. Scale bar, 10 μ m. h, i, Effect of RITA on mesenchymal-endothelial transition in Col1a2-CreERT:R26R^{tdTomato};p53^{CKO} mice after cardiac injury. h, tdTomato-labelled cardiac fibroblasts expressing VECAD in Col1a2-CreERT:R26R^{tdTomato} mice treated with RITA or vehicle and Col1a2-CreERT:R26R^{tdTomato};p53^{CKO} mice treated with RITA or vehicle. Scale bar, 10 μ m. i, Quantification of the percentage of labelled fibroblasts undergoing mesenchymal-endothelial transition for each treatment group (mean \pm s.e.m., * P < 0.05, n.s., not significant, n = 4 mice per group).

10 μ m. f, Higher magnification in injury region demonstrating tdTomato cells (arrowheads) expressing p53 (merged, yellow, arrowheads); p53 staining is not visible in cardiomyocytes (blue). Scale bar, 10 μ m. g, TUNEL staining and quantification to determine p53⁺ apoptotic cells after RITA injection (arrowheads point to p53⁺TUNEL⁺ cells in Col1a2-CreERT:R26R^{tdTomato} mice (left panel) and RITA-injected Col1a2-CreERT:R26R^{tdTomato} mice (right panel); inset shows p53⁺TUNEL⁺ cell in higher magnification (data shown as mean \pm s.e.m., ns, not significant, n = 3). Hearts in both cases were examined 3 days after injury. Scale bar, 10 μ m. h, i, Effect of RITA on mesenchymal-endothelial transition in Col1a2-CreERT:R26R^{tdTomato};p53^{CKO} mice after cardiac injury. h, tdTomato-labelled cardiac fibroblasts expressing VECAD in Col1a2-CreERT:R26R^{tdTomato} mice treated with RITA or vehicle and Col1a2-CreERT:R26R^{tdTomato};p53^{CKO} mice treated with RITA or vehicle. Scale bar, 10 μ m. i, Quantification of the percentage of labelled fibroblasts undergoing mesenchymal-endothelial transition for each treatment group (mean \pm s.e.m., * P < 0.05, n.s., not significant, n = 4 mice per group).



Extended Data Figure 10 | γ H2AX expression in cells expressing p53 after ischaemic cardiac injury. *Coll1a2-CreERT:R26R^{tdTomato}* mice were subjected to ischaemic cardiac injury and immunostaining performed for γ H2AX and p53. **a**, Immunostaining for p53 (green), γ H2AX (red) and DAPI (blue) in region of injury (arrowheads point to nuclei co-expressing γ H2AX and p53).

Scale bar, 10 μ m. **b**, Immunostaining for p53 (green), γ H2AX (blue) and tdTomato (red) to determine co-expression of p53 and γ H2AX in tdTomato-labelled cells. Arrowheads point to tdTomato-positive cells co-expressing γ H2AX and p53. Scale bar, 10 μ m. **c**, Quantification of the fraction of tdTomato⁺p53⁺ cells expressing γ H2AX (mean \pm s.e.m., * $P < 0.01$, $n = 3$).

Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome

Robert K. McGinty¹, Ryan C. Henrici^{1,2} & Song Tan¹

The Polycomb group of epigenetic enzymes represses expression of developmentally regulated genes in many eukaryotes. This group includes the Polycomb repressive complex 1 (PRC1), which ubiquitylates nucleosomal histone H2A Lys 119 using its E3 ubiquitin ligase subunits, Ring1B and Bmi1, together with an E2 ubiquitin-conjugating enzyme, UbcH5c. However, the molecular mechanism of nucleosome substrate recognition by PRC1 or other chromatin enzymes is unclear. Here we present the crystal structure of the human Ring1B–Bmi1–UbcH5c E3–E2 complex (the PRC1 ubiquitylation module) bound to its nucleosome core particle substrate. The structure shows how a chromatin enzyme achieves substrate specificity by interacting with several nucleosome surfaces spatially distinct from the site of catalysis. Our structure further reveals an unexpected role for the ubiquitin E2 enzyme in substrate recognition, and provides insight into how the related histone H2A E3 ligase, BRCA1, interacts with and ubiquitylates the nucleosome.

The Polycomb group (PcG) proteins are transcriptional repressors that correlate with diverse human cancers and constitute key epigenetic regulators of stem-cell self-renewal and lineage development^{1–5}. The PcG proteins assemble into multi-component complexes that remodel chromatin structure both directly and through the establishment and removal of histone post-translational modifications^{6,7}. One of these complexes, PRC1, monoubiquitylates histone H2A Lys 119 (refs 8, 9). Once thought of as a unique complex, it is now clear that human PRC1 exists as a modular subset of complexes with distinct functions and gene targets^{10,11}. All PRC1 complexes share a common core — the RING-type E3 ubiquitin ligase composed of Ring1B (also known as RNF2) or Ring1A, and Bmi1 (also known as PCGF4) or one of five other Polycomb group RING finger (PCGF) paralogues with potential for H2A ubiquitylation. This E3 core combines with accessory subunits to form canonical and variant PRC1s. Variant complexes, especially those containing histone demethylase KDM2B and/or RYBP, seem to be primarily responsible for H2A ubiquitylation^{10,12–14}.

Protein ubiquitylation is accomplished through a cascade of E1, E2 and E3 enzymes¹⁵. First, an E1 activates and transfers ubiquitin to an E2-conjugating enzyme, forming a thioester between ubiquitin and a cysteine in the E2 active site. Then, in an E3-dependent manner, the carboxy terminus of ubiquitin is attached to the ϵ -NH₂ of a target protein lysine side chain through an isopeptide bond. RING family E3 ligases, such as those found in PRC1, bridge an E2 and a target protein, presenting the target lysine to the E2 active site and constraining the geometry of the ubiquityl-E2 thioester, priming it for catalysis^{16–18}. The E3 is considered the primary specificity determinant of ubiquitylation, reflecting the scarcity of E2s (~40 proteins in human) relative to E3s (more than 600 proteins)¹⁹.

Although the cognate E2 for PRC1 *in vivo* is not established, the Ring1B–Bmi1 E3 ligase can pair with one of several E2s, including UbcH5c (also known as UBE2D3), to ubiquitylate nucleosomes *in vitro*²⁰. The nucleosome core particle (NCP) is the fundamental unit of the eukaryotic genome formed by 145–147 base pairs (bp) of DNA wrapped around an octameric scaffold of histone proteins (two copies each of H2A, H2B, H3 and H4)²¹. H2A–H2B dimers and H2A alone are not substrates for Ring1B–Bmi1, suggesting that surfaces formed only in the

context of the nucleosome are required for the E3 ligase activity of PRC1 (refs 22, 23). A minimal RING heterodimer of Ring1B and Bmi1 exhibits comparable activity to the full-length complex, implicating the RING domains in both E2 and substrate binding²⁰. However, we lack a fundamental mechanistic understanding of how the Ring1B–Bmi1 E3 ligase interacts with the nucleosome to direct H2A Lys 119 ubiquitylation, or more generally, for how chromatin enzymes recognize their nucleosome substrate.

We have crystallized the PRC1 ubiquitylation module, an E2–E3 enzyme complex composed of UbcH5c and the minimal Ring1B–Bmi1 RING heterodimer, bound to its NCP substrate, and solved the structure at 3.3 Å resolution (Supplementary Table 1). To our knowledge, this is the first crystal structure of a histone-modifying-enzyme–nucleosome complex or of a ubiquitin E2–E3–substrate complex. The structure shows that in contrast to most chromatin modification enzymes studied so far, the vast majority of interactions between the PRC1 ubiquitylation module and the nucleosome occur away from the catalytic site. Ring1B and Bmi1 bind to the histone surface of the NCP to orient UbcH5c for H2A ubiquitylation. Surprisingly, UbcH5c also makes direct contacts with the NCP that are crucial for nucleosome binding and ubiquitylation. This presents a novel role of ubiquitin E2 enzymes in substrate recognition.

Overview of complex

The structure shows one PRC1 ubiquitylation module bound to each disk face of the nucleosome (Fig. 1a, b). The two PRC1 modules more than double the height of the NCP and with the exception of the C-terminal region of UbcH5c, are entirely contained within a cylinder projecting up from the NCP disk surface. Each PRC1 ubiquitylation module forms a crescent with two halves of the crescent contributed by the E2 and E3, respectively. The Ring1B–Bmi1 E3 heterodimer binds to the central histone surface of the NCP. The E2 UbcH5c completes the crescent, following the inner curvature of the DNA from its interface with Ring1B past the DNA end, ultimately terminating above the nucleosomal dyad. In this manner, the Ring1B–Bmi1 E3 ligase positions the active site of UbcH5c over the C-terminal tail of H2A near the target residue Lys 119 (Fig. 1c). All three components of the PRC1

¹Center for Eukaryotic Gene Regulation, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, Pennsylvania 16802, USA. ²Schreyer Honors College, The Pennsylvania State University, University Park, Pennsylvania 16802, USA.

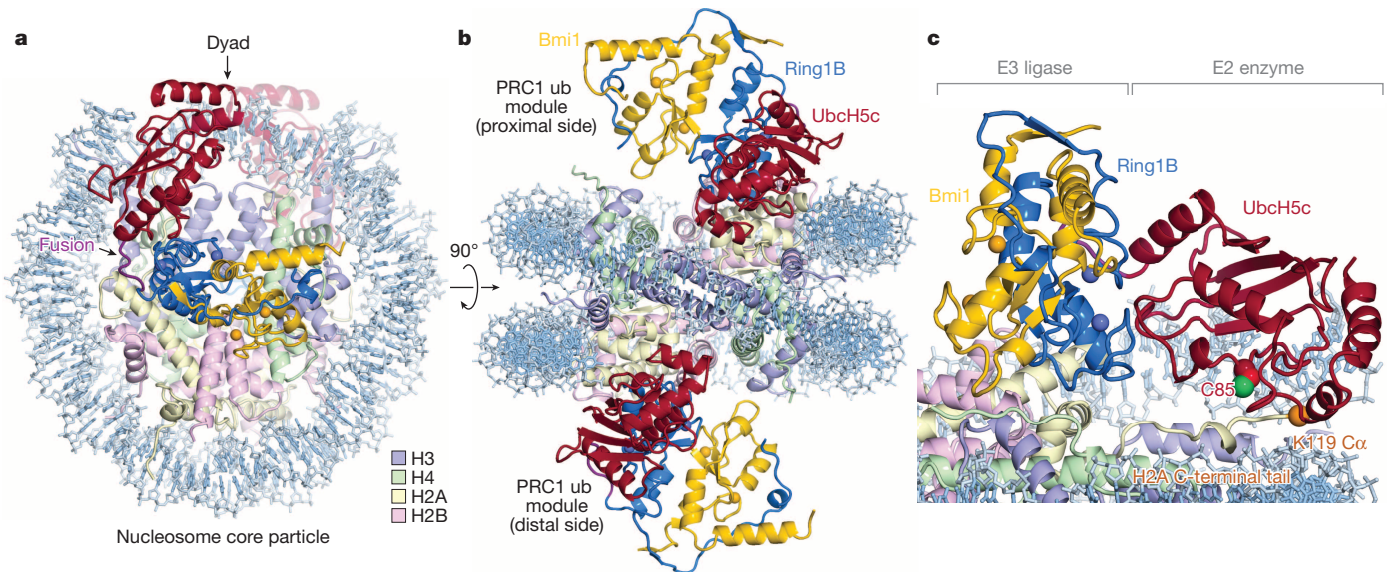


Figure 1 | Crystal structure of PRC1 ubiquitylation module–NCP complex. **a**, View of the complex looking down on the DNA superhelical axis. **b**, Orthogonal view of the complex with proximal and distal halves of the structure indicated. **c**, Zoomed view demonstrating Ring1B–Bmi1 heterodimer

ubiquitylation module contact histone and/or DNA components of the nucleosome burying 1,470 Å² of solvent accessible surface.

We selected UbcH5c for our structural studies owing to its high activity and thorough biochemical and structural characterization with Ring1B–Bmi1. To overcome the low affinity and salt-sensitive E2–E3 interaction²², we genetically fused the C terminus of the Ring1B fragment to the amino terminus of UbcH5c, converting the three-polypeptide PRC1 ubiquitylation module into a stable two-polypeptide version. The two fused PRC1 ubiquitylation modules in our structure align well with the previously determined Ring1B–Bmi1–UbcH5c structure²² (root mean squared deviation (r.m.s.d.) of 1.0 Å over all backbone atoms), suggesting only minor structural changes after NCP binding and confirming the genetic fusion did not markedly alter the structure of the ubiquitylation module (Extended Data Fig. 1a–c, e). Such validation is especially important given that the fused PRC1 ubiquitylation module is inactive in a ubiquitylation assay owing to its inability to undergo E1-mediated trans-thiolation, presumably because the fused E3 clashes with ubiquitin in the E1 adenylation site²⁴ (Extended Data Fig. 2a, b). The NCP is similarly largely unchanged compared to a published NCP structure using the same Widom 601 positioning sequence²⁵ (r.m.s.d. of 0.7 Å over all protein backbone atoms) (Extended Data Fig. 1d, e).

Although the two PRC1 ubiquitylation modules bind their respective nucleosome faces in a similar manner, there are differences particularly at the UbcH5c–nucleosome interfaces. Alignment of the two sides of the structure through the pseudo-two-fold symmetry of the nucleosome reveals near uniformity between the histones on the two sides of the structure (r.m.s.d. of 0.7 Å) (Extended Data Fig. 1f–i). By contrast, the PRC1 ubiquitylation modules from the two faces show more variation (r.m.s.d. of 1.9 Å and 4.6 Å for the E3 and E2 components, respectively), resulting from a flexing of a hinge at the Ring1B–Bmi1–nucleosome interface. Regions of the PRC1 ubiquitylation module close to this hinge, including the Ring1B–Bmi1 nucleosome-binding loops, are very similar on the opposing surfaces of the structure, whereas regions further from the hinge, notably the UbcH5c–nucleosome interfaces, exhibit larger differences. As a result, UbcH5c approaches the NCP more closely on one side (proximal side) of the structure than the other (distal side).

Although we cannot rule out the possibility of plasticity in the UbcH5c–NCP interface, we suspect that the variation in the two halves of the structure result from different environments of each PRC1 module within the crystal lattice. This is not merely a result of the post-crystallization

positioning UbcH5c active site Cys 85 over the H2A C-terminal tail near the H2A Lys 119 Cα atom (orange sphere). Crystals contain the minimal RING heterodimer Ring1B(2–116)–Bmi1(2–109). Ub, ubiquitylation.

dehydration used to obtain 3.3 Å resolution diffraction, as molecular replacement solutions from data collected on less dehydrated crystals with larger unit cells also exhibit similar differences (data not shown). Because the distal side places UbcH5c too far from the nucleosome surface to be consistent with our biochemical studies, we favour the proximal side, which is depicted in the following figures. (Details provided below; the distal side and alignments can be found in Extended Data Fig. 3.) Importantly, the differences between the two sides of the structure influence the extent rather than the presence of interactions at each of the PRC1–NCP interfaces (Extended Data Fig. 1j).

Ring1B–Bmi1–histone interactions

Several basic surfaces of Ring1B and Bmi1 have been shown to be required for nucleosomal ubiquitylation and binding to short duplex DNA²². This informed a structural model docking the Ring1B–Bmi1 heterodimer on DNA near the nucleosome dyad, positioning UbcH5c in proximity to H2A Lys 119 (ref. 22). However, recent work implicating the H2A–H2B acidic patch in Ring1B–Bmi1-mediated H2A ubiquitylation *in vitro* and *in vivo*²⁶ is inconsistent with this docking model. Our structure reveals that the Ring1B–Bmi1 basic residues anticipated to bind DNA actually interact with histones surfaces, including the H2A–H2B acidic patch. The Ring1B–Bmi1 RING heterodimer, which contacts all four histones, forms a saddle over the N-terminal end of the H2B αC helix anchored on each side by Ring1B– and Bmi1–histone interactions (Extended Data Fig. 3a–c). Ring1B binds to the acidic patch of the H2A–H2B dimer. This interaction, which buries nearly 700 Å² of solvent-accessible surface on the proximal side, is mediated by an intricate network of hydrogen bonds and van der Waals contacts and with few exceptions is limited to protein side chains (Fig. 2a and Extended Data Fig. 3d–g, t, u). Consistent with published mutagenesis experiments^{22,26}, H2A Glu 92, Ring1B Lys 97 and Arg 98, and to a lesser extent Ring1B Lys 93 contribute to the Ring1B–nucleosome interface. Ring1B Arg 98 inserts into an acidic pocket generated by H2A residues Glu 61, Asp 90 and Glu 92, making charged hydrogen bonds with each of the H2A side-chain carboxylates. The depth of this pocket is augmented by a ridge composed of H2B residues Glu 105 and His 109, which make van der Waals contacts with the aliphatic region of the Arg 98 side chain. This model of nucleosomal recognition mediated by an ‘arginine-anchor’ binding to the H2A–H2B acidic patch is present in all chromatin factor–NCP crystal structures published so far^{27–30} (Extended Data Fig. 4). Ring1B Lys 97

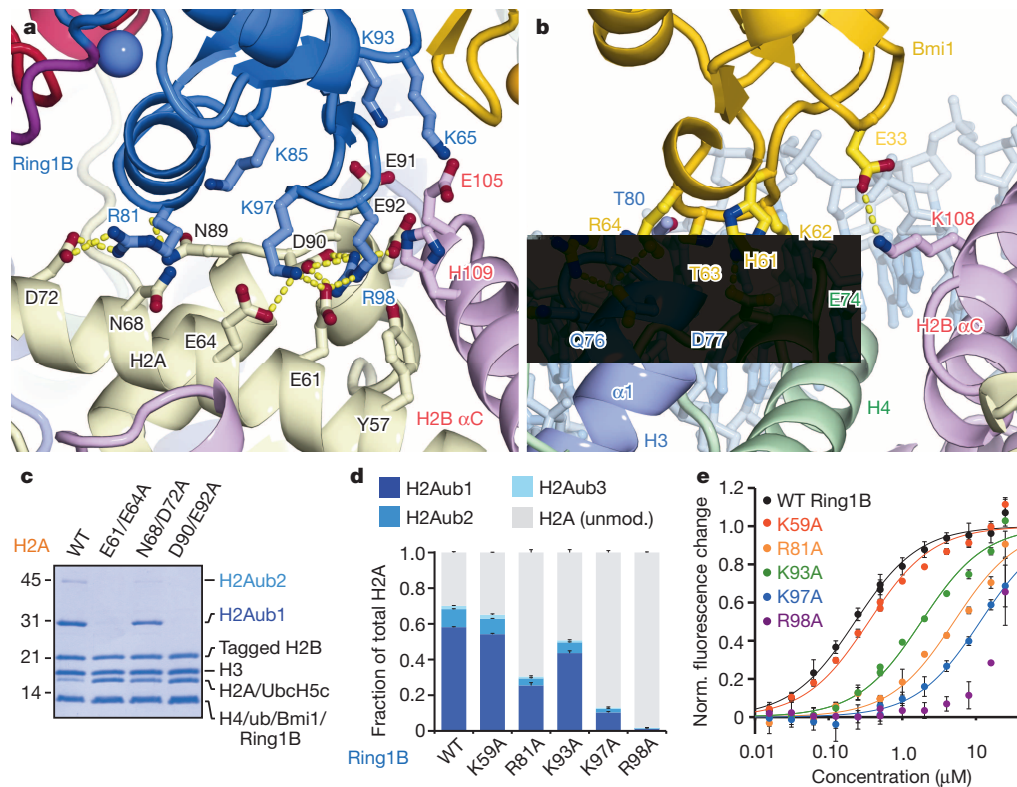


Figure 2 | Ring1B–Bmi1 E3 heterodimer interacts with histone surface.

a, Ring1B–histone acidic patch interactions in cartoon representation with relevant side chains and hydrogen bonds (yellow) highlighted. **b**, Bmi1–histone interface shown similarly. **c**, Coomassie blue stained gel of ubiquitylation assay using H2A mutant NCP as indicated. **d**, Quantified nucleosome ubiquitylation assay using indicated Ring1B

mutants. **e**, Nucleosome binding curves for fused wild-type and Ring1B mutant E2–E3 complexes. Data in **d** and **e** are mean and s.d. for each data point, $n = 3$. Fluorescence is normalized (norm.) to fit values for unbound and saturated NCP. Concentrations depicted using log scale. WT, wild type.

projects into the acidic patch adjacent to Arg 98, hydrogen bonding to H2A residues Glu 61 and Glu 64. Not predicted by previous studies, Ring1B Arg 81 also binds to the acidic patch forming both side-chain and main-chain hydrogen bonds with H2A residues Asp 72 and Asn 89, respectively. This second arginine-binding site is distinct from sites observed in previous chromatin factor–NCP structures. Further contacts are mediated by Ring1B residues Lys 65, Lys 85, Lys 93 and Val95. With minor exceptions the Ring1B–NCP interface is identical on the proximal and distal faces. (Extended Data Fig. 3f, g) The considerable contribution of charged side chains to the Ring1B–nucleosome interface is consistent with the loss of nucleosomal ubiquitylation by the PRC1 ubiquitylation module at higher ionic strength²².

The structural details of the Ring1B–acidic patch interface are validated by our mutagenesis experiments in solution. Tandem alanine mutations of H2A Glu 61 and Glu 64 or Asp 90 and Glu 92 (or the combination thereof) eliminate all detectable nucleosomal ubiquitylation (Fig. 2c and Extended Data Fig. 5a). Nucleosomes containing the H2A Asn68Ala/Asp72Ala double mutant at the secondary arginine-binding site exhibit roughly half wild-type activity levels. To assess the effects of mutations in the PRC1 ubiquitylation module quantitatively, we established fluorescence-based nucleosome ubiquitylation and binding assays³¹ (Extended Data Fig. 6). The Ring1B Arg98Ala mutation results in a 50-fold decrease in both nucleosomal ubiquitylation and the affinity of the fused PRC1 ubiquitylation module for the nucleosome (Fig. 2d, e). Alanine mutations of Ring1B Lys 97, Arg 81 and Lys 93 have similar, yet decreasing effects on nucleosome ubiquitylation and binding, whereas the Ring1B Lys59Ala mutation away from the binding site has little to no effect on PRC1 function.

In contrast to the large Ring1B–NCP interface, the Bmi1–NCP interface is less extensive, burying only 330 Å². Bmi1 forms a cap on the

C-terminal end of the $\alpha 1$ helix of H3 with a Polycomb-specific RING domain loop insertion (Fig. 2b and Extended Data Fig. 3h–k, v, w). This region was previously implicated in PRC1 function given the reduced activity of a Bmi1 (Lys62Ala/Arg64Ala) mutant²². These residues form charged hydrogen bonds with H4 Glu 74 and H3 Asp 77, respectively. The intervening Thr 63 side chain inserts into a small cavity created by the end of the $\alpha 1$ helix and the neighbouring L1 loop of H3, contributing substantial van der Waals contact surface and forming a hydrogen bond with the H3 Gln 76 main-chain carbonyl oxygen. An additional hydrogen bond is made between Bmi1 Glu 33 and H2B Lys 108. The distal side Bmi1–NCP interface shares architectural similarity to its proximal counterpart with the exception of insufficient density to model the Lys 62 side chain (Extended Data Fig. 3j, k).

Mutagenesis of the Bmi1–NCP interface allows us to assess the importance of this surface relative to the Ring1B–NCP interface. The consequences of the Bmi1 (Arg64Ala) mutation are comparable to that of Ring1B (Lys93Ala), the least deleterious mutant tested at the Ring1B–NCP interface (Extended Data Fig. 7d–f and Supplementary Table 2). Meanwhile alanine mutations of His 61 and Lys 62 at the Bmi1–NCP interface and Arg 45 removed from the interface exhibit near wild-type behaviour. A charge reversal mutation of the underlying histone surface (H3 (Asp77Arg)–H4 (Glu74Arg)) decreases ubiquitylation by ~80% (Extended Data Fig. 5a). This reduction in activity is smaller than that resulting from alanine mutations at the Ring1B–NCP interface. Paradoxically, a Bmi1 (Glu33Ala) mutation both markedly decreases nucleosomal ubiquitylation and slightly increases nucleosomal binding affinity. We do not fully understand this phenomenon but it may relate to perturbation of the E3 heterodimer given the location of this side chain at the Ring1B–Bmi1 interface. Overall, our studies suggest a dominant role for Ring1B in the nucleosome binding of PRC1. The 25-fold enhancement

of Ring1B activity by Bmi1 (refs 9, 20) probably represents the modest role for Bmi1 in NCP binding together with a role in constraining ubiquitin at the Bmi1–ubiquitin interface^{16–18}.

In human PRC1 complexes, Bmi1 can be replaced by five paralogues also known as PCGF proteins¹¹. The sequence conservation of these paralogues combined with our structural and biochemical data suggest that each PCGF protein RING domain could support H2A ubiquitylation (Extended Data Fig. 8a). Bmi1 Glu 33 is absolutely conserved mirroring its importance for nucleosomal ubiquitylation. Moreover, Thr 63 and Arg 64 show sequence similarity, maintaining both overall side-chain size and hydrophilicity. Meanwhile, Lys 62, which is dispensable for nucleosomal ubiquitylation, exhibits far less conservation and is replaced by negatively charged side chains in two of the paralogues. The lack of H2A ubiquitylation by canonical PRC1 complexes may reflect masking of the E3 interfaces with either the nucleosome or its E2. Alternatively, the nucleosome surface surrounding H2A Lys 119 may become inaccessible, secondary to canonical PRC1-mediated chromatin compaction. Further structural characterization of canonical and variant PRC1 alone and in complex with oligonucleosomes may shed light on this phenomenon.

UbcH5c–nucleosomal DNA interactions

Nucleosomal binding by the Ring1B–Bmi1 E3 ligase minimal RING heterodimer positions the E2 UbcH5c with its active site directly over the H2A C-terminal tail near the target Lys 119 (Fig. 1c). However, UbcH5c scarcely contacts the underlying nucleosomal histone surface. The lack of E2 active-site–substrate specific interactions may allow UbcH5c to efficiently ubiquitylate several substrates by pairing with different E3 ligases, although we cannot rule out nucleosome-specific contacts by other potential E2s. This contrasts the direct recognition of the target lysine and surrounding consensus sequence by the related E2 Ubc9 in the SUMO–RanGAP1–Ubc9–Nup358 E2–E3 product crystal structure³². On the other hand, UbcH5c does interact with the nucleosome using two non-active site surfaces to bind nucleosomal DNA near the DNA end between superhelical location (SHL) 6 and 7 and at the nucleosomal dyad (Fig. 3a, b and Extended Data Fig. 3l–s). These interfaces vary considerably on the proximal and distal sides of the crystal structure. At the DNA end interface, the UbcH5c antiparallel sheet aligns the side chains of His 32, His 55 and Lys 66, directing them towards the DNA backbone (Fig. 3a). This allows Lys 66 and His 32 to form hydrogen bonds with

adjacent DNA phosphate groups on the proximal side of the structure. His 55 and Thr 58 side chains also approach the neighbouring phosphodiester backbone (3.6 and 3.9 Å, respectively). On the distal side of the crystal structure, each of these residues is ~2 Å farther from the nucleosome surface, limiting direct contact to the Lys 66 side chain that lies 4.1 Å from the DNA backbone (Extended Data Fig. 3n, o). In addition to these backbone interactions, Asn 81 projects into the major groove at the DNA end, although it is unclear in our structure whether base specific recognition occurs.

UbcH5c also contacts the nucleosomal DNA again as its α 3 helix tracks across the nucleosomal dyad (Fig. 3b). On the proximal side, the UbcH5c Arg 125 side chain forms a hydrogen bond with the phosphate in position 1. Although the weaker density in this region of the structure prevents modelling of the Lys 128 side chain, the Lys 128 C β is only 3.9 Å from the phosphate backbone at the dyad, allowing for further contacts. A ~3 Å translation of the distal side UbcH5c α 3 helix along its axis results in a minor rearrangement of the distal interface (Extended Data Fig. 3r, s). Notably, a hydrogen bond is observed between the Lys 128 side chain and the dyad phosphate on the distal side.

Given the novelty of this E2–substrate interaction, we further examined the role of UbcH5c in nucleosomal binding by the PRC1 ubiquitylation module in solution. We initially noted that the fused PRC1 complex binds to the NCP with nearly tenfold increased affinity as compared to Ring1B–Bmi1 alone (Fig. 3c and Supplementary Table 2). Titration of UbcH5c *in trans* incrementally increases the affinity of the Ring1B–Bmi1–NCP interaction, verifying contributions by UbcH5c to nucleosomal binding. At 25 μ M UbcH5c (~2.5 \times the E2–E3 dissociation constant (K_d) under assay conditions²²), the near-saturated, unfused PRC1 ubiquitylation module mimics its fused counterpart both in nucleosome affinity and raw fluorescence change, further validating the use of the fused complex in the structure and to interrogate effects of mutagenesis of all three PRC1 ubiquitylation module components (Fig. 3c and data not shown). Importantly, the Ring1B Asp56Lys mutation that prevents Ring1B–UbcH5c binding²² eliminates affinity changes mediated by UbcH5c *in trans* (Supplementary Table 2).

Having established a role for UbcH5c in nucleosome binding, we next addressed the contributions from the DNA end and DNA dyad interfaces. The UbcH5c(Lys66Glu) charge reversal mutation decreased nucleosomal affinity and ubiquitylation approximately tenfold (Extended Data Fig. 7g–i and Supplementary Table 2). A similar trend was observed

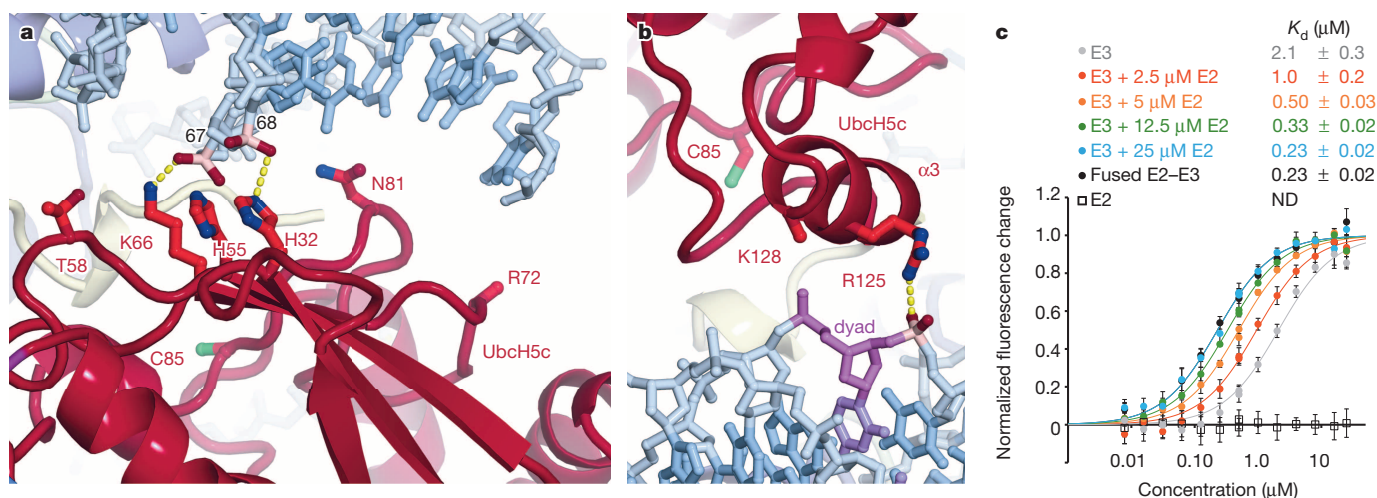


Figure 3 | UbcH5c binds to nucleosomal DNA enhancing Ring1B–Bmi1–NCP affinity. **a**, UbcH5c antiparallel β -sheet–DNA end interactions. Important phosphates are coloured pink. The catalytic UbcH5c Cys 85 is shown. **b**, UbcH5c–dyad interactions. The nucleosome dyad nucleotide is coloured purple. Arg 72 and Lys 128 side chains not modelled owing to limited electron density. **c**, Fluorescence quenching NCP binding curves for E3 Ring1B–Bmi1 alone (grey), fused to E2 UbcH5c (black), or with E2 UbcH5c

added *in trans* (coloured as shown). E2 UbcH5c alone (open squares) is undetectable under assay conditions as it either fails to bind to the NCP or is undetectable owing to the location of the fluorescent probe used to monitor Ring1B–Bmi1 binding. Data are mean and s.d. for each data point, $n = 3$. Fluorescence is normalized to fit values for unbound and saturated NCP. Concentrations depicted using log scale.

with the UbcH5c(Arg125Glu/Lys128Glu) double mutation, although the nucleosomal affinity of this mutant complex could not be calculated owing to aggregation of the complex at high concentrations under assay conditions. The UbcH5c(Thr58Glu) mutant also exhibited significant although less severe defects in binding and ubiquitylation, whereas the Arg136Glu/Arg139Glu negative control mutant away from the NCP surface exhibited wild-type activity. Altogether, these results implicate both DNA end and DNA dyad interfaces in PRC1 function.

The fused PRC1 ubiquitylation module with the UbcH5c(Lys66Glu) charge reversal mutant bound less tightly to the nucleosome than the Ring1B–Bmi1 subcomplex alone, suggesting a repulsive effect in addition to loss of any native interactions. Analysis of alanine mutations provides evidence for the role of UbcH5c side chains in nucleosome binding. At the DNA end interface, the UbcH5c(Lys66Ala) mutation caused a 50% reduction in nucleosome ubiquitylation and a sixfold decrease in nucleosome affinity (Extended Data Fig. 7j–l and Supplementary Table 2). The neighbouring Arg72Ala and His32Ala/His55Ala mutations also showed measurable but more muted effects relative to wild-type proteins. At the dyad interface, the Arg125Ala/Lys128Ala double mutant mimicked the Lys66Ala mutant with a 60% and sixfold reduction in nucleosome ubiquitylation and binding, respectively. Individual alanine mutations of Arg 125 and Lys 128 caused nearly identical intermediate effects. Unlike the charge reversal mutations, all fused complexes bearing UbcH5c alanine mutations bound to the nucleosome with affinities weaker than that of the fused wild-type complex but stronger than the Ring1B–Bmi1 subcomplex alone, implicating the side chains in nucleosome binding. All UbcH5c mutants have similar intrinsic activity as measured by lysine aminolysis of ubiquitin-charged E2s³³ (Extended Data Fig. 2c).

Because the variation between the proximal and distal sides of the structure is greatest at the UbcH5c–DNA end interface, our functional studies present an opportunity to assess the relative authenticity of the two sides. For example, the UbcH5c His 32 and His 55 side chains are 3.3 and 3.6 Å from the DNA backbone on the proximal side, and 5.1 and 7.1 Å on the distal side, respectively. Mutation of both histidines to alanine decreases nucleosomal affinity 1.5-fold while causing a 30% reduction in enzymatic activity. Because the histidine–DNA backbone distances on the proximal side of the structure are more consistent with our biochemical studies, we favour this side as the better representation of the solution state interaction.

Further inspection of the structure suggests potential for additional interactions between UbcH5c and nucleosomal linker DNA (Extended Data Fig. 9a, b). The UbcH5c C-terminal α 4 helix rests across the major groove of linear B-form DNA modelled onto the end of the NCP. The

DNA-adjacent helical face includes two positively charged and two aromatic side chains. Consistent with this hypothesis, we find that the PRC1 ubiquitylation module binds to nucleosomes centred on 185 bp 601 DNA with twice the affinity of those reconstituted with 147 bp 601 DNA (Extended Data Fig. 9c). The absence of linker DNA required for these interactions in our crystals may explain the apparent flexibility of the α 4 helices suggested by weaker electron density and higher B-factors.

In addition to UbcH5c, the Ring1B–Bmi1 E3 ligase can also pair with UbcH5a (also known as UBE2D1), UbcH5b (UBE2D2) and UbcH6 (UBE2E1)²⁰. Meanwhile five other E2s cannot substitute for UbcH5c. These functional classes could result from compatibility differences at the E2–NCP and/or E2–E3 interfaces. Sequence alignment of E2s suggests the possibility that multiple positions including Arg 125 combine to determine PRC1 complementarity (Extended Data Fig. 8b).

Nucleosome recognition by BRCA1

Despite contributions from Bmi1 and UbcH5c, Ring1B is the major contributor to nucleosome recognition as evidenced by buried surface area calculations and the magnitude of effects of Ring1B mutations on nucleosome binding and ubiquitylation. We wondered whether another E3 ligase, BRCA1, recently shown to target nucleosomal H2A Lys 127 and Lys 129 (ref. 34) specifically, might use a similar nucleosome recognition pattern. Much like Ring1B, BRCA1 uses its N-terminal RING domain to pair with the BARD1 RING domain to form a heterodimeric RING-type E3 ligase³⁵. Structural and sequence alignment shows a high level of conservation between the nucleosome-binding loop of Ring1B and the corresponding BRCA1 region, including the Lys–Arg motif (Extended Data Fig. 10a, b). We find that disruption of this motif with the Lys70Ala/Arg71Ala double mutation eliminates nucleosomal ubiquitylation by the BRCA1–BARD1 RING heterodimer (Extended Data Fig. 10c, d). A similar loss of activity was observed after mutation of the H2A–H2B acidic patch (Extended Data Fig. 5b). This suggests the possibility that like Ring1B, BRCA1 uses basic residues in the putative nucleosome-binding loop to bind to the nucleosomal H2A–H2B acidic patch. However, as Ring1B and BRCA1 target different H2A residues (Lys 119 versus Lys 127/Lys 129), additional BRCA1–BARD1 contacts may alter its orientation on the nucleosome.

Nucleosomal H2A specificity

The PRC1 ubiquitylation module requires a nucleosome substrate, showing no activity on an H2A–H2B dimer. Our structure indicates that Ring1B, Bmi1 and UbcH5c interact with nucleosomal H2A–H2B, H3–H4 and DNA, respectively, providing a mechanistic basis for nucleosome specificity. We also believe that the nucleosome has an additional

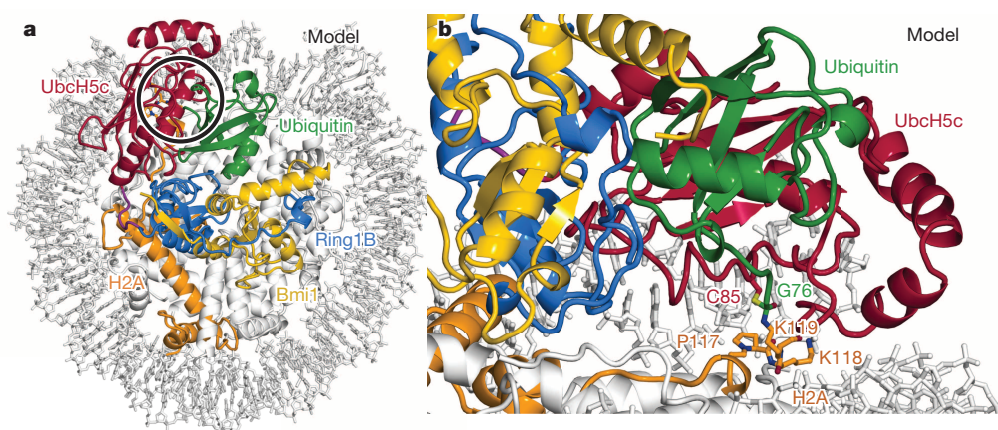


Figure 4 | Model of ubiquitylation transition state complex. **a**, View of transition state model looking down on the DNA superhelical axis with region around catalytic site highlighted with black circle. Ubiquitin modelled from the RNF4–UbcH5a–Ub¹⁶ (PDB accession 4AP4) structure on the basis of the

alignment of its E2 subunit and UbcH5c from the proximal half of the PRC1 ubiquitylation module–NCP structure. **b**, Side view of tetrahedral intermediate model.

role in presenting the H2A C-terminal tail to the UbcH5c active site. In the context of the H2A–H2B dimer, the H2A C-terminal tail is expected to be unstructured beyond residue 105 assigning an entropic cost to ubiquitylation of Lys 119 (ref. 36). However, within the nucleosome, H2A adopts an elongated structure between residues 105 and 117, resting on a platform generated by the underlying H3 surface²¹. Residues beyond this structured region are more flexible, as suggested by the inability to model beyond Lys 118 in many NCP crystal structures^{25,28–30}. As such, the nucleosome both constrains the H2A C-terminal tail trajectory while allowing localized flexibility to reach the UbcH5c active site.

Despite the large number of lysine residues in the nucleosome, the PRC1 ubiquitylation module is specific for H2A Lys 119 and to a lesser extent Lys 118. In our structure, electron density is weak after Pro 117 allowing only main-chain atoms to be built to Lys 119 and Lys 118 on the proximal and distal sides, respectively. Neither main-chain trajectory indicates binding to the UbcH5c active site cleft. To verify that the UbcH5c positions within our structure are compatible with catalysis, we modelled the tetrahedral ubiquitylation transition state using H2A Lys 118 and 119 on each side of the structure (model for proximal side Lys 119 shown in Fig. 4). Aside from the tetrahedral centre, ubiquitin makes no contact with the nucleosome surface in these models. This echoes the lack of SUMO–RanGAP1 interactions in the SUMO–RanGAP1–Ubc9–Nup358 E2–E3 product structure³². We find the position of UbcH5c on both sides of the structure is consistent with H2A Lys 118 and Lys 119 ubiquitylation. Moreover, our structure suggests nucleosome ubiquitylation is limited to these two lysines because UbcH5c is constrained in close proximity to the nucleosome surface, which prevents other lysines from entering the active site. Overall, the multifaceted interaction of the PRC1 ubiquitylation module with the nucleosome and the nucleosomal organization of the H2A C-terminal tail seem to combine to effect H2A ubiquitylation. More generally, our structure illustrates how a chromatin enzyme can use several non-active site surfaces for nucleosome recognition and specificity determination. We anticipate this principle to be shared by many chromatin enzymes that, like PRC1, function as part of large multi-component complexes.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 June; accepted 26 September 2014.

- Boyer, L. A. *et al.* Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature Cell Biol.* **441**, 349–353 (2006).
- Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nature Struct. Mol. Biol.* **20**, 1147–1155 (2013).
- Laugesen, A. & Helin, K. Chromatin repressive complexes in stem cells, development, and cancer. *Cell Stem Cell* **14**, 735–751 (2014).
- Bracken, A. P. & Helin, K. Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nature Rev. Cancer* **9**, 773–784 (2009).
- Crea, F., Paolicchi, E., Marquez, V. E. & Danesi, R. Polycomb genes and cancer: Time for clinical application? *Crit. Rev. Oncol. Hematol.* **83**, 184–193 (2012).
- Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature Rev. Mol. Cell Biol.* **10**, 697–708 (2009).
- Schwartz, Y. B. & Pirrotta, V. A new world of Polycombs: unexpected partnerships and emerging functions. *Nature Rev. Genet.* **14**, 853–864 (2013).
- Wang, H. *et al.* Role of histone H2A ubiquitylation in Polycomb silencing. *Nature* **431**, 873–878 (2004).
- Cao, R., Tsukada, Y.-I. & Zhang, Y. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Mol. Cell* **20**, 845–854 (2005).
- Gao, Z. *et al.* PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344–356 (2012).
- Simon, J. A. & Kingston, R. E. Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol. Cell* **49**, 808–824 (2013).
- Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445–1459 (2014).
- Farcas, A. M. *et al.* KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *eLife* **1**, e00205 (2012).

- Wu, X., Johansen, J. V. & Helin, K. Fbx10/Kdm2b Recruits Polycomb Repressive Complex 1 to CpG Islands and Regulates H2A Ubiquitylation. *Mol. Cell* **49**, 1134–1146 (2013).
- Pickart, C. M. & Pickart, C. M. Mechanisms underlying ubiquitination. *Annu. Rev. Biochem.* **70**, 503–533 (2001).
- Plechanová, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).
- Pruneda, J. N. *et al.* Structure of an E3:E2–Ub Complex Reveals an Allosteric Mechanism Shared among RING/U-box Ligases. *Mol. Cell* **47**, 933–942 (2012).
- Dou, H., Buetow, L., Sibbet, G. J., Cameron, K. & Huang, D. T. BIRC7–E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. *Nature Struct. Mol. Biol.* **19**, 876–883 (2012).
- Deshais, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
- Buchwald, G. *et al.* Structure and E3-ligase activity of the Ring–Ring complex of polycomb proteins Bmi1 and Ring1b. *EMBO J.* **25**, 2465–2474 (2006).
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
- Bentley, M. L. *et al.* Recognition of UbcH5c and the nucleosome by the Bmi1/Ring1b ubiquitin ligase complex. *EMBO J.* **30**, 3285–3297 (2011).
- Mattiroli, F., Uckelmann, M., Sahtoe, D. D., van Dijk, W. J. & Sixma, T. K. The nucleosome acidic patch plays a critical role in RNF168-dependent ubiquitination of histone H2A. *Nat. Commun.* **5**, 3291 (2014).
- Olsen, S. K. & Lima, C. D. Structure of a ubiquitin E1–E2 complex: insights to E1–E2 thioester transfer. *Mol. Cell* **49**, 884–896 (2013).
- Vasudevan, D., Chua, E. Y. D. & Davey, C. A. Crystal structures of nucleosome core particles containing the ‘601’ strong positioning sequence. *J. Mol. Biol.* **403**, 1–10 (2010).
- Leung, J. W., Agarwal, P., Canny, M. D., Gong, F. & Robison, A. D. Nucleosome acidic patch promotes RNF168- and RING1B/BMI1-dependent H2A and H2A ubiquitylation and DNA damage signaling. *PLoS Genet.* **10**, e1004178 (2014).
- Barbera, A. J. *et al.* The nucleosomal surface as a docking station for Kaposi’s sarcoma herpesvirus LANA. *Science* **311**, 856–861 (2006).
- Makde, R. D., England, J. R., Yennawar, H. P. & Tan, S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature* **467**, 562–566 (2010).
- Armache, K. J., Garlick, J. D., Canzio, D., Narlikar, G. J. & Kingston, R. E. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science* **334**, 977–982 (2011).
- Kato, H. *et al.* A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* **340**, 1110–1113 (2013).
- Hieb, A. R., D’Arcy, S., Kramer, M. A., White, A. E. & Luger, K. Fluorescence strategies for high-throughput quantification of protein interactions. *Nucleic Acids Res.* **40**, e33 (2012).
- Reverter, D. & Lima, C. D. Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. *Nature* **435**, 687–692 (2005).
- Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHAR1 to be RING/HECT hybrids. *Nature* **474**, 105–108 (2011).
- Kalb, R., Mallory, D. L., Larkin, C., Huang, J. T. J. & Hiom, K. BRCA1 is a histone–H2A-specific ubiquitin ligase. *Cell Rep.* **8**, 999–1005 (2014).
- Brzovic, P. S., Rajagopal, P., Hoyt, D. W., King, M. C. & Klevit, R. E. Structure of a BRCA1–BARD1 heterodimeric RING–RING complex. *Nature Struct. Mol. Biol.* **8**, 833–837 (2001).
- Zhou, Z. *et al.* NMR structure of chaperone Chz1 complexed with histones H2AZ–H2B. *Nature Struct. Mol. Biol.* **15**, 868–869 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We would like to thank the staff of APS NE-CAT beamlines 24ID-E and 24ID-C for their assistance during synchrotron data collection; H. Yennawar, N. Yennawar and J. Fecko at the Penn State Huck Institute X-ray core facility; A. Minns, M. Moore and J. Malloy for reagent preparation; T. Girish and J. Huang for assistance with data collection; S. Wang for providing U2OS cDNA; R. Makde and members of the Tan laboratory and the Penn State Center for Eukaryotic Gene Regulation for discussions. This work was supported by NIGMS grants GM060489–09S1, GM088236 and GM111651 to S.T. R.K.M. is supported by a Damon Runyon Post-doctoral fellowship (DRG 2107–12).

Author Contributions R.K.M. designed the study, cloned, purified and crystallized macromolecules, collected and processed X-ray data, refined and analysed the structure, performed nucleosome ubiquitylation and binding assays, and wrote the paper. R.C.H. cloned, purified and crystallized macromolecules and performed nucleosome ubiquitylation assays. S.T. designed the study, cloned and purified macromolecules, collected X-ray data, analysed the structure and edited the paper. All authors commented on the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structure have been deposited with the Protein Data Bank under accession code 4R8P. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.T. (sxt30@psu.edu).

METHODS

Preparation of proteins, protein complexes and NCPs. Complete genes for human *RING1B* (also known as *RNF2*), *BMI1*, *UbcH5c* (*UBE2D3*) and *UBA1* were amplified from HeLa complementary DNA. Gene fragments of *BRCA1* and *BARD1* containing the RING domains were amplified from HeLa and U2OS cDNA, respectively. The gene for ubiquitin was synthesized by Invitrogen. UbcH5c(2–147), Uba1(2–1058) and ubiquitin(1–76) were each cloned into pST50Tr vectors with an N-terminal tandem Strep peptide (STR)-decahistidine (His₁₀) affinity tag. STR-His₁₀-Ring1B(2–116)-Bmi1(2–109) and STR-His₁₀-BRCA1(2–103)-BARD1(27–130) were each cloned into pST44 vectors for polycistronic co-expression³⁷. A similar pST44 vector was created to co-express Bmi1(2–109) with Ring1B(2–116) fused in frame to UbcH5c(2–147) through a linker encoding the protein sequence, GSGSRS. Point mutants were generated by site-directed mutagenesis. The fused human PRC1 ubiquitylation module containing Bmi1(2–109) and STR-His₁₀-Ring1B(2–116)-GSGSRS-UbcH5c(2–157) was expressed in Rosetta(DE3)pLysS *Escherichia coli* at 18 °C. The complex was enriched by metal-affinity chromatography with Talon resin (Clontech), the affinity tag removed with tobacco etch virus (TEV) protease and the complex further purified by SourceS cation-exchange chromatography (GE Healthcare). The Ring1B(2–116)-Bmi1(2–109) complex and UbcH5c(2–147) single protein were expressed and purified similarly. STR-His₁₀-Uba1(2–1058) and STR-His₁₀-ubiquitin(1–76) were expressed in CodonPlus(DE3) and BL21(DE3)pLysS *E. coli* at 18 °C and 37 °C, respectively. Uba1 was purified by metal-affinity chromatography as above, followed by SourceQ anion-exchange chromatography (GE Healthcare) without cleavage of the affinity tag. Ubiquitin was purified similarly with the affinity tag intact using SourceS cation-exchange chromatography. The STR-His₁₀-BRCA1(2–103)-BARD1(27–130) complex was expressed in Rosetta(DE3)pLysS *E. coli* at 18 °C and purified by metal-affinity chromatography, TEV cleavage, and SourceQ anion-exchange chromatography. All mutants were expressed and purified identically to their wild-type versions including ion-exchange chromatography. Histone mutants were generated by site-directed mutagenesis. Recombinant *Xenopus* histones were expressed, purified and reconstituted with the 147- or 185-bp Widom 601 sequence³⁸ essentially as described previously³⁹ including anion-exchange purification of NCPs. Fluorescent labelling of histones was performed as described¹.

Characterization of proteins and protein complexes. All proteins and complexes were evaluated by dynamic light scattering using a Protein Solutions DynaPro instrument. Scattering was analysed using the Dynamics v6 software package. All E2 and E3 mutants showed equivalent monodispersity to their wild-type controls (polydispersity <20%). Owing to paradoxical binding and activity results, the secondary structure of Bmi1(Glu33Ala) mutant E3 heterodimer and fused E2–E3 complexes were further analysed and compared to wild-type controls using a Jasco J-1500 Circular Dichroism Spectrometer. Spectra were obtained at 0.1 mg ml⁻¹ protein concentration in 2 mM Tris-Cl, pH 7.6, containing 150 mM NaF. Data were analysed using CDPro.

Complex reconstitution and crystallization. The PRC1 ubiquitylation module was reconstituted with NCPs prepared with 147-bp 601 DNA at a 2.8–3.0:1 module:nucleosome ratio in 20 mM Tris-Cl, pH 7.6, 75 mM NaCl, 10 μM ZnSO₄, 1 mM dithiothreitol (DTT). The complex was purified by Superdex 200 size exclusion chromatography (GE Healthcare) in reconstitution buffer supplemented with 0.2 mM phenylmethylsulphonyl fluoride (PMSF). Pooled fractions were concentrated to ~10 mg ml⁻¹ by VS500 centrifuge ultrafiltration (Vivascience). Crystallization was performed in modified microbatch⁴⁰ by combining 1 μl of the concentrated complex with 1 μl 25 mM HEPES, pH 7.5 (Hampton), 80 mM NH₄NO₃ (Hampton), and 3% PEG 2000-MME (Fluka) at 21 °C overlaid by 70 μl Al³⁺ oil (1:1 mixture of silicon oil (Clearco) and mineral oil (Fisher)). Crystals were collected after 12–60 days. Cryoprotection and dehydration was performed by soaking in 25 mM HEPES, pH 7.5, 120 mM NH₄NO₃, 2.5% ethanol, 0.5 mM DTT, and 5% PEG 2000-MME with increasing amounts of PEG 550-MME (Fluka, 0–28% in 4% increments, every 10–15 min) at room temperature before flash cooling in liquid nitrogen.

Data collection, data processing, model building and refinement. Diffraction data were collected at Advanced Photon Source's NE-CAT beamline 24-ID-E (0.9792 Å; 100 K). Data were processed using XDS⁴¹ and Scala⁴². The structure was solved by molecular replacement with Phaser⁴³ and polyaniline versions of the PRC1 ubiquitylation module²² (PDB code 3RPG) and the NCP with 145-bp Widom 601 DNA²⁵ (PDB code 3LZ0) as rigid body search models (Phaser final log likelihood gain 3,700). Crystallographic refinement was performed with iterative automatic refinement using REFMAC5^{44,45}, PHENIX⁴⁶, and CNS⁴⁷ together with manual model building in COOT⁴⁸. The structure was refined to 3.3 Å resolution with $R_{\text{work}}/R_{\text{free}}$ 19.7%/24.5%. The asymmetric unit contains one copy of the 2:1 PRC1 ubiquitylation module–NCP complex in C222₁ space group. The final model includes Bmi1 residues 3–107 for chain K, residues 6–102 for chain M; Ring1B residues 15–116 for chain L, 16–116 for chain N; UbcH5c residues 2–147 for chain L, 2–147 for chain N (residues for UbcH5c are offset by 200 in the PDB file (that is, 202–347), linker

between Ring1B and UbcH5c in final model residues 117–122 for chains L and N); H3 residues 37–134 for chain A, residues 38–134 for chain E; H4 residues 17–102 for chain B, residues 21–102 for chain F; H2A residues 16–119 for chain C, residues 12–118 for chain G; H2B residues 31–125 for chain D, residues 31–125 for chain H; 601 DNA residues –71 to +72 for chain I, residues –72 to +72 for chain J. The histone tails are typically not observed in NCP structures. Above average B-factors are typical for structures containing the NCP especially at resolutions greater than 3 Å^{30,49–51}. Ramachandran statistics were analysed with PROCHECK⁵² (88.4% of residues in most favoured regions, 11.5% in additional allowed regions, 0.1% in generously allowed regions, 0.1% in disallowed regions — disallowed residue corresponds to Ser 122 in L chain contained in fusion between Ring1B and UbcH5c). Simulated annealing omit maps cumulatively covering the entire structure were generated with CNS. All molecular graphics were prepared with PyMOL⁵³. Buried surface areas and r.m.s.d. values for structure alignments were calculated using PyMOL.

Ubiquitylation assays. Ubiquitylation assays were performed by mixing 2 μM NCP, 5 μM STR-His₁₀-ubiquitin, 30 nM STR-His₁₀-Uba1 (E1), 0.375 μM UbcH5c (E2) and 0.375 μM Ring1B(2–116)-Bmi1(2–109) or BRCA1(2–103)-BARD1(27–130) (E3) in 50 mM HEPES, pH 7.5, 75 mM NaCl, 1 mM DTT, 10 μM ZnSO₄, 2 mM MgCl₂, 3 mM ATP. For histone mutants assays, non-fluorescent nucleosomes with an STR-His₂-tagged H2B were used. For PRC1 ubiquitylation module mutants, wild-type recombinant nucleosomes were doped with 5% nucleosomes containing the H2A(Thr10Cys) mutant labelled with Oregon Green 488 maleimide. Assay samples ranging from 20 to 80 μl were placed in a water bath for 60 min at 30 °C after which assays were quenched with an equivalent volume of 2× SDS–PAGE gel loading dye and boiled for 5 min. Assay samples were separated by SDS–PAGE using 18% gels, scanned with a Typhoon 9410 imager (GE Healthcare; 488 nm excitation; 526 nm emission; PMT 500) and stained with Coomassie brilliant blue. Quantification of assays was performed using Image Quant software by drawing boxes around each band. Each quantified band was background subtracted using an identically sized box in a region of the same lane without signal. As only unmodified H2A, H2Aub1, H2Aub2 and H2Aub3 were observed, the sum of these background-subtracted bands was considered to be the total H2A for each lane. In this manner, fraction of total H2A was calculated for each band. These fractional values could be compared across gels reproducibly (Extended Data Fig. 6). Three technical replicates were performed for each mutant with all mutants of a given type (that is, all Ring1B mutants) executed simultaneously with both positive and negative controls. In most cases, three gels were run each containing one of the three replicates of each of the mutants in a given data set. $n = 3$ was found to be sufficient for reproducible data (Extended Data Fig. 6) with low standard deviations. Assays for E1-mediated transfer of ubiquitin to UbcH5c Cys85Lys mutants were performed as previously described¹⁶. In brief, 100 μM STR-His₁₀-ubiquitin, 1 μM STR-His₁₀-Uba1 (E1), 80 μM UbcH5c or fused PRC1 ubiquitylation module with the Cys85Lys mutant were combined in 50 mM Tris-Cl, pH 10, 150 mM NaCl, 8 mM TCEP, 5 mM MgCl₂ and incubated at 35 °C for 24 h. Assay samples were separated by SDS–PAGE and stained as described above.

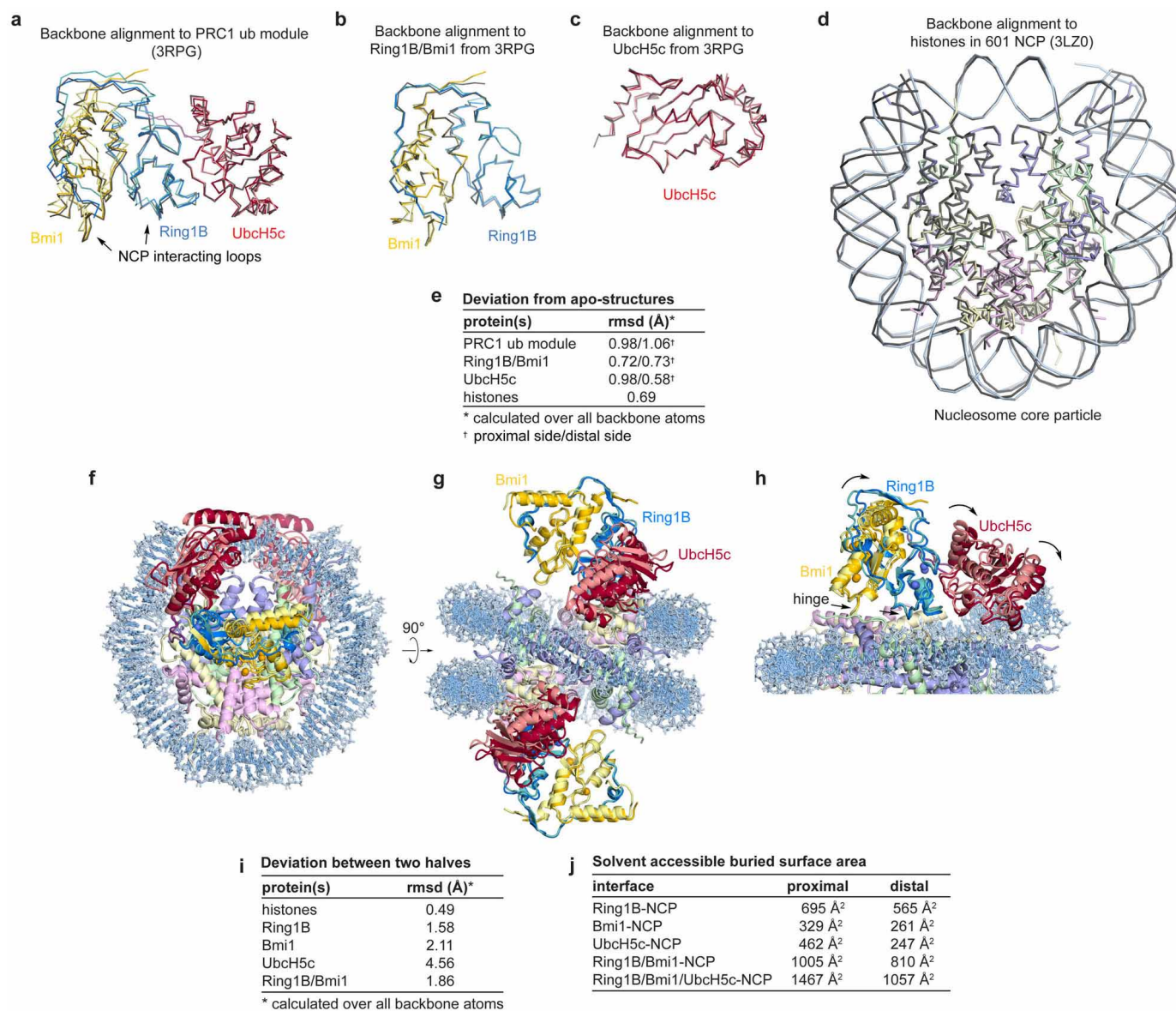
UbcH5c aminolysis assays. To assess intrinsic function of previously uncharacterized UbcH5c mutants, lysine reactivity aminolysis assays were performed basically as previously described³³. In brief, 20 μM UbcH5c (wild-type or mutant) was combined with 20 μM STR-His₁₀-ubiquitin and 0.5 μM STR-His₁₀-Uba1 (E1) in 80 μl of 25 mM sodium phosphate, pH 7.0, 150 mM NaCl, 10 mM MgCl₂ and 10 mM ATP. Following a 20-min incubation at 37 °C, lysine was added to a final concentration of 50 mM. Aminolysis time points were taken at 0, 10 and 30 min and analysed by SDS–PAGE. Reduction of the no-lysine control was performed with 100 mM DTT.

Nucleosome binding assays. Nucleosome binding assays were performed with nucleosomes reconstituted with a H2B(Ser112Cys) mutant labelled with Oregon Green 488 maleimide (Life Technologies) with approximately ~50% efficiency and 147- or 185-bp 601 DNA. Fluorescence quenching assays were performed essentially as previously described³¹. In brief, chromatin binding proteins (wild-type or mutant-fused PRC1 ubiquitylation module or Ring1B–Bmi1 and/or UbcH5c) were diluted in 20 mM Tris-Cl, pH 7.6, 100 mM NaCl, 5 mM DTT, 5% glycerol, 0.01% NP-40, 0.01% CHAPS, 0.1 mg ml⁻¹ BSA to concentrations of 50, 32, 16, 8, 4, 2, 1, 0.5, 0.25, 0.125, 0.062, 0.031 and 0.016 μM. The labelled nucleosome was diluted to 2 nM in identical buffer containing 50 mM NaCl. Twenty microlitres of NCP was added to 20 μl of each protein/complex dilution in triplicate in a passivated 384-well plate (Greiner). Following centrifugation, mixing and 10–20-min incubation, plates were scanned with a Typhoon 9410 imager (488 nm excitation; 526 nm emission; PMT 600). Fluorescence was quantitated with ImageQuant software and data was analysed with ProFit (QuantumSoft). For E2 titration experiments, UbcH5c was added at twice the indicated concentration to each of the Ring1B–Bmi1 dilutions, giving the final indicated concentration following 1:1 mixing with the fluorescent NCP. $n = 3$ was found to be sufficient for reproducible data (Extended

Data Fig. 6) with low standard deviations. One data point of all the data points measured was rejected owing to obvious defects in the microplate well. The fused PRC1 ubiquitylation module was most appropriate for measuring affinity constants for mutants of all three components because the higher affinity of the fused complex (relative to E2 or E3 alone) enabled a larger dynamic range for the assay. This allowed the scope of a mutation's effect to be measured and compared to all other mutants. **BRCA1–BARD1/Ring1B–Bmi1 alignment.** The BRCA1–BARD1/Ring1B–Bmi1 RING heterodimer alignment was performed using the PyMOL pair_fit command with C α atoms of BRCA1 residues 22–55 and 60–76 and the corresponding C α atoms of Ring1B residues 49–79 and 86–102. The alignment gave a r.m.s.d. value of 2.69 Å. The BRCA1 loop containing residues 53–59 (corresponding to Ring1B loop residues 80–85) was excluded owing to length discrepancy. Divergent regions adjacent to the RING domains were also excluded.

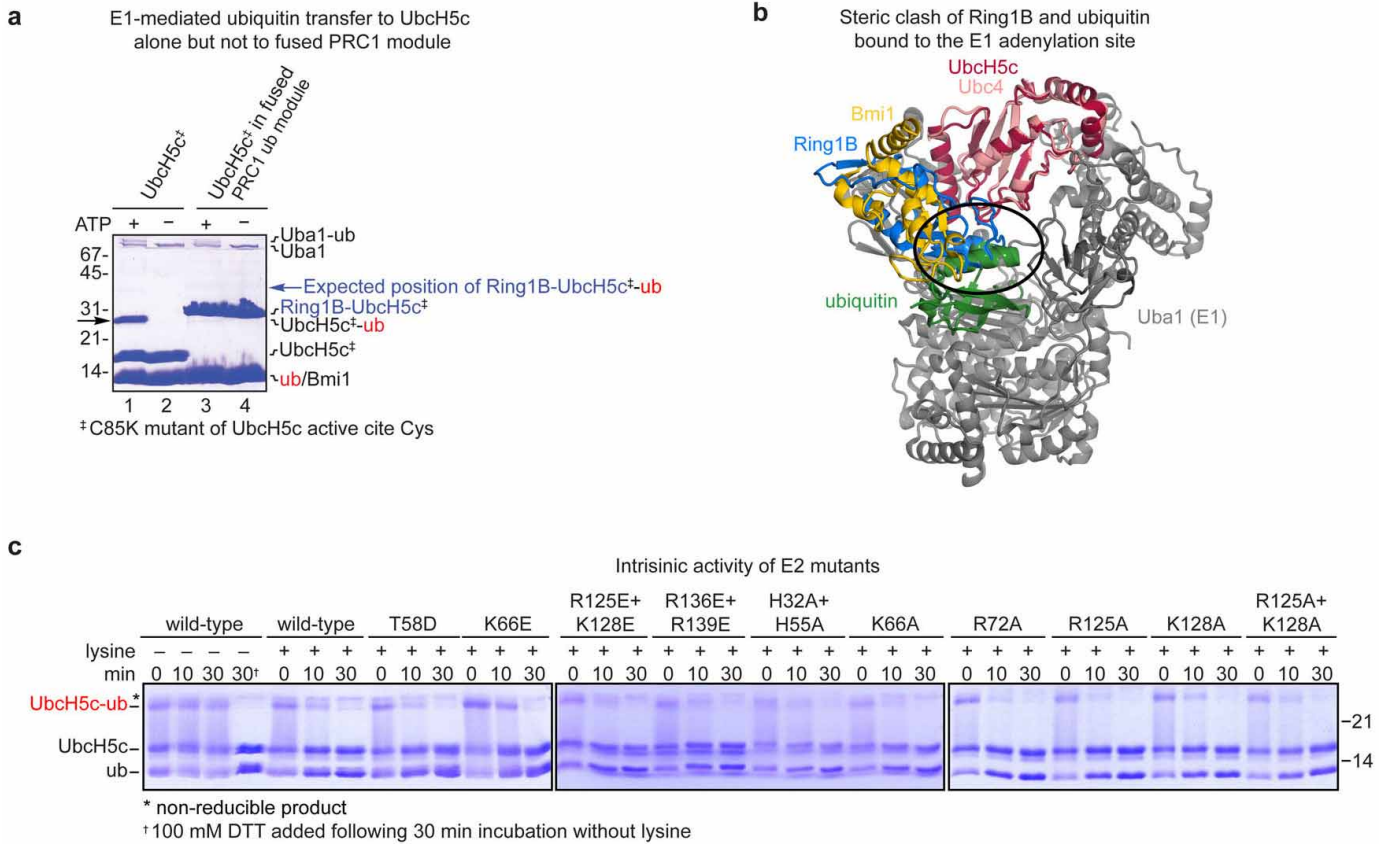
Tetrahedral intermediate model building. To construct the tetrahedral model for the ubiquitylation intermediate, an initial model for ubiquitin was prepared by aligning UbcH5a from the RNF4–UbcH5a–ubiquitin structure¹⁶ (PDB code 4AP4) with UbcH5c in the PRC1 ubiquitylation module–NCP structure. Side chains were built for H2A Lys 118 and Lys 119 allowing the Lys 119 NZ atom to be placed in proximity to the adjacent UbcH5c Cys 85 catalytic residue using COOT. A restraint library was constructed using JLigand⁵⁴ converting ubiquitin Gly 76 to ethanolamine and linking its C β atom to both the NZ atom of H2A Lys 119 and the SG atom of UbcH5c Cys 85. The geometry of the model was idealized using REFMAC5. This was repeated for the proximal and distal sides of the structure for H2A Lys 118 and Lys 119. Each model resulted in minimal perturbation of the PRC1 ubiquitylation module structure outside of H2A Lys 118 and Lys 119 residues. Furthermore, each model allows addition of H2A residues C-terminal to Lys 119 without steric clash.

37. Tan, S., Kern, R. C. & Selleck, W. The pST44 polycistronic expression system for producing protein complexes in *Escherichia coli*. *Protein Expr. Purif.* **40**, 385–395 (2005).
38. Lowary, P. T. & Widom, J. New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. *J. Mol. Biol.* **276**, 19–42 (1998).
39. Luger, K., Rechsteiner, T. J. T. & Richmond, T. J. T. Preparation of nucleosome core particle from recombinant histones. *Methods Enzymol.* **304**, 3–19 (1999).
40. D'Arcy, A., Mac Sweeney, A. & Haber, A. Practical aspects of using the microbatch method in screening conditions for protein crystallization. *Methods* **34**, 323–328 (2004).
41. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
42. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
43. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
44. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
45. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
46. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
47. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
48. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
49. Arnaudo, N. *et al.* The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nature Struct. Mol. Biol.* **20**, 1119–1121 (2013).
50. Tachiwana, H. *et al.* Crystal structure of the human centromeric nucleosome containing CENP-A. *Nature* **476**, 232–235 (2011).
51. Iwasaki, W. *et al.* Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS Open Bio.* **3**, 363–369 (2013).
52. Laskowski, R. A., MacArthur, M. W. & Moss, D. S. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
53. The PyMOL Molecular Graphics System v. 1.7.0.5 (Shrödinger, LLC, 2014).
54. Lebedev, A. A. *et al.* JLigand: a graphical tool for the CCP4 template-restraint library. *Acta Crystallogr. D* **68**, 431–440 (2012).



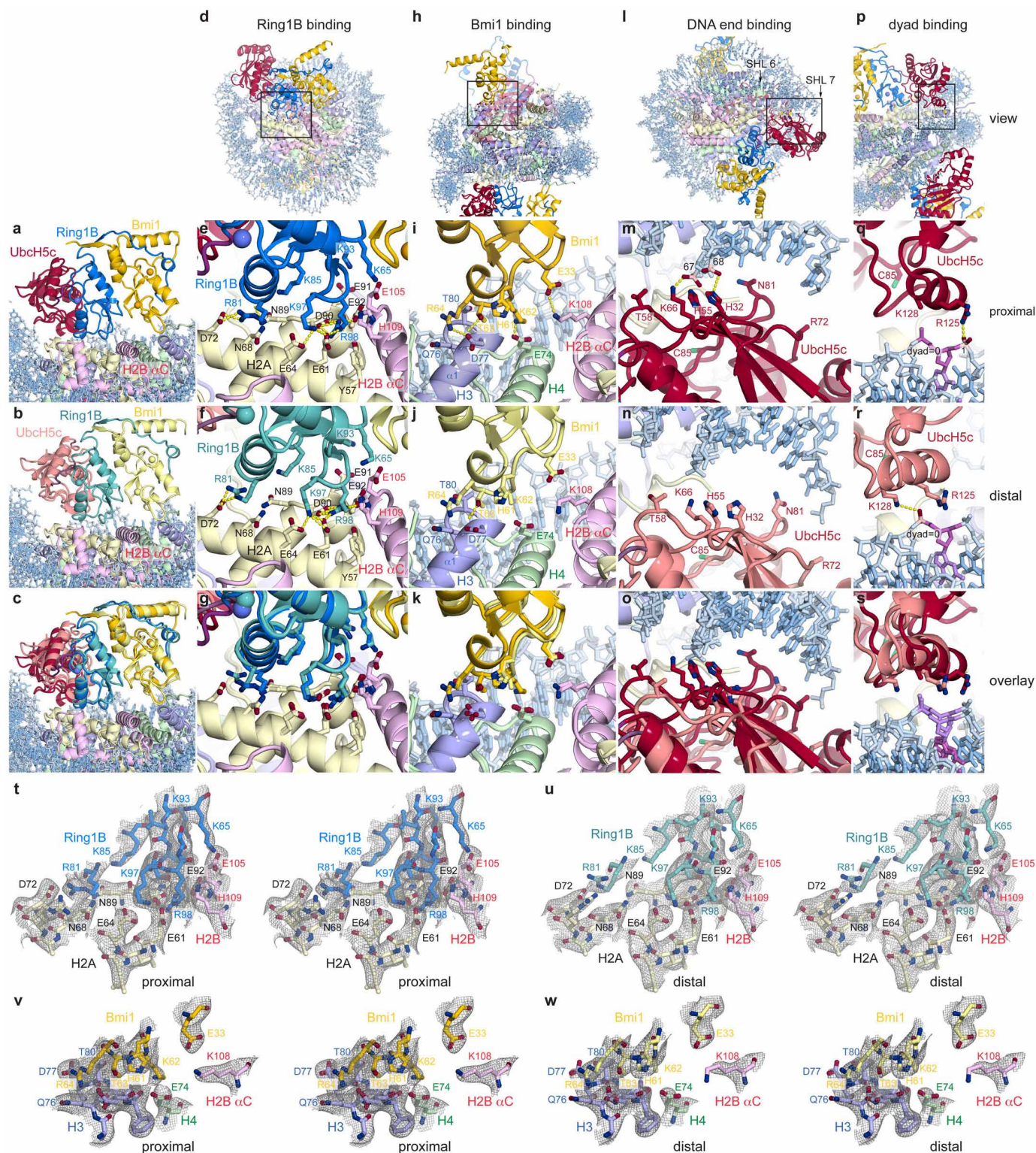
Extended Data Figure 1 | Alignment of PRC1 ubiquitylation module and NCP with previously determined sub-structures and comparison of the two halves of the structure. **a**, Alignment of PRC1 ubiquitylation module from the proximal (darker colours) and distal (lighter colours) halves of the PRC1 ubiquitylation module–NCP structure with the previously determined structure of the PRC1 ubiquitylation module alone²² (grey, PDB code 3RPG) using all backbone atoms. **b, c**, A similar alignment using the Ring1B–Bmi1 subcomplex (**b**) or UbcH5c only (**c**). **d**, Alignment of the NCP from the PRC1 ubiquitylation module–NCP core particle complex (coloured) and the previously determined NCP complex²⁵ (PDB code 3LZ0) containing the identical 601 nucleosome positioning sequence. **e**, R.m.s.d. for each of the alignments calculated over all backbone atoms. **f, g**, Orthogonal views of an

alignment of two copies of the complex following rotation of one copy about the pseudo-two-fold axis of symmetry of the nucleosome. This was accomplished by simultaneously aligning each histone in one copy to its symmetry related counterpart in the other copy. For visualization, the PRC1 ubiquitylation module from one copy of the structure is depicted with lighter colours. **h**, Zoomed view of the alignment showing PRC1 ubiquitylation module–NCP interface. The proximal and distal PRC1 ubiquitylation modules are shown in dark and light colours, respectively. Hinge indicated by arrows. **i**, R.m.s.d. for each of the components following alignment of the two halves as described above. **j**, Buried surface area calculations for the indicated interfaces on the proximal and distal sides of the structure.



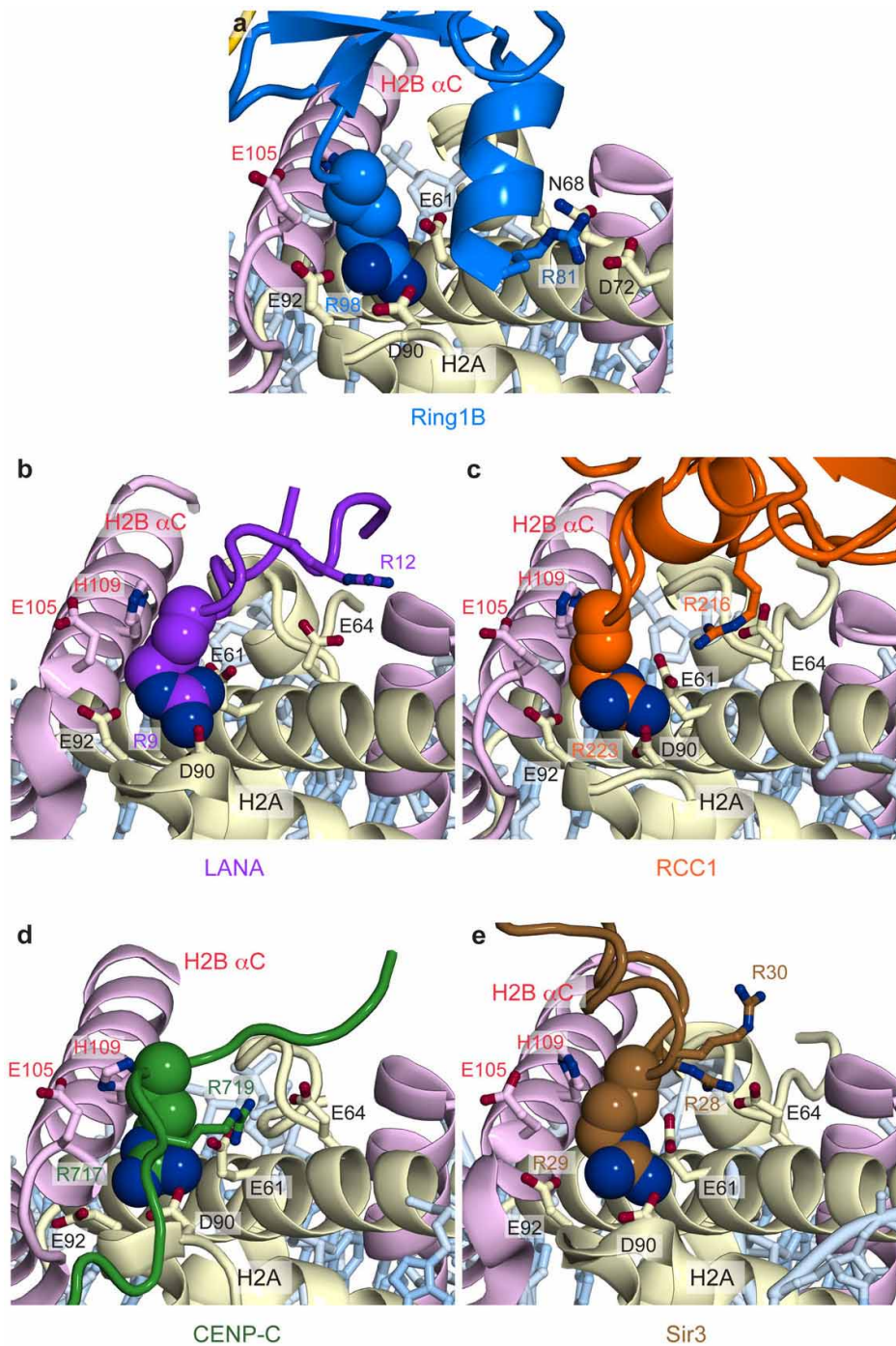
Extended Data Figure 2 | Characterization of fused and unfused UbcH5c. **a**, Coomassie-stained gel of E1-mediated ubiquitin transfer to UbcH5c(Cys85Lys) mutant (UbcH5c[†]) alone and in the context of the fused PRC1 ubiquitylation module. Ubiquitin is transferred by the E1 Uba1 to the UbcH5c mutant in an ATP-dependent manner, black arrow (lanes 1 and 2). No ubiquitin transfer to the fused PRC1 ubiquitylation module is observed (lanes 3 and 4, expected band position for Ring1B-UbcH5c-ubiquitin conjugate indicated by a blue arrow). The Cys85Lys mutant was used to assist in visualization of the UbcH5c-ubiquitin stable isopeptide conjugates. **b**, Ring1B

of the fused PRC1 ubiquitylation module (blue) clashes with ubiquitin in the E1 Uba1 adenylation site (green). Alignment of UbcH5c subunit of PRC1 ubiquitylation module²² (PDB code 3RPG) to Ubc4 in Ubc4-Uba1-ubiquitin structure²⁴ (PDB code 4II2). Ubc4 and Uba1 are shown in pink and grey, respectively. **c**, Intrinsic activity of wild-type and mutant UbcH5c enzymes. Time-course of UbcH5c-ubiquitin thioester aminolysis after lysine addition. Reactions quenched with non-reducing loading buffer unless 100 mM DTT addition is indicated (dagger). Asterisk denotes a non-reducible product.



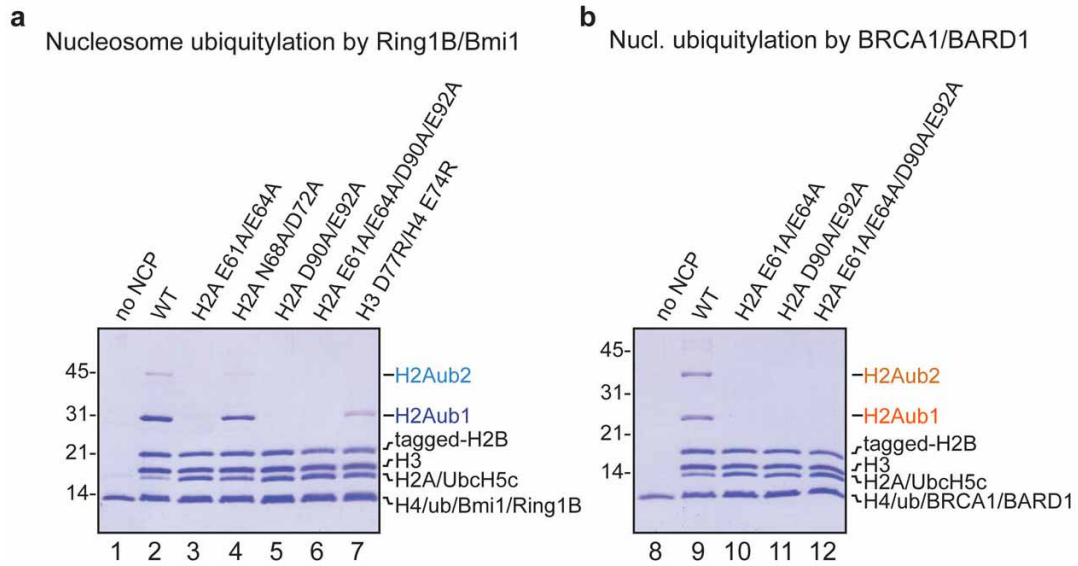
Extended Data Figure 3 | Comparison of the proximal and distal Ring1B-, Bmi1- and UbcH5c-nucleosome interfaces. a–c, Identical views of Ring1B–Bmi1 saddle from proximal (a) and distal (b) halves of structure and overlay (c). d–g, Ring1B–histone interface views, including zoom out (d) with box indicating field of view depicted in magnified panels from proximal (e) and distal (f) halves of structure and overlay (g). h–k, Similar views of Bmi1–histone interface. l–s, Similar views of UbcH5c–DNA interfaces.

t–u, $2mF_o - DF_c$ difference Fourier transform electron density maps for the Ring1B–nucleosome interface contoured at 1.0σ on the proximal (t) and distal (u) sides of the PRC1 ubiquitylation module–NCP structure. v, w, Similarly contoured electron density map for the proximal (v) and distal (w) Bmi1–nucleosome interface. All views are identical to those depicted in Figs 2 and 3.



Extended Data Figure 4 | Chromatin factors use an arginine-anchor to bind the H2A–H2B acidic patch. a–e, Identical views of E3 ligase Ring1B (blue) (a), viral LANA peptide (purple, PDB code 1ZLA; ref. 27) (b), chromatin factor RCC1 (purple, PDB code 3MVD; ref. 28) (c), centromeric protein CENP-C (green, PDB code 4INM; ref. 30) (d) and silencing protein Sir3 (brown, PDB

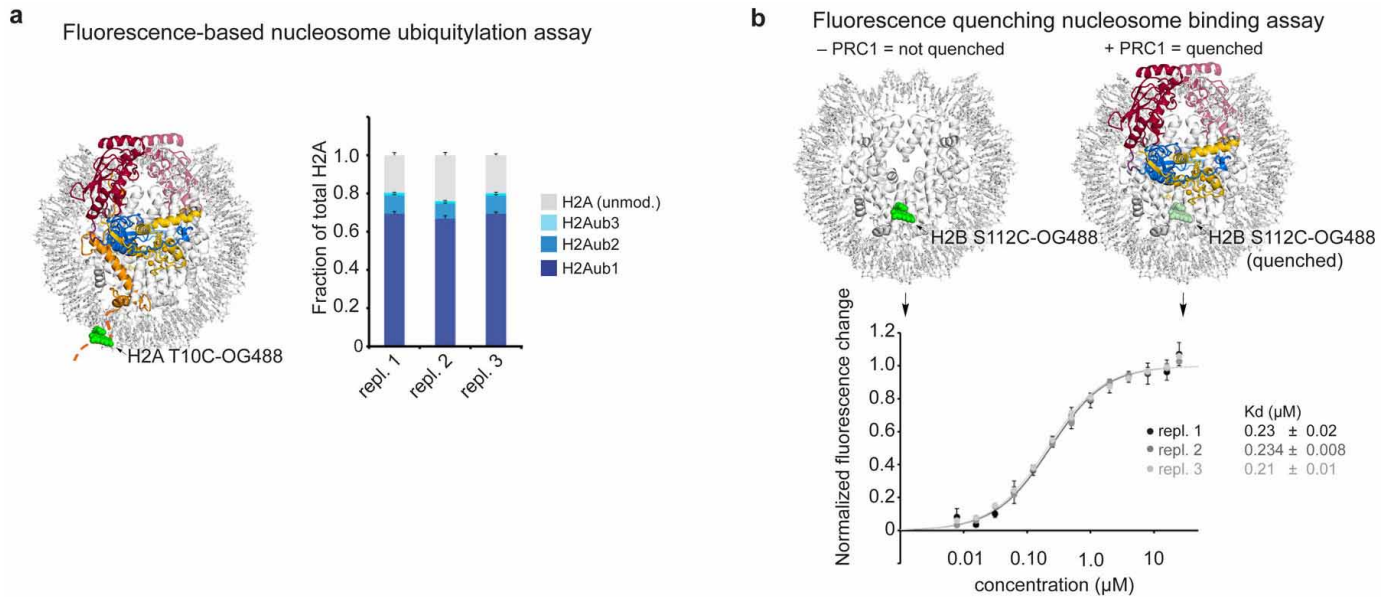
code 3TU4; ref. 29) (e) each bound to the acidic patch. Conserved arginine is shown in space-filling representation. Other arginines involved in binding are shown as sticks. UbcH5c and the Ring1B loop between residues 81 and 98 are not shown for figure clarity.



Extended Data Figure 5 | Ring1B–Bmi1- and BRCA1–BARD1-mediated ubiquitylation of nucleosomes containing histone mutations.

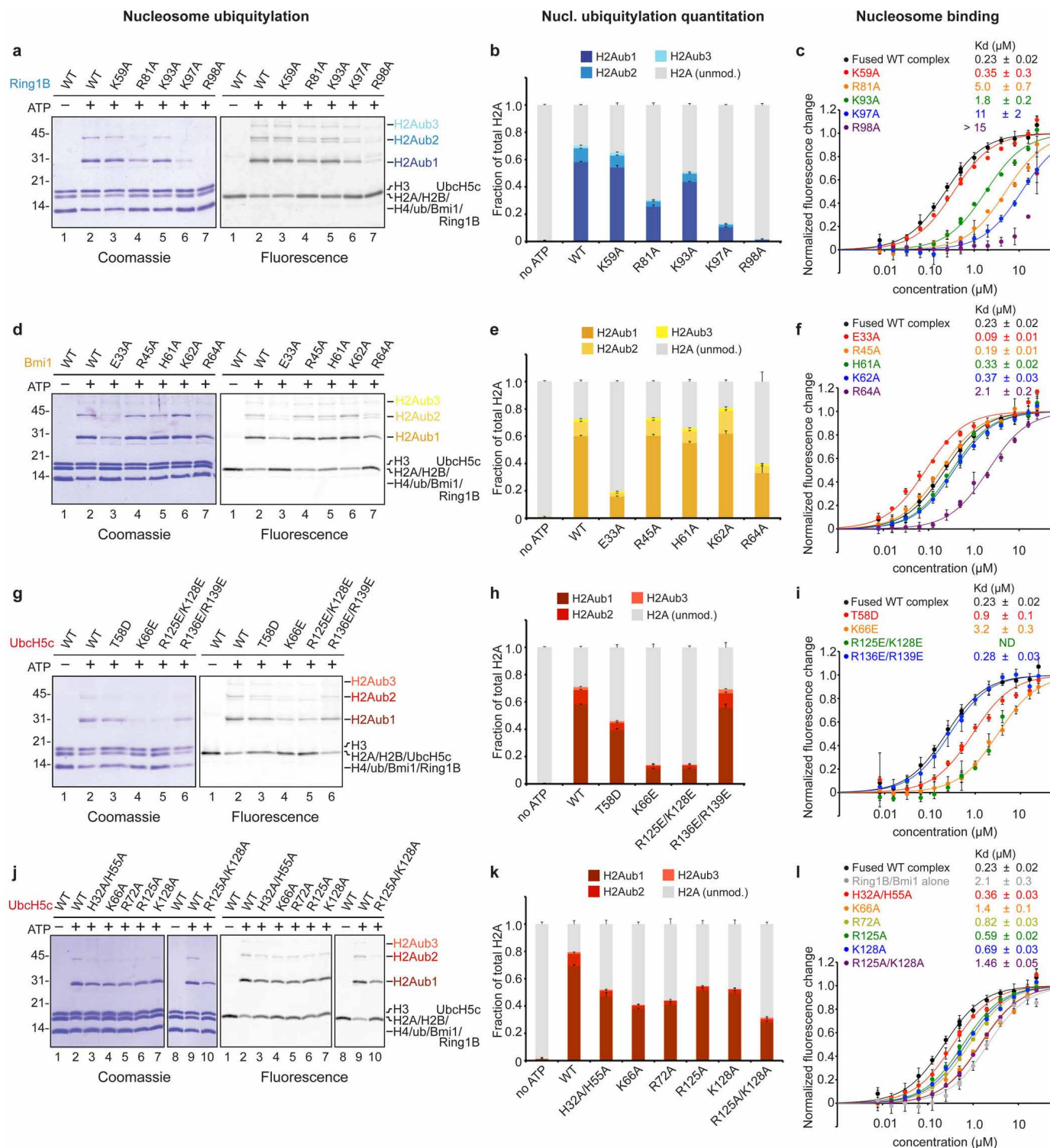
a, b, Coomassie-stained gels of ubiquitylation assays using nucleosomes with the specified histone mutants, E1 Uba1, E2 UbcH5c, STR-His₁₀ tagged

ubiquitin and Ring1B–Bmi1 (**a**) or BRCA1–BARD1 RING heterodimers (**b**). Tagged H2B (STR-His₆) was used to prevent electrophoretic comigration with unmodified H2A. Experiment was repeated twice.



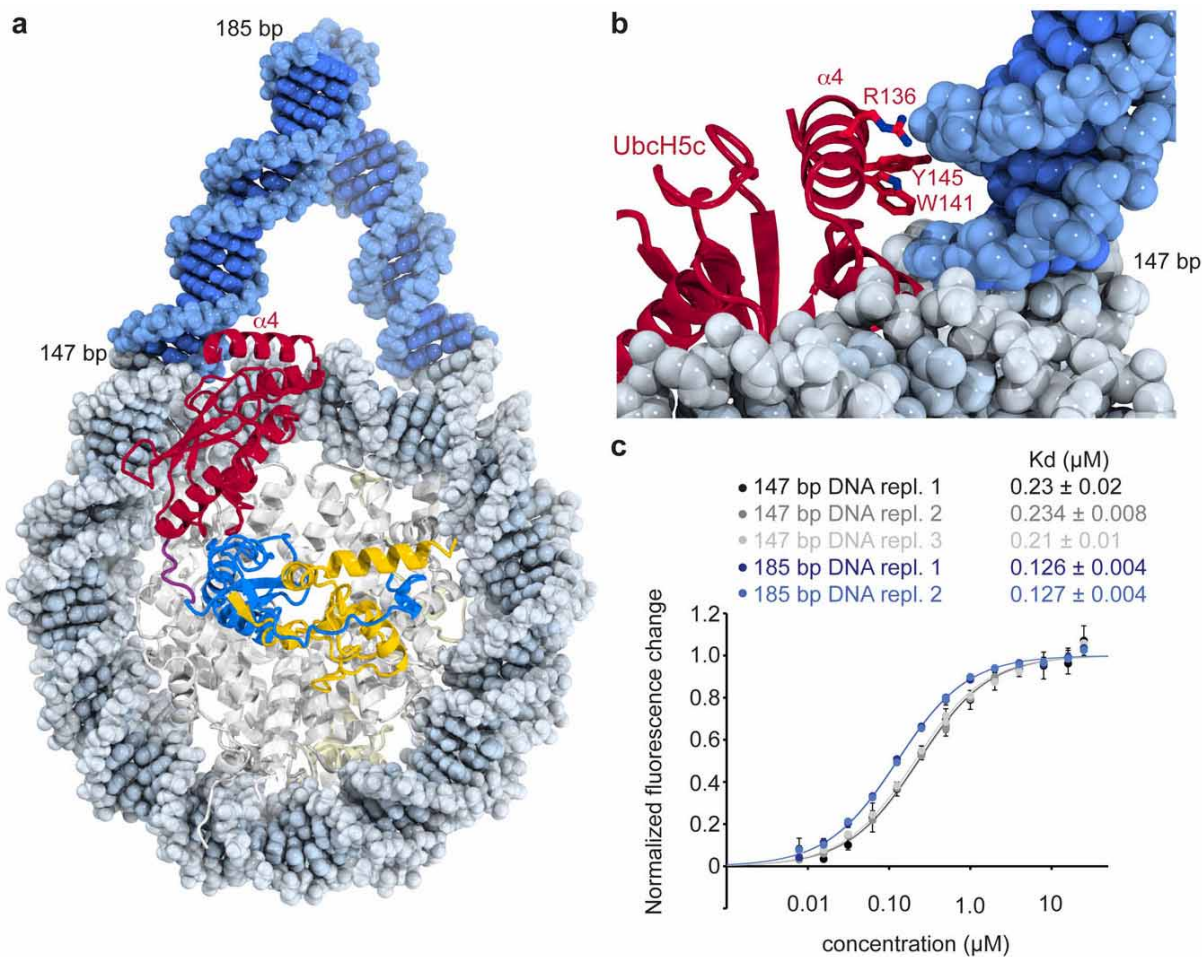
Extended Data Figure 6 | Fluorescence-based nucleosome ubiquitylation and binding assays. **a**, Fluorescence-based nucleosome ubiquitylation assay. Left, ubiquitylation assays performed with nucleosomes labelled with Oregon Green 488 maleimide on H2A(Thr10Cys) mutant (H2A T10C-OG488). Right, replicates 1, 2 and 3 of wild-type assay run on four different gels and quantified to demonstrate reproducibility across gels. Data are mean and s.d., $n = 4$. **b**, Fluorescence-based nucleosome binding assay. Top left, nucleosomes

labelled with Oregon Green 488 on H2B(Ser112Cys) mutant (H2B S112-OG488). Top right, binding of PRC1 ubiquitylation module leads to partial quenching of the fluorophore allowing affinity measurements to be made. Bottom, three technical replicates of fused wild-type PRC1 ubiquitylation module are shown to demonstrate reproducibility. Mean and s.d. shown for each data point, $n = 3$.



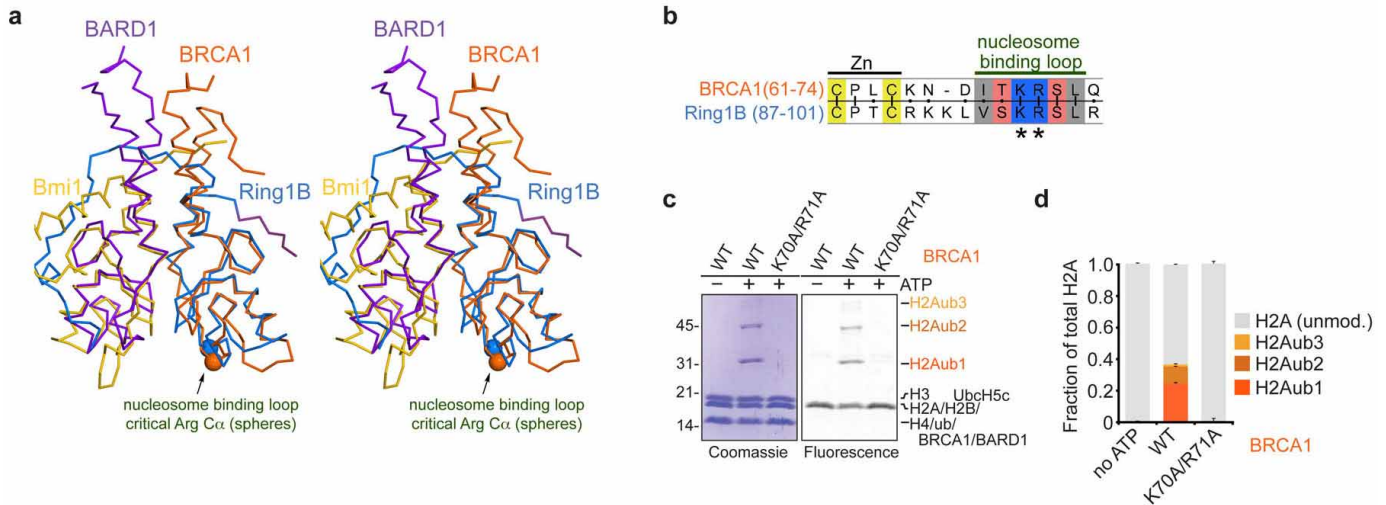
Extended Data Figure 7 | Effects of PRC1 ubiquitylation module mutants on nucleosome ubiquitylation and binding. **a**, Representative gel of one replicate of ubiquitylation assay using E1 Uba1, UbcH5c, STR-His₁₀ tagged ubiquitin, NCPs, and E3 Ring1B–Bmi1 with the indicated Ring1B mutants stained with Coomassie (left) and scanned for fluorescent H2A (right). NCPs in the experiment are doped with NCPs containing the H2A(Thr10Cys) mutant labelled with Oregon Green 488 maleimide. **b**, Quantitation of mono-ubiquitylated (H2Aub1, dark blue), di-ubiquitylated (H2Aub2, blue) and tri-ubiquitylated (H2Aub3, light blue) H2A as a fraction of total

H2A. Data are mean and s.d. from three technical replicates. Samples from the same experiment were analysed on different gels, processed in parallel. **c**, Fluorescence quenching binding curves for fused PRC1 ubiquitylation modules containing the indicated mutations of Ring1B (coloured as shown). Mean and s.d. are shown for each data point, *n* = 3. Fluorescence is normalized to fit values for unbound and saturated NCPs. Concentrations depicted using log scale. **d–l**, Experiments as described above for indicated Bmi1 mutants (**d–f**), UbcH5c charge reversal mutants (**g–i**) and UbcH5c alanine mutants (**j–l**). Triplicate ubiquitylation assays were repeated at least twice.



Extended Data Figure 9 | Linker DNA increases the affinity of the PRC1 ubiquitylation modules for the nucleosome. **a**, Linear B-form duplex DNA modelled onto the DNA ends of the PRC1 ubiquitylation module–NCP structure. **b**, The UbcH5c $\alpha 4$ helix occupies the major groove of modelled linker DNA. Several basic and aromatic side chains line the DNA-adjacent face of the $\alpha 4$ helix. **c**, Linker DNA enhances nucleosomal binding of fused PRC1 ubiquitylation module. Fluorescence quenching binding curves for the

wild-type fused PRC1 ubiquitylation module binding to nucleosomes centred on 147-bp (black, grey and light grey) or 185-bp (blue and light blue) 601 DNA. Technical replicates performed on different days are depicted in different shades of grey and blue. Mean and s.d. are shown for each data point, $n = 3$. Fluorescence is normalized to fit values for unbound and saturated NCPs. Concentrations depicted using log scale.



Extended Data Figure 10 | BRCA1 requires similar loop region and H2A–H2B acidic patch for nucleosome ubiquitylation. **a**, Alignment of the BRCA1–BARD1 and Ring1B–Bmi1 heterodimers using the RING domains of BRCA1 from the BRCA1–BARD1 NMR structure³⁵ (PDB code 1JM7) and Ring1B. BRCA1 and BARD1 are shown in orange and purple, respectively. C α atoms of essential arginine residues are indicated by spheres. **b**, Sequence alignment of Ring1B and BRCA1 shows conserved nucleosome interacting loop. **c**, BRCA1(Lys70Ala/Arg71Ala) mutation eliminates E3 ligase activity of BRCA1–BARD1 RING heterodimer. Representative gel of one replicate of

ubiquitylation assay using E1 Uba1, E2 UbcH5c, STR-His₁₀ ubiquitin, NCPs, and E3 BRCA1–BARD1 with the indicated mutants stained with Coomassie and scanned for fluorescent H2A. NCPs in the experiment are doped with NCPs containing the H2A(Thr10Cys) mutant labelled with Oregon Green 488 maleimide. **d**, Quantification of mono-ubiquitylated (H2Aub1, dark orange), di-ubiquitylated (H2Aub2, orange) and tri-ubiquitylated (H2Aub3, light orange) H2A are shown as a fraction of total H2A. Mean and s.d. from three technical replicates are depicted. Ubiquitylation assay was repeated twice.

Spatially resolved magnetic field structure in the disk of a T Tauri star

Ian W. Stephens^{1,2}, Leslie W. Looney², Woojin Kwon³, Manuel Fernández-López^{2,4}, A. Meredith Hughes⁵, Lee G. Mundy⁶, Richard M. Crutcher², Zhi-Yun Li⁷ & Ramprasad Rao⁸

Magnetic fields in accretion disks play a dominant part during the star formation process^{1,2} but have hitherto been observationally poorly constrained. Field strengths have been inferred on T Tauri stars³ and possibly in the innermost part of their accretion disks⁴, but the strength and morphology of the field in the bulk of a disk have not been observed. Spatially unresolved measurements of polarized emission (arising from elongated dust grains aligned perpendicularly to the field⁵) imply average fields aligned with the disks^{6,7}. Theoretically, the fields are expected to be largely toroidal, poloidal or a mixture of the two^{1,2,8–10}, which imply different mechanisms for transporting angular momentum in the disks of actively accreting young stars such as HL Tau (ref. 11). Here we report resolved measurements of the polarized 1.25-millimetre continuum emission from the disk of HL Tau. The magnetic field on a scale of 80 astronomical units is coincident with the major axis (about 210 astronomical units long¹²) of the disk. From this we conclude that the magnetic field inside the disk at this scale cannot be dominated by a vertical component, though a purely toroidal field also does not fit the data well. The unexpected morphology suggests that the role of the magnetic field in the accretion of a T Tauri star is more complex than our current theoretical understanding.

HL Tau is located 140 pc away¹³, in the Taurus molecular cloud. Although HL Tau is a T Tauri star, it is considered to be in an early stage of development owing to its bipolar outflow¹⁴ and possible residual envelope¹⁵. Observations and modelling of this protostar assuming a thick, flared disk suggest a stellar mass of ~ 0.55 times the solar mass (M_{\odot}) and a disk mass of $0.14M_{\odot}$ (ref. 12). A possible planet forming in the disk of HL Tau has been observed¹⁶, though this detection was not confirmed¹². However, the disk of HL Tau is gravitationally unstable, which could favour fragmentation into Jupiter-mass planets^{12,16}. HL Tau has the brightest disk of any T Tauri star at millimetre wavelengths, allowing observations of the fractional polarization (P) to have the best possible sensitivity. Previous observations of the polarization of the disk of HL Tau yielded marginally significant, spatially unresolved polarization detections with the James Clerk Maxwell Telescope (JCMT; polarization $P = 3.6 \pm 2.4\%$ at $14'' = 1,960$ astronomical units (AU) resolution)⁶ and the Submillimetre Array (SMA; polarization $P = 0.86 \pm 0.4\%$ at $2'' = 280$ AU resolution, archival observations released in this Letter). In addition, observations of HL Tau with the Combined Array for Millimeter-wave Astronomy (CARMA) have shown that the interferometric emission comes entirely from the disk with no contamination from large-scale envelope emission¹². HL Tau is therefore a very promising source to search for a spatially resolved polarization detection.

Only through observations of polarized dust emission can the morphology of the magnetic field be ascertained; however, higher-resolution dust polarimetric observations of T Tauri star disks do not detect polarization and place stringent upper limits ($P < 1\%$) on the polarization fraction^{17,18} that disagree with theoretical models of high efficiency grain

alignment with a purely toroidal field⁸. There is a clear discrepancy between theoretical models of the magnetic fields in disks and the observations to date, requiring more sensitive observations of the dust polarization. The SMA recently detected the magnetic field morphology in the circumstellar disk of the Class 0 (that is, the earliest protostellar stage) protostar IRAS 16293–2422 B¹⁹, but since this disk is nearly face-on, observations cannot detect the vertical component of the magnetic field. Moreover, this source is perhaps one of the youngest of the known Class 0 sources²⁰, increasing the chances that the polarized flux could be from the natal environment. Nevertheless, since the disk is the brightest component at the probed scale, polarization most probably comes from the disk, and the magnetic field morphology hints at toroidal wrapping¹⁹.

We obtained 1.25-mm CARMA polarimetric maps of HL Tau at $0.6''$ (84 AU) resolution and plot the magnetic field morphology in Fig. 1. This is a spatially resolved detection (with approximately three independent beams) of the magnetic field morphology in the circumstellar disk of a T Tauri star. The central magnetic field vector has a measured position angle (PA, measured anticlockwise from north) of $\theta_B = 143.6$

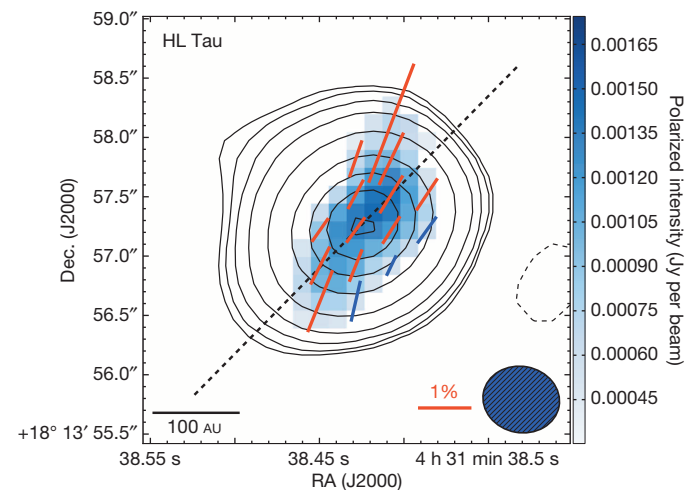


Figure 1 | Detected magnetic field morphology of HL Tau at $0.6''$ resolution. The polarization vectors (short coloured lines) have been rotated by 90° to show the inferred magnetic field orientation. Red vectors are detections $> 3\sigma_P$ while blue vectors are detections between $2\sigma_P$ and $3\sigma_P$, where σ_P is the r.m.s. noise of P . We do not show vectors when the signal-to-noise ratio for Stokes I is below 2. The sizes of the vector are proportional to the fractional polarization, P , with the red scale bar corresponding to $P = 1\%$. Stokes I contours are shown for $[-3, 3, 4, 6, 10, 20, 40, 60, 80, 100] \times \sigma_I$, where $\sigma_I = 2.1$ mJy per beam and is the r.m.s. noise of I . The dashed line shows the major axis of PA = 136° (ref. 12). The synthesized beam is shown at bottom right and has a size of $0.65'' \times 0.56''$ and PA = 79.5° . Dec., declination; RA, right ascension.

¹Institute for Astrophysical Research, Boston University, Boston, Massachusetts 02215, USA. ²Department of Astronomy, University of Illinois, Urbana, Illinois 61801, USA. ³SRON Netherlands Institute for Space Research, Landjeven 12, 9747 AD Groningen, The Netherlands. ⁴Instituto Argentino de Radioastronomía, CCT-La Plata (CONICET), CC5, 1894 Villa Elisa, Argentina. ⁵Van Vleck Observatory, Astronomy Department, Wesleyan University, Middletown, Connecticut 06459, USA. ⁶Astronomy Department and Laboratory for Millimeter-wave Astronomy, University of Maryland, College Park, Maryland 20742, USA. ⁷Astronomy Department, University of Virginia, Charlottesville, Virginia 22904, USA. ⁸Institute of Astronomy and Astrophysics, Academia Sinica, Hilo, Hawaii 96720, USA.

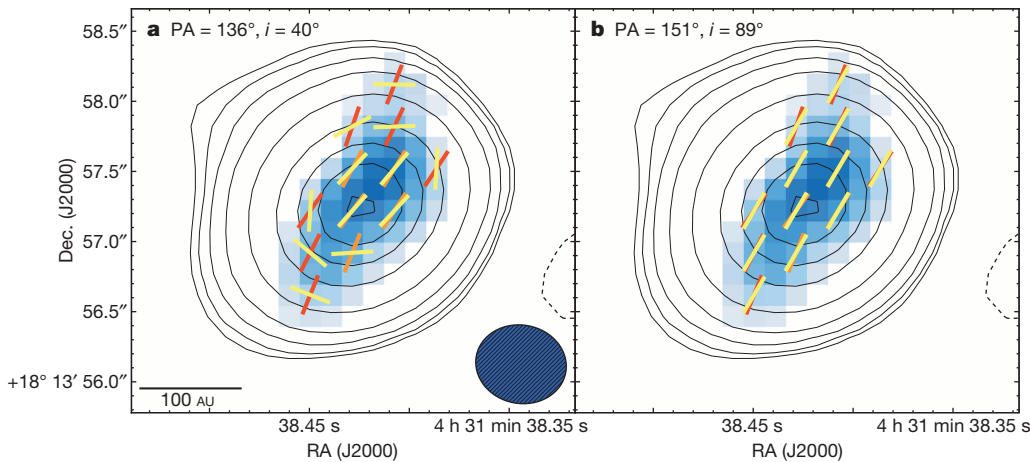


Figure 2 | Observed magnetic field morphology compared with models. **a**, 100% toroidal field model with PA = 136° and $i = 40^\circ$. **b**, As **a** but with PA = 151° and $i = 89^\circ$. The observed 3σ and 5σ detections are shown in red and orange, respectively, while modelled vectors are shown in yellow. Black contours, colour scale and beam size are the same as Fig. 1.

$\pm 4.4^\circ$, which is within 9° of the previously measured PA of the major axis of the HL Tau disk¹². This angle is in agreement with that measured by the JCMT ($\theta_B = 140 \pm 20^\circ$)⁶ and archival observations from the SMA analysed here ($\theta_B = 137 \pm 13^\circ$). The central vector has a fractional polarization of $P = 0.59 \pm 0.09\%$, and P varies over the disk between $0.54 \pm 0.13\%$ and $2.4 \pm 0.7\%$ with an average and median of 0.90% and 0.72% respectively, in agreement with the upper limits ($P < 1\%$) of other T Tauri star polarimetric observations^{17,18}. This median value is significantly less than in IRAS 16293–2422 B (1.4%), which could indicate disk evolution; the dust grains could be growing larger and becoming more spherical with time, and/or the magnetic field is becoming more turbulent.

To constrain the intrinsic magnetic field configuration inside the disk of HL Tau, we made a simple model which incorporates a combination of a toroidal and a vertical component; any radial field component inside the bulk of the disk (most probably probed by our polarization observation) is expected to be sheared quickly into a toroidal configuration by differential rotation on the short timescale of the disk rotation. We use the best estimates of disk parameters¹² (PA = 136° and inclination $i = 40^\circ$, both accurate to within a few degrees) and vary the relative strength of the two field components from 100% toroidal and 0% vertical to 0% toroidal and 100% vertical in 1% increments (see Methods for modelling details). Using the $\geq 3\sigma$ detections in Fig. 1, we find that a completely toroidal field is the best-fit model. However, the reduced χ^2 value is high (69, where a value of 1 indicates an optimal fit). If we do not constrain the magnetic field to be aligned with the best-fit disk¹² we can achieve a better fit. For a disk with a PA of 151°, the best-fit model to the observations is almost completely edge on, with an inclination of $i = 89^\circ$ (reduced χ^2 value of 1). Figure 2 shows the observed and the modelled parameters for a PA of 136° and $i = 40^\circ$ as well as a PA of 151° and $i = 89^\circ$. Although a PA of 151° is not very different from the major axis of the disk (15° difference), $i = 89^\circ$ is inconsistent with our disk dimensions and the well-constrained continuum observations of other studies^{12,21}. There is also no detected polarized flux in the northeast (upper-left) and southwest (bottom-right) of the disk where P for a toroidal field should be highest due to less beam smearing. This fact, along with the poor fit with the constrained disk parameters, indicates that although there is probably a toroidal component, the field apparently has substantial contributions from other components.

We note that adding a vertical component to all our models makes the fit worse. This lack of any vertical component in our model suggests that dominant poloidal fields are probably absent inside the disk because, as we argued earlier on dynamical grounds, a predominantly radial field is unlikely inside the disk because of rapid differential rotation. Moreover, although points with $>5\sigma$ polarization detections fit a toroidal morphology even for the $i = 40^\circ$ model, we cannot be sure that the disk field is only toroidal; this is because a straight, uniform field in the plane of the disk, although physically unlikely for a disk system, also fits the data well.

If the disk of HL Tau has a dominant toroidal component, then it is uncertain what causes the outer vectors to not fit the toroidal morphology based on the best-fit disk dust model¹² (Fig. 2a). Perhaps the grains in some parts of the disk do not efficiently align with the disk field. Another possibility is that the magnetic field towards the inner disk is toroidal and beam-averaged to be aligned with the major axis of the disk. However, towards the edge of the disk, where the disk field may be less tightly wound and weaker, the magnetic field may be dominated by external field lines that are already toroidal (for example, due to a rotating envelope). The incoming fields may strongly influence or even dominate the magnetic field in the outer disk. In any case, the discrepancy indicates that, at least for this particular source, one needs to go beyond the simplest magnetorotational instability² disk models that do not include any external influence or distortion. Both theoretical studies tailored to HL Tau and higher-resolution polarization observations are needed to resolve this puzzle.

At the 1,000 AU scale, structured magnetic fields are observed around young, low-mass protostars^{22,23}, but the fields appear to be randomly aligned with respect to the inferred disk^{24,25}. Misalignment of the field lines with the rotation axis can help overcome magnetic braking to create a centrifugally supported disk at the 100 AU scale⁰. Further disk evolution can be driven by magnetorotational instability² or a magnetocentrifugal wind¹⁹; a toroidally dominant disk field is expected in the former scenario, and a significant poloidal field is required for the latter (at least near the wind launching surface). Both processes can possibly contribute to the disk accretion at the same time.

Until now, we have been unable to observationally constrain the magnetic field morphology in disks. Along with the Class 0 source IRAS 16293–2422 B, our observations of HL Tau show that a toroidal field component may last from the low-mass protostar's initial formation to the T Tauri star stage—approximately the first 10^6 years of a protostar's life²⁶. The apparent absence of vertical fields in these observations implies that magnetocentrifugal winds driven along large-scale poloidal magnetic fields²⁷ are probably not the dominant mechanism for redistributing the disk's angular momentum during the accretion process of a star at the 80 AU size-scale of our observations. However, the morphology detected in HL Tau also cannot be fully explained by a simple mix of toroidal and vertical components, requiring future observations at both large scale and small scale to truly understand the role of magnetic fields in the formation of solar systems like our own.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 11 April; accepted 9 September 2014.

Published online 22 October 2014.

1. Blandford, R. D. & Payne, D. G. Hydromagnetic flows from accretion discs and the production of radio jets. *Mon. Not. R. Astron. Soc.* **199**, 883–903 (1982).

2. Balbus, S. A. & Hawley, J. F. Instability, turbulence, and enhanced transport in accretion disks. *Rev. Mod. Phys.* **70**, 1–53 (1998).
3. Johns-Krull, C. M. The magnetic fields of classical T Tauri stars. *Astrophys. J.* **664**, 975–985 (2007).
4. Donati, J.-F., Paletou, F., Bouvier, J. & Ferreira, J. Direct detection of a magnetic field in the innermost regions of an accretion disk. *Nature* **438**, 466–469 (2005).
5. Lazarian, A. Tracing magnetic fields with aligned grains. *J. Quant. Spectrosc. Radiat. Transf.* **106**, 225–256 (2007).
6. Tamura, M., Hough, J. H. & Hayashi, S. S. 1 millimeter polarimetry of young stellar objects: low-mass protostars and T Tauri stars. *Astrophys. J.* **448**, 346–355 (1995).
7. Tamura, M. *et al.* First detection of submillimeter polarization from T Tauri stars. *Astrophys. J.* **525**, 832–836 (1999).
8. Cho, J. & Lazarian, A. Grain alignment and polarized emission from magnetized T Tauri disks. *Astrophys. J.* **669**, 1085–1097 (2007).
9. Königl, A. & Pudritz, R. E. in *Protostars and Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S. S.) 759–787 (Univ. Arizona Press, 2000).
10. Hennebelle, P. & Ciardi, A. Disk formation during collapse of magnetized protostellar cores. *Astron. Astrophys.* **506**, L29–L32 (2009).
11. Robitaille, T. P., Whitney, B. A., Indebetouw, R. & Wood, K. Interpreting spectral energy distributions from young stellar objects. II. Fitting observed SEDs using a large grid of precomputed models. *Astrophys. J. Suppl. Ser.* **169**, 328–352 (2007).
12. Kwon, W., Looney, L. W. & Mundy, L. G. Resolving the circumstellar disk of HL Tauri at millimeter wavelengths. *Astrophys. J.* **741**, 3 (2011).
13. Rebull, L. M., Wolff, S. C. & Strom, S. E. Stellar rotation in young clusters: the first 4 million years. *Astron. J.* **127**, 1029–1051 (2004).
14. Movsessian, T. A., Magakian, T. Y. & Moiseev, A. V. Kinematics and the origin of the internal structures in HL Tauri jet (HH 151). *Astron. Astrophys.* **541**, A16 (2012).
15. Welch, W. J., Hartmann, L., Helfer, T. & Briceño, C. High-resolution, wide-field imaging of the HL Tauri environment in ^{13}CO (1–0). *Astrophys. J.* **540**, 362–371 (2000).
16. Greaves, J. S., Richards, A. M. S., Rice, W. K. M. & Muxlow, T. W. B. Enhanced dust emission in the HL Tau disc: a low-mass companion in formation? *Mon. Not. R. Astron. Soc.* **391**, L74–L78 (2008).
17. Hughes, A. M. *et al.* A spatially resolved inner hole in the disk around GM Aurigae. *Astrophys. J.* **698**, 131–142 (2009).
18. Hughes, A. M., Hull, C. L. H., Wilner, D. J. & Plambeck, R. L. Interferometric upper limits on millimeter polarization of the disks around DG Tau, GM Aur, and MWC 480. *Astron. J.* **145**, 115 (2013).
19. Rao, R., Girart, J. M., Lai, S.-P. & Marrone, D. P. Detection of a magnetized disk around a very young protostar. *Astrophys. J.* **780**, L6 (2014).
20. Loinard, L. *et al.* ALMA and VLA observations of the outflows in IRAS 16293–2422. *Mon. Not. R. Astron. Soc.* **430**, L10–L14 (2013).
21. Lay, O. P., Carlstrom, J. E. & Hills, R. E. Constraints on the HL Tauri protostellar disk from millimeter- and submillimeter-wave interferometry. *Astrophys. J.* **489**, 917–927 (1997).
22. Girart, J. M., Rao, R. & Marrone, D. P. Magnetic fields in the formation of sun-like stars. *Science* **313**, 812–814 (2006).
23. Stephens, I. W. *et al.* The magnetic field morphology of the class 0 protostar L1157-mm. *Astrophys. J.* **769**, L15 (2013).
24. Hull, C. L. H. *et al.* Misalignment of magnetic fields and outflows in protostellar cores. *Astrophys. J.* **768**, 159 (2013).
25. Hull, C. L. H. *et al.* TADPOL: A 1.3 mm survey of dust polarization in star-forming cores and regions. *Astrophys. J. Suppl. Ser.* **213**, 13 (2014).
26. Evans, N. J. II *et al.* The Spitzer c2d legacy results: star-formation rates and efficiencies; evolution and lifetimes. *Astrophys. J. Suppl. Ser.* **181**, 321–350 (2009).
27. Wardle, M. & Koenigl, A. The structure of protostellar accretion disks and the origin of bipolar flows. *Astrophys. J.* **410**, 218–238 (1993).

Acknowledgements We thank R. L. Plambeck and C. L. H. Hull for consultation during the data reduction process and C. F. Gammie for discussions. This research made use of APLpy, an open-source plotting package for Python hosted at <http://aplpy.github.com>. Work at the Universities of Illinois and Maryland was supported by NSF AST-1139950 and AST-1139998, respectively. Support for CARMA construction was derived from the states of California, Illinois and Maryland, the James S. McDonnell Foundation, the Gordon and Betty Moore Foundation, the Kenneth T. and Eileen L. Norris Foundation, the University of Chicago, the Associates of the California Institute of Technology, and the National Science Foundation. Ongoing CARMA development and operations are supported under a cooperative agreement (NSF AST 08-38226) and by the CARMA partner universities.

Author Contributions Data acquisition and reduction were performed by I.W.S., L.W.L. and M.F.-L. Polarization modelling was performed by W.K. and fitted by I.W.S. All authors analysed and discussed the observations and manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to I.W.S. (ianws@bu.edu).

METHODS

Data reduction. The CARMA dual polarization receivers allow for the measurement of polarized dust emission. Polarimetric maps can thus be created, which provide the flux density (Stokes I), the PA of the dust polarization, and P at every point within the map. CARMA continuum observations in Full Stokes mode were taken at 237 GHz in two different resolutions with each resolution having multiple tracks. The C-array observations ($\sim 0.79''$ resolution) consisted of four tracks in October and November 2013 while the B-array observations ($\sim 0.37''$ resolution) consisted of three more tracks in November and December 2013.

We have reduced the CARMA observations using the MIRIAD package²⁸. The dual-polarization receivers of CARMA measure right- (R) and left- (L) circular polarization and the four cross-polarization (RR, LL, LR, RL) terms. In order to calibrate CARMA Full Stokes observations, apart from the usual interferometric calibrations (bandpass, phase and flux corrections), two additional calibrations are required: XYPhase (due to L and R channel phase differences) and leakage (due to L and R channels cross-coupling). These calibrations were done in the typical manner for CARMA²⁵. The leakage terms for each antenna were consistent from track to track, and the overall accuracy of the leakage calibrations are expected to be approximately 0.1%. Day to day consistency in polarization observations was tested by measuring the phase calibrator, 0510+180. The polarization angle changed slightly from day to day, with a steady increase of a few degrees with each newer track. Intraday variability is a well-known phenomenon which affects the total flux density, the linearly polarized flux density, and the polarization angle²⁹ and probably explains the variations of a few degrees seen in 0510+180 from track to track. Since the variation of the polarization in 0510+180 was not very large, the consistency of polarization measurements of 0510+180 between tracks made us confident that our calibration is accurate. For B-array tracks, we also saw that the bandpass calibrator, 3C454.3, showed consistency for polarization measurements for all the tracks. Other calibrators observed did not have polarization detected, signifying that our polarization detection of HL Tau is not a spurious detection. We also note that there was a slight difference in the polarization angle between the lower and upper sideband; this difference may be due to Faraday rotation and was almost constant for all tracks.

For bandpass calibration, C-array observations used 3C84 and B-array observations used 3C454.3. The phase calibrator used for both arrays was 0510+180. Observations of MWC349, with an adopted flux of 2.1 Jy (ref. 30), provided the absolute scale for the flux calibration at 237 GHz in most of the tracks. The bootstrapped flux of 0510+180 using MWC349 was consistent within 10% and 15% with bootstrapped fluxes using Mars and Uranus, respectively, in other tracks. Since planets are resolved at these resolutions, MWC349 is likely to have a more accurate flux calibration and was bootstrapped for all tracks. The absolute flux uncertainty is estimated to be 15%, but only statistical uncertainties are discussed in this work. When imaging, natural weighting was used to maximize the sensitivity.

Detections of polarization were found in every HL Tau track with consistent polarization angles and measurements. Since polarization is calculated from Stokes Q and U and can only have positive values, there exists a bias in the polarized intensity; hence, all our polarization measurements have been de-biased²⁵. The sensitivity in Stokes I is limited by dynamic range rather than the flux sensitivity of the observations. The uncertainty in the absolute position angle of CARMA is approximately 3° (ref. 25).

These high resolution interferometric observations are insensitive to large scale structure. Single dish and interferometric HL Tau observations find very similar compact fluxes, suggesting that the envelope dust continuum is negligible²¹, and the flux appears to come entirely from the disk¹².

Also reported in this study are unpublished polarization observations (PA and P) of HL Tau from the SMA. These observations were taken in the compact configuration

in October 2005 in the 345 GHz atmospheric window with the Local Oscillator tuned to 341.5 GHz. The polarization data reduction process was done in the typical manner employed by the SMA³¹.

Linear polarization modelling. We employed a flared viscous accretion disk model that was constrained by high angular resolution data and a broad spectral energy distribution, detailed in another paper¹². The accretion disk model has a power-law density distribution with an exponential tapering, and the vertical density distribution is assumed to be 1.5 times thicker than the hydrostatic equilibrium case. The temperature distributions are calculated by interpolation of two power-law distributions at the cold mid-plane, comparable to the results of the Monte-Carlo radiative code RADMC-3D³² and at the surface based on radiation equilibrium: $T_m[K] = 190(r/\text{AU})^{-0.43}$ and $T_s[K] = 600(r/\text{AU})^{-0.43}$, where T_m and T_s are the temperatures of the disk at the mid-plane and surface, respectively. The disk parameters employed were the volume density power-law index, taken as 1.064, the dust opacity spectral index $\beta = 0.729$, the disk mass $M_d = 0.1349M_\odot$, the inner radius $R_{in} = 2.4$ AU, and the characteristic radius $R_c = 78.9$ AU. As described in the main text, we investigated various cases of different inclination and position angles.

Our polarization modelling consists of toroidal and vertical magnetic fields. Instead of constraining the detailed morphology of magnetic fields, we intended to constrain which morphology is preferred. In order to achieve this goal, we examined 101 cases spanning relative polarization fractions of the two orthogonal fields in steps of 1% (that is, 100% toroidal, 99% toroidal and 1% vertical, 98% toroidal and 2% vertical, ...). For constructing linear polarization information, we built Stokes I , Q and U maps by numerically solving radiative transfer (necessary for a thick disk). In individual integral elements of radiative transfer along the line of sight, we compute the intensity for Q and U . The fractional intensities added up to the Q and U maps by an integral element are:

$$\Delta Q = \Delta I f_p (f_{\text{tor}} q_{\text{tor}} + f_{\text{ver}} q_{\text{ver}}) \quad (1)$$

$$\Delta U = \Delta I f_p (f_{\text{tor}} u_{\text{tor}} + f_{\text{ver}} u_{\text{ver}}) \quad (2)$$

where f_p is a total polarization fraction ($\sqrt{Q^2 + U^2}/I$), f_{tor} and f_{ver} are relative fractions of the toroidal and vertical fields (for example, $f_{\text{tor}} = 0.7$ and $f_{\text{ver}} = 0.3$ for 70% toroidal and 30% vertical fields), and q_{tor} and u_{tor} are $\cos(2\chi_{\text{tor}})$ and $\sin(2\chi_{\text{tor}})$ respectively. χ_{tor} is the angle, χ , of the toroidal magnetic field measured anticlockwise from the north. Similarly, q_{ver} and u_{ver} for vertical fields are $\cos(2\chi_{\text{ver}})$ and $\sin(2\chi_{\text{ver}})$, where χ_{ver} is the angle of the vertical magnetic field. Since the disk is optically thin and we are only concerned with the morphology, f_p can be given an arbitrary value (for example, 0.01 or 0.1). Note that the toroidal and vertical magnetic field vectors at each integral element have been tilted and rotated based on the inclination and the position angle of the model disk, before the calculation of the fractions. Q and U maps are convolved with the synthesized beam from the polarization observations, and the modelled position angles of the magnetic field morphology are created using $\chi = 0.5 \tan^{-1}(U/Q)$.

28. Sault, R. J., Teuben, P. J. & Wright, M. C. H. in *Astronomical Data Analysis Software and Systems IV* (eds Shaw, R. A., Payne, H. E. & Hayes, J. J. E.) 433–436 (Astron. Soc. Pacif. Conf. Ser. Vol. 77, 1995).
29. Witzel, A., Heeschen, D. S., Schalinski, C. & Krichbaum, T. Kurzzeit-Variabilität extragalaktischer Radioquellen. *Mitt. Astron. Ges. Hamburg* **65**, 239 (1986).
30. Tafaya, D., Gómez, Y. & Rodríguez, L. F. Imaging MWC 349 from 7 millimeters to 90 centimeters. *Astrophys. J.* **610**, 827–834 (2004).
31. Marrone, D. P. & Rao, R. The submillimeter array polarimeter. *Proc. SPIE* **7020**, 70202B (2008).
32. Dullemond, C. P. & Dominik, C. Flaring vs. self-shadowed disks: the SEDs of Herbig Ae/Be stars. *Astron. Astrophys.* **417**, 159–168 (2004).

Possible planet formation in the young, low-mass, multiple stellar system GG Tau A

Anne Dutrey^{1,2}, Emmanuel Di Folco^{1,2}, Stéphane Guilloteau^{1,2}, Yann Boehler³, Jeff Bary⁴, Tracy Beck⁵, Hervé Beust⁶, Edwige Chapillon^{1,7}, Frédéric Gueth⁷, Jean-Marc Huré^{1,2}, Arnaud Pierens^{1,2}, Vincent Piétu⁷, Michal Simon⁸ & Ya-Wen Tang⁹

The formation of planets around binary stars may be more difficult than around single stars^{1–3}. In a close binary star (with a separation of less than a hundred astronomical units), theory predicts the presence of circumstellar disks around each star, and an outer circumbinary disk surrounding a gravitationally cleared inner cavity around the stars^{4,5}. Given that the inner disks are depleted by accretion onto the stars on timescales of a few thousand years, any replenishing material must be transferred from the outer reservoir to fuel planet formation (which occurs on timescales of about one million years). Gas flowing through disk cavities has been detected in single star systems⁶. A circumbinary disk was discovered around the young low-mass binary system GG Tau A (ref. 7), which has recently been shown to be a hierarchical triple system⁸. It has one large inner disk⁹ around the single star, GG Tau Aa, and shows small amounts of shocked hydrogen gas residing within the central cavity¹⁰, but other than a single weak detection¹¹, the distribution of cold gas in this cavity or in any other binary or multiple star system has not hitherto been determined. Here we report imaging of gas fragments emitting radiation characteristic of carbon monoxide within the GG Tau A cavity. From the kinematics we conclude that the flow appears capable of sustaining the inner disk (around GG Tau Aa) beyond the accretion lifetime, leaving time for planet formation to occur there. These results show the complexity of planet formation around multiple stars and confirm the general picture predicted by numerical simulations.

The triple stellar system GG Tau A consists of a single star GG Tau Aa and a close binary GG Tau Ab (with individual stars named Ab1 and Ab2). The system is 1–5 million years old^{12,13}, and is located 140 pc from Earth in a hole of the Taurus molecular cloud. Its molecular emission is free of contamination¹⁴ and there is no known outflow or jet associated with the source. GG Tau Aa and Ab have an apparent separation of 35 astronomical units (AU) while the separation of GG Tau Ab1 and Ab2 is 4.5 AU (ref. 8). The outer Keplerian disk of gas and dust surrounding GG Tau A (called the circumbinary disk for simplicity) consists of a ring extending from radius $r \approx 190$ AU to $r \approx 280$ AU and an outer disk extending up to 800 AU from the central stars with a total mass of $\sim 0.15 M_{\odot}$ (ref. 14; here M_{\odot} is the solar mass).

Using the Atacama Large Millimetre Array (ALMA), we observed GG Tau A in the dust thermal emission at 0.45 mm wavelength and in the CO $J = 6-5$ line (Fig. 1a–c) with an angular resolution of $\theta \approx 0.25''$ or ~ 35 AU. The continuum image shows cold dust emission from only one circumstellar disk-like structure associated with GG Tau Aa^{9,15}. We estimate the minimum dust disk size to be ~ 7 AU while the minimum mass of gas and dust is roughly $10^{-3} M_{\odot}$, about Jupiter's mass. The complex CO $J = 6-5$ spectral line shape at the location of GG Tau Aa also reveals the existence of a CO circumstellar disk of outer radius ~ 20 AU (Methods and Extended Data Fig. 1). We do not detect cold dust emission around GG Tau Ab, even though the existence of an inner dust disk (or disks) has been inferred from unresolved infrared emission¹⁶.

Our 0.45-mm flux upper limits (Methods) are compatible with tidal truncation, which would prevent any circumstellar disk extending beyond about 2 AU (ref. 8). The ALMA CO $J = 6-5$ image (Fig. 1a–c and Extended Data Figs 1 and 2) also clearly resolves CO gas within the central cavity with a structure indicative of the streamer-like features which have been hinted at by hydrodynamic simulations in binary systems^{5,17}. The CO gas appears to be inhomogeneous, existing as a series of fragments, and the structure is dominated by an east–west extension. No northern feature is seen, contrary to the very low level (a signal-to-noise ratio of 2) extension seen in continuum at 270 GHz (ref. 15).

The Institut de RadioAstronomie Millimétrique (IRAM) image (Fig. 1d–f and Extended Data Fig. 2) reveals, at lower angular resolution, that the CO $J = 2-1$ emission peaks are located near the inner edge of the outer ring (at radius $\sim 100-150$ AU). In contrast, the CO $J = 6-5$ emission peaks near GG Tau Aa and GG Tau Ab, close to the bright regions of near-infrared $\nu = 1-0$ S(1) H₂ emission (Fig. 1c) that have been interpreted as shock-excited gas at the interface between the streamer and gas associated with the inner disks¹⁰. A study of the excitation conditions (Methods) reveals that CO $J = 2-1$ and $J = 6-5$ emissions arise in different physical conditions: the former correspond to extended, cold (~ 35 K) optically thick areas, while the latter trace optically thin, warmer gas (~ 70 K) particularly at the interface between the streamer and the inner disk of Aa. The mass of each CO $J = 6-5$ fragment is about $5 \times 10^{-5} M_{\odot}$ (Methods). With a minimum accretion rate of $10^{-8} M_{\odot} \text{ yr}^{-1}$, a fragment reaching the Aa disk may disappear in at most 5,000 years (a few tens of times the orbital period of the binary Aa–Ab). The Aa disk mass currently represents about 20 such fragments; in 1 Myr, at least 200 fragments of similar mass must have been accreted to sustain such a disk. The corresponding minimum mass, which accreted from the circumbinary disk, represents about 10% of the current outer disk mass ($\sim 0.15 M_{\odot}$). The morphology of the gas in the CO $J = 6-5$ and $J = 2-1$ maps reveals departures from symmetry, unlike hydrodynamical simulations which predict symmetric streamers for an equal-mass, low-eccentricity binary system⁵. GG Tau Aa and Ab each have a mass of about $0.65 M_{\odot}$ (ref. 13) and their orbital eccentricity is constrained to $e \leq 0.35$ (ref. 18). The origin of this asymmetry might either be found in this eccentricity^{19,20} or in the triple nature of GG Tau A, as the binarity of Ab breaks the symmetry.

The change in velocity of the CO $J = 6-5$ emission along the major axis of the dust ring is similar to that of CO $J = 2-1$ (Fig. 1b, e and Extended Data Figs 1 and 2) and the velocity gradient is that of a rotating disk^{14,21}. At radius 200 AU, we find that the velocity of the CO $J = 6-5$ gas agrees with the known Keplerian speed (Methods and Extended Data Table 1) derived from existing ¹³CO maps^{14,21} and corresponds to the canonical dynamical mass of the triple star GG Tau A ($1.28 M_{\odot}$)^{14,21}. This is still true down to a radius of $\sim 70-80$ AU. Closer to the stars, the velocity pattern of the CO $J = 6-5$ emission becomes dominated by the individual gravitational fields of GG Tau Aa and GG Tau Ab. Limited

¹Université de Bordeaux, LAB, UMR 5804, F-33270 Floirac, France. ²Centre National de la Recherche Scientifique (CNRS), LAB, UMR 5804, F-33270 Floirac, France. ³Centro de Radioastronomía y Astrofísica (CRA), University of Mexico, Apartado Postal 3-72, 58089 Morelia, Michoacan, Mexico. ⁴Department of Physics and Astronomy, Colgate University, 13 Oak Drive, Hamilton, New York 13346, USA. ⁵Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. ⁶Institut de Planetologie et d'Astrophysique de Grenoble (IPAG), UMR 5274, BP 53, F-38041 Grenoble Cedex 9, France. ⁷Institut de RadioAstronomie Millimétrique (IRAM), 300 rue de la Piscine, F-38046 Saint Martin d'Hères, France. ⁸Stony Brook University, Stony Brook, New York 11794-3800, USA. ⁹Academia Sinica Institute of Astronomy and Astrophysics, PO Box 23-141, Taipei 106, Taiwan.

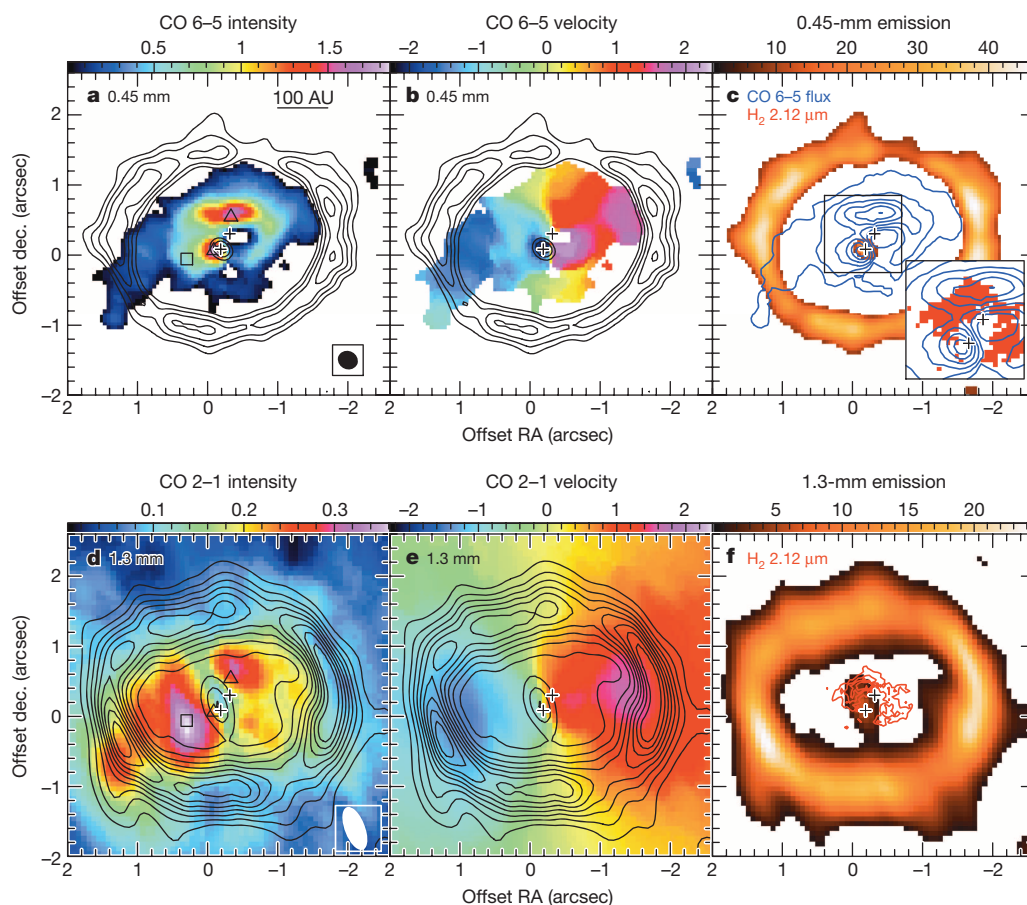


Figure 1 | ALMA and IRAM images of GG Tau A. **a–c,** ALMA; **d–f,** IRAM. **a,** 0.45-mm emission (black contours) and CO 6–5 flux (colour: see colour scale at top). **b,** 0.45-mm emission (black contours) and CO velocity field (colour). **c,** 0.45-mm emission (colour) with CO 6–5 flux (blue contours) and in inset H₂ intensity (red). **d, e,** as **a, b,** but for 1.3-mm emission (contours) and CO 2–1 flux (colour). **f,** 1.3-mm emission (colour) and H₂ intensity (contours).

Positions are relative to right ascension (RA) 04 h 32 min 30.359 s and declination (dec.) 17° 31' 40.38" (J2000). Crosses are the locations of Aa (south) and Ab (north) components, and triangles and squares show the locations of the CO $J = 6-5$ and $J = 2-1$ peaks, respectively. Units on the colour scales are Jy per beam km s⁻¹ (**a, d**), km s⁻¹ (**b, e**) and mJy per beam (**c, f**). In **a** and **d**, the beam size is given in the inset.

angular resolution (35 AU) precludes detailed kinematical analysis, but study of the northern CO $J = 6-5$ peak vicinity reveals that its velocity gradient is dominated by rotation, with departure from rotation compatible with infall motion (Methods and Extended Data Fig. 3).

Another result concerns the dust and gas circumbinary reservoir. The dust emission from the circumbinary ring is very well resolved (Fig. 1 and Extended Data Fig. 4) but appears uniform within the noise. The combination of the 0.45-mm ALMA image with existing 1.3- and 3.4-mm continuum maps^{14,21} allows for a direct measurement of the radial dependence of the temperature in the circumbinary dust disk. Using a radiative transfer model (Methods), we determined the dust temperature profile $T_D(r) = 13.8 \times (r/200 \text{ AU})^{-1.1} \text{ K}$, which gives a temperature of $\sim 8.5 \text{ K}$ at 300 AU. Thanks to the 0.45-mm image, the derived dust temperature is almost independent of the dust properties. This represents a significant improvement on the knowledge of the mid-plane physical conditions (Methods, Extended Data Table 2 and Extended Data Fig. 4). Since our measurement is the result of vertical averaging of the temperature, and the upper layers of the disk are hotter owing to direct irradiation by the stellar ultraviolet flux, the mid-plane must be even colder. The limited grain growth (Methods) helps explain the low temperature, since small grains are most efficient at attenuating ultraviolet radiation coming from the central stars. Furthermore, at the ring inner edge, a puffed-up rim, caused by direct stellar heating, should cast a shadow on the outer region of the disk. Such a shielding effect could explain the steep slope of the temperature profile. Previous ¹³CO measurements and analysis yielded a similarly steep slope (exponent ~ 0.9) for the gas temperature profile in the circumbinary disk but a warmer

gas temperature ($\sim 20 \text{ K}$, at 300 AU)²¹. This is consistent with CO molecules being trapped onto grains in the mid-plane, whose temperature is below the CO freeze-out point ($\sim 17 \text{ K}$), and with CO gas only present in the heated upper layers of the disk²².

Figure 1 shows that, unlike the continuum emission of dust, little CO $J = 6-5$ line emission is detected near the circumbinary disk inner radius at $\sim 200-250 \text{ AU}$. This could be attributable to both an H₂ density insufficient to thermalize the upper energy level of the CO transition and lower sensitivity of the interferometer to relatively large angular scale ($>1''$) structure (Methods). Figure 1 also shows strong excess of CO $J = 6-5$ and 2–1 emission at position angle $\sim 120^\circ$ (measured east of north) and radius $\sim 240-270 \text{ AU}$. From these CO data alone we cannot discriminate between a temperature and/or density increase (Methods). We suggest that this localized emission is attributable to a progenitor of a sub-stellar body reminiscent of the planetary-mass companion imaged around the low-mass binary 2MASS 0103AB²³. Such a companion would provide a natural explanation for the confinement of $\sim 80\%$ of the circumbinary mass within the $\sim 90 \text{ AU}$ breadth of the ring²¹. Simulations of the evolution of binary systems in which Kepler detected exoplanets demonstrate that Saturn-mass proto-planets can remain at the ring outer edge in some cases²⁰. We do not, however, detect a gap near the outer edge of the gas and dust circumbinary ring but existing ¹²CO, ¹³CO and thermal dust images are not incompatible with its existence (Methods).

Our observations demonstrate that active replenishment from the outer disk can sustain the circumstellar disk surrounding GG Tau Aa for a time exceeding the accretion lifetime, increasing its potential for planet formation. The presence of a condensation at the outer edge of

the ring is puzzling, and needs further investigation to determine its links with accretion processes and possible planet formation. Since almost half of Sun-like stars were born in multiple systems²⁴, our observations provide a step towards understanding the true complexity of protoplanetary disks in multiple stellar systems and unveiling planet formation mechanisms for a significant fraction of stellar systems in our Galaxy.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 March; accepted 1 September 2014.

- Nelson, A. F. Planet formation is unlikely in equal-mass binary systems with $A \sim 50$ AU. *Astrophys. J.* **537**, L65–L68 (2000).
- Mayer, L., Wadsley, J., Quinn, T. & Stadel, J. Gravitational instability in binary protoplanetary discs: new constraints on giant planet formation. *Mon. Not. R. Astron. Soc.* **363**, 641–648 (2005).
- Thébault, P., Marzari, F. & Scholl, H. Relative velocities among accreting planetesimals in binary systems: the circumprimary case. *Icarus* **183**, 193–206 (2006).
- Artymowicz, P. & Lubow, S. H. Dynamics of binary-disk interaction. 1: Resonances and disk gap sizes. *Astrophys. J.* **421**, 651–667 (1994).
- Bate, M. R. & Bonnell, I. A. Accretion during binary star formation — II. Gaseous accretion and disc formation. *Mon. Not. R. Astron. Soc.* **285**, 33–48 (1997).
- Casassus, S. *et al.* Flows of gas through a protoplanetary gap. *Nature* **493**, 191–194 (2013).
- Skrutskie, M. F. *et al.* Detection of circumstellar gas associated with GG Tauri. *Astrophys. J.* **409**, 422–428 (1993).
- Di Folco, E. *et al.* GG Tauri: the fifth element. *Astron. Astrophys.* **565**, L2 (2014).
- Andrews, S. M. *et al.* Resolved multifrequency radio observations of GG Tau. *Astrophys. J.* **787**, 148 (2014).
- Beck, T. L. *et al.* Circumbinary gas accretion onto a central binary: infrared molecular hydrogen emission from GG Tau A. *Astrophys. J.* **754**, 72 (2012).
- Guilloteau, S. & Dutrey, A. G. G. in *The Formation of Binary Stars* (eds Zinnecker, H. & Mathieu, R.) 229–233 (IAU Symposium, Vol. 200, 2001).
- Hartigan, P. & Kenyon, S. J. A spectroscopic survey of subarcsecond binaries in the Taurus-Auriga dark cloud with the Hubble Space Telescope. *Astrophys. J.* **583**, 334–357 (2003).
- White, R. J., Ghez, A. M., Reid, I. N. & Schultz, G. A test of pre-main-sequence evolutionary models across the stellar/substellar boundary based on spectra of the young quadruple GG Tauri. *Astrophys. J.* **520**, 811–821 (1999).
- Dutrey, A., Guilloteau, S. & Simon, M. Images of the GG Tauri rotating ring. *Astron. Astrophys.* **286**, 149–159 (1994).
- Piétu, V., Gueth, F., Hily-Blant, P., Schuster, K.-F. & Pety, J. High resolution imaging of the GG Tauri system at 267 GHz. *Astron. Astrophys.* **528**, A81 (2011).
- Skemer, A. J. *et al.* Dust grain evolution in spatially resolved T Tauri binaries. *Astrophys. J.* **740**, 43 (2011).
- Artymowicz, P., Clarke, C. J., Lubow, S. H. & Pringle, J. E. The effect of an external disk on the orbital elements of a central binary. *Astrophys. J.* **370**, L35–L38 (1991).
- Beust, H. & Dutrey, A. Dynamics of the young multiple system GG Tauri. I. Orbital fits and inner edge of the circumbinary disk of GG Tau A. *Astron. Astrophys.* **439**, 585–594 (2005).
- Pierens, A. & Nelson, R. P. On the migration of protoplanets embedded in circumbinary disks. *Astron. Astrophys.* **472**, 993–1001 (2007).
- Pierens, A. & Nelson, R. P. Migration and gas accretion scenarios for the Kepler 16, 34, and 35 circumbinary planets. *Astron. Astrophys.* **556**, A134 (2013).
- Guilloteau, S., Dutrey, A. & Simon, M. G. G. Tauri: the ring world. *Astron. Astrophys.* **348**, 570–578 (1999).
- Dutrey, A. *et al.* Physical and chemical structure of planet-forming disks probed by millimeter observations and modeling. Preprint at <http://arXiv.org/abs/1402.3503> (2014).
- Delorme, P. *et al.* Direct-imaging discovery of a 12–14 Jupiter-mass object orbiting a young binary system of very low-mass stars. *Astron. Astrophys.* **553**, L5 (2013).
- Duchêne, G. & Kraus, A. Stellar multiplicity. *Annu. Rev. Astron. Astrophys.* **51**, 269–310 (2013).

Acknowledgements ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. IRAM is supported by INSU/CNRS (France), MPG (Germany) and IGN (Spain). A.D. thanks the French programmes PNP, PCMI, PNPS and ASA for providing funding for this study.

Author Contributions A.D. led the project and participated in data reduction. All authors contributed to the data analysis, discussed the results and commented on the manuscript.

Author Information This paper makes use of the following ALMA data: ADS/JAO.ALMA2011.0.00059. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.D. (Anne.Dutrey@obs.u-bordeaux1.fr).

METHODS

ALMA observations. GG Tau was observed with ALMA in Cycle 0 (project 2011.0.00059.S) in Band 9 on August 13 2012. The spectral set-up was of 4 spectral windows of total width 0.938 GHz, centred at 691.485, 689, 676 and 674 GHz respectively. The correlator observed in dual polarization, yielding a channel spacing of 244 kHz, or 0.11 km s^{-1} . With Hanning smoothing, the spectral resolution is twice those values. The array of 23 antennas was in its extended configuration. The water vapour was about 0.25 mm and system temperatures ranged from 500 to 1,000 K. Comparison of the CO data with existing $^{12}\text{CO } J = 2-1$ IRAM data (Fig. 1) revealed a frequency shift of the ALMA observations whose origin could not be identified. We corrected the velocity offset by correlating with the IRAM data. The velocity difference was $0.97 \pm 0.01 \text{ km s}^{-1}$. Since it is an offset, it does not introduce any error on the velocity gradient. The absolute flux scale was determined using Ceres. The ALMA data were calibrated using CASA then exported through UVFITS to be imaged and analysed using GILDAS. Extended Data Fig. 1 presents the resulting channel maps.

The integrated ring+disk continuum flux is in good agreement with the previous unresolved detection²⁵: no more than 20% of the flux is lost.

For the CO $J = 6-5$ line, we only recover a fraction of the integrated line flux of $\sim 38 \pm 8 \text{ Jy km s}^{-1}$ (ref. 26). The ALMA antenna configuration exhibits a lack of short baselines which induces a partial loss of flux for the extended structures that deconvolution cannot recover, particularly in the CO circumbinary disk which extends up to 800 AU. We ran several disk simulations of the CO $J = 6-5$ emission using the ALMA simulator²⁷. If emission only arises from the ring, our simulation shows that about 20% of the emission would be filtered out by the configuration of ALMA. A much larger fraction is lost if CO $J = 6-5$ extends into the whole outer disk (up to 800 AU); this fraction depends on the signal to noise ratio. The expected extension of the CO $J = 6-5$ emission depends on the excitation conditions in the outer disk, so uncertainties on the amount of flux loss preclude any quantitative analysis of the large scale disk emission in CO $J = 6-5$ data.

IRAM data. We complement ALMA data with CO $J = 2-1$ and continuum images from the IRAM interferometer. The compact configuration was used for observations on 5–6 December 1997, and long baselines up to 400 m were used on 28–31 January 1998 and 18 March 1998¹¹. The unpublished extended configuration data were obtained on 16 February 2008. System temperatures were 120 K. Flux calibration, based on MWC349, gives an absolute flux accuracy of about 10%. Imaging was performed using GILDAS software. The resulting beam size is $0.65'' \times 0.29''$ at position angle PA = 21° . The spectral resolution is 0.10 km s^{-1} , and the effective noise is about 14 mJy per beam (up to 20 mJy per beam on the channels with the most extended emission due to deconvolution limits). The continuum image at 1.3 mm was produced using line-free channels. Its effective noise is limited by dynamic range to about 1 mJy per beam. Extended Data Fig. 2 shows the CO $J = 2-1$ channel maps.

Proper motions. Accounting for the expected orbital motion, we adopted proper motions of $(17, -19) \text{ mas yr}^{-1}$ for the centre of mass of the system²⁸ to merge all data.

Inner dust disk properties. For the disk of GG Tau Aa, the total fluxes of 55 mJy at 0.45 mm and 10 mJy at 1.3 mm are consistent with optically thick dust emission at a mean temperature of $\sim 35 \text{ K}$ (ref. 29) and an outer radius of the order of 7 AU, a size in agreement with tidal truncation by GG Tau Ab. Assuming a standard dust opacity of $0.02 \text{ cm}^2 \text{ g}^{-1}$ at 1.3 mm, the minimum disk mass is roughly $10^{-3} M_\odot$ or a Jupiter's mass worth of gas and dust. The rate at which matter is accreting from the disk onto the surface of GG Tau Aa is in the range $(\sim 10^{-7.56} - 10^{-8.02}) M_\odot \text{ yr}^{-1}$ (ref. 12). Such a disk would dissipate in less than $\sim (3 \times 10^4) - 10^5 \text{ yr}$ without external replenishment. We place an upper limit on the circumstellar disk mass surrounding GG Tau Ab of $\sim 10^{-5} M_\odot$ at 3σ (a little more than one Earth mass) or an upper limit on its radius of $\sim 1 \text{ AU}$ by extrapolating the dust opacity to 0.45 mm, $0.067 \text{ cm}^2 \text{ g}^{-1}$, and using a similar temperature as in the disk of GG Tau Aa. This is consistent with the separation of Ab1 and Ab2 of $\sim 4.5 \text{ AU}$ (ref. 8).

Inner cavity. CO excitation conditions. We performed escape probability calculations using the code RADEX³⁰ (<http://www.sron.rug.nl/~vdtak/radex/radex.php>), a radiative transfer code dedicated to calculations of molecular line excitations. At the location of the eastern CO $J = 2-1$ peak ($\sim 80 \text{ AU}$ from the stars), the CO lines are optically thick, H_2 densities must be greater than 10^5 cm^{-3} , and the gas is cold with a temperature in the range $\sim 30-40 \text{ K}$. Closer to the stars, the peaks of the bright CO $J = 6-5$ regions correspond to optically thin emission, with a warmer temperature of about 70 K and a H_2 density of $\sim 5 \times 10^5 \text{ cm}^{-3}$. These densities agree with the pre-shock gas densities derived from shock models for the 2.12- μm H_2 emission maps. With a standard CO/ H_2 ratio of 10^{-4} , the mass of each optically thin clump is of the order of $\sim (3-7) \times 10^{-5} M_\odot$.

CO velocity field. We studied the velocity variations of the CO $J = 6-5$ emission. The spectra displayed in Extended Data Fig. 3, at the locations of the CO $J = 6-5$ clumps, also show that the velocity is dominated by rotation. The complexity of the

system (which is triple) and the quality of the data preclude any detailed modelling of the velocity field, but some points can be addressed. From a compilation of existing values in the literature, we derive for GG Tau Aa a velocity of $5.8 \pm 0.2 \text{ km s}^{-1}$ (ref. 31). This implies a difference of the order of $0.6 \pm 0.2 \text{ km s}^{-1}$ from the known systemic (barycentric) velocity of the whole stellar system²¹. Extended Data Fig. 3 reveals that the measured CO gas velocity near GG Tau Aa is in very good agreement with Aa velocity as expected for bound motions. Around the disk minor axis, we note that the isovelocity contours are twisted at a radius of about 70–80 AU from the mass centre. Inside this radius, the isocontours are closely packed, and suddenly spread out beyond. Furthermore, the isovelocity contours are aligned on the Aa–Ab line between GG Tau Aa and Ab. This is exactly what is expected when the gravitation field, externally dominated by the total stellar mass ($1.28 M_\odot$), becomes dominated by the individual mass of the star Aa and binary Ab. At the location of the CO $J = 6-5$ clumps, the spectra displayed in Extended Data Fig. 3 also exhibit a large velocity dispersion ($\delta v \approx 2-2.5 \text{ km s}^{-1}$) which could be partly due to infall. We check this by analysing the velocity field along the minor axis at the northern clump. In an inclined disk, projected velocities due to infall are at a maximum along the minor axis, contrary to rotation velocities which are at a maximum along the major axis^{32,33}. Taking into account the beam size of 35 AU, we find that the maximum velocity dispersion along the minor axis in a spectrum would be of the order 1.2 km s^{-1} if it is only due to Keplerian rotation around central objects of $0.65 M_\odot$ (because this CO $J = 6-5$ clump is located in the area where the velocity field is dominated by GG Tau Ab). This is a factor two lower than the observed line widths of $2-2.5 \text{ km s}^{-1}$. In disks, turbulence cannot explain such a large dispersion³⁴. However, infall velocities can be $\sqrt{2}$ larger than Keplerian velocities. It is therefore reasonable to conclude that the observed velocities and line widths are compatible with a combination of rotation and infall. For the southern clump, located near GG Tau Aa, the velocity profile is very complex. For instance, the spectrum of the Aa star shows two double peaked profiles, perfectly centred on the velocity of Aa (5.8 km s^{-1}), but with significantly different widths (full-width at half-maximum, FWHM, ~ 1.5 and 6 km s^{-1}). The total velocity dispersion of the broader double peaked profile ($\sim 6 \text{ km s}^{-1}$) is compatible with a circumstellar disk of outer radius of the order of $\sim 20 \text{ AU}$ (ref. 35), if the inclination of the system is the same as that of the outer ring (35°). Such a CO outer radius is compatible with the value derived from the dust (provided a small part of the dust emission is optically thin) and tidal truncation due to GG Tau Ab (the physical star separation being about 42 AU).

Outer disk. Dust ring properties. We use Diskfit, a radiative transfer code dedicated to the modelling of the disk images in mm/submm regime³⁶, to analyse the dust properties inside the outer ring using the best data available (3 mm and 1.3 mm PdBI data and 0.45 mm ALMA observations). A modified Levenberg-Marquardt scheme with step adjustment is used to derive the best fit³⁷. The analysis is performed in the Fourier plane to avoid non linear effects due to the deconvolution. We explore large ranges for all parameters to avoid local minima. Extended Data Fig. 4 and Extended Data Table 2 present the results, with 1σ formal error bars derived from the covariance matrix. The errors are computed on the basis of the difference between data and model, so include dynamic range limitations in addition to thermal noise. We use simple truncated power laws to model the surface density and temperature. We allow dust emissivity κ_v to vary with frequency as $\kappa_v = \kappa_0 \times (v/v_0)^\beta$ (β being the emissivity index). Such a simple model is valid for the dust distribution since the emission is known to be confined in a ring with sharp edges^{9,21}. The derived geometrical parameters are in very good agreement with the previous results²¹. Since the dust temperature is mostly constrained by the ratio of the flux density at 0.45 mm and 1.3 mm, we checked that the results remain robust even assuming an absolute flux calibration error as high as 20% at 0.45 mm. The density, temperature and dust spectral index are robust with respect to all minimizations. However, we find a slight dependence between β and the surface density. This is not surprising because in the optically thin case, the brightness temperature is proportional to $\tau \propto \kappa(0.45 \text{ mm}) \times \Sigma(r)$ (with $\kappa(0.45 \text{ mm})$ the absorption coefficient of the dust at 0.45 mm). A better removal of this dependence would require higher angular resolution at 1.3 mm. We found a value for the emissivity index $\beta = 1.0 \pm 0.1$, which indicates moderate grain growth compared to interstellar dust. Such a grain growth is compatible with a recent analysis of the dust ring performed between 7 and 1.3 mm⁹. We found a marginal dependency of β with radius, suggesting that larger dust particles reside near the inner edge of the outer disk (that is, closer to the central stars), as observed for some circumstellar disks orbiting young single stars³⁸. So far, dust temperatures have been derived either from Monte-Carlo dust disk models³⁹ or one-dimensional thermal models associated with pure mm data⁹. In our case, we are able to directly derive the dust temperature thanks to the deviation from the Rayleigh Jeans regime at 0.45 mm, combined with sufficiently high spatial resolution at 0.45 mm and 1.3 mm. This permits the removal of the degeneracy between the temperature and the opacity. This direct

measurement is of great interest for a better characterization of disk mid-planes where planet formation is expected to occur.

CO excitation conditions. Analysis of the physical conditions in the outer ring at radius around 200 AU shows that the mid-plane H₂ density ($\sim 10^9 \text{ cm}^{-3}$) is high enough to thermalize CO transitions up to $J = 6-5$ ²¹. Excitation of the CO $J = 6-5$ line requires a H₂ critical density of $\sim 3 \times 10^5 \text{ cm}^{-3}$. We also derive a dust temperature lower than the CO freeze-out temperature of 17 K. Therefore the observed large scale emissions from the CO $J = 1-0$ and $J = 2-1$ transitions must arise from a warmer molecular layer located above the mid-plane, as predicted by layered disk models²² and revealed by ALMA observations of HD 163296⁴⁰.

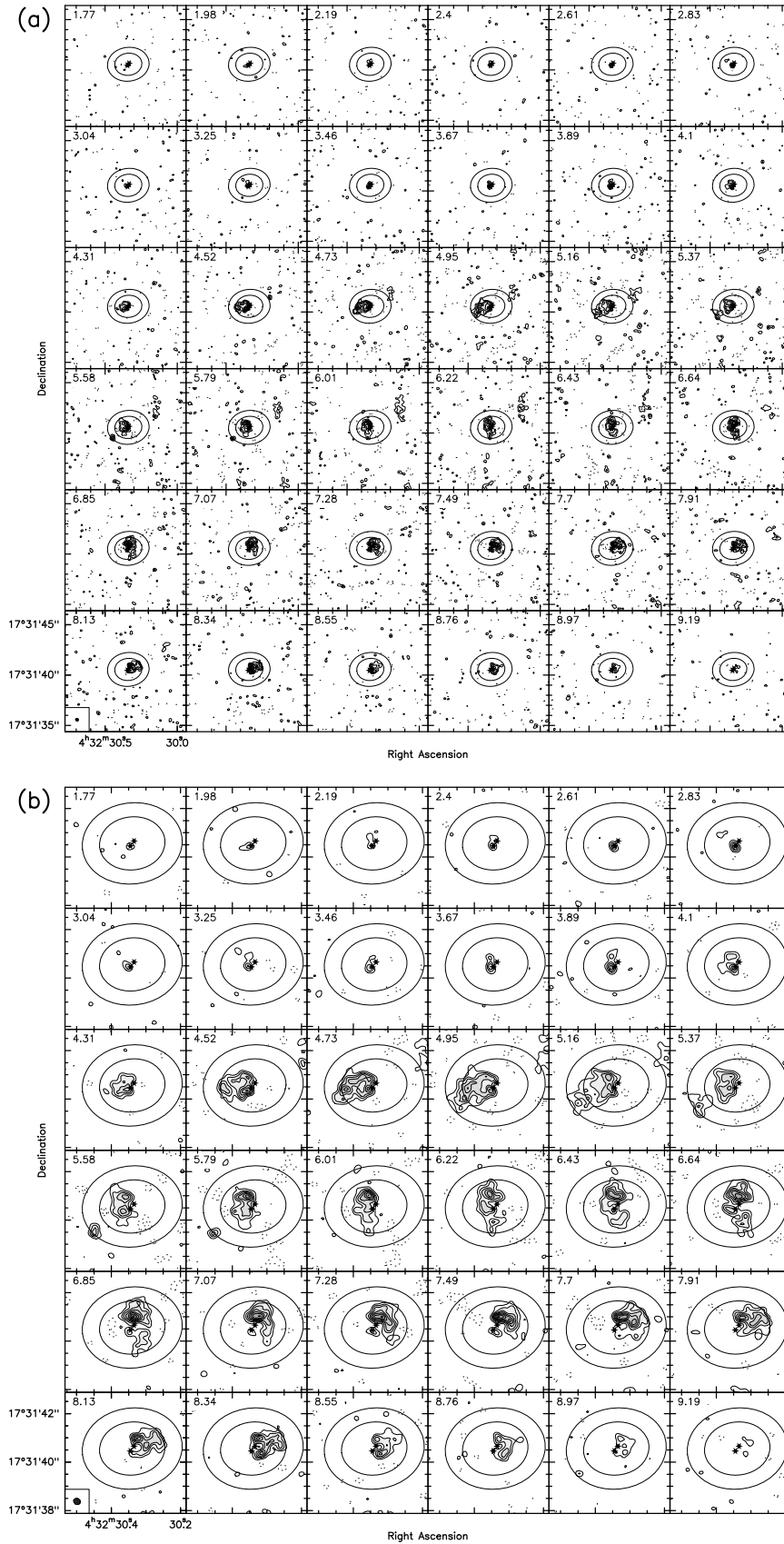
Using our measurement of the dust temperature, we find that the hydrostatic scale height of the gas and dust is about ~ 25 and ~ 35 AU at radii of 200 and 300 AU, respectively. If the CO gas originates in a layer located at about 3 scale heights (in agreement with chemical models), and taking into account the vertical temperature gradient, the H₂ density would be $\sim 10^5$ and $\sim 10^4 \text{ cm}^{-3}$ at radii of 200 and 300 AU, respectively. This is sufficient to thermalize the $J = 1-0$ and $J = 2-1$ lines, but not the $J = 6-5$ transition, which can be partially subthermally excited outside the ring. Note that any gap at the outer edge of the ring created by an embedded protoplanet (see section below) would imply some subthermal excitation there.

A CO hotspot at the outer edge of the ring. The CO $J = 6-5$ and $J = 2-1$ maps reveal a brightness enhancement of unknown origin near the outer edge of the dust ring (Fig. 1 and spectrum in Extended Data Fig. 3). From the CO $J = 2-1$ emission which is optically thick and thermalized, we derive a minimum temperature of ~ 40 K in this clump. This indicates a temperature enhancement of at least 20 K compared to the surrounding optically thick CO gas. The feature is clearly dominated by the gas emission, unlike the recent detection of a dusty vortex-like structure, or dust trap, in the circumstellar disk associated with Oph IRS 48⁴¹. In the high angular resolution continuum maps at 0.45 mm and 1.3 mm, there is no dust counterpart to this bright CO peak. The dust absorption coefficient being small, only a sufficiently large and massive dust clump would be detectable. Given the effective (dynamic range limited) noise, the (dust) mass limit is a few times $10^{-7} M_{\odot}$ at 3σ , or the size must be smaller than about 4 AU, assuming a dust temperature of 40 K. In Extended Data Fig. 3, there is a marginal excess emission at 1.3 mm (4σ level) in the difference map (right panel) at the location of the blob. This has to be taken with caution since there are other 4σ peaks which are likely to trace some extended cold emission beyond the ring outer radius²¹. Analysis of the GG Tau A dust ring in term of a dust trap, based on resolved images between 7 and 1.3 mm, has been recently achieved but remains inconclusive⁹. The presence $10''$ away of the binary star GG Tau B could induce gravitational disturbances in the circumbinary disk especially if its orbit is eccentric⁴². However, the total mass of GG Tau B is very low ($\sim 0.17 M_{\odot}$; ref. 13) and it is located far away from the GG Tau A circumbinary ring ($\sim 1,400$ AU). Therefore, we conclude that it is unlikely that GG Tau B would have induced such a sizeable disturbance in the circumbinary ring. We explore here another interesting possibility; this feature would be indirect evidence for an unknown (already formed) embedded companion that is still actively accreting material from the circumbinary disk⁴³. In this scenario, the planet remains undetectable at millimetre wavelengths but the warmer accreting envelope can be seen^{44,45}. The presence of such a companion would provide an explanation for the mass confinement in a narrow ring ($\sim 80\%$ of the circumbinary mass is located within the ~ 90 AU breadth ring²¹). This situation is similar to the gas free, dusty ring orbiting the more evolved star Fomalhaut, where confinement by shepherding planets appears to be the most consistent explanation⁴⁶.

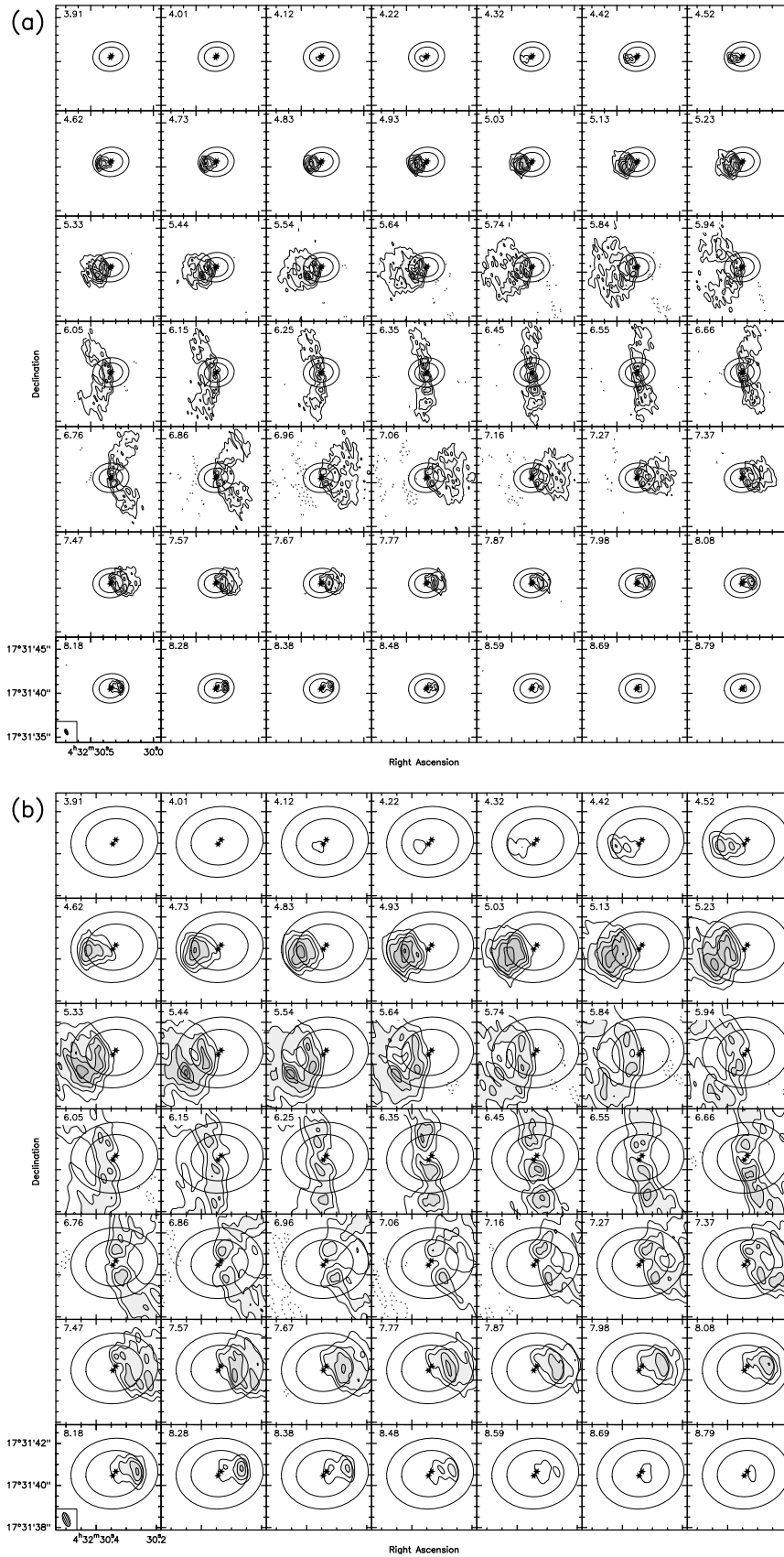
Simulations of the evolution of Kepler systems where exoplanets have been observed reveal that Saturn-like planets are expected to migrate to reach the 4:1 resonance²⁰ without opening a gap, except if the disk is very cold, as is the case for the GG Tau ring. In such cold disks, a protoplanet can remain at the ring's outer edge and the size of the gap will mainly depend on the planet mass²⁰. So far, we have not detected a gap in the gas and dust circumbinary distribution of GG Tau A. Our CO $J = 6-5$ map shows that there is almost no CO $J = 6-5$ emission located outside the outer ring; this is expected if there is a gap or a strong density decrease which implies subthermal excitation. Gap opening depends on several physical parameters, such as the mass of the planet, and the disk viscosity and aspect ratio⁴⁷. For example, a Saturn-like body located at 250 AU in the circumbinary disk would open a gap of about 35 AU assuming standard viscosity ($\alpha = 10^{-2}$) and aspect ratio $h(r)/r = 0.1$. Under the same physical conditions, the gap would be only 12 AU wide for a Neptune-like planet⁴⁸. Observing a gap in the existing CO $J = 2-1$

map would be difficult because the CO $J = 2-1$ line is very optically thick, making the expected ring/gap brightness contrast difficult to detect, except if the gap is very deep with a high ring/gap density contrast (larger than the CO opacity), and is large enough to be directly seen (about 40 AU given the beam size). ¹³CO is less optically thick (by a factor of 70) but the existing ¹³CO $J = 2-1$ map²¹ has an angular resolution ($0.88 \times 0.56''$ or 120×80 AU) insufficient for direct imaging. However, a ring/outer disk density contrast of ~ 25 at 260 AU has been reported²¹. This contrast could be the signature of an unresolved gap in the ¹³CO $J = 2-1$ map. A spectroscopic detection⁴⁹ would be possible only by using very high sensitivity data, which do not exist so far. Very sensitive observations of optically thinner CO isotopologues at high angular resolution ($\ll 0.1 - 0.2''$) are needed to study the hot spot's properties (including the viewing of a gap) and its implications for planet formation scenarios.

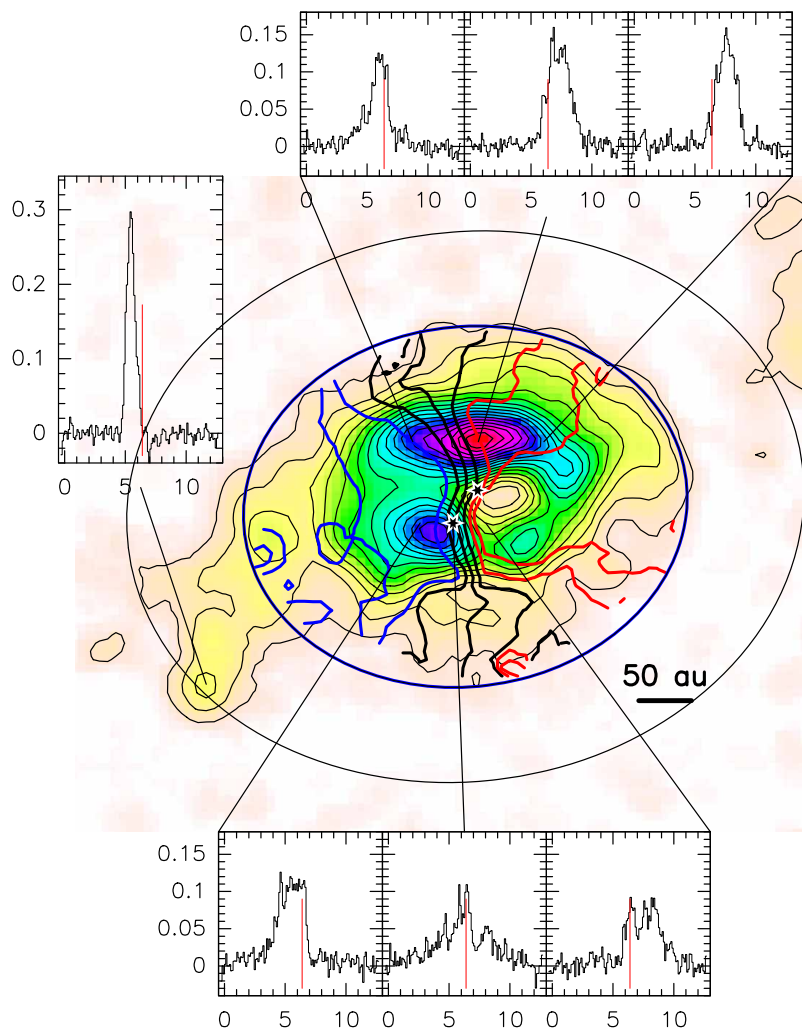
25. Moriarty-Schieven, G. H. & Butner, H. M. A submillimeter-wave "flare" from GG Tauri? *Astrophys. J.* **474**, 768–773 (1997).
26. Thi, W.-F., van Dishoeck, E. F., Blake, G. A., van Zadelhoff, G.-J. & Hogerheijde, M. R. Detection of H₂ pure rotational line emission from the GG Tauri binary system. *Astrophys. J.* **521**, L63–L66 (1999).
27. Pety, J., Gueth, F. & Guilloteau, S. ALMA+ACA simulation tool. *ALMA Memo* **386**, 1–10 (2002).
28. Ducourant, C. *et al.* Pre-main sequence star proper motion catalogue. *Astron. Astrophys.* **438**, 769–778 (2005).
29. Piétu, V., Guilloteau, S., Di Folco, E., Dutrey, A. & Boehler, Y. Faint disks around classical T Tauri stars: small but dense enough to form planets. *Astron. Astrophys.* **564**, A95 (2014).
30. van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J. & van Dishoeck, E. F. A computer program for fast non-LTE analysis of interstellar line spectra. With diagnostic plots to interpret observed line intensity ratios. *Astron. Astrophys.* **468**, 627–635 (2007).
31. Nguyen, D. C., Brandeker, A., van Kerkwijk, M. H. & Jayawardhana, R. Close companions to young stars. I. A large spectroscopic survey in Chamaeleon I and Taurus-Auriga. *Astrophys. J.* **745**, 119 (2012).
32. Guilloteau, S. & Dutrey, A. Physical parameters of the Keplerian protoplanetary disk of DM Tauri. *Astron. Astrophys.* **339**, 467–476 (1998).
33. Rosenfeld, K. A., Chiang, E. & Andrews, S. M. Fast radial flows in transition disk holes. *Astrophys. J.* **782**, 62 (2014).
34. Guilloteau, S. *et al.* Chemistry in disks. VIII. The CS molecule as an analytic tracer of turbulence in disks. *Astron. Astrophys.* **548**, A70 (2012).
35. Beckwith, S. V. W. & Sargent, A. I. Molecular line emission from circumstellar disks. *Astrophys. J.* **402**, 280–291 (1993).
36. Boehler, Y., Dutrey, A., Guilloteau, S. & Piétu, V. Probing dust settling in proto-planetary discs with ALMA. *Mon. Not. R. Astron. Soc.* **431**, 1573–1586 (2013).
37. Piétu, V., Dutrey, A. & Guilloteau, S. Probing the structure of protoplanetary disks: a comparative study of DM Tau, LkCa 15, and MWC 480. *Astron. Astrophys.* **467**, 163–178 (2007).
38. Guilloteau, S., Dutrey, A., Piétu, V. & Boehler, Y. A dual-frequency sub-arcsecond study of proto-planetary disks at mm wavelengths: first evidence for radial variations of the dust properties. *Astron. Astrophys.* **529**, A105 (2011).
39. Madlener, D., Wolf, S., Dutrey, A. & Guilloteau, S. The circumstellar disk of HH 30. Searching for signs of disk evolution with multi-wavelength modeling. *Astron. Astrophys.* **543**, A81 (2012).
40. de Gregorio-Monsalvo, I. *et al.* Unveiling the gas-and-dust disk structure in HD 163296 using ALMA observations. *Astron. Astrophys.* **557**, A133 (2013).
41. van der Marel, N. *et al.* A major asymmetric dust trap in a transition disk. *Science* **340**, 1199–1202 (2013).
42. Beust, H. & Dutrey, A. Dynamics of the young multiple system GG Tauri. II. Relation between the stellar system and the circumbinary disk. *Astron. Astrophys.* **446**, 137–154 (2006).
43. Gressel, O., Nelson, R. P., Turner, N. J. & Ziegler, U. Global hydromagnetic simulations of a planet embedded in a dead zone: gap opening, gas accretion, and formation of a protoplanetary jet. *Astrophys. J.* **779**, 59 (2013).
44. Wolf, S., Gueth, F., Henning, T. & Kley, W. Detecting planets in protoplanetary disks: a prospective study. *Astrophys. J.* **566**, L97–L99 (2002).
45. Wolf, S. & D'Angelo, G. On the observability of giant protoplanets in circumstellar disks. *Astrophys. J.* **619**, 1114–1122 (2005).
46. Boley, A. C. *et al.* Constraining the planetary system of Fomalhaut using high-resolution ALMA observations. *Astrophys. J.* **750**, L21 (2012).
47. Crida, A., Morbidelli, A. & Masset, F. On the width and shape of gaps in protoplanetary disks. *Icarus* **181**, 587–604 (2006).
48. Takeuchi, T., Miyama, S. M. & Lin, D. N. C. Gap formation in protoplanetary disks. *Astrophys. J.* **460**, 832–847 (1996).
49. Dutrey, A. *et al.* Cavities in inner disks: the GM Aurigae case. *Astron. Astrophys.* **490**, L15–L18 (2008).



Extended Data Figure 1 | ALMA CO $J=6-5$ velocity channel map. For each spectroscopic channel, the velocity is given at top left. **a**, Full map; **b**, inner zoom. The beam size is $0.29'' \times 0.25''$ at PA 68° . The level step is 100 mJy per beam or 3.51 K corresponding to 3.4σ .

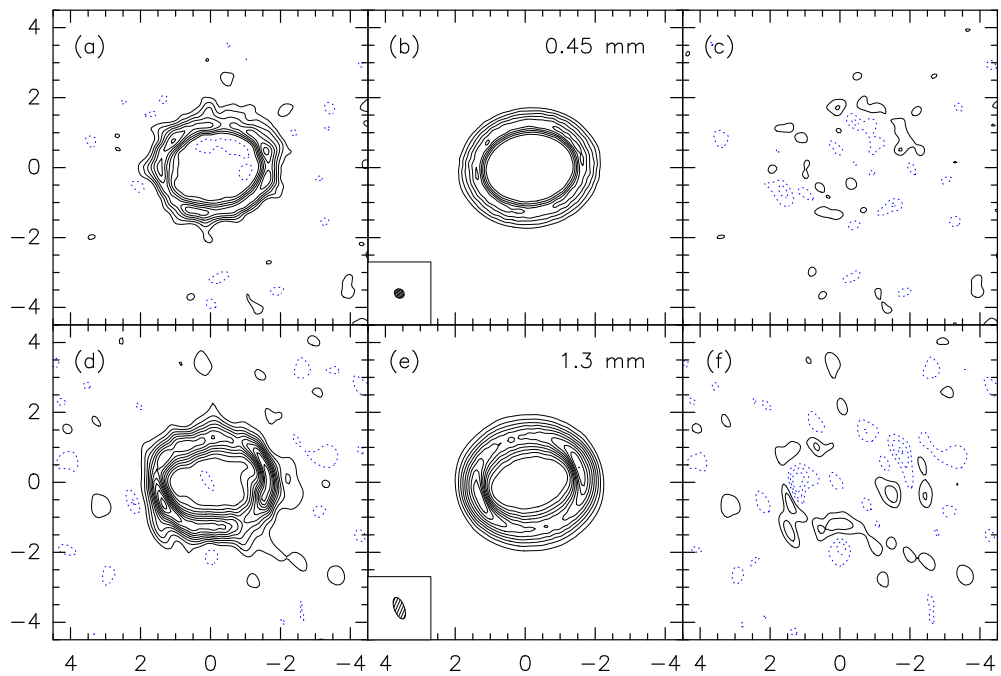


Extended Data Figure 2 | IRAM CO $J=2-1$ velocity channel map. For each spectroscopic channel, the velocity is given at top left. **a**, Full map; **b**, inner zoom. The beam size is $0.68'' \times 0.31''$ at PA 21° . The level step is 50 mJy per beam or 5.48 K corresponding to 3.85σ .



Extended Data Figure 3 | Montage of the CO $J=6-5$ data. False colours and black contours show the integrated area. The velocity gradient is given in thick contours: blue (gas approaching), black (systemic velocity) and red (gas receding). Stars show the location of Aa (south) and Ab (north). The two large ellipses show the ring edges. The three spectra sets (y axis, intensity in units of J_y per beam; x axis, velocity in units of km s^{-1}) show the velocity gradient

along the northern/southern CO $J=6-5$ clump, respectively (dominated by rotation). On spectra, the red line is the systemic velocity (6.4 km s^{-1}). From east to west, the black contours correspond to velocity contours of 6.0 , 6.4 and 6.8 km s^{-1} . The systemic velocity contour passes between the two stars (barycentre). The single spectrum corresponds to the location of the hotspot.



Extended Data Figure 4 | Dust ring best model. **a**, ALMA continuum data at 0.45 mm (emission from Aa circumstellar disk has been removed). **b**, Best model at 0.45 mm, same contour levels. **c**, Difference between the observations and the best model, contour levels correspond to 2σ . **d**, **e**, **f**, As **a**, **b**, **c** but for the IRAM continuum data at 1.3 mm.

Extended Data Table 1 | Parameters relevant to the analysis of the ALMA data

CO J=2-1 analysis		
Systemic velocity	V_{LSR}	$6.38 \pm 0.02 \text{ km.s}^{-1}$
Orientation	PA	$7 \pm 2^\circ$
Inclination	i	$37 \pm 1^\circ$
Disk outer radius	R_{out}	$\sim 800 \text{ au}$
Velocity law:	$V(r) = V_{100} \left(\frac{r}{100 \text{ au}} \right)^{-v}$	
velocity at 100 au	V_{100}	$3.4 \pm 0.1 \text{ km.s}^{-1}$
exponent	v	0.5 ± 0.1

Values are adopted from previous work (see Methods).

Extended Data Table 2 | Best fit results for the GG Tau circumbinary dust disk, as derived from the whole continuum data set

Geometry		
Orientation	PA	$6.5 \pm 0.2^\circ$
Inclination	i	$35.0 \pm 0.2^\circ$
Size		
Ring inner radius	R_i	$193 \pm 1 \text{ au}$
Ring outer radius	R_o	$285 \pm 1 \text{ au}$
Surface density law: $\Sigma(r) = \Sigma_{200} \left(\frac{r}{200\text{au}}\right)^{-p}$		
H ₂ surface density	Σ_{200}	$9 \pm 1 \text{ } 10^{24} \text{ cm}^{-2}$
exponent	p	1.3 ± 0.2
Temperature law: $T(r) = T_{200} \left(\frac{r}{200\text{au}}\right)^{-q}$		
Dust Temperature	T_{200}	$13.8 \pm 0.3 \text{ K}$
exponent	q	1.1 ± 0.1
Gas-to-Dust Ratio	G/D	100 -
Absorption Coefficient (gas+dust)	$\kappa(1.3 \text{ mm})$	$0.02 \text{ cm}^2/\text{g}$ -
Dust Spectral Index	β	1.0 ± 0.1

Quantum tomography of an electron

T. Jullien^{1*}, P. Roulleau^{1*}, B. Roche¹, A. Cavanna², Y. Jin² & D. C. Glattli¹

The complete knowledge of a quantum state allows the prediction of the probability of all possible measurement outcomes, a crucial step in quantum mechanics. It can be provided by tomographic methods¹ which have been applied to atomic^{2,3}, molecular⁴, spin^{5,6} and photonic states. For optical^{7–9} or microwave^{10–13} photons, standard tomography is obtained by mixing the unknown state with a large-amplitude coherent photon field. However, for fermions such as electrons in condensed matter, this approach is not applicable because fermionic fields are limited to small amplitudes (at most one particle per state), and so far no determination of an electron wavefunction has been made. Recent proposals involving quantum conductors suggest that the wavefunction can be obtained by measuring the time-dependent current of electronic wave interferometers¹⁴ or the current noise of electronic Hanbury-Brown/Twiss interferometers^{15–17}. Here we show that such measurements are possible despite the extreme noise sensitivity required, and present the reconstructed wavefunction quasi-probability, or Wigner distribution function¹⁷, of single electrons injected into a ballistic conductor. Many identical electrons are prepared in well-controlled quantum states called levitons¹⁸ by repeatedly applying Lorentzian voltage pulses to a contact on the conductor^{19–21}. After passing through an electron beam splitter, the levitons are mixed with a weak-amplitude fermionic field formed by a coherent superposition of electron–hole pairs generated by a small alternating current with a frequency that is a multiple of the voltage pulse frequency¹⁶. Antibunching of the electrons and holes with the levitons at the beam splitter changes the leviton partition statistics, and the noise variations provide the energy density matrix elements of the levitons. This demonstration of quantum tomography makes the developing field of electron quantum optics with ballistic conductors a new test-bed for quantum information with fermions^{20,22–24}. These results may find direct application in probing the entanglement of electron flying quantum bits²⁵, electron decoherence¹⁷ and electron interactions. They could also be applied to cold fermionic (or spin-1/2) atoms²⁶.

A quantum state ψ contains all the information about a particle or a system. Disregarding spin for simplicity, a practical representation is given by the product of the wavefunction φ which contains the spatial information and an occupation state in a Fock space representation. Although the determination of the latter requires similar experimental resources for a fermion and for a single boson, the complete determination of the wavefunction by tomographic methods is fundamentally more demanding for fermions than for bosons. To emphasize this let us consider a single particle propagating in a single spatial mode (an electron in an effectively one-dimensional ballistic conductor or a photon in an optical medium). In the space and time domain representation, the electron quantum state at position x and time t is $\varphi(t - x/v_F)|f\rangle$, where v_F is the velocity, φ is the wavefunction and $|f\rangle$ is the fermionic occupation state, $|1\rangle$ or $|0\rangle$. Similarly, a photonic state is characterized by $E_0 u(t - x/c)|b\rangle$, where E_0 is the electric field amplitude of a single photon, u is the spatial mode which plays the role of φ , $|b\rangle$ is, for example, a number state or a Glauber coherent state, and c is the speed of light. In quantum optics, the determination of the mode u can be done

by mixing with a coherent field (local oscillator) with amplitude $\sqrt{\bar{N}_{\text{LO}}}E_0 u(t - x/c)$, where \bar{N}_{LO} is the mean photon number. Then a classical measurement of u can be made because the fundamental quantum measurement uncertainty $\sim 1/\sqrt{\bar{N}_{\text{LO}}}$ vanishes for large $\sqrt{\bar{N}_{\text{LO}}}$. Only the determination of the bosonic state $|b\rangle$ requires purely quantum tomography. For fermions, however, the state occupation being limited to 1, no classical amplitude level can be reached and a quantum tomography is necessary for both the wavefunction φ and the fermion state $|f\rangle$. In the present work, we prepare many electrons, each in the known fermionic state $|f\rangle = |1\rangle$. We show that, although extremely demanding, a quantum tomography of the wavefunction of a time-resolved single electron is possible using shot noise, and we determine the first-order coherence $\tilde{\varphi}^*(\varepsilon')\tilde{\varphi}(\varepsilon)$ in the energy representation (where $\tilde{\varphi}(\varepsilon)$ is the Fourier transform of $\varphi(t)$ and the asterisk denotes complex conjugation).

We first explain how we prepare the single-electron state in a quantum conductor and then how the quantum tomography is performed. The quantum conductor is a quantum point contact (QPC) placed at low temperature (electron temperature, $T_e \approx 35$ mK). It is formed in a high-mobility two-dimensional electron gas using split gates deposited on top of a GaAs/GaAlAs heterojunction (Fig. 1a and Supplementary Information). Applying negative voltage V_G to the gates creates a constriction transmitting a limited number of electronic modes. Here we select a single mode whose transmission $D(V_G)$ is obtained by measuring the conductance $G = (2e^2/h)D$, where e is the electron charge and h is Planck's constant. The on-demand injection of charges into the conductor is obtained by applying a voltage pulse $V_L(t)$ on one of the ohmic contacts adjacent to the QPC, say the left. If the voltage flux satisfies $e \int_{-\infty}^{+\infty} V_L(t) dt = h$ then the charge is unity. Moreover, a Lorentzian pulse $V(t) = 2\hbar w/e(t^2 + w^2)$, where $\hbar = h/2\pi$ and $2w$ is the pulse width at mid height, creates a clean single-electron state¹⁹ called a leviton¹⁸ with an energy just greater than the Fermi energy of the conductor. Remarkably, although the leviton originates from a collective displacement of the Fermi sea towards positive energy, it can be considered a single-particle state^{20,21} disentangled from the Fermi sea²¹. A wavefunction representation of the leviton in the energy domain is $\tilde{\varphi}(\varepsilon) = \sqrt{2w/\hbar\theta}(\varepsilon) e^{-\varepsilon w/\hbar}$, where θ is the Heaviside step function. Experimentally, periodic trains of levitons have been generated¹⁸. When partitioned by a QPC¹⁸ they generate a current shot noise corresponding to electrons prepared in the occupation state $|1\rangle$. Using shot-noise spectroscopy and electron Hong–Ou–Mandel noise correlation techniques, their wavefunction probability in the energy and time domain was found to be consistent with expectations. Thus, levitons provide a clean way of testing wavefunction quantum tomography in a quantum conductor.

We now discuss how to probe the first-order coherence from which the full wavefunction can be extracted. A first approach could be wave interferometry. In quantum optics, this gives $G_1(t - t')$, which is the first-order coherence $\langle \psi^+(t')\psi(t) \rangle$ averaged over the mean time $\bar{t} = (t + t')/2$ (here ψ^+ is the Hermitian conjugate of ψ). The information is incomplete, however, and other approaches must be used²⁷. Wave interferometry is more promising with electrons: measuring the time-dependent current of an electronic Mach–Zehnder interferometer has been proposed¹⁴ as an elegant way to determine the full coherence $\langle \psi^+(t')\psi(t) \rangle$,

¹Service de Physique de l'Etat Condensé, IRAMIS/DSM (CNRS URA 2464), CEA Saclay, F-91191 Gif-sur-Yvette, France. ²CNRS, Laboratoire de Photonique et de Nanostructures, Route de Nozay, 91460 Marcoussis, France.

*These authors contributed equally to this work.

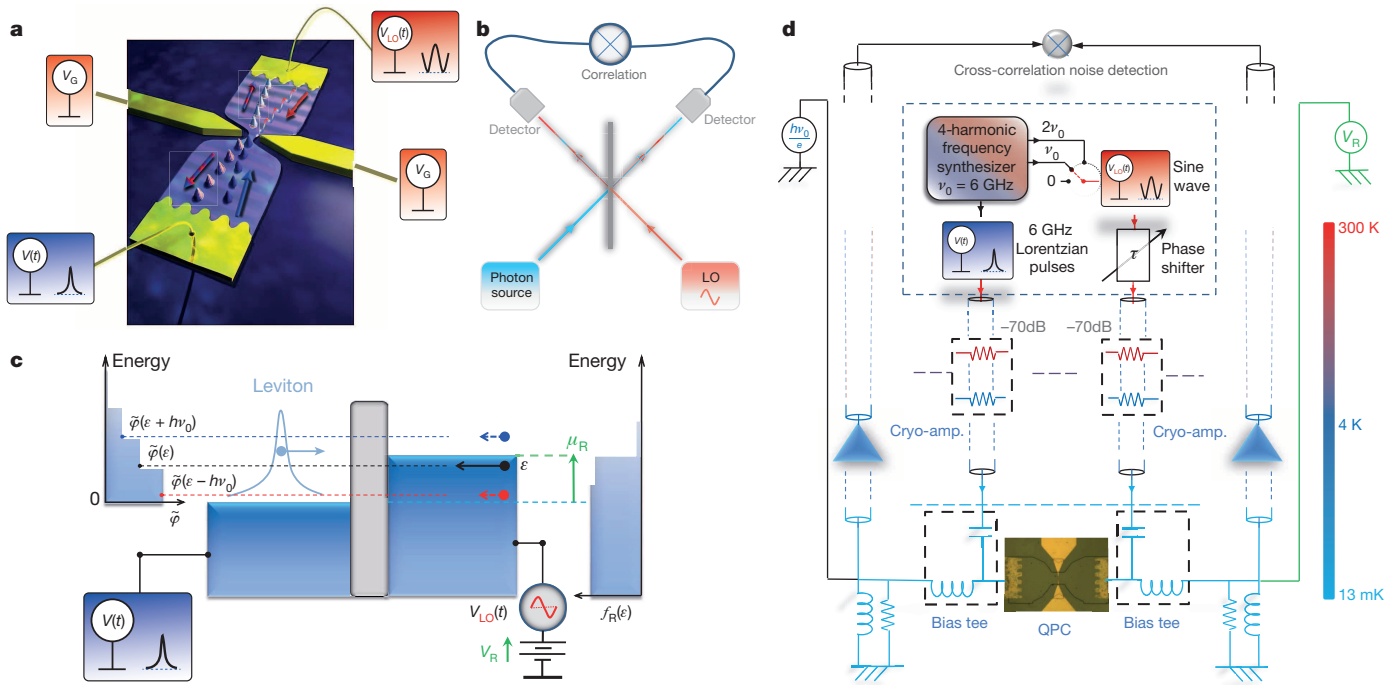


Figure 1 | Schematics of quantum wave tomography. **a**, Experiment principle. Periodic Lorentzian voltage pulses $V(t)$ applied to the left contact of a two-dimensional electron gas inject unit charge pulses called levitons which are partitioned by an electronic beam splitter called a quantum point contact (QPC). The split-gate voltage V_G controls the transmission D of the one-dimensional electronic mode formed at the QPC. A d.c. voltage V_R (not shown) and a weak a.c. voltage $V_{LO}(t) = (k\eta_{LO}h\nu_0/e)\cos(2\pi k\nu_0(t - \tau))$ are applied to the right contact. The latter generates a small flux of electrons and holes (red–blue wavy line) which interferes with the incoming levitons (blue bumps) and is analogous to the local oscillator used in quantum optics. The red and blue arrows indicate the directions of electron–hole and leviton excitation propagation, respectively. Bicolour symbols indicate the outgoing states made of levitons mixed with the small flux of electron and hole excitations. Measuring the low-frequency current shot noise while varying V_R for different harmonics ($k = 0, 1$ and 2) provides a tomographic measurement of the leviton energy density matrix. **b**, Optical analogue. The local oscillator (LO) emits a monochromatic photon field, here of high amplitude, which interferes in the beam splitter with photons emitted by the source to be analysed. The photons detected provide a tomographic measurement of the photonic state.

but so far no such experiments have been done. Another approach is performing shot-noise measurements. This is the approach followed in the present experimental work. Photon shot noise has been exploited using adaptive homodyne tomography²⁷ to infer the unknown spectrotemporal properties of a single photon mode u . A large local oscillator field is mixed with the single photon in a beam splitter and its temporal shape is tuned to maximize the output noise, providing a classical copy of the unknown mode (Fig. 1b). For electrons, a large fermion field cannot be used but the principle of homodyne mixing can be kept. A recent proposal is to reverse the amplitude hierarchy¹⁶. Because it is the single-electron source corresponding to the maximum fermionic amplitude (as $|f\rangle = |1\rangle$), this is the local oscillator fermionic probe source whose amplitude has to be weak so that simple information can be provided by the noise variation at the beam-splitter output when mixing particles coming from the two sources. For simplicity, let us consider zero temperature. When a flux-quantized Lorentzian voltage pulse is applied to the left contact, a leviton is sent towards the QPC beam splitter. Repeating the experiment and recording the charge at the right contact gives the mean charge $\langle Q \rangle = De$ and the charge partition noise $\langle \Delta Q^2 \rangle = D(1 - D)e^2$. Now let us apply a d.c. voltage V_R to the right contact, rising the electrochemical potential by eV_R . The part of the leviton partition noise proportional to $\int_0^{eV_R} \tilde{\varphi}(\varepsilon)^* \tilde{\varphi}(\varepsilon) d\varepsilon$, corresponding to the energy range $\varepsilon \in [0, eV_R]$, is

c, Experiment principle in energy representation. All quantities are step functions of the energy with energy scale $h\nu_0$. Left: the energy dependence of the wavefunction of periodic levitons emitted from the left contact. Right: the energy distribution of the right reservoir when the voltage $V_R + V_{LO}(t)$ is applied. For small η_{LO} , an electron emitted with energy ε arrives at the QPC in a superposition of states of energy $\varepsilon \pm kh\nu_0$. Its interference with the levitons changes the leviton partition noise by a quantity proportional to $\eta_{LO}(\tilde{\varphi}(\varepsilon)\tilde{\varphi}(\varepsilon + kh\nu_0) - \tilde{\varphi}(\varepsilon - kh\nu_0)\tilde{\varphi}(\varepsilon))$, giving information on the non-diagonal energy density matrix. $f_R = |\varphi(\varepsilon)|^2$. **d**, The 6 GHz periodic Lorentzian voltage pulses are obtained from a frequency synthesizer by combining four harmonics with appropriate phases and amplitudes and sent to the left ohmic contact of the sample via 40 GHz d.c. transmission lines. To generate V_{LO} harmonics 1 and 2 are derived from the synthesizer, attenuated and sent to the right contact. A time delay is added using a phase shifter. Bias tees separate the high-frequency components from the d.c. voltage bias and the detected low-frequency current noise. The latter is converted into voltage fluctuations. A computer performs fast Fourier transform cross-correlations after cryogenic amplification and fast acquisition. Appropriate d.c. voltage biases are applied to the left and right contacts.

replaced by the partition noise of the right electrons reduced by the factor $1 - \int_0^{eV_R} \tilde{\varphi}(\varepsilon)^* \tilde{\varphi}(\varepsilon) d\varepsilon$ because of antibunching. Differentiating the noise with respect to V_R provides the diagonal part of the energy density matrix of the leviton $|\tilde{\varphi}(\varepsilon = eV_R)|^2$. Remarkably, very little modification is necessary to access the non-diagonal part $\langle \psi^+(\varepsilon' \neq \varepsilon)\psi(\varepsilon) \rangle$. Let us superimpose on the right d.c. voltage the a.c. voltage $V_{LO}(t) = \eta_{LO}(h\nu/e)\cos(2\pi\nu(t - \tau))$ with small amplitude $\eta_{LO} \ll 1$, frequency ν and time delay τ referred to the leviton emission time. This generates at the QPC a weak fermionic field made of electrons in a superposition of states with their initial energies shifted by $\pm h\nu$ and with probability amplitudes $\eta_{LO}e^{i2\pi\nu\tau}/2$ and $-\eta_{LO}e^{-i2\pi\nu\tau}/2$, respectively, the probability of keeping the initial energy being $1 - \eta_{LO}^2 \approx 1$ (Fig. 1c). Their mixing with the levitons at the beam splitter gives the antibunching probability amplitude

$$a(\eta_{LO}, \varepsilon) = \tilde{\varphi}^*(\varepsilon)(1 - \eta_{LO}^2) + \frac{1}{2}\tilde{\varphi}^*(\varepsilon + h\nu)\eta_{LO}e^{i2\pi\nu\tau} - \frac{1}{2}\tilde{\varphi}^*(\varepsilon - h\nu)\eta_{LO}e^{-i2\pi\nu\tau}$$

Differentiating the noise with respect to eV_R , and keeping only the a.c. contribution, gives a noise variation proportional to

$$\eta_{\text{LO}} \text{Re}[\tilde{\varphi}^*(eV_{\text{R}})\tilde{\varphi}(eV_{\text{R}} + \hbar\nu)e^{-i2\pi\nu\tau} - \tilde{\varphi}^*(eV_{\text{R}})\tilde{\varphi}(eV_{\text{R}} - \hbar\nu)e^{i2\pi\nu\tau}]$$

Fully quantum tomography can then be done by varying the voltage frequency and time delay τ , providing the real and imaginary parts of the first-order coherence¹⁶. Because the leviton wavefunction in the energy domain is real, we get

$$\frac{d\langle\Delta Q^2\rangle}{dV_{\text{R}}} \propto \eta_{\text{LO}} \cos(2\pi\nu\tau) [\tilde{\varphi}^*(eV_{\text{R}})\tilde{\varphi}(eV_{\text{R}} + \hbar\nu) - \tilde{\varphi}^*(eV_{\text{R}})\tilde{\varphi}(eV_{\text{R}} - \hbar\nu)] \quad (1)$$

Because only positive energy is concerned, equation (1) gives $\tilde{\varphi}(eV_{\text{R}})\tilde{\varphi}(eV_{\text{R}} + \hbar\nu)$ for $0 \leq eV_{\text{R}} < \hbar\nu$, $\tilde{\varphi}(eV_{\text{R}})\tilde{\varphi}(eV_{\text{R}} + \hbar\nu) - \tilde{\varphi}(eV_{\text{R}} - \hbar\nu)\tilde{\varphi}(eV_{\text{R}})$ for $\hbar\nu \leq eV_{\text{R}} < 2\hbar\nu$ and so on, providing a full determination of the energy density matrix.

In our experiment, single-electron levitons are injected at a frequency ν_0 by applying periodic Lorentzian voltage pulses from the left contact. The full-width at mid-height is $2w = 30$ ps and the period is $T = \nu_0^{-1} = 166$ ps. To extract the diagonal and off-diagonal parts of the energy density matrix, we follow the experimental approach discussed above with current noise measured instead of charge fluctuations. With $V_{\text{R}} = 0$ and in the absence of a.c. voltage, the partitioning of levitons generates the low-frequency current noise spectral density $S_I^0(0) = 4\nu_0 \langle\Delta Q^2\rangle = 4\nu_0 e^2 D(1 - D)$ where double counting of the mode from spin degeneracy is included. Because of the periodicity, $\langle\psi^\dagger(\varepsilon')\psi(\varepsilon)\rangle = \sum_k \delta(\varepsilon' - \varepsilon - kh\nu_0)\tilde{\varphi}^*(\varepsilon + kh\nu_0)\tilde{\varphi}(\varepsilon)$, where δ is the Dirac delta function and k is an integer. Only energies separated by multiple of $h\nu_0$ contribute to off-diagonal terms, and the frequency of $V_{\text{LO}}(t)$ must be $\nu = \pm\nu_0$. Because high harmonics give too small a signal, measurements are restricted to $k = 0, 1$ and 2 .

The diagonal part ($k = 0$) is obtained by measuring the difference between the partition noise of single-electron levitons injected at frequency ν_0 , using the periodic Lorentzian voltage $V_{\text{L}}(t)$, and that of continuously injected electrons provided by a constant voltage equal to the mean value of the voltage pulse $\bar{V}_{\text{L}}(t) = h\nu_0/e$. The zero-temperature noise difference is^{18,28}

$$\Delta S_I^{k=0}(\mu_{\text{R}}) = \frac{S_I^0}{h\nu_0} \left(\int_0^{\mu_{\text{R}}} (1 - \tilde{\varphi}(\varepsilon)^2) d\varepsilon - \int_{\mu_{\text{R}}}^{\infty} \tilde{\varphi}(\varepsilon)^2 d\varepsilon - |\mu_{\text{R}} - h\nu_0| \right) \quad (2)$$

On the right-hand side, the first term represents the noise of right electrons emitted at energy below $\mu_{\text{R}} = eV_{\text{R}}$, whose shot noise is reduced by antibunching with the leviton. The second term is the partition noise of levitons for energy above μ_{R} . It is followed by the subtracted d.c. shot noise. Because $\varphi(\varepsilon)$ varies by steps in $h\nu_0$ intervals, equation (2) displays linear variation by parts. Introducing the notation $\tilde{\varphi}(\varepsilon)^2 = \rho_{l,l}$ for $lh\nu_0 < \varepsilon < (l+1)h\nu_0$, this yields 0 for $\mu_{\text{R}} < 0$, $2(1 - \rho_{0,0})\mu_{\text{R}} + h\nu_0 - \mu_{\text{R}}$ for $0 < \mu_{\text{R}} < h\nu_0$, $2(1 - \rho_{0,0}) - 2\rho_{1,1}(\mu_{\text{R}} - h\nu_0)/h\nu_0$ for $h\nu_0 < \mu_{\text{R}} < 2h\nu_0$, and so on. The finite electron temperature $T_{\text{e}} \approx 35$ mK $\approx 0.12h\nu_0/k_{\text{B}}$ smoothens the singularities separating linear variations. A finite-temperature expression, obtained in a Floquet scattering approach^{29,30}, is given in Supplementary Information. Figure 2a shows measurements of $\Delta S_I^{k=0}$ versus μ_{R} for levitons (open blue circles). From the data we extract $\rho_{l,l}$ directly from a fit including the small-temperature effects. The solid red curve corresponds to the best four-parameter fit, from which the four first $\rho_{l,l}$ (that is, $l = 0$ to 3) are extracted (Supplementary Information). Self-heating including a small temperature increase (35 mK for $V_{\text{R}} \approx 0$ to 47 mK for $V_{\text{R}} \approx 3h\nu_0/e$) has been included. The $\rho_{l,l}$ are plotted in red in Fig. 2b. Comparison with theoretical values based on a Floquet scattering modelling (in black) is good within the experimental uncertainties.

For the measurement of the off-diagonal part ($k \neq 0$), we measure the noise difference when switching on and off the amplitude of

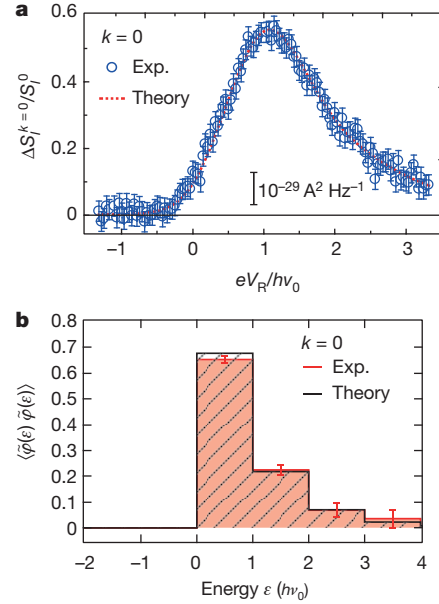


Figure 2 | Measurement of the diagonal part of the energy density matrix.

a, The data (blue circles) display the shot noise measured when applying a d.c. voltage $h\nu_0/e$, minus the shot noise measured when sending single-charge levitons at frequency ν_0 . The data are plotted versus the d.c. voltage V_{R} applied on the opposite contact. The red curve represents the best fit to equation (2), including finite temperature, from which the parameters $\rho_{l,l}$ are extracted. The temperature T_{e} is ~ 35 mK at $V_{\text{R}} = 0$ and ~ 48 mK at $V_{\text{R}} = 3h\nu_0/e$. The transmission is $D = 0.19$ and $S_I^0 = 9.48 \times 10^{-29} \text{ A}^2 \text{ Hz}^{-1}$. Error bars, s.e.m. calculated from the cross-correlation noise spectrum standard deviation in a 800 kHz bandwidth and after $n \approx 100,000$ acquisitions. **b**, Diagonal part of the energy density matrix versus energy using the $\rho_{l,l}$ extracted from the fit. Comparison with theory for levitons is in black. Error bars, s.e. calculated from the best fit.

$V_{\text{LO}}(t) = \eta_{\text{LO}}(h\nu/e)\cos(2\pi\nu(t - \tau))$ superimposed on the right d.c. voltage V_{R} and synchronized with the Lorentzian pulses. The zero-temperature shot noise difference is

$$\Delta S_I^k(\mu_{\text{R}}) = \frac{S_I^0}{h\nu_0} 2k\eta_{\text{LO}} \cos(2\pi k\nu_0\tau) \times \int_0^{\mu_{\text{R}}} (\tilde{\varphi}(\varepsilon)\tilde{\varphi}(\varepsilon + kh\nu_0) - \tilde{\varphi}(\varepsilon)\tilde{\varphi}(\varepsilon - kh\nu_0)) d\varepsilon \quad (3)$$

This expression also varies linearly by parts because $\tilde{\varphi}(\varepsilon)\tilde{\varphi}(\varepsilon + kh\nu_0) = \rho_{l,l+k}$ for $lh\nu_0 < \varepsilon < (l+1)h\nu_0$ is a step function of the energy. For example, for $k = 1$ the integral in equation (3) is 0 for $\mu_{\text{R}} < 0$, $\rho_{0,1}\mu_{\text{R}}/h\nu_0$ for $0 < \mu_{\text{R}} < h\nu_0$, $\rho_{0,1} - \rho_{1,2}(\mu_{\text{R}} - h\nu_0)/h\nu_0$ for $h\nu_0 < \mu_{\text{R}} < 2h\nu_0$, and so on. Finite-temperature expressions are given in Supplementary Information.

Figure 3a, b shows measurements of ΔS_I^k versus τ for $k = 1$, $\eta_{\text{LO}} = 0.095$ and $k = 2$, $\eta_{\text{LO}} = 0.0425$ at fixed $\mu_{\text{R}} \approx h\nu_0$. Clear oscillations are observed. The delay period in Fig. 3b is twice that observed in Fig. 3a. Before going further, we emphasize that observation of these oscillations demonstrates off-diagonal coherence in an electronic orbital state, here a leviton. No experiment has given direct evidence of off-diagonal coherence of electrons in a conductor yet. We now compare the off-diagonal terms extracted from the measurements with those predicted for a leviton. Figure 3c, d shows the experimental variation of ΔS_I^k with μ_{R} for $k = 1$ and 2 (open blue circles), fixing $\tau = 0$ to maximize the signal. In both plots, the solid black curves represent the theoretical signal expected for a leviton. The red dashed curves are the best fits, from which the off-diagonal terms $\rho_{l,l+k}$ can be extracted. For $k = 2$ a four-parameter fit is used and $l = 0$ to 3, and for $k = 1$ a five-parameter fit is used because a finite, but very small, hole excitation contribution

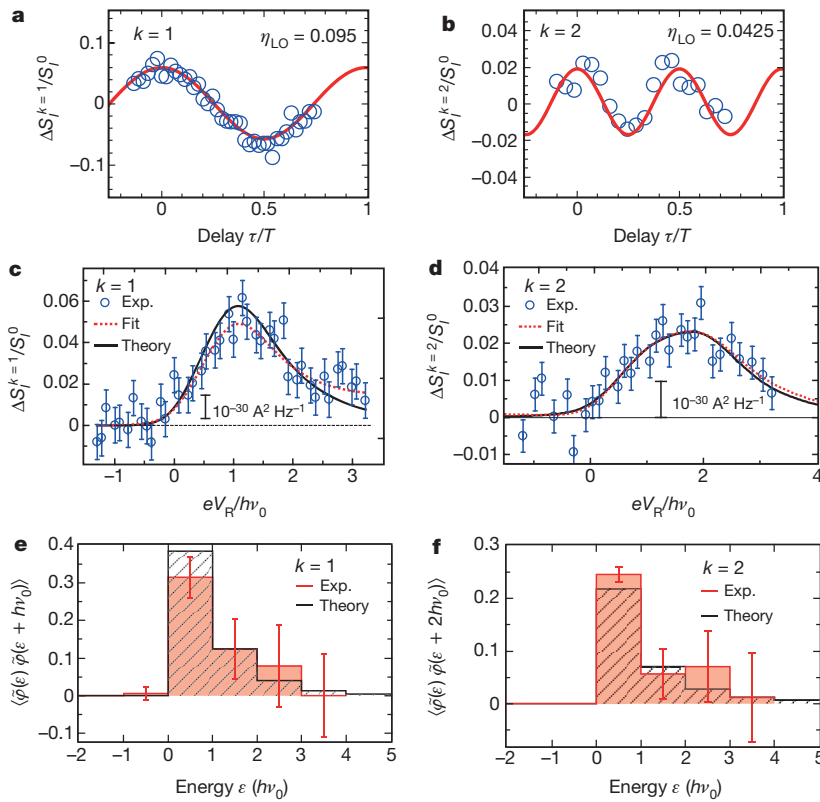


Figure 3 | Off-diagonal part of energy density matrix. **a, b,** Weak sine-wave voltages of frequencies and amplitudes ν_0 and $\eta_{LO} = 0.095$ (**a**) and $2\nu_0$ and $\eta_{LO} = 0.0425$ (**b**) superimposed on a fixed d.c. voltage $eV_R = h\nu_0$ applied on the right contact. The data (blue circles) show the noise difference obtained by switching on and off the sine-wave amplitude, plotted versus the time delay τ between the sine-wave and the periodic Lorentzian pulses. The oscillation of the noise with τ and the period doubled in **b** relative to **a** is the signature of off-diagonal coherence. The red curves are sinusoidal fits. **c, d,** Same shot noise difference measurements (blue circles) but plotted versus the d.c. right voltage with $\tau = 0$, chosen to maximize the signal. T_e is ~ 35 mK at $V_R = 0$ and ~ 48 mK at $V_R = 3h\nu_0/e$. The red curves are the best fit of equation (3) with $\rho_{l,l+k}$ as free parameters, including finite-temperature effects, with $k = 1$ (**c**) and $k = 2$ (**d**). The black curves are comparison with theory. Error bars, s.e.m. calculated as in Fig. 2a but with n four times larger. **e, f,** Plot of $\tilde{\varphi}(\epsilon)\tilde{\varphi}(\epsilon + h\nu_0)$ (red) using the parameters $\rho_{l,l+k}$ extracted from the best fits of **c** and **d** with $k = 1$ (**e**) and $k = 2$ (**f**), respectively. Expected values from theory are shown in black. Error bars, s.e. calculated from the best fit.

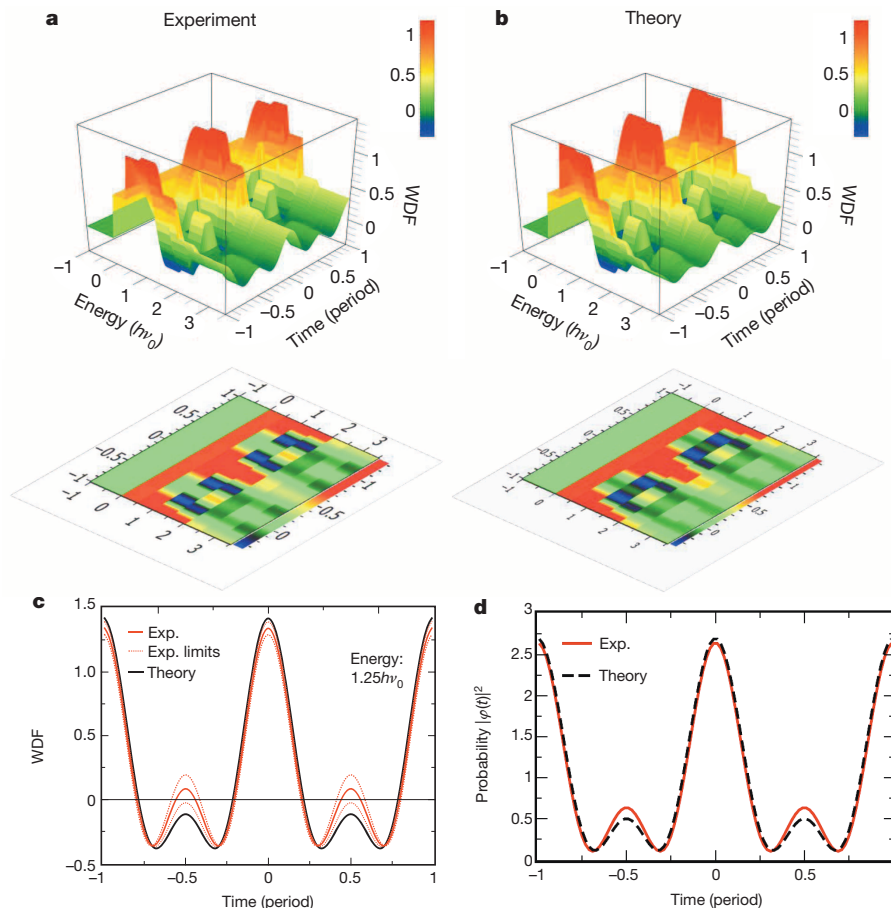


Figure 4 | Wigner function and leviton wavefunction in the time domain. **a,** The zero-temperature Wigner distribution function of the levitons is reconstructed from the $\rho_{l,l+k}$ values obtained from tomography. The WDF is limited to the first two temporal harmonics and to the energy domain $\epsilon < 3.5h\nu_0$. Negative parts are characteristic of a wavefunction which cannot be described by a classical probability distribution in the time and energy domains. **b,** Theoretical WDF. The full WDF has been truncated to the first two harmonics and to the energy range probed experimentally, for better comparison. The two-dimensional colour plots are shown to better display the region of negative values of the WDF. **c,** Cut of the experimental and theoretical WDFs at energy $\epsilon = 1.25h\nu_0$ (for energies $1.5h\nu_0 > \epsilon > h\nu_0$ both the first and second harmonics contribute to the WDF). Using the experimental errors in Fig. 3e, f, we have generated the two red dashed curves between which the experimental WDF is expected to lie. Robust negative values of the WDF are found, a hallmark of interference effects. **d,** Projecting the WDF along the energy direction gives access to the temporal variation of the electron probability distribution of the leviton $|\varphi(t)|^2$. The experimental data are (red) are compared with the theoretical truncated WDF (black).

($l = -1$) was found necessary to improve the fit. They are shown in Fig. 3e, f. The errors bars are obtained from the standard error of the fit. They are much larger than those in Fig. 2b because the measured noise is about ten times smaller.

The noise oscillations with phase in Fig. 3a, b and the noise variations with voltage in Fig. 3c, d are well reproduced by a model including only the 35–50 mK thermal rounding. Supplementary Fig. 2 provides additional tomographic noise measurements performed on two-electron sine-wave pulses. Here again good agreement is found between data and theory. Deviation arising from decoherence effects due to electron–electron interaction is not expected here because the two-dimensional Landau quasi-particle lifetime is longer than the (50 mK) thermal time for the typical energies probed here (Supplementary Information).

From the experimental values of $\tilde{\varphi}(\varepsilon)\tilde{\varphi}(\varepsilon + khv_0)$ obtained above we can now reconstruct part of the zero-temperature Wigner distribution function¹⁷ (WDF) of the levitons: $W(\tilde{t}, \varepsilon) = \int_{-\infty}^{+\infty} \langle \psi^+(\varepsilon + \delta/2)\psi(\varepsilon - \delta/2) \rangle e^{-i\delta\tilde{t}/h} d\delta$. The WDF gives the full information on the complex wavefunction. It has been calculated for a single Lorentzian pulse¹⁴ and for periodic pulses¹⁷. In the latter case, the WDF is also periodic. Its expression in terms of the quantities $\rho_{l,l+k}$ is given in Supplementary Information. Figure 4a and Fig. 4b respectively give three-dimensional plots of the experimental and theoretical WDFs. Because the experimental WDF is limited to the first two harmonics and the energy range $-hv_0 \leq \varepsilon < 3.5hv_0$, the theoretical WDF is truncated accordingly for comparison. To better show the agreement between experiment and theory, we provide in Fig. 4c a cut of the WDF at energy $\varepsilon = 1.25hv_0$ (red, experiment; black, theory). Using the experimental errors of $\rho_{l,l+k}$ ($l = 0$) in Fig. 3e, f, we have generated the two red dashed curves in Fig. 4c between which the experimental WDF is expected to be located. For reduced times t/T around ~ 0.3 and 0.7 we clearly find robust regions of negative WDF, a hallmark of a quantum state.

Finally, as a particular application of the WDF, the integration over energy ε gives the probability $|\varphi(t)|^2$. Its experimental (red solid curve) and theoretical (black dashed curve) variations with time are given in Fig. 4d. The excellent agreement validates our experimental determination of the WDF from the tomographic noise procedure of ref. 16. This provides better information on the temporal shape of the wavefunction than the convolution $|\langle \varphi(t) | \varphi(t + \tau) \rangle|^2$ given by the electronic Hong–Ou–Mandel method used for levitons in ref. 18 or for single electrons in ref. 31.

Using the known leviton state, we have demonstrated that the quantum wave tomography of itinerant electrons can be done experimentally using shot noise. Probing the off-diagonal energy density matrix opens the possibility of studying more complex electron quantum states such as entangled electron quantum bits in ballistic conductors or complex electron quantum states built by interaction. Examples of the latter include Kondo electronic states, fractional quantum Hall quasi-particles or electron Andreev pairs built by superconducting proximity effects.

Received 2 June; accepted 29 August 2014.

- Vogel, K. & Risken, H. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A* **40**, 2847–2849 (1989).
- Leibfried, D. *et al.* Experimental determination of the motional quantum state of a trapped atom. *Phys. Rev. Lett.* **77**, 4281–4285 (1996).
- Kurtsiefer, C., Pfau, T. & Mlynek, J. Measurement of the Wigner function of an ensemble of helium atoms. *Nature* **386**, 150–153 (1997).
- Dunn, T. J., Walmsley, I. A. & Mukamel, S. Experimental determination of the quantum-mechanical state of a molecular vibrational mode using fluorescence tomography. *Phys. Rev. Lett.* **74**, 884–887 (1995).

- Shulman, M. D. *et al.* Demonstration of entanglement of electrostatically coupled singlet-triplet qubits. *Science* **336**, 202–205 (2012).
- Medford, J. *et al.* Self-consistent measurement and state tomography of an exchange-only spin qubit. *Nature Nanotechnol.* **8**, 654–659 (2013).
- Smithey, D. T., Beck, M., Raymer, M. G. & Faridani, A. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomography: application to squeezed states and the vacuum. *Phys. Rev. Lett.* **70**, 1244–1247 (1993).
- Breitenbach, G., Schiller, S. & Mlynek, J. Measurement of the quantum states of squeezed light. *Nature* **387**, 471–475 (1997).
- Lvovsky, A. I. *et al.* Quantum state reconstruction of the single-photon Fock state. *Phys. Rev. Lett.* **87**, 050402 (2001).
- Bertet, P. *et al.* Direct measurement of the Wigner function of a one-photon Fock state in a cavity. *Phys. Rev. Lett.* **89**, 200402 (2002).
- Houck, A. A. *et al.* Generating single microwave photons in a circuit. *Nature* **449**, 328–331 (2007).
- Hofheinz, M. *et al.* Synthesizing arbitrary quantum states in a superconducting resonator. *Nature* **459**, 546–549 (2009).
- Eichler, C. *et al.* Experimental tomographic state reconstruction of itinerant microwave photons. *Phys. Rev. Lett.* **106**, 220503 (2011).
- Haack, G., Moskalets, M. & Büttiker, M. Glauber coherence of single-electron sources. *Phys. Rev. B* **87**, 201302(R) (2013).
- Samuelsson, P. & Büttiker, M. Quantum state tomography with quantum shot noise. *Phys. Rev. B* **73**, 041305(R) (2006).
- Grenier, C. *et al.* Single-electron quantum tomography in quantum Hall edge channels. *New J. Phys.* **13**, 093007 (2011).
- Ferraro, D. *et al.* Wigner function approach to single electron coherence in quantum Hall edge channels. *Phys. Rev. B* **88**, 205303 (2013).
- Dubois, J. *et al.* Minimal-excitation states for electron quantum optics using levitons. *Nature* **502**, 659–663 (2013).
- Levitov, L. S., Lee, H. & Lesovik, G. Electron counting statistics and coherent states of electric current. *J. Math. Phys.* **37**, 4845–4886 (1996).
- Lebedev, A. V., Lesovik, G. V. & Blatter, G. Generating spin-entangled electron pairs in normal conductors using voltage pulses. *Phys. Rev. B* **72**, 245314 (2005).
- Keeling, J., Klich, I. & Levitov, L. Minimal excitation states of electrons in one-dimensional wires. *Phys. Rev. Lett.* **97**, 116403 (2006).
- Beenakker, C. W. J., Emary, C., Kindermann, M. & van Velsen, J. L. Proposal for production and detection of entangled electron-hole pairs in a degenerate electron gas. *Phys. Rev. Lett.* **91**, 147901 (2003).
- Beenakker, C. W. J., Titov, M. & Trauzettel, B. Optimal spin-entangled electron-hole pair pump. *Phys. Rev. Lett.* **94**, 186804 (2005).
- Sherkunov, Y. B., d’Ambrumenil, N., Samuelsson, P. & Büttiker, M. Optimal pumping of orbital entanglement with single-particle emitters. *Phys. Rev. B* **85**, 081108 (2012).
- Yamamoto, M. *et al.* Electrical control of a solid-state flying qubit. *Nature Nanotechnol.* **7**, 247–251 (2012).
- Brantut, J. P. *et al.* Conduction of ultracold fermions through a mesoscopic channel. *Science* **337**, 1069–1071 (2012).
- Polycarpou, C., Cassemiro, K. N., Venturi, G., Zavatta, A. & Bellini, M. Adaptive detection of arbitrarily shaped ultrashort quantum light states. *Phys. Rev. Lett.* **109**, 053602 (2012).
- Dubois, J. *et al.* Integer and fractional charge Lorentzian voltage pulses analyzed in the framework of photon-assisted shot noise. *Phys. Rev. B* **88**, 085301 (2013).
- Pedersen, M. H. & Büttiker, M. Scattering theory of photon-assisted electron transport. *Phys. Rev. B* **58**, 12993 (1998).
- Dasenbrook, D., Flindt, C. & Büttiker, M. Floquet theory of electron waiting times in quantum-coherent conductors. *Phys. Rev. Lett.* **112**, 146801 (2014).
- Bocquillon, E. *et al.* Coherence and indistinguishability of single electrons emitted by independent sources. *Science* **339**, 1054–1057 (2013).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge the ERC Advanced Grant 228273 MeQuaNo and thank P. Jacques for technical help, P. Pari, P. Forget and M. de Combarieu for cryogenic support, and P. Degiovanni and C. Grenier for discussions improving the manuscript.

Author Contributions D.C.G. designed the project. T.J. and P.R. made the measurements and did the data analysis. B.R. contributed to the data analysis. P.R., T.J., B.R. and D.C.G. wrote the article. The sample was provided by Y.J. on wafer from A.C.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.C.G. (christian.glatli@cea.fr).

Room-temperature magnetic order on zigzag edges of narrow graphene nanoribbons

Gábor Zsolt Magda¹, Xiaozhan Jin², Imre Hagymási^{3,4}, Péter Vancsó¹, Zoltán Osváth¹, Péter Nemes-Incze¹, Chanyong Hwang², László P. Biró¹ & Levente Tapasztó¹

The possibility that non-magnetic materials such as carbon could exhibit a novel type of s - p electron magnetism has attracted much attention over the years, not least because such magnetic order is predicted to be stable at high temperatures¹. It has been demonstrated that atomic-scale structural defects of graphene can host unpaired spins^{2,3}, but it remains unclear under what conditions long-range magnetic order can emerge from such defect-bound magnetic moments. Here we propose that, in contrast to random defect distributions, atomic-scale engineering of graphene edges with specific crystallographic orientation—comprising edge atoms from only one sub-lattice of the bipartite graphene lattice—can give rise to a robust magnetic order. We use a nanofabrication technique⁴ based on scanning tunnelling microscopy to define graphene nanoribbons with nanometre precision and well-defined crystallographic edge orientations. Although so-called ‘armchair’ ribbons display quantum confinement gaps, ribbons with the ‘zigzag’ edge structure that are narrower than 7 nanometres exhibit an electronic bandgap of about 0.2–0.3 electronvolts, which can be identified as a signature of interaction-induced spin ordering along their edges. Moreover, upon increasing the ribbon width, a semiconductor-to-metal transition is revealed, indicating the switching of the magnetic coupling between opposite ribbon edges from the antiferromagnetic to the ferromagnetic configuration. We found that the magnetic order on graphene edges of controlled zigzag orientation can be stable even at room temperature, raising hopes of graphene-based spintronic devices operating under ambient conditions.

The intrinsic magnetism of graphite has a long and controversial history¹. The origin of the measured magnetic signal is generally attributed to atomic-scale structural defects locally breaking the sub-lattice balance of the bipartite hexagonal lattice^{5,6}. However, unambiguous identification of the structural sources of the measured magnetic signal has proved challenging because the sources are buried inside the bulk of the material. The isolation of single graphene layers⁷ opens up new prospects in this direction^{8,9} because their atomic structure is fully accessible for imaging and controlled modification. In particular, the graphene edges of a specific crystallographic orientation comprising carbon atoms from only one sub-lattice of the bipartite hexagonal lattice (zigzag) are predicted to host magnetic order¹⁰, in striking contrast to graphene edges that incorporate an equal number of carbon atoms from both sub-lattices (armchair).

The strong influence of edge orientation on the electronic structure of graphene nanoribbons was predicted¹¹ long before graphene was isolated⁷. However, a systematic experimental verification of this fundamental prediction is still lacking, owing to very limited control over the edge orientation of the graphene ribbons investigated. Opening of a gap inversely proportional to the ribbon width has been experimentally demonstrated by electrical transport measurements^{12,13}; however, no evidence on the effect of edge orientation has been revealed. This is attributed mainly to the random edge orientations and the presence of

a considerable, nanometre-scale edge roughness, suppressing orientation effects¹⁴. Scanning tunnelling microscopy (STM) study of irregularly shaped graphene ribbons revealed that structures with more zigzag edges display smaller bandgaps than those with more armchair edges¹⁵. This is clearly indicative of edge-specific physics; however, the lack of orientation control did not allow a more systematic insight. In theory, there is a broad consensus that graphene nanoribbons with armchair edges are semiconductors, owing to the quantum confinement of the charge carriers, while zigzag ribbons host peculiar flat-band edge states^{10,16}. In the picture neglecting electron–electron interactions, these edge states localized near the Fermi level render all zigzag graphene nanoribbons metallic. However, such one-dimensional metallic edge states with a high local density of states at the Fermi level become unstable upon electron–electron interactions. To lower the energy of the system, it is favourable to split the flat band (open a gap) by ordering the spins along the two ribbon edges with antiferromagnetic coupling between opposite edges. Consequently, the emergence of magnetic order is closely linked to altering the electronic structure of the ribbons, through opening a bandgap in the otherwise metallic zigzag nanoribbons^{10,16,17}. This enables us to detect the signature of edge magnetism on individual graphene nanostructures by investigating their electronic structure (directly measuring magnetic signals would require a macroscopic amount of such ribbons). This edge-state magnetism and the associated bandgap opening in zigzag ribbons is consistently predicted by various theoretical models, including first principles density functional theory¹⁶, and mean-field-theory-based Hubbard¹⁸ and quantum Monte Carlo calculations¹⁹, indicating that edge magnetism is a robust property of zigzag graphene nanoribbons, not sensitive to the specific details of the models. However, the stability of the magnetic order on real graphene edges and experimental conditions is strongly debated. Experimental indications that edge magnetism can indeed occur at low temperatures (7 K) have been provided by tunnelling spectroscopy measurements on graphene ribbons obtained by unzipping carbon nanotubes²⁰. However, the random orientation of the edges and the influence of a possible strong edge–substrate hybridization²¹ did not allow edge magnetism in graphene to be properly characterized. Although the magnetic order is expected to persist to some extent on zigzag segments of randomly oriented graphene edges, the mixing of different edge types is expected to substantially weaken the effect^{19,22}. Therefore, the lack of experimental control over edge orientation seems to be one of the main reasons that the magnetic graphene edge states consistently predicted by various theoretical models have remained experimentally so elusive.

To realize graphene nanoribbons with precisely controlled crystallographic edge orientations, we used a nanofabrication technique based on STM⁴. In atomic-resolution images, the crystallographic directions of the graphene lattice can be identified and matched with the desired cutting directions. The direct cutting of graphene is done by beam-induced chemical etching activated by the sub-nanometre-wide channel of tunnelling electrons, locally breaking the carbon–carbon bonds

¹Research Centre for Natural Sciences, Institute of Technical Physics and Materials Science, Nanotechnology Department, 2D Nanoelectronics “Lendület” Research Group, Budapest 1121, Hungary.

²Korea Research Institute of Standards and Science, Center for Nanometrology, Daejeon 305-340, South Korea. ³Wigner Research Centre for Physics, Institute of Solid State Physics and Optics, Strongly Correlated Systems “Lendület” Research Group, Budapest 1121, Hungary. ⁴Department of Theoretical Physics, University of Szeged, Szeged 6720, Hungary.

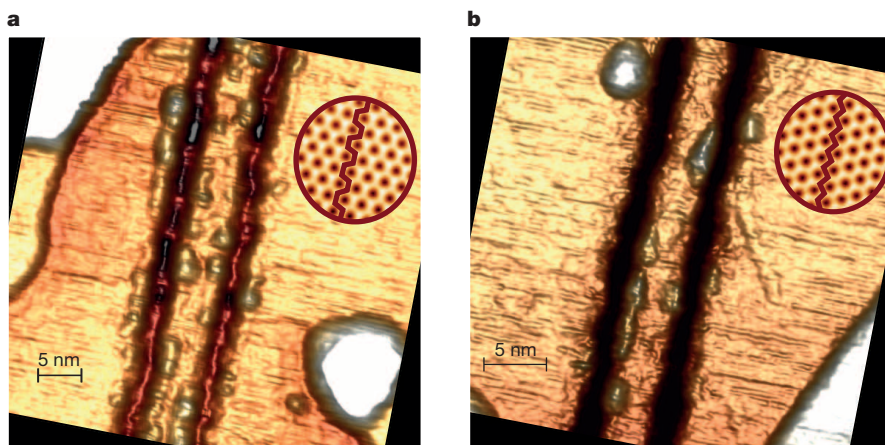


Figure 1 | Fabrication of graphene nanoribbons with precisely defined crystallographic edge orientations. STM image (500 mV, 0.8 nA) of a 5-nm-wide graphene nanoribbon with armchair edge orientation (a) and STM image of a 6.5-nm-wide ribbon with edges of precisely zigzag orientation (300 mV, 2 nA) (b) patterned by scanning tunnelling lithography in a graphene sheet

underneath the atomically sharp STM tip apex. More details on the nanolithographic process are given in the Methods.

Using the STM nanolithographic technique, we defined nanoribbons into graphene sheets grown by chemical vapour deposition and transferred onto gold (Au) substrates with large Au (111) terraces. Au (111) has been predicted to be one of the few substrates preserving edge magnetism in supported graphene ribbons²¹. Nanoribbons with pre-defined armchair or zigzag edge orientation and widths ranging down to 3 nm were defined (Fig. 1). The as-fabricated ribbons display regular edges of subnanometre roughness. The bright protrusions near the edges in Fig. 1 are due to a known imaging instability²³, rather than edge defects or impurities. Upon optimizing the STM imaging, we found the edges defined by STM lithography to be straight and close to atomically smooth, free of detectable impurities, reconstructions or curvature (Extended Data Fig. 1).

Tunnelling spectra acquired on armchair ribbons displayed semi-conducting current–voltage (I – V) characteristics, in contrast to spectra taken on graphene outside the ribbons (see Extended Data Fig. 2). The measured bandgap as a function of the corresponding ribbon width is plotted in Fig. 2a. A clear inverse proportionality is revealed, in excellent quantitative agreement with our calculations based on the Hubbard model (equation (1)) considering hydrogen-saturated and relaxed edges, as well as first-principles DFT calculations¹⁶. The theoretical data points correspond to three classes of armchair ribbons: with $3n$, $3n + 1$

and $3n + 2$ (for $n = 1, 2, 3, \dots$) rows of carbon dimers across their width. The measurement error and the emerging atomic-scale edge irregularities do not allow us to distinguish clearly between ribbons belonging to the $3n$ and $3n + 1$ classes; however, armchair ribbon segments with no detectable gap have been observed in a few cases, which probably belong to the $3n + 2$ class. Consequently, our measurements systematically verify the predicted mechanism of bandgap opening caused by the lateral confinement of the charge carriers in ribbons of precisely armchair edge orientation.

A strikingly different behaviour is revealed in zigzag graphene nanoribbons (Fig. 2b). For zigzag graphene nanoribbons narrower than 7 nm, we observed a fairly large bandgap of about 200–300 meV. As discussed above, the broadly predicted origin of bandgap opening in zigzag ribbons is electron–electron interactions and the predicted many-body ground state implies that the edges are magnetized^{16,18,19}.

However, in contrast to first principles theoretical predictions¹⁶, the measured gap suddenly vanishes for zigzag ribbons wider than 8 nm. We attribute this discrepancy to the fact that these calculations were performed at zero temperature and without doping, whereas the experimental data was acquired at room temperature and finite doping (the corresponding shift of the Fermi level ΔE_F is about 50–100 meV; see Extended Data Fig. 2c). When the effects of temperature and doping were not taken into consideration, we found the ground state to be always semiconducting (antiferromagnetic) and no steep transition to a metallic

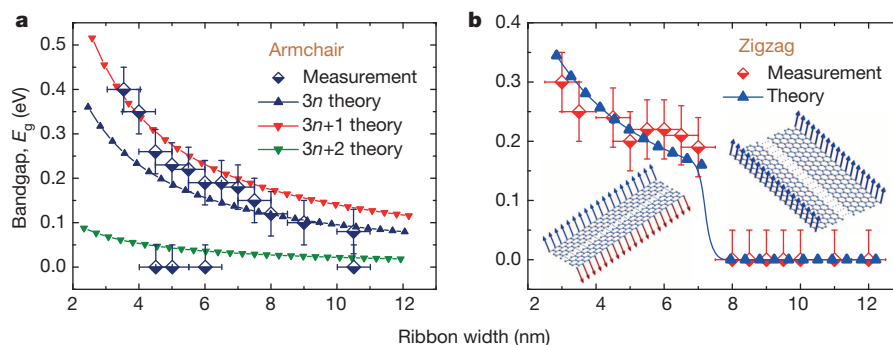


Figure 2 | Edge-specific electronic and magnetic properties of graphene nanoribbons. The bandgap measured by tunnelling spectroscopy as a function of ribbon width in armchair (a) and zigzag (b) ribbons. Armchair ribbons display a quantum confinement gap inversely proportional to their width (where $n = 1, 2, 3, \dots$; see main text). In zigzag ribbons the band structure is governed by the emerging edge magnetism and a sharp semiconductor

(antiferromagnetic) to metal (ferromagnetic) transition is revealed. Theoretical data points were calculated using the mean field Hubbard model (continuous lines are only guides to the eye). Error bars of the measured gap values originate from thermal broadening and substrate effects; error bars for ribbon widths (horizontal) arise from tip convolution effects.

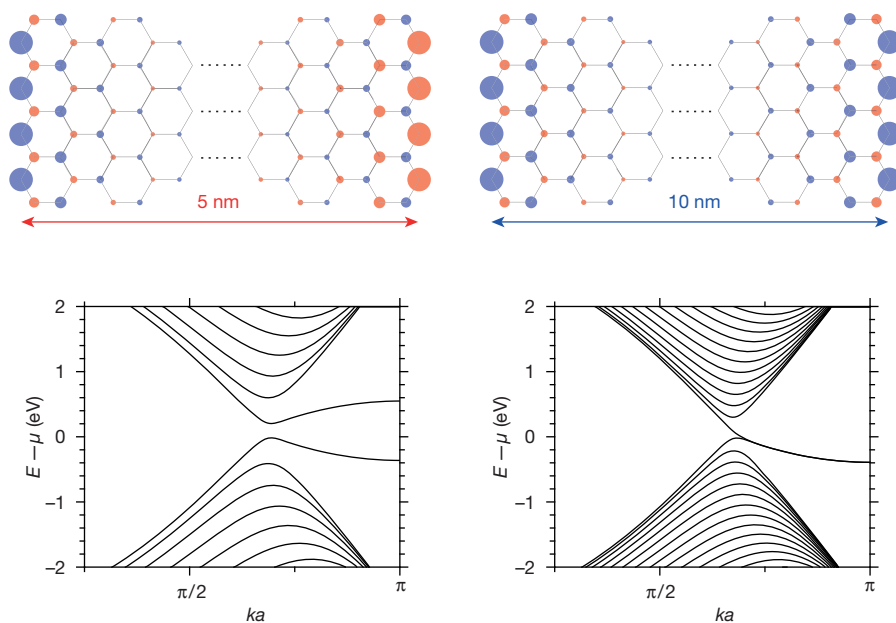


Figure 3 | Correlating electronic and magnetic properties of zigzag graphene nanoribbons. Spin density distribution (\uparrow , blue; \downarrow , red) in 5-nm-wide (left) and 10-nm-wide (right) zigzag graphene nanoribbons calculated in the mean field Hubbard model for $T = 300$ K and $\Delta E_F \approx 100$ meV. The lower panels display the corresponding band structure, clearly indicating that narrow zigzag ribbons are antiferromagnetic semiconductors, whereas the wider (>8 nm) zigzag ribbons display a ferromagnetic inter-edge coupling and no bandgap. ka is wavenumber times the lattice constant.

state is expected. To understand the origin of the experimentally observed, sharp semiconductor–metal transition, we performed calculations based on the mean-field approximation of the Hubbard Hamiltonian that accounts for both finite temperature and doping:

$$\mathcal{H} = - \sum_{ij} t_{ij} \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma} + U \sum_j \hat{n}_{j\uparrow} \hat{n}_{j\downarrow} - \mu \hat{N} \quad (1)$$

where the first term is the nearest-neighbour tight-binding Hamiltonian, and the second term stands for the onsite Coulomb repulsion. The effects of temperature and doping have been included using the grand canonical ensemble (third term), where μ is the chemical potential, and \hat{N} is the particle number operator. The electron density and the chemical potential were determined self-consistently. (For more details on calculations see the Methods.)

The main result of our extended theoretical model is that it reproduces the observed steep semiconductor–metal transition, quantitatively accounting for both the critical ribbon width (about 7 nm) at which the transition occurs and the magnitude of the measured bandgaps, using the strength of the on-site repulsion U as the single unknown parameter (Fig. 2b). By calculating the spin density distribution for both semiconducting (<7 nm wide) and metallic (>8 nm wide) ribbons the origin of the observed semiconductor–metal transition can be identified as a transition from an antiferromagnetic (semiconducting) state, where magnetic moments on opposite ribbon edges are aligned antiparallel, to a ferromagnetic (metallic) state, with parallel spin alignment on opposite edges (Fig. 3). The metallic nature of zigzag graphene nanoribbons with ferromagnetic inter-edge coupling has already been predicted¹⁸, together with a complex magnetic phase diagram as a function of doping. Our theoretical results are consistent with these findings in the $T \rightarrow 0$ K limit. We note that the observation of the gap opening alone in zigzag graphene nanoribbons would not unambiguously indicate edge magnetism, given that reconstructed zigzag edge configurations (for example, z_{211}) leading to semiconducting but non-magnetic zigzag graphene nanoribbons have been predicted²⁴. However, only magnetic zigzag edges can account for both gap opening and the observed semiconductor–metal transition. This, together with the quantitative agreement between our calculations and experiments provide indirect but compelling evidence that magnetic order emerges on zigzag graphene edges of precisely engineered crystallographic orientation. Moreover, it is remarkable that the signature of edge magnetism can be experimentally detected at room temperature, although this is in agreement with the expected high-temperature stability of s - p electron magnetism^{25,26},

as well as experiments reporting hints of room-temperature magnetism in defective graphitic samples^{5,27–29}.

Since the only free parameter used in the calculations for zigzag ribbons is the magnitude of the on-site repulsion U , the best quantitative agreement with our experimental findings provides an experimental estimate for the strength of the electron–electron interaction in graphene of $U = 3.24$ eV. Theoretical works have predicted the magnitude of U to range from 2 eV to 6 eV (see, for example, ref. 30). The experimentally estimated value of U is even more reliable, because a single U value can reproduce both the magnitude of the gaps and the semiconductor–metal transition ribbon width.

That the experimental results can be quantitatively interpreted by theoretical calculations on perfect zigzag (or armchair) edges confirms the experimentally found high edge quality. However, deviations from the ideal edge structure can be present on the atomic scale. To estimate the effect of atomic-scale edge irregularities, we performed calculations on a model system having a high density of such defects (Extended Data Fig. 3). The results indicate that, even in the presence of a high defect density, for overall zigzag oriented edges the qualitative picture holds but the strength of the effect (gap size, spin polarization) decreases. To fit the experimental data, the value of electron–electron interactions estimated based on the ideal zigzag edge has to be increased from $U = 3.24$ eV to $U = 4.32$ eV.

Our findings demonstrate that graphene nanoribbons display strongly edge-orientation-specific behaviour and that engineering the crystallographic orientation of graphene edges allows us unprecedented control over both electronic and magnetic properties of graphene nanostructures. This opens the way towards the realization of electronic and spintronic devices that are based on robust quantum mechanical effects and that can operate at room temperature.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 14 April; accepted 1 September 2014.

1. Magnetism, C.-B. *An Overview of Metal Free Carbon-Based Compounds and Materials* (eds Makarova, T. & Palacio, F.) (Elsevier, 2005).
2. Nair, R. R. *et al.* Spin-half paramagnetism in graphene induced by point defects. *Nature Phys.* **8**, 199–202 (2012).
3. Yazayev, O. V. & Helm, L. Defect-induced magnetism in graphene. *Phys. Rev. B* **75**, 125408 (2007).
4. Tapasztó, L., Dobrik, G., Lambin, P. & Biro, L. P. Tailoring the atomic structure of graphene nanoribbons by scanning tunneling microscopy lithography. *Nature Nanotechnol.* **3**, 397–401 (2008).

5. Esquinazi, P. *et al.* Induced magnetic ordering by proton irradiation in graphite. *Phys. Rev. Lett.* **91**, 227201 (2003).
6. Palacios, J. J. *et al.* Vacancy-induced magnetism in graphene and graphene ribbons. *Phys. Rev. B* **77**, 195428 (2008).
7. Novoselov, K. S. *et al.* Two-dimensional gas of massless Dirac fermions in graphene. *Nature* **438**, 197–200 (2005).
8. Wang, Y. *et al.* Room-temperature ferromagnetism of graphene. *Nano Lett.* **9**, 220–224 (2009).
9. Chen, L. *et al.* Towards intrinsic magnetism of graphene sheets with irregular zigzag edges. *Sci. Rep.* **3**, 2599 (2013).
10. Nakada, K. *et al.* Edge state in graphene ribbons: nanometer size effect and edge shape dependence. *Phys. Rev. B* **54**, 17954–17961 (1996).
11. Fujita, M. *et al.* Peculiar localized state at zigzag graphite edge. *J. Phys. Soc. Jpn* **65**, 1920–1923 (1996).
12. Han, M. *et al.* Energy band-gap engineering of graphene nanoribbons. *Phys. Rev. Lett.* **98**, 206805 (2007).
13. Li, X. *et al.* Chemically derived, ultrasoft graphene nanoribbon semiconductors. *Science* **319**, 1229–1232 (2008).
14. Mucciolo, E. R. *et al.* Conductance quantization and transport gaps in disordered graphene nanoribbons. *Phys. Rev. B* **79**, 075407 (2009).
15. Ritter, K. A. & Lyding, J. W. The influence of edge structure on the electronic properties of graphene quantum dots and nanoribbons. *Nature Mater.* **8**, 235–242 (2009).
16. Son, Y. W. *et al.* Energy gaps in graphene nanoribbons. *Phys. Rev. Lett.* **97**, 216803 (2006).
17. Son, Y. W. *et al.* Half-metallic graphene nanoribbons. *Nature* **444**, 347–349 (2006).
18. Jung, J. & Macdonald, A. H. Carrier density and magnetism in graphene zigzag nanoribbons. *Phys. Rev. B* **79**, 235433 (2009).
19. Golor, M. *et al.* Quantum Monte Carlo studies of edge magnetism in chiral graphene nanoribbons. *Phys. Rev. B* **87**, 155441 (2013).
20. Tao, C. *et al.* Spatially resolving edge states of chiral graphene nanoribbons. *Nature Phys.* **7**, 616–620 (2011).
21. Li, Y. *et al.* Electronic and magnetic properties of zigzag graphene nanoribbons on the (111) Surface of Cu, Ag, and Au. *Phys. Rev. Lett.* **110**, 216804 (2013).
22. Yazyev, O. *et al.* Theory of magnetic states in chiral graphene nanoribbons. *Phys. Rev. B* **84**, 115406 (2011).
23. Mashoff, T. *et al.* Bistability and oscillatory motion of natural nanomembranes appearing within monolayer graphene on silicon dioxide. *Nano Lett.* **10**, 461–465 (2010).
24. Wassmann, T. *et al.* Structure, stability, edge states, and aromaticity of graphene ribbons. *Phys. Rev. Lett.* **101**, 096402 (2008).
25. Wang, W. L. *et al.* Graphene nanoflakes with large spin. *Nano Lett.* **8**, 241–245 (2008).
26. Edwards, D. M., & Katsnelson, M. I. High-temperature ferromagnetism of *sp* electrons in narrow impurity bands. *J. Phys. Condens. Matter* **18**, 7209–7225 (2006).
27. Joly, V. L. J. *et al.* Observation of magnetic edge state in graphene nanoribbons. *Phys. Rev. B* **81**, 245428 (2010).
28. Chen, L. *et al.* Towards intrinsic magnetism of graphene sheets with irregular zigzag edges. *Sci. Rep.* **3**, 2599 (2013).
29. Cervenka, J. *et al.* Room-temperature ferromagnetism in graphite driven by two-dimensional networks of point defects. *Nature Phys.* **5**, 840–844 (2009).
30. Jung, J. Nonlocal exchange effects in zigzag-edge magnetism of neutral graphene nanoribbons. *Phys. Rev. B* **83**, 165415 (2011).

Acknowledgements The experimental work was conducted within the framework of the Korea Hungary Joint Laboratory for Nanosciences through the Korean Research Council of Fundamental Science and Technology and the “Lendület” programme of the Hungarian Academy of Sciences. L.T. acknowledges OTKA grant K108753 and the Bolyai Fellowship. L.P.B. acknowledges OTKA grant K101599. C.H. is supported in part by the Nano-Material Technology Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2012M3A7B4049888). I.H. was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of the TÁMOP-4.2.4.A/ 2-11/1-2012-0001 National Excellence Program and OTKA grant K100908. I.H. acknowledges discussions with K. Itai. L.T. and P.V. acknowledge discussions with Y.-S. Kim.

Author Contributions L.T. and C.H. conceived and designed the experiments. G.Z.M. performed the lithography and STM experiments. I.H. and P.V. provided the theoretical results. X.J. and C.H. performed the graphene growth experiments. Z.O. and P.N.-I. carried out preliminary experiments. L.T., P.V., I.H., G.Z.M. and L.P.B. analysed the data. L.T. wrote the paper. All the authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.T. (tapaszto@mfa.kfki.hu).

METHODS

The graphene samples were grown by chemical vapour deposition on Cu foil, and transferred to a stripped gold substrate with large, atomically flat Au (111) terraces. The lithographic process and tunnelling microscopy/spectroscopy investigations were performed using a Nanoscope E Scanning Tunneling Microscope operating under ambient conditions. First, atomic-resolution STM images of the graphene lattice were acquired (bias voltage $U_{\text{bias}} = 5\text{--}50$ mV and tunnelling current $I_{\text{tunnel}} = 1\text{--}2$ nA), to precisely identify the zigzag and armchair directions of the graphene lattice. For patterning, a bias voltage of 2.0–2.3 V (tip negative) is applied between the tip (90% Pt, 10% Ir) and the sample while the tip is slowly ($1\text{--}5$ nm s^{-1}) moved along the desired cutting direction. The humidity of the cutting atmosphere (70%–75%) can be precisely controlled inside an atmospheric hood. Before the tunnelling spectroscopy measurements on ribbons, tunnelling spectra were acquired on gold to ensure that the spectra were linear and confirm that the tip was good quality and contamination-free. We acquired tunnelling spectra inside the ribbons to confirm the presence or absence of a bandgap. This method has proved more efficient and reproducible under ambient conditions than measuring strongly edge-localized peaks in the dI/dV spectra. We define the gap of the ribbon as the width of the plateau around zero bias in tunnelling spectra (Extended Data Fig. 2b). These plateaus are not entirely flat, but have a shallow slope, which we attribute to the presence of gold substrate underneath the atomically thin graphene nanostructures. The uncertainty (error bars) in the measured gap values mainly comes from the thermal broadening and the influence of the substrate, while for ribbon widths the uncertainty is due to the STM tip convolution effects. As control experiments, both before and after measuring the $I\text{--}V$ characteristics of a ribbon, tunnelling spectra were acquired on the graphene sheet outside the ribbons. The (nonlinear) spectra measured on the ribbons were only taken as reliable if both before and after their measurement, the spectra of the unpatterned graphene displayed the expected closely linear $I\text{--}V$ characteristics. The tunnelling $I\text{--}V$ characteristics of graphene and metallic graphene nanoribbons on gold substrate are close to linear owing to the contribution of the high local density of state of the Au substrate. Within (along) the ribbons the measured spectra were fairly homogeneous (Extended Data Fig. 4), apart from a few exceptions, which we attribute to local edge defects or impurities.

To interpret the experimental results we considered the Hamiltonian of equation (1). We applied the $\hat{n}_{i\sigma} = \hat{n}_{i\sigma} - \langle \hat{n}_{i\sigma} \rangle + \langle \hat{n}_{i\sigma} \rangle$ identity and neglected the term that contains the fluctuations. Quantum Monte Carlo simulations including fluctuations found the agreement between mean-field theory and quantum Monte Carlo to be remarkably accurate for moderate Coulomb interactions, justifying the application of mean-field theory for the description of realistic ribbon geometries³¹. We arrived at a single-particle problem, which can be diagonalized by a generalized Bogoliubov transformation in k space:

$$\mathcal{H}_{\text{mean field}} = \sum_{k,\sigma,n} (\epsilon_{nk\sigma} - \mu) \hat{C}_{n,k,\sigma}^\dagger \hat{C}_{n,k,\sigma} - U \sum_j \langle \hat{n}_{j\uparrow} \rangle \langle \hat{n}_{j\downarrow} \rangle$$

where $\hat{C}_{n,k,\sigma}^\dagger$ ($\hat{C}_{n,k,\sigma}$) are the transformed operators that destroy (or create) a particle with wavenumber k and spin σ in band n . The energy bands are given by $\epsilon_{nk\sigma}$, which depends on the yet unknown electron densities and chemical potential. The densities are calculated self-consistently, while the chemical potential is determined by the conservation of the number of particles. The self-consistent iteration was stopped when the difference between the electron densities was smaller than 10^{-12} . It is known from the Hubbard-model theory that the half-filled case shows antiferromagnetic

correlations for $U > 0$, and that at a certain U_c a Mott transition can take place. However, in the doped case both antiferromagnetic and ferromagnetic ‘ground’ states can be obtained. The solution which has the lowest free energy was accepted as the ‘ground’ state.

To provide accurate bandgap values for the graphene nanoribbons, we first calculated the relaxed edge geometries with hydrogen passivation of narrow (2 nm) zigzag and armchair ribbons. We used density-functional-theory-based molecular dynamics calculations within the framework of local density approximation using the VASP^{32,33} simulation package. With the help of the relaxed edge coordinates of the atoms we were able to parameterize the tight-binding hopping elements in our Hubbard model. These hopping elements were applied to ribbons of various widths.

In the VASP calculations projector augmented wave pseudo-potentials^{34,35} were used and the kinetic energy cut-off for the plane wave expansion was 400 eV. In all ribbon geometries the atomic positions were relaxed using the conjugate-gradient method until the forces of the atoms were reduced to 0.02 eV \AA^{-1} . We used a rectangular super-cell with $dx = 40$ \AA in the x direction and $dy = 2.46$ \AA in the y direction for the zigzag graphene nanoribbon (4.26 \AA for the armchair graphene nanoribbon). The Brillouin zone was sampled using 2 k -points along the x axis and approximately 8/ dy k -points along the y axis (where dy is in nanometres). Vacuum layers of 20 \AA in the ribbon plane and 24.6 \AA in the normal direction were applied to avoid interactions between nanoribbons in different unit cells.

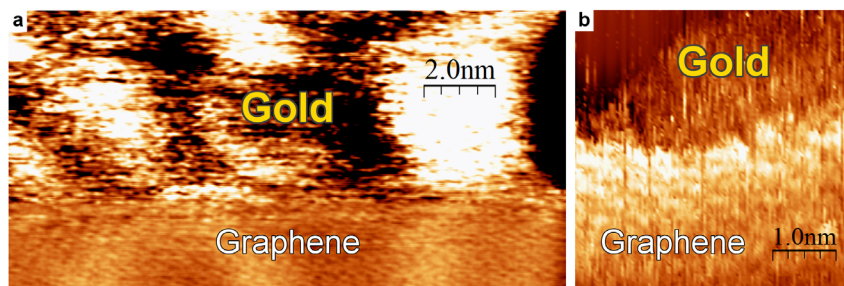
From the relaxed carbon–carbon distances, we computed the hopping amplitudes using the parameterization³⁶

$$t(r_{ij}) = \left(\frac{r_{ij}}{a_0}\right)^{-\alpha_2} \exp\left[-\alpha_3\left(\frac{r_{ij}^{\alpha_4}}{a_0^{\alpha_4}}\right)\right]$$

where r_{ij} is the distance between the atoms i and j , $a_0 = 1.42$ \AA (the carbon–carbon distance in the bulk), $\alpha_2 = 1.2785$, $\alpha_3 = 0.1383$ and $\alpha_4 = 3.4490$.

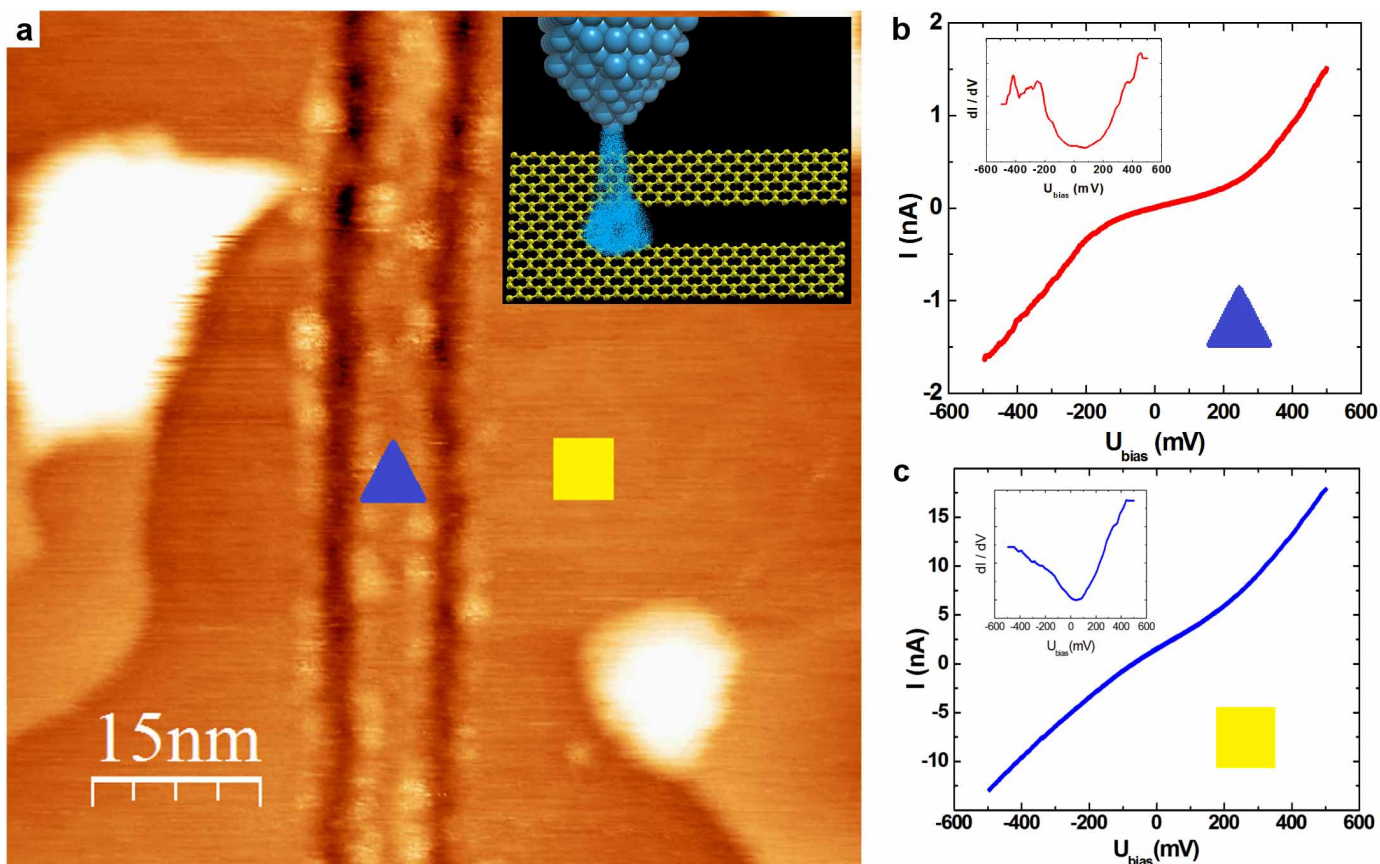
To estimate the maximum possible quantitative effect of edge irregularities on the magnetism, we performed the calculations presented here on a model system with edges of overall zigzag orientation, but containing the maximum amount of atomic-scale edge defects that does not completely destroy edge magnetism (Extended Data Fig. 3). We found that the qualitative picture of edge magnetism including the semiconducting (antiferromagnetic) to metallic (ferromagnetic) transition holds true for the defective zigzag edges, but the effect is substantially weakened compared to perfect edges. In particular, the calculated gap values are reduced. Therefore, the best fit to the experimental data occurs for a higher parameter value of the onsite repulsion parameter of $U = 4.32$ eV. Also, the width dependence of the gap becomes weaker, but this is still in good agreement with the measurements (Extended Data Fig. 3b) for the increased U parameter value.

31. Golor, M. *et al.* Quantum Monte Carlo studies of edge magnetism in chiral graphene nanoribbon. *Phys. Rev. B* **87**, 155441 (2013).
32. Kresse, G. & Hafner, J. *Ab initio* molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys. Rev. B* **49**, 14251 (1994).
33. Kresse, G. & Furthmüller, J. Efficient iterative schemes for *ab initio* total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169 (1996).
34. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758 (1999).
35. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953 (1994).
36. Tang, M. S. *et al.* Environment-dependent tight-binding potential model. *Phys. Rev. B* **53**, 979 (1996); erratum **54**, 10982 (1996).



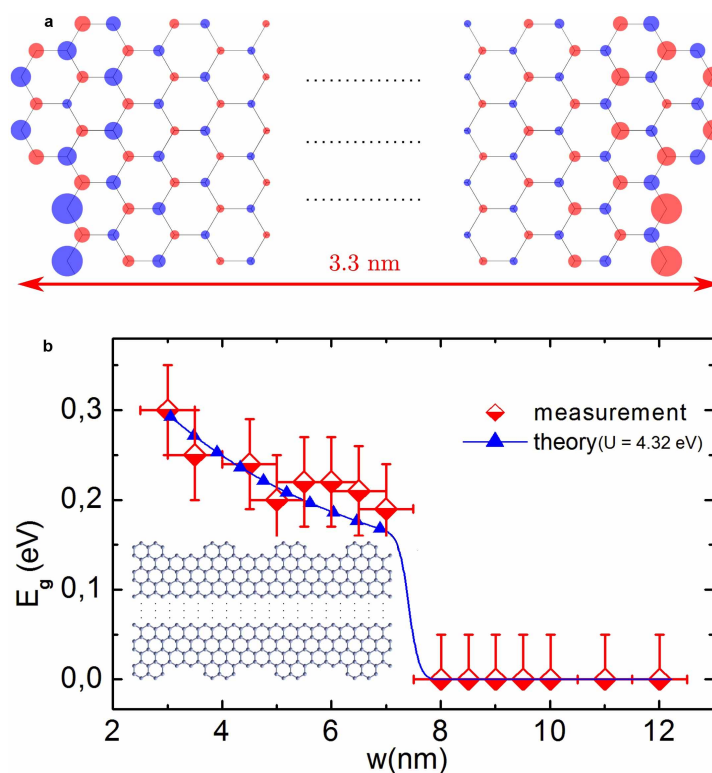
Extended Data Figure 1 | The structure of graphene edges defined by STM lithography. **a**, STM image (5 mV, 2 nA) of 15-nm-long edge segments cut by STM lithography, revealing edges that are close to atomically smooth ($<5 \text{ \AA}$ edge roughness) free of detectable reconstructions, contaminations or curvature. **b**, The increased local density of states on zigzag edges observed

under specific imaging conditions (200 mV, 2 nA) can be attributed to the presence of edge states that rule out the possibility of sp^3 -type edge terminations (for example, di-hydrogenated edges), because no edge states are expected to occur for such edge configurations.



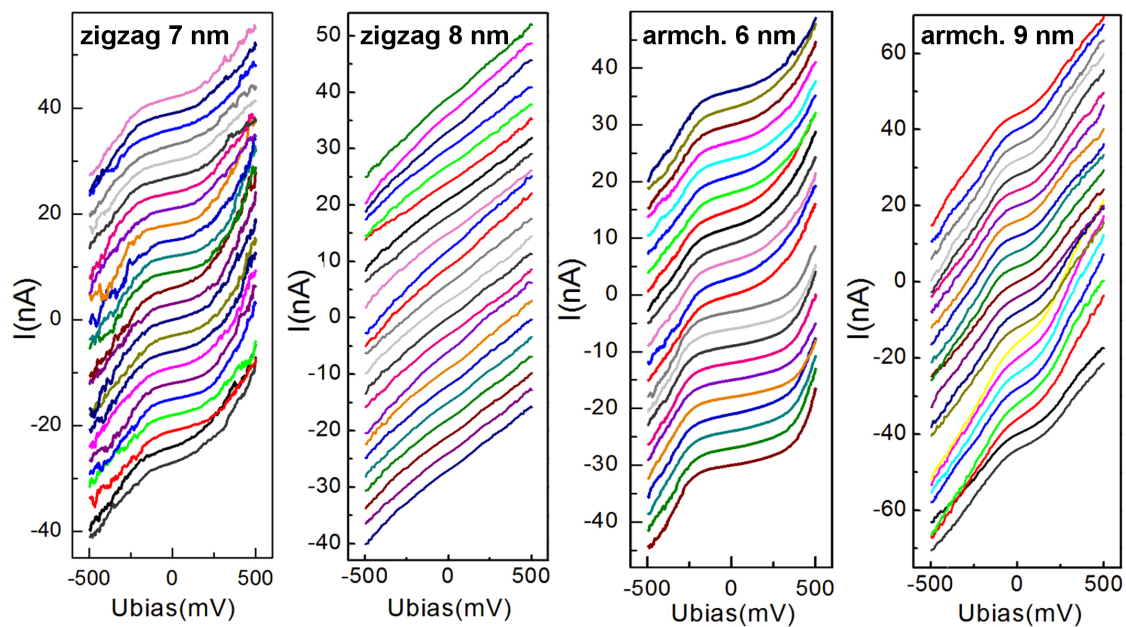
Extended Data Figure 2 | Scanning tunnelling spectroscopy of graphene nanoribbons on Au (111). Tunnelling I - V characteristics acquired on a 5-nm-wide armchair ribbon (a) displaying nonlinear I - V spectra corresponding to a gap of about 250 meV (b). Outside the ribbon a close-to-linear characteristic of the unpatterned graphene is revealed (c). The blue triangle and yellow square mark the positions of the corresponding tunnelling spectra. The insets show

the schematics of STM lithography (a), and the differential tunnelling conductance (dI/dV) obtained as numerical derivatives of the measured I - V curves (b and c). The ~ 70 -mV shift of the Dirac point (curve minimum) from the Fermi level (zero bias) observed on graphene (inset to c) is due to the doping from the Au(111) substrate and the ambient atmosphere.



Extended Data Figure 3 | The effect of edge irregularities on edge magnetism. Calculated spin density distribution in the unit cell of a 3.3-nm-wide zigzag ribbon with a high density of atomic-scale defects, revealing the substantial decrease of the emerging spin polarization (to about a third of that

of defect-free zigzag edges). The experimental width dependence can be fitted for defective ribbon edges by using higher values of the on-site repulsion parameter of $U = 4.32$ eV.



Extended Data Figure 4 | Individual tunnelling I - V spectra. Tunnelling I - V characteristics recorded on various ribbons (the spectra have been shifted along the vertical axis for clarity). Each individual spectrum was recorded as

the average of ten voltage sweeps between -500 mV and $+500$ mV. The metallic (close to linear) or semiconducting (strongly nonlinear) nature of the ribbons is clearly apparent from the individual tunnelling I - V characteristics.

Stochastic transport through carbon nanotubes in lipid bilayers and live cell membranes

Jia Geng^{1,2,3*}, Kyunghoon Kim^{1,3,4*}, Jianfei Zhang², Artur Escalada⁵, Ramya Tunuguntla^{1,3}, Luis R. Comolli⁶, Frances I. Allen^{7,8,9}, Anna V. Shnyrova⁹, Kang Rae Cho¹, Dayannara Munoz¹, Y. Morris Wang¹⁰, Costas P. Grigoropoulos⁴, Caroline M. Ajo-Franklin^{3,8,11}, Vadim A. Frolov^{5,12} & Aleksandr Noy^{1,2,3}

There is much interest in developing synthetic analogues of biological membrane channels¹ with high efficiency and exquisite selectivity for transporting ions and molecules. Bottom-up² and top-down³ methods can produce nanopores of a size comparable to that of endogenous protein channels, but replicating their affinity and transport properties remains challenging. In principle, carbon nanotubes (CNTs) should be an ideal membrane channel platform: they exhibit excellent transport properties^{4–8} and their narrow hydrophobic inner pores mimic structural motifs typical of biological channels¹. Moreover, simulations predict that CNTs with a length comparable to the thickness of a lipid bilayer membrane can self-insert into the membrane^{9,10}. Functionalized CNTs have indeed been found to penetrate lipid membranes and cell walls^{11,12}, and short tubes have been forced into membranes to create sensors¹³, yet membrane transport applications of short CNTs remain underexplored. Here we show that short CNTs spontaneously insert into lipid bilayers and live cell membranes to form channels that exhibit a unitary conductance of 70–100 picosiemens under physiological conditions. Despite their structural simplicity, these ‘CNT porins’ transport water, protons, small ions and DNA, stochastically switch between metastable conductance substates, and display characteristic macromolecule-induced ionic current blockades. We also show that local channel and membrane charges can control the conductance and ion selectivity of the CNT porins, thereby establishing these nanopores as a promising biomimetic platform for developing cell interfaces, studying transport in biological channels, and creating stochastic sensors.

We created short CNTs using sonication-assisted cutting¹⁴ of purified long CNTs with an average inner diameter of 1.51 ± 0.21 nm in the presence of 1,2-dioleoyl-*sn*-glycero-3-phosphocholine (DOPC) lipid (Fig. 1a). This procedure produced a stable dark suspension (Fig. 1b, inset) containing short CNTs stabilized by the lipid coating. Transmission electron microscopy (TEM) images of uncut and cut CNTs are shown in Extended Data Fig. 1, and CNT length and diameter distributions are shown in Extended Data Figs 2 and 3a, b, respectively. Raman spectroscopy confirmed that the cutting procedure preserved the inner diameter and rolled-up graphene-sheet structure of the nanotube (Extended Data Fig. 3c). TEM images of the purified cut CNTs (Fig. 1b, Extended Data Fig. 1) also indicate the presence of a significant population of short CNT fragments with lengths in the 5–15 nm range (Fig. 1d, Extended Data Fig. 2).

Cryogenic TEM (cryo-TEM) images of the CNTs reconstituted into 200-nm-diameter liposomes (Methods) reveal short CNT fragments inserted into the lipid membrane and spanning both membrane leaflets (Fig. 1c; see also Extended Data Fig. 4 for wider field-of-view cryo-TEM

images, and Extended Data Fig. 5 for additional higher-magnification images). The presence of CNTs does not affect the integrity of the lipid bilayer, and statistical analysis of the cryo-TEM images reveals a relatively wide length distribution of the inserted tubes (Fig. 1d) that indicates that insertion is not selective to a particular length. The perpendicular orientation of the CNTs to the membrane plane is strongly preferred, with the majority of the nanotubes tilting only 15° or less from the membrane normal (Fig. 1e). This observation disagrees with molecular dynamics simulations that predict tilting of CNTs longer than the thickness of the bilayer (4.6 ± 0.2 nm for the DOPC membranes) to maximize the interactions of the hydrophobic bilayer core with the hydrophobic CNT walls⁹. Finally, for most inserted CNTs at least one tube end abutted the hydrophilic head groups of the lipid bilayer, allowing for energetically favourable interactions of the hydrophilic groups on the tube rim with the hydrophilic lipid headgroups. Because the short inserted CNTs have a graphitic structure and functionality resembling the β -barrel structure of porins, the water-filled transmembrane proteins that act as pores through which molecules can diffuse, we refer to them as CNT porins.

We explored the proton transport behaviour of CNT porins by placing liposomes containing an encapsulated pH-sensitive dye (8-hydroxypyrene-1,3,6-trisulphonic acid trisodium salt, HPTS) at pH 8.0 into isotonic buffer solution at pH 3.0. With liposomes lacking CNT porins, the background proton leakage through the vesicle membrane led to a slow HPTS fluorescence decay with a characteristic timescale of 187 s (Fig. 1f). Conversely, samples that contained CNT porins in the bilayer showed much faster initial fluorescence decay with a 35 s timescale, indicating the opening of a second pathway for protons to enter the vesicle lumen. The overall decay kinetics was described by a sum of two exponential terms, with the second term matching the 187 s timescale obtained in a control experiment (Fig. 1f, inset), indicating that a small subpopulation of the liposomes did not contain CNT porins.

To explore transport through the CNT porins further, we placed liposomes containing pure water into dilute NaCl and K_2SO_4 solutions and used dynamic light scattering to monitor liposome size changes resulting from osmotically driven water expulsion¹⁵. Control liposomes without CNT porins showed minimal (<2.3%) size changes, whereas liposomes incorporating CNT porins shrank by up to 20% (Fig. 1g). CNT porin vesicle shrinkage depends strongly on the ionic strength of the outside solution, being highest at low ionic strength and reducing dramatically at high ionic strengths to approach values measured for the control samples. This behaviour can be understood by considering that at pH 7.0, the carboxylic acid groups at the CNT porin entrance are negatively charged, creating an electrostatic barrier for the anion passage; this barrier slows down the ion leakage into the liposome lumen

¹Biology and Biotechnology Division, Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550, USA. ²School of Natural Sciences, University of California, Merced, California 95340, USA. ³The Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁴Mechanical Engineering Department, University of California, Berkeley, California 94720, USA. ⁵Biophysics Unit (CSIC, UPV/EHU) and Department of Biochemistry and Molecular Biology, University of the Basque Country, 48940 Leioa, Spain. ⁶Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁷Department of Materials Science and Engineering, University of California, Berkeley, California 94720, USA. ⁸Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁹National Center for Electron Microscopy, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ¹⁰Materials Science Division, Lawrence Livermore National Laboratory, Livermore, California 94550, USA. ¹¹Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ¹²ikerbasque, Basque Foundation for Science, 48011 Bilbao, Spain.

*These authors contributed equally to this work.

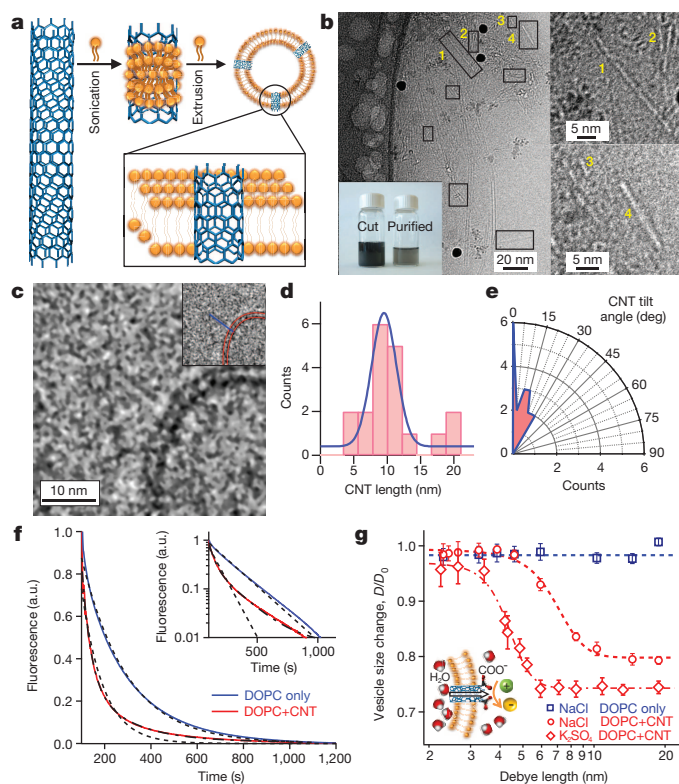


Figure 1 | Synthesis, characterization, and bulk transport properties of CNT porins. **a**, Schematic showing CNT porin preparation and incorporation into liposomes (see text for details). **b**, Left, TEM image of cut CNTs stabilized by lipid coating; boxed areas highlight cut CNTs; boxes labelled 1–4 are shown magnified at right. Inset, a photograph of cut CNT suspensions before and after centrifugation. **c**, Cryo-TEM image of a CNT porin. Inset, overlaid outlines of the bilayer (red) and CNT (blue). Image was processed to enhance contrast (see Extended Data Fig. 5 for unfiltered images of this and other CNT porins). **d**, Histogram of the lengths of the CNTs inserted into the lipid membrane and a Gaussian fit to the data ($n = 20$). **e**, Radius plot of the histogram of CNT tilt angles measured relative to the axis normal to the bilayer plane ($n = 20$). **f**, Time traces of the fluorescence of HPTS dye encapsulated in liposomes with (red trace) and without (blue trace) CNT porins after outside pH was lowered from pH 8 to pH 3. Dashed lines, single exponential fits; dash-dotted line (red trace), a sum of two exponential terms. a.u., arbitrary units. Inset shows the plot on a log-linear scale. **g**, Plot of the liposome size change D/D_0 (where D stands for liposome diameter and D_0 for the initial diameter) after exposure to NaCl (red circles and blue squares) and K_2SO_4 (red diamonds) osmotic gradients. Dashed lines, fits to sigmoidal function. Error bars, s.d. ($n = 4-5$). Inset illustrates ion rejection by the carboxylic acid groups at the CNT porin entrance.

but not water outflow, allowing the osmotic pressure to shrink the vesicles (Fig. 1g, inset). At higher ionic strength these entrance charges are screened, allowing ions to pass through the CNT pore to neutralize the osmotic gradient. Indeed, the transition from high to low ion rejection in our measurements occurs at the same Debye length ranges (6–7 nm for NaCl and 4–5 nm for K_2SO_4) that we previously observed for bulk CNT membranes¹⁶, confirming that the observed effects are closely related to ion diffusion through the CNT pores.

We observed spontaneous membrane incorporation and characterized transport properties of individual CNT porins using single-channel conductance measurements (Fig. 2a). When we added CNT porins to the *cis*-chamber of the setup, the bilayer conductance showed a series of characteristic amplitude jumps (Fig. 2b), indicating membrane incorporation events. Conductance traces contained a mixture of well-separated events and multiple steps corresponding to a relatively quick succession of incorporation events. The histogram of conductance step values (Fig. 2c) shows clear evidence of quantization, with the first peak

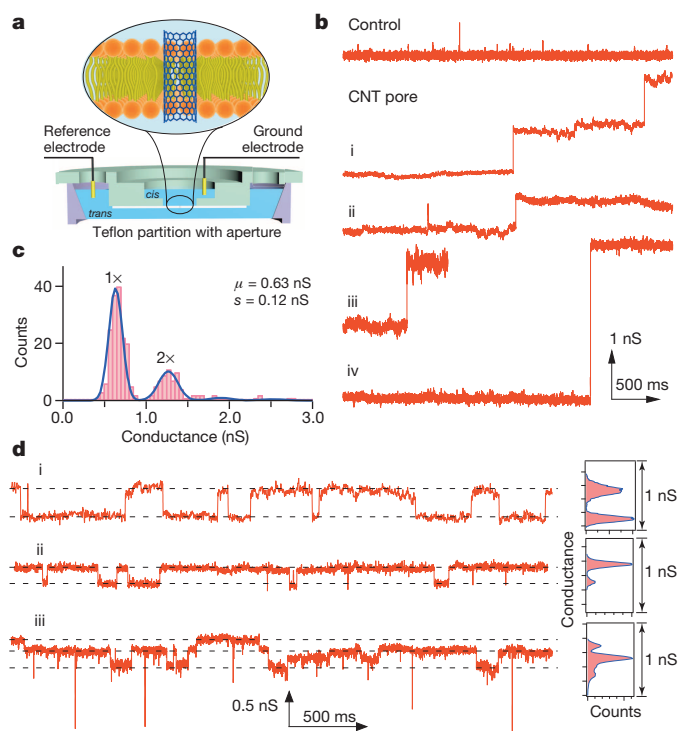


Figure 2 | Transport through individual CNT porins. **a**, Set-up for single-channel recording of CNT porin conductance where a small patch of lipid bilayer separates two solution chambers (labelled *cis* and *trans*). The transmembrane potential is set by the reference electrode (see Methods for details). **b**, Conductance traces showing incorporation of single (i, ii) and multiple (iii, iv) CNT porins into the lipid bilayer. Control trace was recorded in absence of CNTs. **c**, Histogram of CNT porin conductance values ($n = 236$). Solid blue line is a sum of two Gaussian peaks where the position and width of the second peak are determined from the position, μ , and width, σ , of the first peak as 2μ and $\sqrt{2}\sigma$. **d**, Conductance traces and normalized conductance histograms showing stochastic ‘gating’ transitions for one (i, ii) and two (iii) CNT porins.

corresponding to the 0.63 ± 0.12 nS value for a single CNT pore conductance and a subsequent peak at exactly double that value. Occasionally, we recorded conductance jumps at higher multiples of the single pore conductance value, but the frequency of those events was low. Remarkably, despite a relatively wide distribution of CNT porin lengths, their unitary conductance is well defined. Cryo-TEM images argue strongly against the possibility that only a narrow subset of short nanotubes incorporates into the lipid membrane. Thus, we conclude that the conductance of CNT porins is a relatively weak function of the CNT length and instead is determined mainly by the interfacial resistances of the pore ends¹⁷. Interestingly, the 0.63 nS conductance of a CNT porin is close to the 0.65–0.73 nS conductance of an α -haemolysin channel¹⁸, which has a 10 nm-long pore with a width that varies from 1.4 nm in the neck region to 4.6 nm in the vestibule region. A much wider range of conductance values reported for acid-cut CNTs¹³ could be attributed to the microinjection protocol used to insert these pores into the membrane.

CNT porin conductance scaled linearly with salt concentration in the 0.5–2 M range both at neutral pH and at pH 2 (Extended Data Fig. 6), strongly suggesting that the dominant conductance mechanism through the CNT porin is ionic. The unitary conductance of a CNT porin at pH 2 was reduced by almost a factor of 2 compared to neutral pH, from 0.63 nS to 0.33 nS (Extended Data Fig. 7b). This observation rules out proton conductivity as the dominant current-carrying mechanism in these pores and suggests that the conductance of the CNT porins is sensitive to the charges at the pore mouth (which are neutralized at pH 2). A similar decrease in channel conductance at acidic pH due to the changes

in the charge state of the pore residues has been reported for OmpF porins¹⁹.

For less than 30% of the CNT porins, the conductance signal oscillated between ‘open’ and ‘partially-closed’ sub-states after insertion, showing distinct peaks in the histograms of the trace amplitude (Fig. 2d). Notably, we did not observe these low conductance sub-states at pH 2 (Extended Data Fig. 7a, c). Computer simulations of pore insertion into lipid bilayers indicate that rapid removal and re-insertion of the CNT porins from the bilayer cannot cause these oscillations⁹. The ionic nature of the conductance in CNT porins argues against the possibility that the oscillations originate from the ion blocking of the proton-based current as previously reported for macroscopically-long CNT channels⁶. Transient blocking by nanoprecipitation²⁰ requires the presence of divalent ions and produces distinct triangular shaped blockades different from the ones we observed. Another possible cause, spontaneous wetting/dewetting transitions in hydrophobic nanopores²¹, should have produced a clear zero-conductance ‘off’ state; however, our system frequently shows only a partially-blocked low conductance sub-state (Extended Data Fig. 8b, c), suggesting that the CNT interior remains filled with water⁸.

Conductance oscillations in CNT porins are similar to the stochastic ‘gating’ behaviour of biological ion channels. Unlike biological channels, however, where conformation transitions are commonly invoked as a cause of conductance sub-states, CNT porins contain no movable parts and the entire pore structure is rigid. Similar transitions have been reported previously for ion-track-etched polyethylene terephthalate solid-state nanopores²² and glass-polymer nanogaps²³. One possibility is that these sub-states represent a spontaneous transition between ionic-penetration and ionic-exclusion states, as predicted²⁴ for weakly charged or neutral nanopores embedded in a dielectric medium. This model predicts the existence of low-conductance sub-states for charged nanopores and on/off transitions for uncharged pores. Consistent with this prediction, the low conductance states disappear at pH 2. Although we find some parallels with the behaviour of conductance sub-states reported for OmpF porin channels¹⁹, further experimental work is necessary to understand the detailed nature of conductance fluctuations.

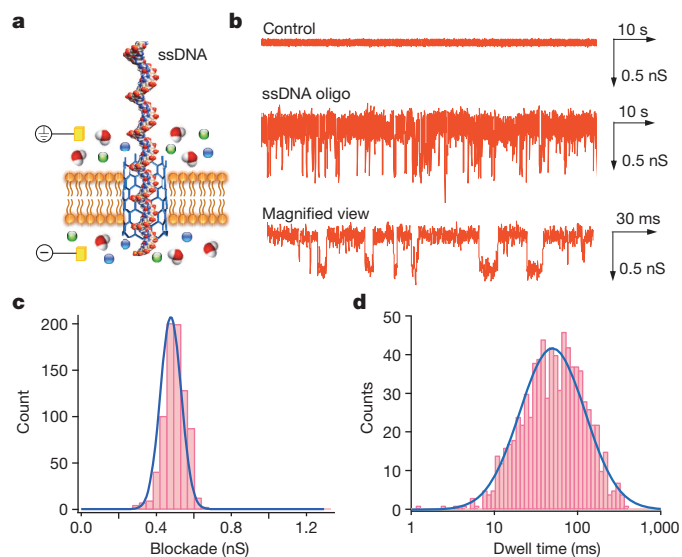


Figure 3 | DNA translocation through CNT ion channels. **a**, Schematic showing the translocation of single-stranded DNA through a CNT porin in the lipid bilayer (see text for details). Symbols in circles indicate ground electrode (top) and reference electrode (bottom). **b**, Current trace showing multiple transient blockades caused by 81-nt ssDNA translocation through the CNT channel (middle) with magnified view (bottom). The top trace shows the control trace recorded in absence of ssDNA. The applied voltage was -50 mV. **c**, **d**, Histograms of conductance blockade levels (**c**) and dwell times (**d**) for 806 ssDNA translocation events.

When exploring electrophoretic translocation of macromolecules through individual CNT porins with 81-nucleotide-long single-stranded DNA (81-nt ssDNA) oligomers added to the *trans*-chamber of our set-up (Fig. 3a), we observed rapid conductance blockades as individual ssDNA strands passed through the CNT pore (Fig. 3b). The conductance blockade values followed a normal distribution centred at 0.48 ± 0.08 nS (Fig. 3c), which falls between the 0.8 nS blockades reported for wild-type α -haemolysin pores²⁵ with a diameter similar to that of CNT porins and the 0.3 nS blockades reported for a hybrid α -haemolysin/solid-state nanopore structure²⁶. DNA translocation dwell times followed a log-normal distribution (Fig. 3d) centred around a value of 53 ms, which corresponds to an average DNA translocation speed of 1.5 nt per ms. This value is comparable to the speed achieved by biological nanopores and falls within the range suitable for DNA analysis applications²⁷.

Finally, we used patch-clamp measurements to demonstrate that CNT porins can spontaneously incorporate into the membranes of live Chinese hamster ovary (CHO) and human embryonic kidney (HEK293T) cell lines. For both CHO cells (in 5 out of 6 trials) and HEK293T cells (in 13 out of 15 trials), we detected channel-like activity emerging 60–90 s after establishing a tight contact between a cell and patch-pipette containing CNT porins (Fig. 4a, see Extended Data Fig. 9 for examples of longer traces). In both cases the dominant activity corresponded to a unitary conductance of approximately 70 pS (68.1 ± 0.8 pS for HEK293T cells and 67.6 ± 1.4 pS for CHO cells) and its multiples (Fig. 4b, c, also see Extended Data Fig. 9). Control experiments performed without CNT porins showed no endogenous channel activity in CHO cells (10 out of 10 cells). HEK293T cells showed endogenous channel activity with unitary conductance of 30.6 ± 0.3 pS (Fig. 4b, Extended Data Fig. 9c, 5 out of 5 cells), which was also present as a 21.9 ± 0.7 pS peak in the data obtained for interactions of HEK293T cells with CNT porins (Fig. 4b). Importantly, CNT porins incorporated into cell membranes display low-conductance sub-states (Extended Data Fig. 9e), completely mimicking the conductance dynamics observed in planar lipid membranes.

We detected similar single-channel activity using patch-clamp experiments on lipid bilayers, corroborating spontaneous incorporation

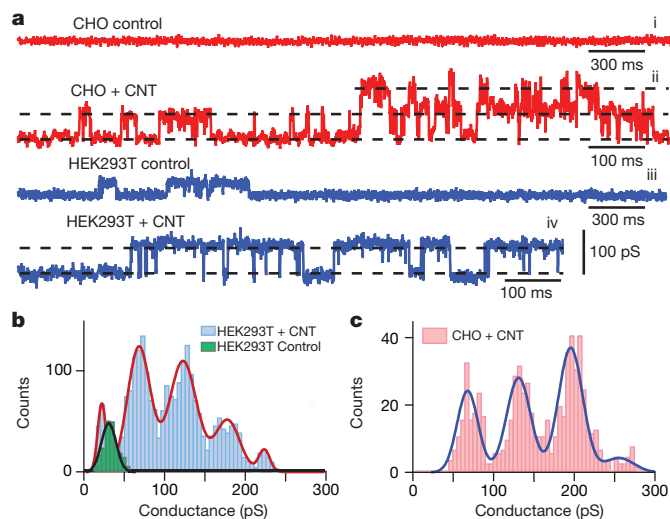


Figure 4 | CNT porin incorporation into live cell membranes. **a**, Representative conductance traces for CNT porin incorporation and gating in CHO (trace ii) and HEK293T (trace iv) cells (-60 mV holding potential, 150 mM KCl). Traces i and iii were recorded in the absence of CNT porins. Dashed lines are added as a guide to the eye. **b**, **c**, Histograms of conductance steps for CNT porin incorporation into HEK293T (**b**; blue bars, $n = 2,467$, 13 cells) and CHO (**c**; red bars, $n = 657$, 5 cells) cell membranes fitted to a sum of Gaussian peaks (solid lines). See Methods for peak positions. Native channel activity for HEK293T cells (green bars and black solid line, $n = 227$, 5 cells) is also seen with CNT porins.

of the CNT porins into the lipid membrane patch inside the patch-pipette. Notably, the unitary conductance detected in these experiments depended on the lipid composition (Extended Data Fig. 10a). In charged membranes containing the core ingredients of the plasma membrane, the measured conductance was 64.7 ± 0.5 pS (at 150 mM KCl), closely corresponding to the dominant unitary conductance detected in cellular membranes. We observed similar conductances for CNT porins pre-constituted into giant unilamellar vesicles (Extended Data Fig. 10b, e), demonstrating that the activity of CNT porins does not depend on the membrane incorporation pathway. However, when we excluded charged lipid species from the bilayer composition, the unitary conductance shifted to 92.5 ± 0.6 pS (at 150 mM KCl). This conductance level corresponds to a conductance of 0.62 nS after correcting for the difference in electrolyte concentration (150 mM versus 1 M KCl), in excellent agreement with the value from the planar lipid bilayer experiments (Fig. 2c). The origins of the modulation in conductance caused by the membrane surface charge need to be clarified by follow-up studies.

Taken together, our findings establish CNT porins as a promising prototype of a synthetic membrane channel with inherent robustness towards biological and chemical challenges and exceptional biocompatibility that should prove valuable for bionanofluidic and cellular interface applications. We expect that our CNT porins could be modified with synthetic 'gates'²⁸ to dramatically alter their selectivity, opening up possibilities for their use in synthetic cells, drug delivery, and biosensing.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 October 2013; accepted 26 August 2014.

- Sui, H., Han, B.-G., Lee, J. K., Walian, P. & Jap, B. K. Structural basis of water-specific transport through the AQP1 water channel. *Nature* **414**, 872–878 (2001).
- Langecker, M. *et al.* Synthetic lipid membrane channels formed by designed DNA nanostructures. *Science* **338**, 932–936 (2012).
- Li, J. *et al.* Ion-beam sculpting at nanometre length scales. *Nature* **412**, 166–169 (2001).
- Hinds, B. J. *et al.* Aligned multiwalled carbon nanotube membranes. *Science* **303**, 62–65 (2004).
- Liu, H. *et al.* Translocation of single-stranded DNA through single-walled carbon nanotubes. *Science* **327**, 64–67 (2010).
- Lee, C. Y., Choi, W., Han, J.-H. & Strano, M. S. Coherence resonance in a single-walled carbon nanotube ion channel. *Science* **329**, 1320–1324 (2010).
- Zimmerli, U. & Koumoutsakos, P. Simulations of electrophoretic RNA transport through transmembrane carbon nanotubes. *Biophys. J.* **94**, 2546–2557 (2008).
- Hummer, G., Rasaiah, J. C. & Noworyta, J. P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* **414**, 188–190 (2001).
- Lopez, C. F., Nielsen, S. O., Moore, P. B. & Klein, M. L. Understanding nature's design for a nanosyringe. *Proc. Natl Acad. Sci. USA* **101**, 4431–4434 (2004).
- Dutt, M., Nayhouse, M. J., Kuksenok, O., Little, S. R. & Balazs, A. C. Interactions of end-functionalized nanotubes with lipid vesicles: spontaneous insertion and nanotube self-organization. *Curr. Nanosci.* **7**, 699–715 (2011).
- Lacerda, L. *et al.* How do functionalized carbon nanotubes land on, bind to and pierce through model and plasma membranes. *Nanoscale* **5**, 10242–10250 (2013).
- Kam, N. W. S., O'Connell, M., Wisdom, J. A. & Dai, H. Carbon nanotubes as multifunctional biological transporters and near-infrared agents for selective cancer cell destruction. *Proc. Natl Acad. Sci. USA* **102**, 11600–11605 (2005).
- Liu, L., Yang, C., Zhao, K., Li, J. & Wu, H.-C. Ultrashort single-walled carbon nanotubes in a lipid bilayer as a new nanopore sensor. *Nature Commun.* **4**, <http://dx.doi.org/10.1038/ncomms3989> (2013).
- Sun, X. *et al.* Optical properties of ultrashort semiconducting single-walled carbon nanotube capsules down to sub-10 nm. *J. Am. Chem. Soc.* **130**, 6551–6555 (2008).
- Le Duc, Y. *et al.* Imidazole-quartet water and proton dipolar channels. *Angew. Chem. Int. Ed.* **50**, 11366–11372 (2011).
- Fornasiero, F. *et al.* Ion exclusion by sub 2-nm carbon nanotube pores. *Proc. Natl Acad. Sci. USA* **105**, 17250–17255 (2008).
- Walther, J. H., Ritos, K., Cruz-Chu, E. R., Megaridis, C. M. & Koumoutsakos, P. Barriers to superfast water transport in carbon nanotube membranes. *Nano Lett.* **13**, 1910–1914 (2013).
- Gu, L.-Q. & Bayley, H. Interaction of the noncovalent molecular adapter, β -cyclodextrin, with the staphylococcal α -hemolysin pore. *Biophys. J.* **79**, 1967–1975 (2000).
- Nestorovich, E. M., Rostovtseva, T. K. & Bezrukov, S. M. Residue ionization and ion transport through OmpF channels. *Biophys. J.* **85**, 3718–3729 (2003).
- Powell, M. R. *et al.* Nanoprecipitation-assisted ion current oscillations. *Nature Nanotechnol.* **3**, 51–57 (2007).
- Powell, M. R., Cleary, L., Davenport, M., Shea, K. J. & Siwy, Z. S. Electric-field-induced wetting and dewetting in single hydrophobic nanopores. *Nature Nanotechnol.* **6**, 798–802 (2011).
- Lev, A. A. *et al.* Rapid switching of ion current in narrow pores: implications for biological ion channels. *Proc. R. Soc. Lond. B* **252**, 187–192 (1993).
- Shimizu, S. *et al.* Stochastic pore blocking and gating in PDMS–glass nanopores from vapor–liquid phase transitions. *J. Phys. Chem. C* **117**, 9641–9651 (2013).
- Buyukdagli, S., Manghi, M. & Palmeri, J. Ionic capillary evaporation in weakly charged nanopores. *Phys. Rev. Lett.* **105**, 158103 (2010).
- Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl Acad. Sci. USA* **93**, 13770–13773 (1996).
- Hall, A. R. *et al.* Hybrid pore formation by directed insertion of α -haemolysin into solid-state nanopores. *Nature Nanotechnol.* **5**, 874–877 (2010).
- Venkatesan, B. M. & Bashir, R. Nanopore sensors for nucleic acid analysis. *Nature Nanotechnol.* **6**, 615–624 (2011).
- Majumder, M., Chopra, N. & Hinds, B. J. Effect of tip functionalization on transport through vertically oriented carbon nanotube membranes. *J. Am. Chem. Soc.* **127**, 9062–9070 (2005).

Acknowledgements Parts of this work were supported by the US Department of Energy, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering (characterization and transport studies), and the LDRD programme at LLNL, 12-ERD-073 (synthesis). R.T. acknowledges support from the LSP programme at LLNL. D.M. acknowledges support from an ROTC summer fellowship. V.A.F. acknowledges partial support by the Spanish Ministry of Economy and Competitiveness, grant BFU2012-34885, co-financed with European FEDER funds, and the Basque Government, grant IE12-332. Work at LLNL was performed under the auspices of the US Department of Energy under contract DE-AC52-07NA27344. Work at the Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract DE-AC02-05CH11231.

Author Contributions J.G. performed and analysed (with D.M.) the planar lipid bilayer transport measurements; K.K. performed bulk transport studies; J.Z., K.K., J.G. and R.T. performed the synthesis and purification, K.R.C. performed AFM analysis, Y.M.W., L.R.C., F.I.A. and K.K. performed TEM analysis; L.R.C. and F.I.A. performed cryo-TEM analysis; A.N. conceived and directed the research, and wrote the manuscript draft; A.E., A.V.S. and V.A.F. designed the cell and GUV reconstitution experiments, and A.E. and A.V.S. performed them. All authors contributed to the data analysis, discussion, and manuscript preparation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.N. (roy1@llnl.gov).

METHODS

Materials. 1.5 nm diameter single-walled carbon nanotubes were purchased from NanoLab, Inc. (Waltham, Massachusetts, USA, product purity >95%). 1,2-dioleoyl-*sn*-glycerol-3-phosphocholine (DOPC), 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine (DOPE), 1,2-dioleoyl-*sn*-glycero-3-phospho-L-serine (DOPS), cholesterol (Chol), 1,2-dioleoyl-*sn*-glycero-3-phosphoinositol-4,5-bisphosphate (PI(4,5)P₂) and 1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine-N-(Lissamine Rhodamine B sulphonyl) (Rh-DOPE) were purchased from Avanti Polar Lipids, Inc. (Alabaster, Alabama, USA). N-decane was purchased from Fisher Scientific. 8-Hydroxypyrene-1,3,6-trisulphonic acid trisodium salt (HPTS), hexane, and gramicidin A from *Bacillus brevis* were purchased from Sigma-Aldrich (St Louis, Missouri, USA) and Invitrogen (Carlsbad, California, USA). Short single-stranded DNA oligonucleotides (81-nt) were purchased from Integrated DNA Technologies, Inc. (Coralville, Iowa, USA).

CNT cutting. To remove amorphous carbon and impurities from the CNT sample, 1.6 to 2 mg of as-received CNTs were first purified in a thermal gravimetric analysis (TGA) system (Q5000 IR TGA-MS, TA Instruments, New Castle, Delaware, USA) in 25 ml min⁻¹ stream of air while the temperature was ramped at a rate of 5 °C min⁻¹ from 20 °C to 450 °C. In a typical CNT porin preparation procedure, 2 ml of DOPC (10 mg ml⁻¹ in chloroform) was added to a 20 ml glass vial and the solvent was evaporated using a V10 evaporator (Biotage, Uppsala, Sweden) to form a thin lipid film on the walls of the glass vial. 1.6 mg of purified CNTs and 20 ml DI water (Milli-Q, Millipore, Boston, Massachusetts, USA) were added to the same vial. Following 1 h of bath-sonication (Model 1510, Branson Ultrasonics, Danbury, Connecticut, USA), the sample was probe-sonicated for 16 h (Model VC 100, Sonics & Materials Inc., Newtown, Connecticut, USA) at 100 W power delivered in 3 s pulses separated by a 1 s pause. During sonication, the vial was continuously cooled in an ice-water bath and, if necessary, DI water was added to compensate for loss in sample volume due to evaporation.

Purification and incorporation of CNT porins into liposomes. To separate the cut CNTs from the uncut nanotube material, 10 ml of suspension processed by sonication cutting was added to a 15 ml centrifuge tube and centrifuged at 4,000 r.p.m. for 1 h using an Allegra X22 centrifuge (Beckman-Coulter, Indianapolis, Indiana, USA). The dark supernatant solution containing lipid-stabilized cut CNTs was stable and could be dried and re-suspended in DI water, suggesting that the lipid coating helped to prevent the aggregation of the cut CNTs.

To incorporate CNTs into DOPC liposomes, a 0.2 ml aliquot of 10 mg ml⁻¹ DOPC in chloroform was added to a 20 ml glass vial and the solvent was evaporated to leave a dried lipid film. 1 ml of purified CNT/DOPC complex solution was added to the vial and the mixture was allowed to hydrate for one hour. The solution was then extruded through a 200-nm-diameter pore polycarbonate membrane using a hand-held Mini-Extruder device (Avanti Polar Lipids) for 10 cycles.

Cryo-TEM specimen preparation and instrumentation. 5 µl aliquots of liposomes incorporating CNT porins (as well as vesicle-only control samples) were placed onto 200 mesh carbon-stabilized lacey-Formvar copper grids (#01881, Ted Pella, Inc., Redding, California, USA), manually blotted with filter paper, and plunged into liquid ethane at liquid nitrogen temperature. For control studies of isolated cut and uncut CNTs with and without lipids, 5 µl aliquots were placed on 200 mesh carbon-stabilized Formvar-film copper grids (#01800, Ted Pella, Inc.), blotted with filter paper and then air-dried. All TEM grids were glow discharged. Samples were transferred to a JEOL-3100-FFC transmission electron microscope equipped with a field emission gun (FEG) operating at 300 kV, an Omega energy filter, a Gatan 795 4K × 4K CCD camera (Gatan Inc., Pleasanton, California, USA) mounted at the exit of an electron decelerator operated at 200 kV resulting in images formed by a 100 kV electron beam at the CCD, and a cryo-transfer stage. The stage was cooled with liquid nitrogen to 80 K during the acquisition of all data sets.

Cryo-TEM image acquisition and analysis. Images were recorded at a nominal magnification of 40,000× giving a pixel size of 0.14 nm at the specimen, either without camera binning or using a camera binning factor of two. Underfocus values ranged from 2 ± 0.25 µm to 3.6 ± 0.25 µm, and energy filter widths were typically around 30 eV. The surveys of cryo samples and the selection of suitable targets were performed in low-dose defocused diffraction mode. Images were acquired under low-dose conditions, with doses of ~20–40 e⁻ Å⁻² per image. For the control samples of cut and uncut CNTs with and without lipids, images were acquired at 80,000× giving a pixel size of 0.07 nm at the specimen. Several dozen images were acquired with defocus values ranging from minimum contrast to 2 µm underfocus. Through-focus series were also acquired. Images acquired at 2 µm underfocus with camera binning factor of two, giving a pixel size of 0.14 nm, agree well with the images obtained for the cryo vesicle-CNT samples acquired at the same defocus. The package of imaging tools Priism (<http://msg.ucsf.edu/IVE/>) was used for noise reduction and contrast enhancement. ImageJ 1.38x software (NIH, <http://rsb.info.nih.gov/ij/>) was used for analysis and measurements of the 2D image projections. Since the images represent a 2D projection of a 3D structure, the measured CNT length

values represent the lower boundary estimate of the true CNT lengths. To characterize the length- and tilt-angle-distribution (Fig. 1d and e, respectively) of short CNTs before and after insertion into lipid bilayers, we manually measured the length of identifiable individual CNTs from a series of high-resolution TEM images (Fig. 1c, Extended Data Figs 1 and 2a). Only CNTs with relatively high signal-to-noise were selected for analysis.

Conventional TEM specimen preparation and imaging. Control samples of isolated cut and uncut CNTs were also investigated (Extended Data Figs 1a, b and 2b, c) with a room-temperature stage using a Philips CM300 TEM, operated at 300 kV with the field-emission gun extraction voltage set at 4.2 kV. Specimens were prepared by first dispersing the CNTs in ethanol using 30 min of bath sonication, and then drop-casting onto lacey-carbon TEM copper grids. The average diameter for the uncut CNTs was 1.51 ± 0.21 nm, as measured over a dozen individual tubes.

Atomic force microscopy (AFM) imaging. AFM imaging was performed using a Multimode Nanoscope IIIA, Nanoscope VIII AFM (Digital Instruments, Santa Barbara, California, USA), and Asylum Research MFP-3D (Asylum Research, Santa Barbara, California, USA), both operating in tapping mode. For imaging in air, we used silicon tips (PPP-FMR, Nanosensors GMBH, Germany) with nominal probe radius of less than 7 nm, force constant of 2.8 N m⁻¹, and resonance frequency of 75 kHz. 10 µl of purified CNT solution was deposited onto freshly cleaved mica pre-treated with poly-L-lysine solution (Ted Pella) to facilitate adsorption of negatively-charged CNTs, allowed to incubate for 2 min, rinsed with DI water, and dried with nitrogen. AFM images typically represent a true height (diameter) value of the CNTs, and the average CNT height observed in the AFM images, 1.54 ± 0.44 nm (Extended Data Fig. 3a), matches the CNT diameter value determined from the cryo-TEM and Raman measurements (Extended Data Figs 3b and c, respectively).

Raman spectroscopy. A droplet of CNT/DOPC suspension was placed onto a glass cover slip, dried and placed into a Renishaw micro-Raman spectrometer (Renishaw, Hoffman Estates, Illinois, USA) using 488 nm laser illumination. The Raman spectrum for short CNTs showed the typical CNT bands at ~1,600 cm⁻¹ (G band) and ~1,300 cm⁻¹ (D band), and show a G/D ratio of 8 (Extended Data Fig. 3c), indicating that the CNT walls maintain a graphitic structure after sonication cutting.

Proton transport assay. To prepare the CNT-containing liposomes loaded with a pH-sensitive fluorescent dye, a 0.1 ml aliquot of 10 mg ml⁻¹ DOPC in chloroform was added to a 20 ml glass vial, and dried into a film by evaporating the solvent. 1 ml of CNT porin (1 mg ml⁻¹) solution in 1 mM HPTS and 150 mM KCl (titrated to pH 8), were added to the dried lipid film, allowed to rehydrate for 1 h and then extruded through a 200 nm diameter pore polycarbonate membrane. For preparation of a control sample, 0.2 ml aliquot of 10 mg ml⁻¹ DOPC in chloroform was dried to a film, rehydrated with 1 ml of 1 mM HPTS and 150 mM KCl (pH = 8), and extruded the same way. Unencapsulated external dye was removed using a size exclusion chromatography (SEC) column packed with Sepharose CL-6B resin (Sigma-Aldrich). After we inserted the CNT porins into the liposome walls we did not observe any leakage of the HPTS dye from the liposomes, indicating that the CNT porins were impermeable to HPTS. All solutions used for these experiments contained a 150 mM isosmotic concentration of background KCl electrolyte to ensure the absence of osmotic pressure gradients.

HPTS dye emission (excitation 450 nm, emission 514 nm) was monitored using a FluoroMax-4 fluorometer (Horiba Inc. Albany, New York, USA). For kinetic measurements, 0.2 ml CNT-liposome sample (or 0.2 ml control liposome sample) was added to the cuvette containing 2 ml of 150 mM KCl (pH = 3). Solution was continuously stirred with a magnetic stir bar during measurements.

Osmotic-pressure-induced liposome size change measurement. 0.2 ml of 10 mg ml⁻¹ DOPC solution in chloroform was added to a vial and dried to a film using a Biotage V-10 evaporator. 1 ml of CNT porin solution was added to the vial and extruded through polycarbonate filter membrane with 200 nm diameter pores using a hand-held extrusion device (mini-extruder, Avanti Polar Lipids). A control sample was prepared in a similar way, except that 0.3 ml of the DOPC solution was used to prepare a lipid film and water was added instead of CNT porin solution.

Hydrodynamic radius of the liposomes was measured using a dynamic light scattering set-up (Zetasizer, Malvern Instruments, Malvern, UK). In a typical measurement, we mixed 10 µl of the liposome solution with 70 µl NaCl solution of known concentration in a low-volume cuvette cell (Malvern Instruments).

Single-channel recordings of CNT porin conductance. A planar lipid bilayer was formed over a 200 µm diameter aperture in a Teflon film partition installed in a two-chamber BLM cell (Eastern Scientific LLC, Rockville, Maryland, USA) using a painting technique²⁹. The *trans*-chamber (connected to a reference Ag/AgCl electrode) and *cis*-chamber (connected to a ground electrode Ag/AgCl) were filled with 2 ml and 0.2 ml conducting buffer (KCl or NaCl with 5 mM HEPES, pH 7.4) respectively. During the measurement a 2 µl solution of DOPC liposomes containing CNT porins or the stock CNT porin solution was added to the *cis*-chamber.

A holding potential between -200 mV to 200 mV was applied to the reference electrode, and the *trans*-membrane current signal was recorded by a patch clamp

set-up, which consisted of an Axopatch 200B patch clamp amplifier and an Axon DigiData 1322A analogue-digital converter (Axon Instruments, Milpitas, California, USA) connected to a computer system running Clampex 10.3 software (Axon Instruments). Traces were acquired at a sampling frequency of 10–100 kHz and were low band-pass filtered at a frequency of 5 kHz or 2 kHz to increase signal-to-noise ratio. The data were analysed and exported using PClamp 10.3 software (Axon Instruments), and further analysed using Origin Pro 8.0 and Igor Pro 6.31.

As a positive control, the conductance of gramicidin channels was measured using the following procedure: 0.1 mg gramicidin A was incubated with 200 μ l of the DOPC liposome solution (150 mM KCl, 4 mg ml⁻¹ lipid concentration, pH 7.4). After a lipid bilayer was formed over the testing set-up aperture (in the presence of 1 M KCl solution in both chambers), 2 μ l of gramicidin A/liposome solution was added into the *cis*-chamber and the conductance traces were recorded using a transmembrane potential of 75 mV. The conductance measured for gramicidin was 0.058 \pm 0.008 nS, which is close to the reported values of 0.0459 \pm 0.005 pS³⁰.

DNA translocation measurements. We used a similar set-up and sample preparation technique to that described in the preceding section. ssDNA oligonucleotide (5'-/5Phos/GCG GCC GCT ACT AGT CTT ACC GCC ACC CAG AGG GCC ATA ACG GGT ACG GTA TTG GCT TAC ACG GTT ACG CAG ACG CTG TAC-3', 81 nt length) was pre-mixed into the solution (1 M KCl, pH 7.4) in the chambers to a final concentration of 100 pM. After a planar lipid bilayer was formed, 2 μ l of CNT-containing liposomes was added into the solution in the *cis*-chamber by pipette. A holding or ramping potential of 50 mV maximum amplitude was applied across the bilayer. For the control experiment no ssDNA was added to the chambers. Translocation events from current traces of a total of 338 s duration were counted either manually or using custom-written software in Igor Pro (WaveMetrics, Inc., Lake Oswego, Oregon, USA).

Cell culture and on cell patch-clamp measurements. CHO and HEK 293T cells were cultured according to the prescribed protocol (ATCC website), at 37 °C, 5% CO₂, in 25 cm² flasks, in DMEM (HyClone SH30243.01 High Glucose) supplemented with gentamicin antibiotic (5 mg l⁻¹) and 10% serum (DE14-801F, Lonza, Allendale, New Jersey, USA). These cells are known for low endogenous channel activity and, consequently, widely used for heterologous expression of ion channels. 24 h before the experiment, the cells were plated in 35 mm Petri dishes (with plastic bottom observation window, 35 mm low μ -Dish cat#80136, Ibidi, Martinsried, Germany) using PBS without Ca²⁺/Mg²⁺ (D8537, Sigma, St Louis, Missouri, USA). For the patch-clamp experiment the culturing media was substituted with external solution (PBS with Ca²⁺/Mg²⁺, Sigma, D8662), and the experiments were conducted within the next 45 min, at room temperature.

Patch-pipettes were produced from borosilicate glass (GB150-10, Science Products, Hofheim, Germany) using a Sutter P-2000 puller (Sutter Instruments, Novato, California, USA) and filled with the solution containing (in mM): 150 KCl, 10 HEPES, and 1 EDTA pH 7.0; the pipette resistance was 3–5 M Ω . CNT porin solution was added to the pipette solution before each experiment. We used 0.15% dilution of the CNT porin stock solution in the pipette buffer, a higher amount of the CNT interfered with the gigaseal formation. The data were obtained and pre-processed using an EPC10 patch-clamp amplifier (HEKA Elektronik, Lambrecht/Pfalz, Germany), at 5 kHz sampling rate with the 2.9 kHz internal filter of the EPC10. At the beginning of each record we applied a series of holding potential steps (from -60 mV to +60 mV) to estimate the reversal potential. For data acquisition, we used cells with approximately symmetric current response (Extended Data Fig. 10e), using -30 mV and -60 mV holding potentials which induced minimal endogenous channel activity. Stable conductance levels were detected using FITMASTER software (HEKA). The background (reference) conductance was determined from an all-point histogram. The Gaussian fits to the conductance histograms for HEK293T cells (Fig. 4b, blue bars, red solid line) give the peak values of 68.1 \pm 0.8, 122.7 \pm 1.0, 177.8 \pm 2.1, 223.4 \pm 2.7 pS. For CHO cells (Fig. 4c, red bars, blue solid line) the peak values were 67.6 \pm 1.4, 130.8 \pm 1.3, 195.3 \pm 1.1, and 255.3 \pm 10 pS.

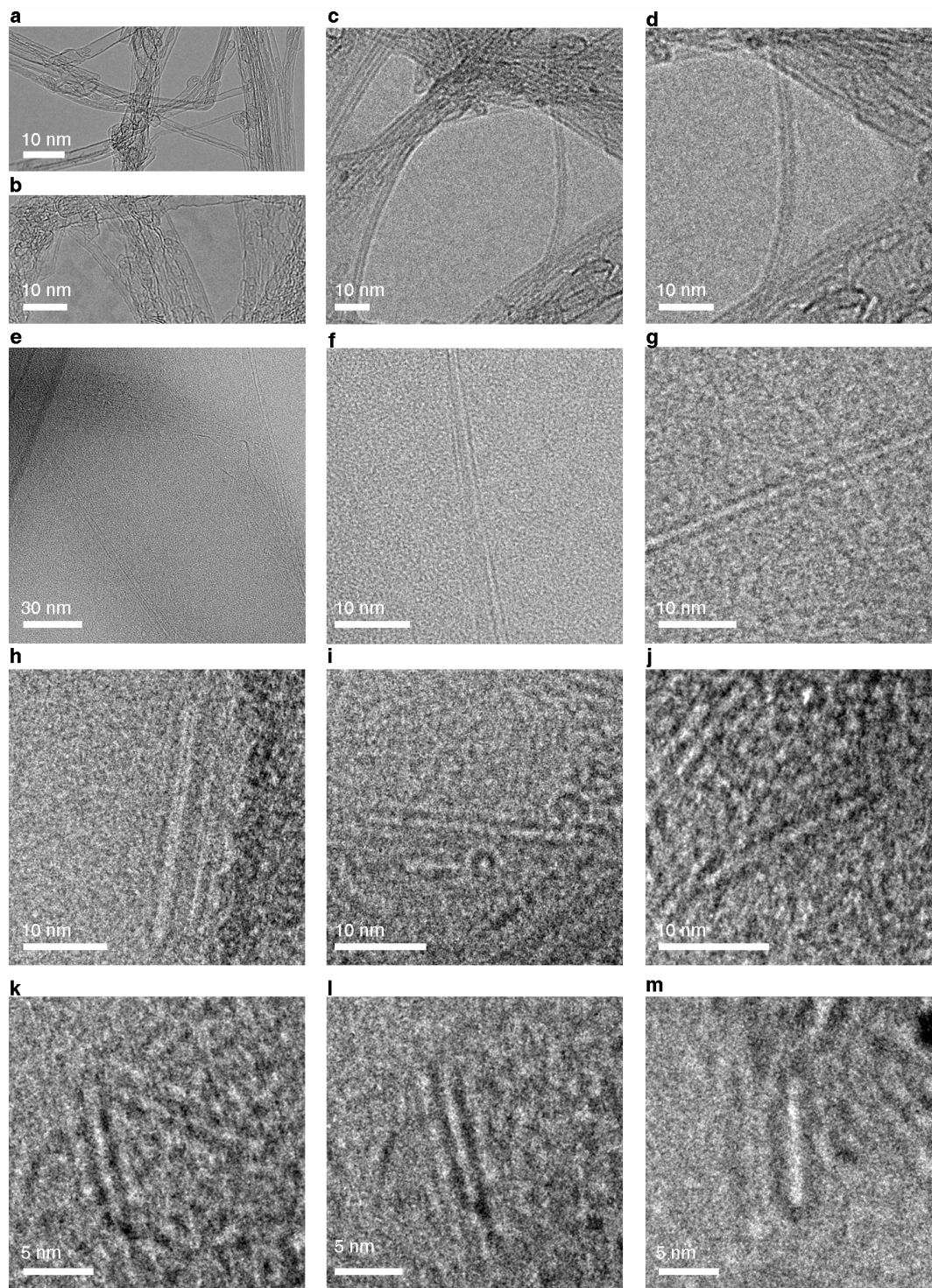
Patch-clamp measurements on planar lipid bilayer and giant unilamellar vesicles. Planar lipid bilayers were prepared using a similar methodology as above, based on the modified Montal-Mueller 'painting' procedure described elsewhere³¹. The bilayers were produced from 20 g l⁻¹ solution of lipids in squalane; charged membranes contained (in mol%): 28% DOPC, 25% DOPE, 30% Chol, 15% DOPS, and 2% PI(4,5)P₂; non-charged membranes contained (in mol%) 50% DOPC, 20% DOPE, and 30% Chol. Membranes were formed on 300x mesh copper TEM grids (Pelco, IGC300) mounted on the bottom of the Petri dish using adhesive spacers (Grace BioLabs Inc., Oregon, USA) as described previously³². Grids were pretreated with the appropriate combination of lipids dissolved in 1:1 mol/mol octane:decane, 20 g l⁻¹ (ref. 31). The bilayers were formed in the working Buffer 1, containing (in mM): 150 KCl, 10 mM Hepes, and 1 mM EDTA, pH 7.

Giant unilamellar vesicles (GUVs) were produced by spontaneous swelling of lipid films deposited on silica beads³². The 'charged membrane' composition was used, supplemented with 0.5 mol% of Rh-DOPE for fluorescence microscopy detection of GUVs. The GUV preparation was modified to allow for the CNT incorporation into lipid films before GUV formation. Briefly, the initial lipid film (total 0.05 mg of the appropriate lipid mixture in chloroform/methanol (9:1 vol/vol) mixture) was formed in a round bottom flask using a rotary evaporator. 1 ml of CNT stock solution was added and mixed thoroughly to resuspend the lipid film. The mixture was dried in vacuum and then 50 μ l of chloroform was added and rapidly dried in a rotary evaporator in order to form a homogeneous lipid-CNT film. 100 μ l of distilled water was added for a 30 min pre-hydration of the film at 50 °C. The film was vortexed and the resulting mixture was added to solution containing 40 μ m plain silica beads (Corpuscular Inc., Cold Spring, New York, USA): a small (2.5 μ l) drop of CNT/lipid mixture was added to a 0.5 μ l drop of the bead solution (10% vol/vol) deposited on Teflon film. The drops were dried in a vacuum chamber drier (15 min, room temperature). CNT-GUVs formation was detected within 10 min after the addition of 300 mM sucrose buffer at room temperature.

CNT-GUVs remained attached to the beads transferred to the observation chamber (with a glass-bottom observation window pretreated with bovine serum albumin as described previously)³² containing Buffer 1. The chamber was mounted in an inverted microscope equipped with an Andor iXon+ camera (Andor Technology, Belfast, UK) operating with the ImageJ- μ Manager open source software³³ and a LUCPLFL 40x/0.60NA objective. 550/590 nm excitation/emission wavelengths were used. Conductance measurements were performed 30 min after the GUV transfer to Buffer 1 to allow for the equilibration between the GUVs' interior solution and the buffer.

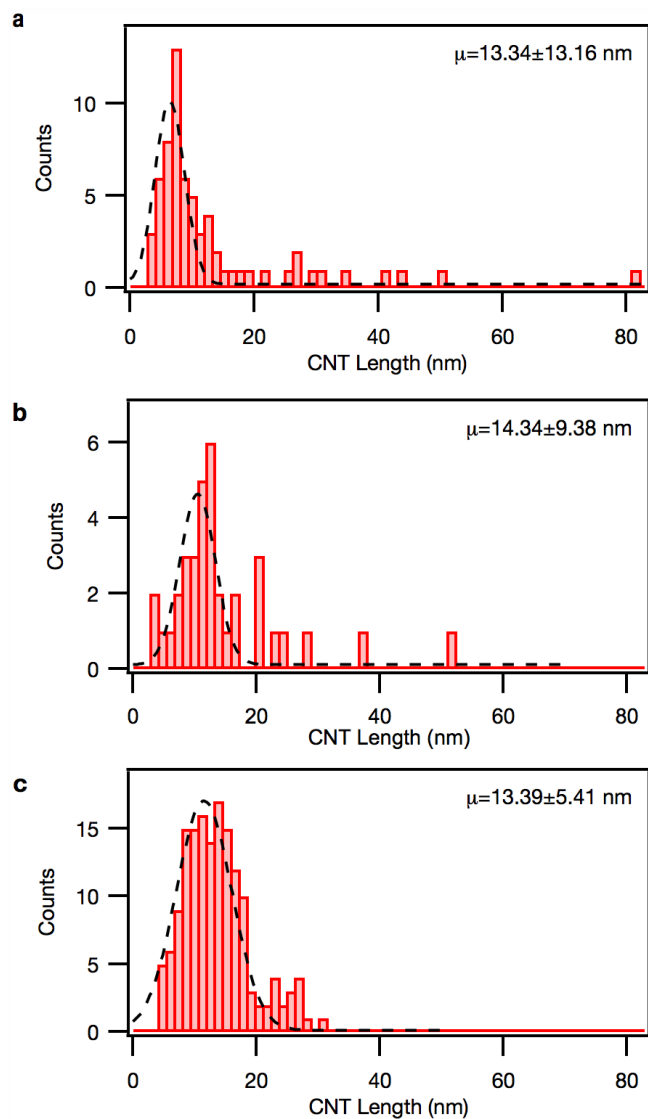
For patch-clamp experiments on planar bilayers and GUVs we used borosilicate glass pipettes (Science Products, GB150-10) filled with Buffer 1. The pipette resistance was 3–5 M Ω and 0.15% to 0.30% of the CNT porin solution was added to the pipette solution. Experiments were conducted at room temperature. The pipette approach to the bilayer and the formation of a gigaseal was performed as described earlier^{31,32}. The patch conductance was measured at +50 and -50 mV holding potentials using an Axopatch 200B (Axon Instruments) amplifier and a PC-44 acquisition board (Signallogix). The sampling frequency f was 1 kHz and signals were passed through 8-pole Bessel filters (Frequency Devices) set at $f/2$ corner frequency. The data were processed using FITMASTER software (HEKA).

29. Haque, F., Geng, J., Montemagno, C. & Guo, P. Incorporation of a viral DNA-packaging motor channel in lipid bilayers for real-time, single-molecule sensing of chemicals and double-stranded DNA. *Nature Protocols* **8**, 373–392 (2013).
30. Neher, E., Sandblom, J. & Eisenman, G. Ionic selectivity, saturation, and block in gramicidin A channels. *J. Membr. Biol.* **40**, 97–116 (1978).
31. Frolov, V. A., Lizunov, V. A., Dunina-Barkovskaya, A. Y., Samsonov, A. V. & Zimmerberg, J. Shape bistability of a membrane neck: a toggle switch to control vesicle content release. *Proc. Natl Acad. Sci. USA* **100**, 8698–8703 (2003).
32. Shnyrova, A. V. et al. Geometric catalysis of membrane fission driven by flexible dynamins rings. *Science* **339**, 1433–1436 (2013).
33. Edelstein, A., Amodaj, N., Hoover, K., Vale, R. & Stuurman, N. Computer control of microscopes using μ Manager. *Curr. Protoc. Molec. Biol.* **92**, 14.20.1–14.20.17 (2010).

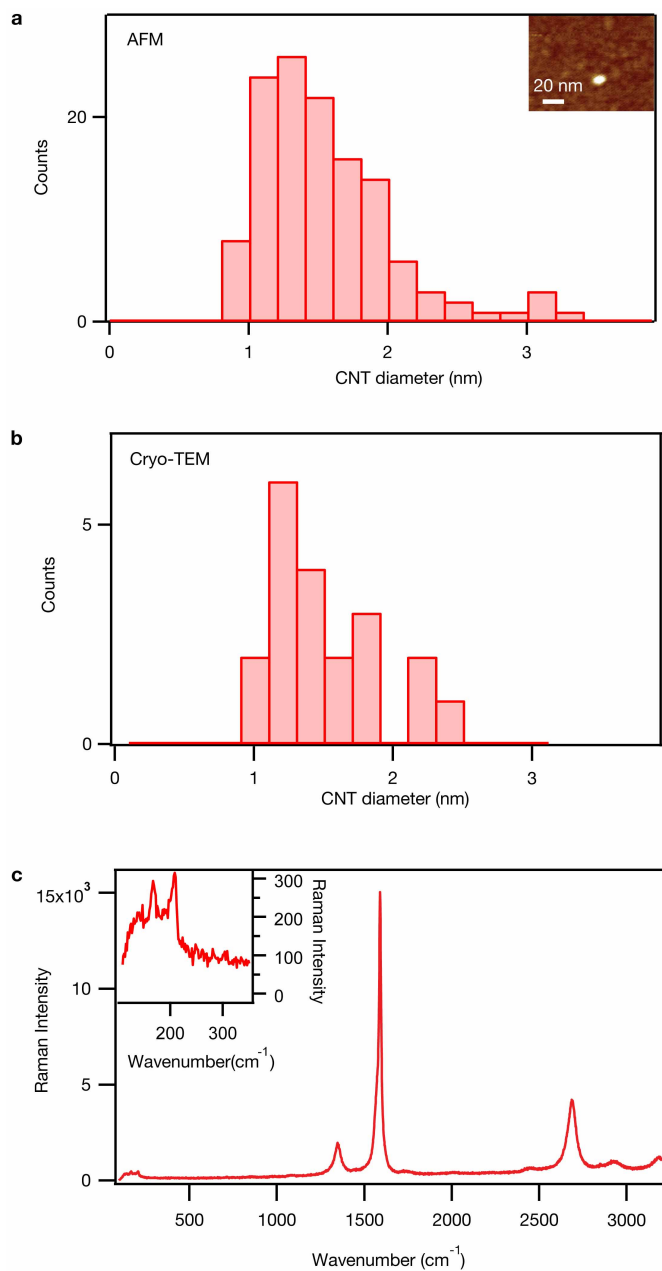


Extended Data Figure 1 | TEM images of isolated CNTs. **a–c**, Purified CNTs before cutting (**a** and **b**, conventional TEM, **c** imaged on a cryo stage; all air-dried samples). **d**, Magnified image of one CNT from **c**. **e–g**, CNTs mixed

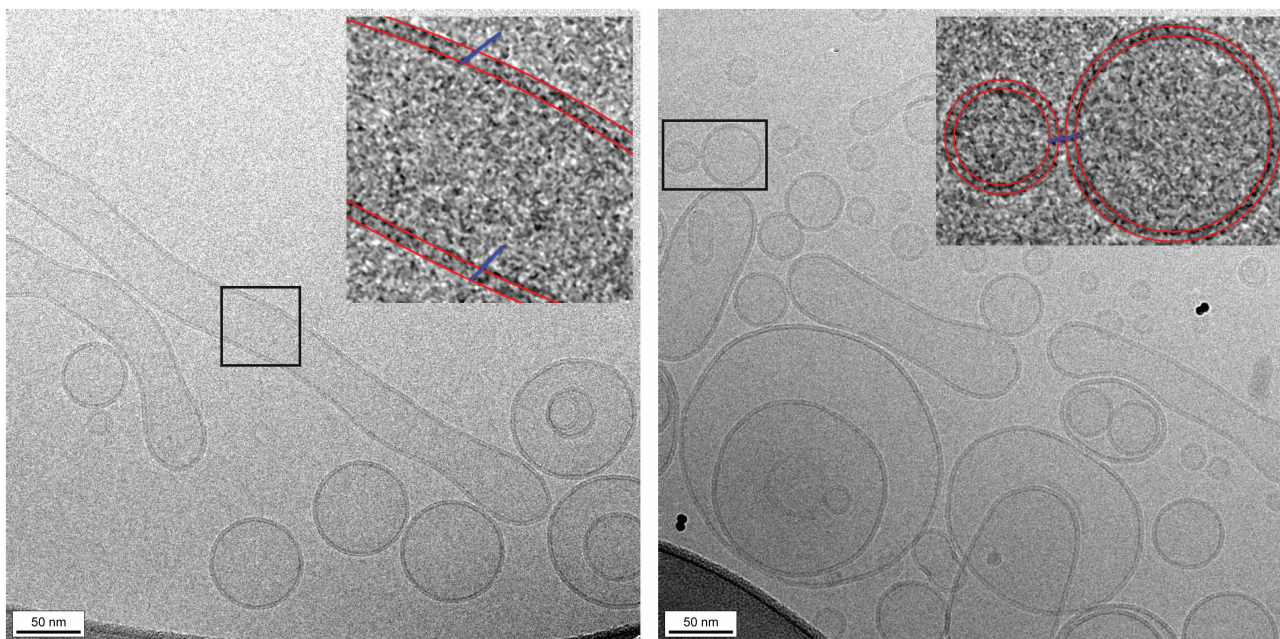
with lipid and lightly sonicated to aid nanotube dispersion. **h–m**, Isolated CNT porins after the cutting procedure.



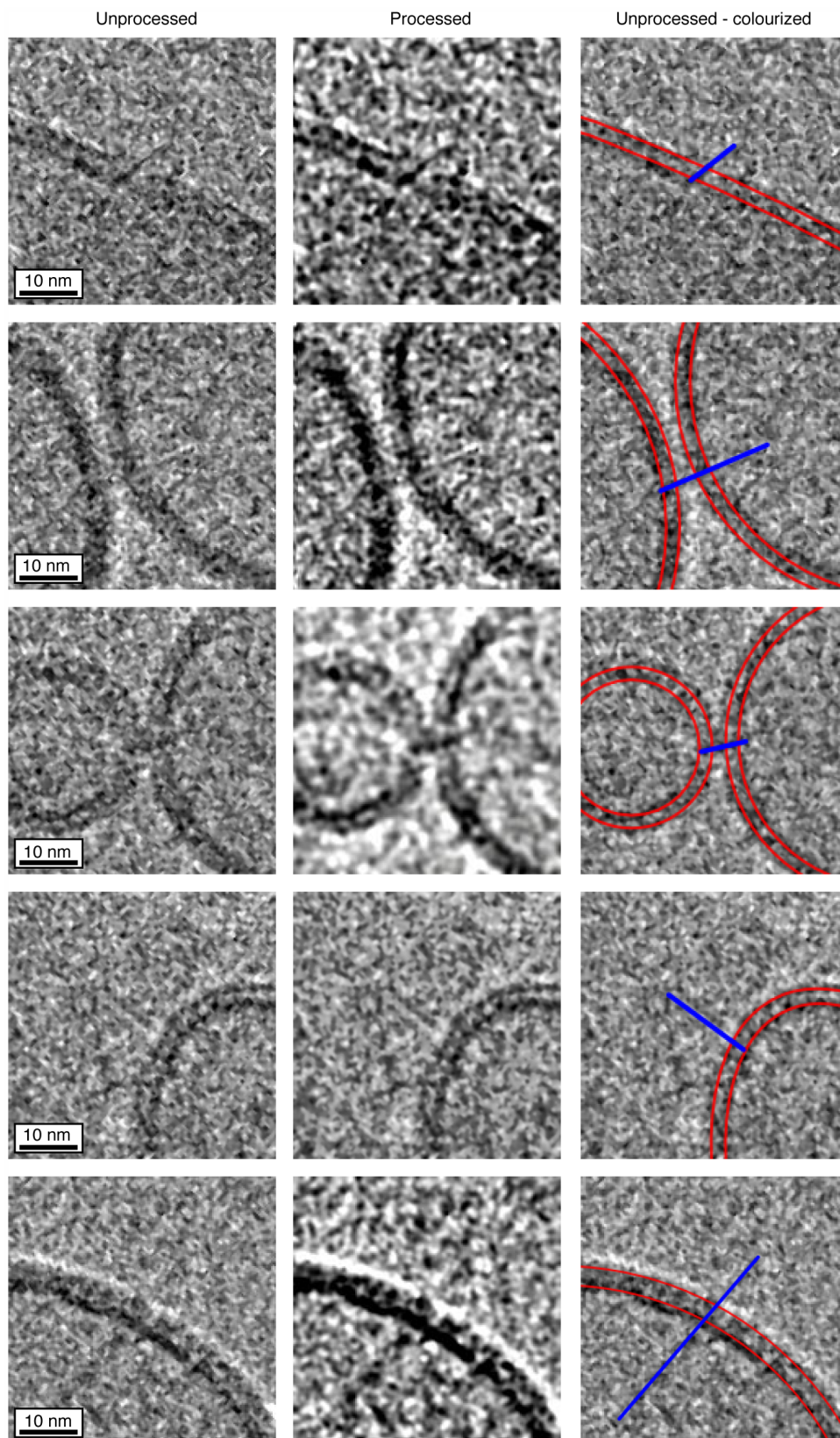
Extended Data Figure 2 | Length distributions of CNT porins measured from TEM images. **a**, CNT porin sample after cutting and purification procedures ($n = 65$). **b**, Cut and purified sample after passing through a 100 nm pore filter membrane ($n = 36$). **c**, As **b** but after passing through a 50 nm pore filter membrane ($n = 156$). μ indicates the mean CNT porin length.



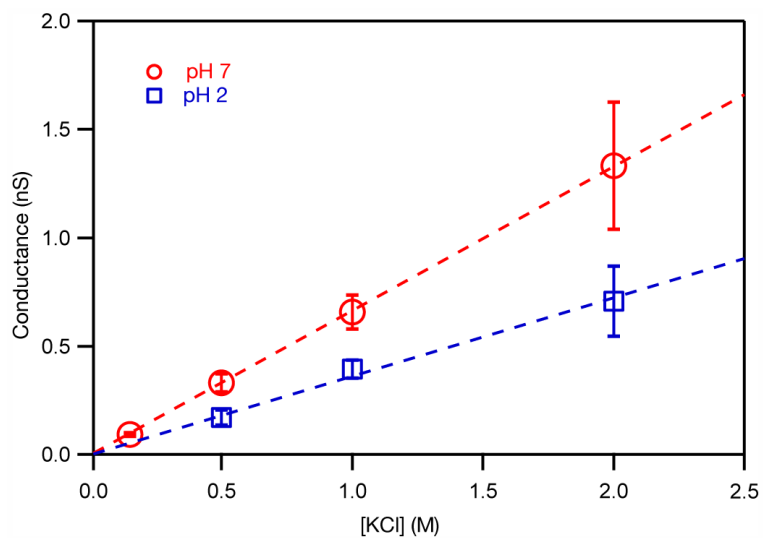
Extended Data Figure 3 | Characterization of CNT porins. **a**, Histogram of the diameters of short CNT fragments measured with AFM. Most CNT porins show diameters between 1 and 2 nm. Inset, high-magnification AFM image of a single CNT on a bare mica surface. **b**, Histogram of the diameters of CNT porins inserted into the lipid vesicles measured by cryo-TEM. **c**, Raman spectrum of the short-CNT/lipid complex after 16 h of sonication-assisted cutting. Inset, magnified view of the radial breathing mode region of the CNT spectrum ($100\text{--}300\text{ cm}^{-1}$).



Extended Data Figure 4 | Cryo-TEM images of CNT-liposome complexes. Original large field-of-view images (main panels); insets, magnified views of the selected structures, with false-colour lines added to highlight the lipid bilayer (red) and CNT porins (blue).

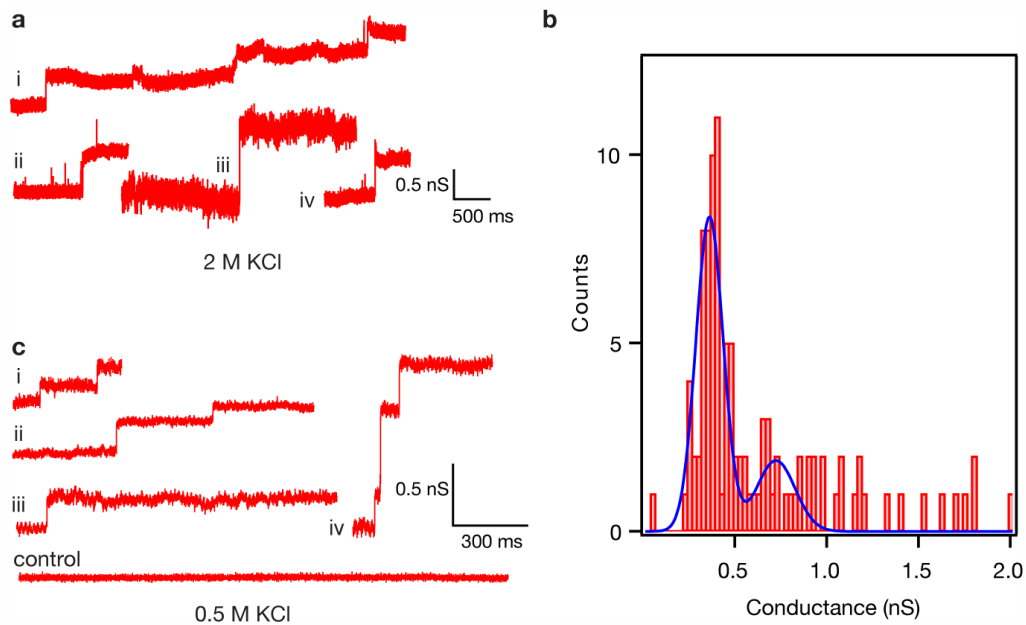


Extended Data Figure 5 | High magnification cryo-TEM imaged of CNT-liposome complexes. Panels (left to right) show unprocessed, image-processed, and colourized unprocessed examples. See Methods for details of image processing.



Extended Data Figure 6 | Single CNT porin conductance at different electrolyte concentrations. Plots of the average single CNT porin conductance at different values of background electrolyte (KCl) concentrations at pH 7 (red circles) and pH 2 (blue squares). The slopes of the linear fit through

the data (dashed lines) correspond to 0.66 nS M^{-1} at pH 7 and 0.36 nS M^{-1} at pH 2. Error bars, s.d.; n values as follows (pH = 7: 0.15 M, $n = 6$; 0.5 M, $n = 161$; 1 M, $n = 39$; 2 M, $n = 8$. pH = 2: 0.5 M, $n = 18$; 1 M, $n = 3$; 2 M, $n = 36$).

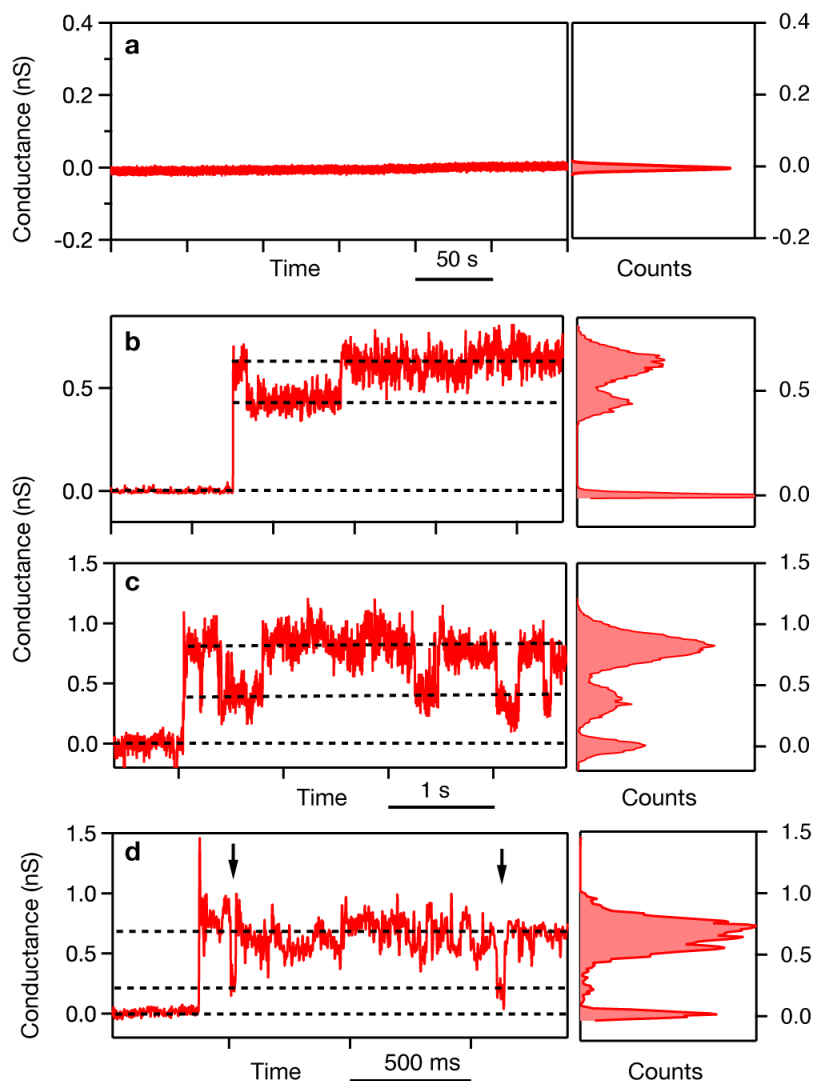


Extended Data Figure 7 | Single CNT porin conductance at pH 2.

a, Representative conductance traces showing individual CNT porin incorporation at pH 2 and 2 M KCl background electrolyte concentration. Traces i–iv represent incorporation from 4 independent experiments.

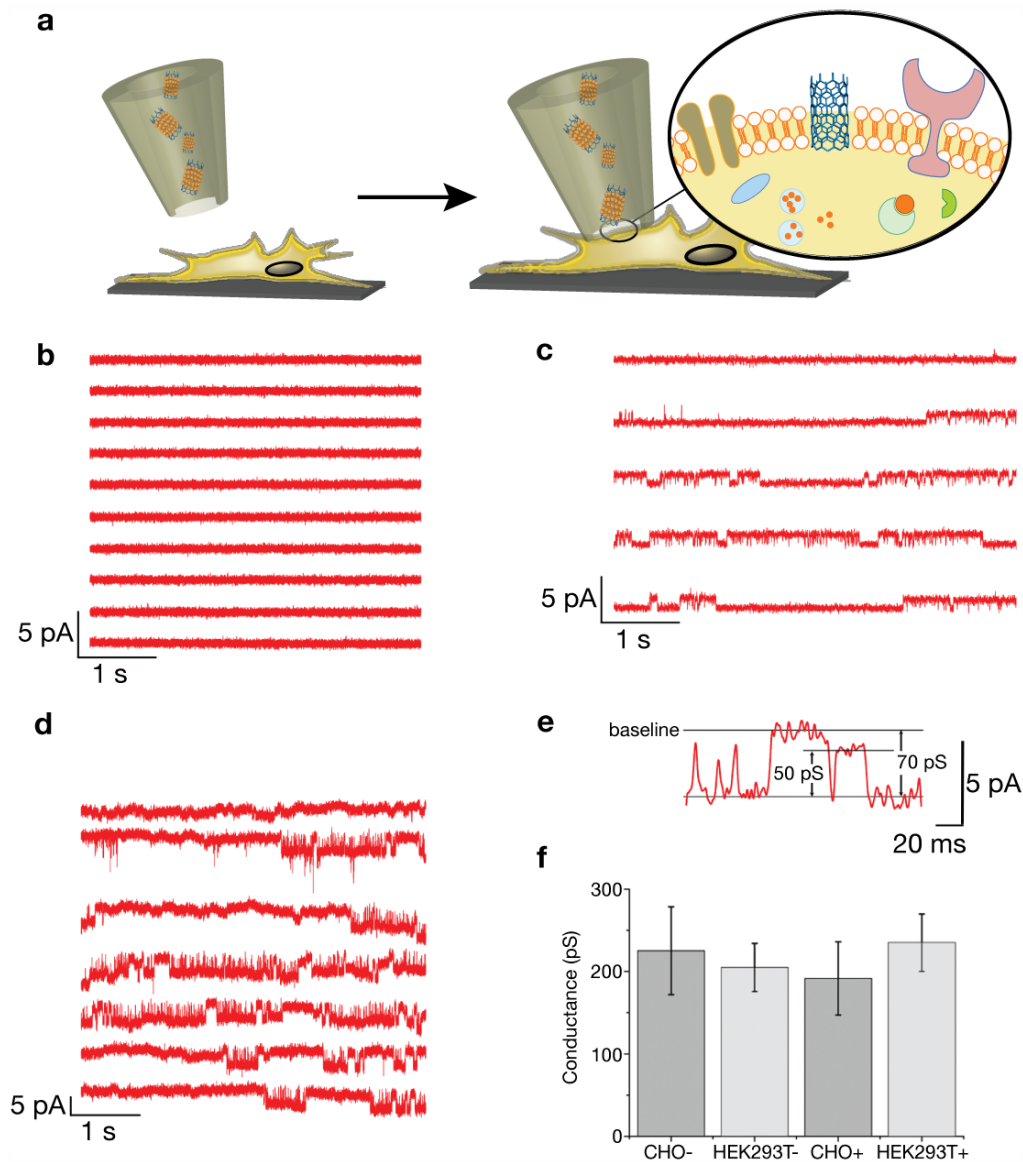
b, Histogram of conductance values measured for 109 individual CNT channel incorporation events ($n = 109$). Solid blue line corresponds to the fit of the data to a distribution expected for multiple channel incorporation, where the

position and width of the second peak are determined from the position, μ (0.36 ± 0.006 nS), and width, σ (0.10 ± 0.008 nS), of the first peak as 2μ and $\sqrt{2}\sigma$. **c**, Representative conductance traces showing individual CNT porin incorporation at pH 2 and 0.5 M KCl background electrolyte. Traces i–iv represent incorporation from 4 independent experiments. Control trace shows a conductance trace collected at the same conditions without adding CNT porins.



Extended Data Figure 8 | Representative traces and corresponding current level histograms of CNT porin incorporation, gating and DNA translocation in lipid bilayers at pH 7. **a**, Control conductance trace recorded in the absence of CNT porin. Although only the first 5 min are shown, the current level remained stable for up to 80 min ($n = 3 \times 10^6$). **b**, **c**, Conductance traces and histograms showing single CNT porin incorporation events and subsequent current fluctuations between gating sub-levels ($n > 5 \times 10^4$).

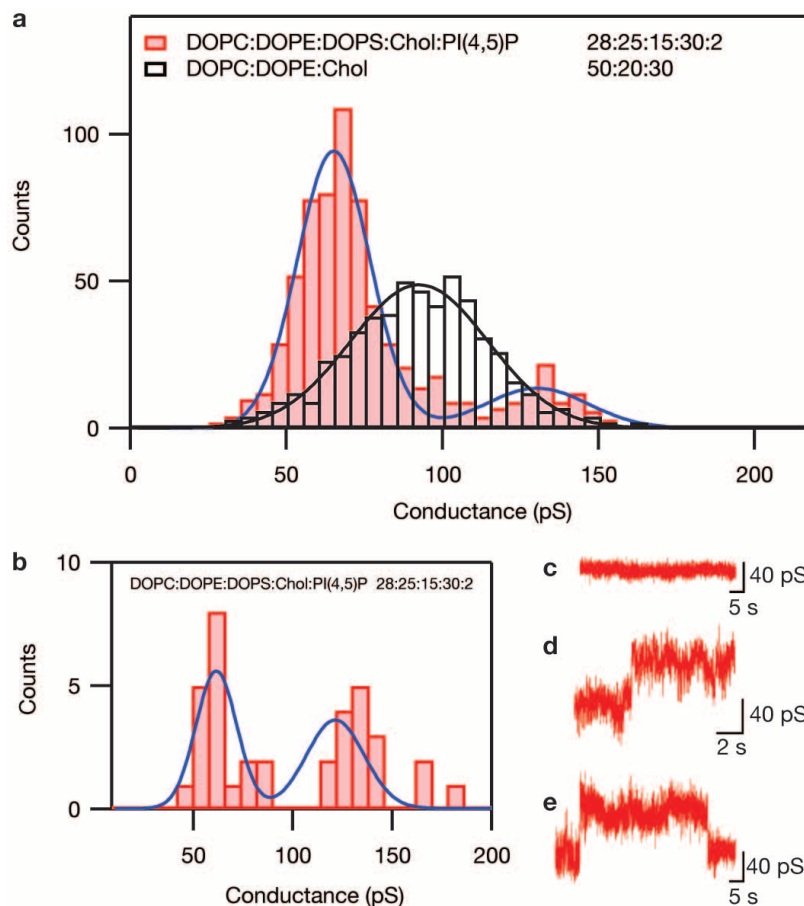
d, Trace and conductance value histogram showing CNT porin incorporation event and subsequent two DNA translocation events (indicated by black arrows). Here $n = 3.6 \times 10^5$. The peaks on the histogram correspond to 0.01 nS for pre-incorporation baseline, 0.22 nS for DNA blockade, and 0.67 nS for open pore conductance, which gives the DNA blockade of 0.45 nS and a blockade ratio of 67%. Dashed lines are added to all traces as a guide to the eye.



Extended Data Figure 9 | CNT porin conductance in cellular membranes.

a, Schematic of the on-cell patch clamp measurement. **b**, **c**, On-cell conductance measurements demonstrate no channel activity (**b**) and appearance of a small flickering channel (**c**) in control experiments with HEK293T cells. Holding potentials were -60 mV (**b**) and $+60$ mV (**c**), the time order is from left to right, top to bottom. **d**, In the presence of CNT porin (0.15% dilution of the stock), large channel activity is detected, holding potential -60 mV. **e**, Metastable conductance sub-states of 50 pS detected with CNT

porin incorporated into the plasma membrane of CHO cell, holding potential -60 mV. **f**, Mean patch conductance (on cell configuration, 0.15% CNT stock solution in the patch pipette) detected in CHO and HEK293T cells at positive and negative holding potentials. Error bar, s.d.; 4 cells). In $\sim 70\%$ of HEK293T cells and 95% of CHO cells we did not observe significant dependence of the conductance on the holding potential; only those cells were used for the conductance analysis.



Extended Data Figure 10 | CNT porin conductance in patches isolated from planar lipid bilayers and giant unilamellar vesicles. **a**, Histogram of conductance values measured for CNT porin incorporation into a planar lipid bilayer mimicking the composition of a charged plasma membrane (red bars, $n = 681$, 5 membranes, 3 controls without CNT porins showed no events) and a similar lipid bilayer without charged species (black bars, $n = 540$, 5 membranes, 3 controls without CNT porins showed no events). The background electrolyte concentration was kept at 150 mM. Solid lines correspond to the fit of the data to a distribution expected for multiple channel incorporation. The primary peak positions are 65.1 ± 0.4 pS (charged bilayer)

and 92.5 ± 0.6 pS (uncharged bilayer). **b**, Histogram of conductance values measured for CNT porins reconstituted into giant unilamellar vesicles (GUVs) mimicking the composition of charged plasma membrane (red bars, $n = 40$, 3 membranes). Solid lines correspond to the fit of the data to a distribution expected for multiple channel incorporation with the primary conductance peak at 60 ± 2 pS. **c**, Representative control experiment demonstrating lack of spontaneous pore formation in the bilayer patch. **d**, **e**, Patch-clamp measurements of the CNT porin incorporation events in planar lipid bilayer (**d**) and GUV membrane (**e**).

Centennial-scale changes in the global carbon cycle during the last deglaciation

Shaun A. Marcott^{1,2}, Thomas K. Bauska¹, Christo Buizert¹, Eric J. Steig³, Julia L. Rosen¹, Kurt M. Cuffey⁴, T. J. Fudge³, Jeffery P. Severinghaus⁵, Jinho Ahn⁶, Michael L. Kalk¹, Joseph R. McConnell⁷, Todd Sowers⁸, Kendrick C. Taylor⁷, James W. C. White⁹ & Edward J. Brook¹

Global climate and the concentration of atmospheric carbon dioxide (CO₂) are correlated over recent glacial cycles^{1,2}. The combination of processes responsible for a rise in atmospheric CO₂ at the last glacial termination^{1,3} (23,000 to 9,000 years ago), however, remains uncertain^{1–3}. Establishing the timing and rate of CO₂ changes in the past provides critical insight into the mechanisms that influence the carbon cycle and helps put present and future anthropogenic emissions in context. Here we present CO₂ and methane (CH₄) records of the last deglaciation from a new high-accumulation West Antarctic ice core with unprecedented temporal resolution and precise chronology. We show that although low-frequency CO₂ variations parallel changes in Antarctic temperature, abrupt CO₂ changes occur that have a clear relationship with abrupt climate changes in the Northern Hemisphere. A significant proportion of the direct radiative forcing associated with the rise in atmospheric CO₂ occurred in three sudden steps, each of 10 to 15 parts per million. Every step took place in less than two centuries and was followed by no notable change in atmospheric CO₂ for about 1,000 to 1,500 years. Slow, millennial-scale ventilation of Southern Ocean CO₂-rich, deep-ocean water masses is thought to have been fundamental to the rise in atmospheric CO₂ associated with the glacial termination⁴, given the strong covariation of CO₂ levels and Antarctic temperatures⁵. Our data establish a contribution from an abrupt, centennial-scale mode of CO₂ variability that is not directly related to Antarctic temperature. We suggest that processes operating on centennial timescales, probably involving the Atlantic meridional overturning circulation, seem to be influencing global carbon-cycle dynamics and are at present not widely considered in Earth system models.

Ice cores from Greenland^{6,7} provide unique records of rapid climate events of the past 120 kyr. However, because of relatively high concentrations of impurities⁸, Greenlandic ice cores do not provide reliable atmospheric CO₂ records. Antarctic ice cores, which contain an order-of-magnitude fewer impurities, provide reliable records⁸, but existing data from the deglacial period from Antarctica either provide only low temporal resolution⁹ or are of relatively low precision¹⁰.

The West Antarctic Ice Sheet Divide ice core (WDC) (79.467° S, 112.085° W, 1,766 m above sea level) was drilled to a depth of 3,405 m in 2011 and spans the past ~68 kyr. At present, the site has a mean annual snow accumulation of 22 cm ice equivalent per year and a surface temperature of -30 °C. Annual layer counting to 2,800 m depth (~30 kyr ago) provides a very accurate timescale for comparison with data from other archives¹¹. The difference in age (Δ age) between the ice and the gas trapped within it, which is critical for developing a gas-age chronology, is 205 ± 10 yr at present and was 525 ± 100 yr at the last glacial maximum (LGM) (Extended Data Fig. 1). Given the high accumulation at the site, minimal smoothing due to gas transport and gradual occlusion, and precise chronological constraints, WDC is the best Antarctic

analogue to central Greenlandic deep ice cores, with a substantially better-dated gas chronology during the glacial period, and is able to resolve atmospheric CO₂ at sub-centennial resolution.

We exploit the unique aspects of WDC to reconstruct atmospheric CO₂ and CH₄ concentrations at high resolution over the last deglaciation. The atmospheric CH₄ record is important in this context because CH₄ varies in phase with rapid Greenlandic climate changes during the deglaciation and therefore provides a means of synchronizing our Antarctic record with Northern Hemisphere climate¹². Atmospheric CH₄ is primarily produced by land-based sources, principally tropical and boreal wetlands¹³ where emissions are driven by changes in temperature and hydrology, and CH₄ therefore serves as an indicator of processes in the terrestrial biosphere.

The WDC deglacial CO₂ record is characterized by a long-term ~80 p.p.m. increase, similar to other cores^{9,14}, which begins at 18.1 kyr ago, occurs in several discrete steps throughout the deglacial transition, and ends 7 kyr later, during the early Holocene epoch. Given the precise WDC chronology, the phasing of CO₂ changes with respect to Antarctic and global temperature² can be more precisely evaluated. The timing of the initial deglacial CO₂ rise leads the initial rise in global temperature² (17.2 kyr ago) by several centuries, and continues to lead across the entire termination (Extended Data Fig. 7). The timing of the initial CO₂ rise is consistent with the newest gas chronology of EPICA Dome C⁵ (EDC), but more highly resolved and precise given the smaller gas-age uncertainty in WDC than in EDC (Extended Data 4). The WDC methane record shows in detail the abrupt changes at the onset of the Bølling-Allerød and Younger Dryas stadials and the start of the Holocene. We observe a smaller, but prominent, methane excursion at 16.3 kyr ago not previously reported from other records (Fig. 1). Our record also resolves the beginning of the deglacial methane rise at 17.8 kyr ago (Extended Data Table 1).

The WDC CO₂ record demonstrates that CO₂ varied in three distinct modes during the deglaciation. The first mode is relatively gradual change (~10 p.p.m. kyr⁻¹): such changes in CO₂ began at 18.1 and 13.0 kyr ago and were broadly coincident with a reduction in the strength of the Atlantic meridional overturning circulation (AMOC; Fig. 2i), a cold North Atlantic² and warming in the Southern Hemisphere². The second mode is rapid increase: 10–15 p.p.m. increases in CO₂ occurred in three short (100–200 yr) intervals at 16.3, 14.8 and 11.7 kyr ago, the latter two at times of rapid resumption of the AMOC and warming in the Northern Hemisphere. The rapid changes at 14.8 and 11.7 kyr ago were first noted at EDC^{9,14}, but the magnitude, duration and timing are now more fully resolved because of the unique site conditions at WDC. The third mode is no change in atmospheric CO₂. These apparent plateaus in the WDC CO₂ record lasted for 1,000–1,500 yr and occurred directly after the rapid, century-scale increases.

¹College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA. ²Department of Geoscience, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA. ³Department of Earth and Space Sciences, University of Washington, Seattle, Washington 98195, USA. ⁴Department of Geography, University of California, Berkeley, California 94720, USA. ⁵Scripps Institution of Oceanography, University of California, San Diego, California 92037, USA. ⁶School of Earth and Environmental Sciences, Seoul National University, Seoul 151-742, South Korea. ⁷Desert Research Institute, Nevada System of Higher Education, Reno, Nevada 89512, USA. ⁸Earth and Environmental Systems Institute, Pennsylvania State University, University Park, Pennsylvania 16802, USA. ⁹INSTAAR, University of Colorado, Boulder, Colorado 80309, USA.

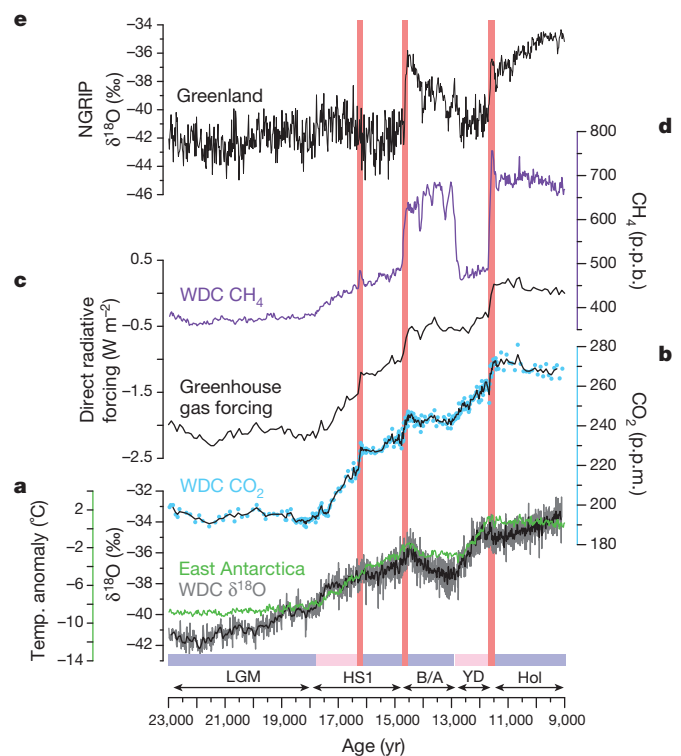


Figure 1 | Greenhouse gas and stable water isotope measurements from Antarctica and Greenland. **a**, Oxygen isotopes from WDC¹¹ (grey; black line is 11-point weighted average) and water-isotope-derived temperature composite from East Antarctica⁵ (green). $\delta^{18}\text{O} = (^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{VSMOW}} - 1$; VSMOW, Vienna Standard Mean Ocean Water. **b**, Atmospheric CO_2 concentrations (this study; black line is 5-point weighted average). **c**, Direct radiative forcing of CO_2 , CH_4 and N_2O (ref. 31) using a simplified expression³². **d**, Atmospheric CH_4 concentrations (this study and ref. 11). **e**, Oxygen isotope measurements from the North Greenland Ice Project⁶ (NGRIP). Coloured bands at bottom indicate times when CO_2 is stable (blue), slowly increasing (pink) or rapidly increasing (red), as described in the text. LGM, Last Glacial Maximum; HS1, Heinrich stadial 1 (~18.0–14.6 kyr ago); B/A, Bolling–Allerød; YD, Younger Dryas; Hol, Holocene.

The increase in CO_2 from glacial to interglacial states is typically explained by a combination of changes in deep-ocean ventilation, Antarctic sea-ice cover, ocean salinity and temperature, marine biological surface nutrient utilization, changes in ocean alkalinity, and carbonate compensation in the ocean¹⁵. These processes are generally thought to take place at millennial or longer timescales. However, our new observations from WDC demonstrate that in several instances the observed atmospheric CO_2 change was rapid, on centennial timescales; that the most rapid changes correlate in time with ice-raftered debris events in the North Atlantic or with abrupt changes in the AMOC (Fig. 2); and that the abrupt changes in CO_2 do not always have a centennial counterpart in the Antarctic $\delta^{18}\text{O}$ record (Fig. 1).

Given these new observations, we suggest that the AMOC had a critical role in the glacial–interglacial rise of atmospheric CO_2 and may have acted as a modulator of both the timing and the magnitude of CO_2 and CH_4 change during the last termination, either directly, through changes in ocean circulation, or indirectly, through the redistribution of heat and the associated climate response. Elements of this hypothesis have been outlined previously: a reduction in the AMOC is thought to cause increased overturning in the Southern Ocean, which allows CO_2 in the deep ocean to enter the ocean mixed layer more rapidly, where it can exchange with the atmosphere¹⁶. For methane, a cooling in the Northern Hemisphere resulting from a reduction in the AMOC and a southward expansion of sea-ice extent may have displaced the intertropical convergence zone southwards¹⁷ resulting in decreased production of

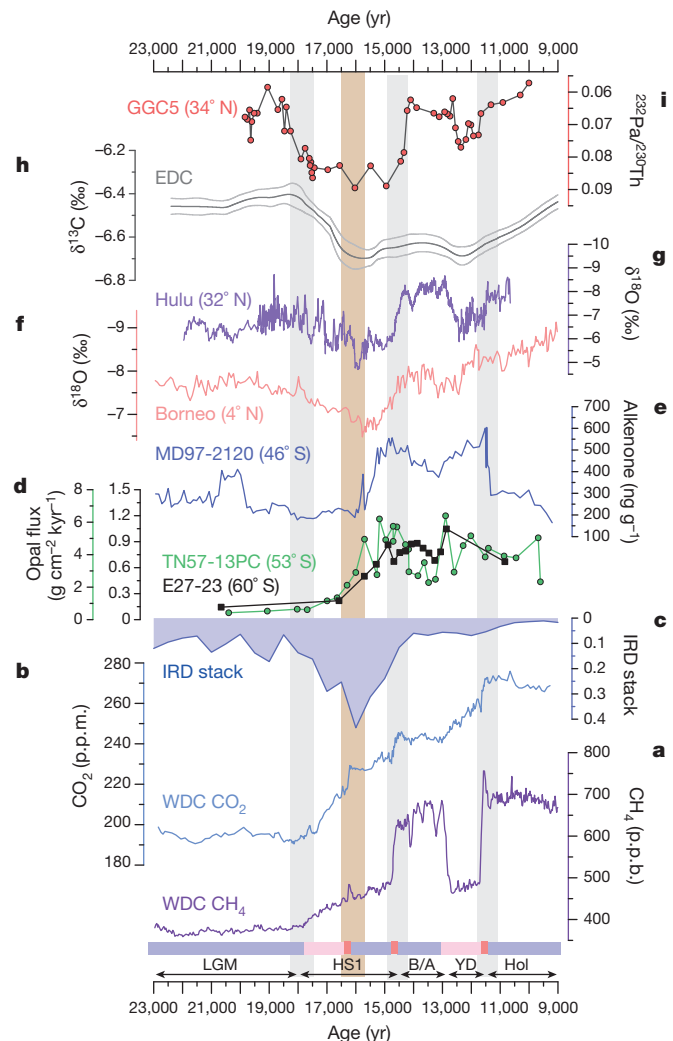


Figure 2 | WAIS Divide CO_2 and CH_4 data plotted against multiple environmental proxies. **a**, **b**, WAIS Divide CH_4 (**a**) and CO_2 (**b**). **c**, Iceberg discharge indices from the North Atlantic²³ (15-record composite; normalized units). IRD, ice-raftered debris. **d**, **e**, Biological flux proxies from the Southern Ocean¹⁸ (**d**; TN57-13PC and E27-23) and off New Zealand²⁵ (**e**; MD97-2120). **f**, **g**, Precipitation indices in speleothem records from Borneo²⁶ (**f**) and Hulu²⁷ (**g**). **h**, Carbon isotopes of CO_2 (line with 1σ uncertainty band) from a multi-core composite⁴ placed on an updated chronology⁵. $\delta^{13}\text{C} = (^{13}\text{C}/^{12}\text{C})_{\text{sample}} / (^{13}\text{C}/^{12}\text{C})_{\text{VPDB}} - 1$; VPDB, Vienna Pee Dee Belemnite. **i**, An AMOC proxy¹⁹ (GGC5). All age models that use radiocarbon dates were recalibrated with IntCal13 and updated (Supplementary Data). Coloured bands at bottom of plot indicate times when CO_2 is stable (blue), slowly increasing (pink) or rapidly increasing (red), as described in the text. The grey and brown bands denote the timing of notable climatic transitions (for example LGM to HS1).

methane from mid- and low-latitude sources, and vice versa when the Northern Hemisphere warmed. Several studies have proposed other complementary mechanisms (see, for example, ref. 18) for the CO_2 changes across the last termination; here we guide the discussion with a simple working hypothesis that the AMOC had a critical role, either directly or indirectly, in the rise of atmospheric CO_2 .

The gradual changes in CO_2 that begin at 18.1 and 13.0 kyr ago are accompanied by a reduction in the AMOC strength¹⁹ and warming in Antarctica¹¹ (Figs 1 and 2). Like CO_2 , methane begins to increase gradually at ~18 kyr ago (50 p.p.b. kyr^{-1}), yet at 13.0 kyr ago it is decoupled from CO_2 and instead decreases by 200 p.p.b. in two centuries before stabilizing during the Younger Dryas (Fig. 1). The CO_2 rise and abrupt CH_4 drop at 13 kyr ago are synchronous within the resolution of the data (20–60 yr), and both rise almost synchronously at 18 kyr ago (Extended Data Table 1), suggesting that a common forcing drives the

changes in both gases. Freshwater hosing experiments with a coupled atmosphere–ocean–biogeochemical model²⁰ show that atmospheric CO₂ begins to rise because of a decrease in the ocean carbon inventory due to a collapse in the AMOC, and continues rising until the AMOC resumes. The rate of CO₂ rise in the model²⁰ (25 p.p.m. in 1,700 yr) is similar to rates observed from WDC, and the hypothesis that CO₂ fluctuations are controlled by ocean circulation change is consistent with the timing of the AMOC reduction based on ocean proxies (Fig. 2). Two significant depletion events in stable carbon isotope ratios of CO₂ from a composite of multiple ice cores⁴, by 0.3‰ and 0.1‰, are consistent with the source of the CO₂ rises at 18.1 and 13.0 kyr ago being from upwelling of respired carbon from the deep ocean (Fig. 2h).

The most abrupt changes in CO₂ during the last termination occur at 16.3, 14.8 and 11.7 kyr ago and are synchronous with abrupt increases in CH₄ within the 20–60 yr resolution of the data (Fig. 3). No relationship between the carbonate chemistry of the ice and CO₂ concentration is observed that would suggest *in situ* production of CO₂ as observed in Greenland⁸ (Extended Data Fig. 2), and the sharp increases are resolved by multiple data points, each multiply replicated. Moreover, the WDC CO₂ record is consistent with results from the EDC ice core^{5,9,14} after accounting for the greater smoothing of the EDC gas record by the increased diffusion in the firn and the gradual bubble closure characteristic of low-accumulation sites (Extended Data Fig. 5b). For all three of the abrupt CO₂ increases, the concomitant CH₄ rise is even more sudden, suggesting that the rate of the CO₂ increase in the WDC record is not limited by firn smoothing processes, but reflects the true atmospheric rate of change.

For the abrupt increases at 14.8 and 11.7 kyr ago, both CO₂ and CH₄ increased by ~13 p.p.m. and ~200 p.p.b., respectively, in 100–200 yr, and then remained at the new levels. For these two events, the increases are coincident with proxy evidence for an abrupt recovery of the AMOC and a rapid warming of the North Hemisphere as recorded by Greenlandic ice cores (Fig. 2). As previously noted, modelling suggests that slow increases in CO₂ can occur when the AMOC is reduced²⁰. Conversely, similar modelling supports the hypothesis that a rapid resumption of the AMOC can result in an abrupt increase in atmospheric CO₂ and CH₄ through changes in carbon storage on land²¹ and increased precipitation in wetland regions. Other modelling studies²² suggest that the CO₂ increase at the onset of the Bølling warming could have been much higher

given the large smoothing at EDC. Such an overshoot in CO₂ is not recorded at the Bølling onset at WDC, but smaller overshoots at 11.7 and 16.3 kyr ago do lend partial support to this mechanism.

At 16.3 kyr ago, CO₂ and CH₄ also abruptly increased by 12 p.p.m. and 50 p.p.b. in 100 yr, respectively, although the elevated methane levels were short lived (Fig. 3). However, unlike the changes at 14.8 and 11.7 kyr ago, these increases show no apparent evidence for an abrupt AMOC resumption or warming in the Northern Hemisphere. We propose that the abrupt rise in CO₂ at 16.3 kyr ago is directly related to an iceberg discharge event recorded in several ocean cores from the North Atlantic²³, which occurred while the AMOC was already reduced (Fig. 2c). One potential cause of the CO₂ increase at this time could be a southward shift in the position of the Southern Hemisphere westerlies²⁴, leading to increased ocean upwelling and outgassing of respiration-derived CO₂ (ref. 18). Other archives preserve geochemical signals attributed to the iceberg discharge event at 16.3 kyr ago, for instance an increase in upwelling in the Southern Ocean¹⁸ (Fig. 2d) and off the coast of New Zealand²⁵ (Fig. 2e), drying in northern Borneo²⁶ (Fig. 2f), and an inferred sharp decrease in monsoon intensity in central China²⁷ (Fig. 2g). Cooling and drying of large regions of the globe may also be a consequence of iceberg discharge and freshwater input into the North Atlantic²⁸, which could have initiated a rapid release of carbon from land stocks²⁹.

The small increase in methane at the time of the CO₂ rise at 16.3 kyr ago is somewhat puzzling because a southward shift in the intertropical convergence zone related to an iceberg discharge event would be expected to suppress northern tropical sources of CH₄. However, observations from cave deposits suggest wetter conditions at southern equatorial sites during Heinrich stadials²⁶, which could compensate for reduced northern sources²⁷ and explain the small CH₄ increase.

Finally, the apparent plateaux in CO₂ concentration that occur directly after each of the abrupt CO₂ increases at 16.3, 14.8 and 11.7 kyr ago (Fig. 1) deserve explanation. These plateaux last for 1,000–1,500 yr, and the plateaux following the abrupt changes at 14.8 and 11.7 kyr ago are contemporaneous with reinvigorated AMOC¹⁹, cooling in Antarctica⁵ and a decreased north–south gradient in global temperature². One possible explanation is stratification of the Southern Ocean, perhaps in response to large freshwater discharges from the Antarctic ice sheet³⁰, which would lead to a reduction in outgassing of CO₂ to the atmosphere through decreased exchange of deep water with the surface ocean and

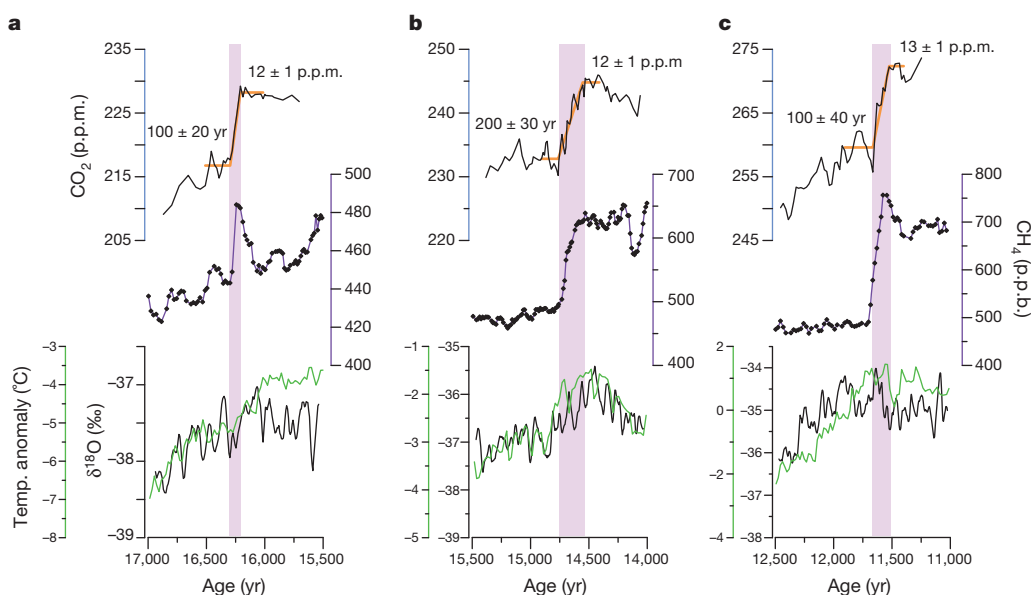


Figure 3 | Detailed view of greenhouse gas and stable isotope measurements from WDC. Oxygen isotope measurements from WDC¹¹ (black line is 11-point weighted average), water-isotope-derived temperature composite from East Antarctica⁵ (green), and atmospheric CH₄ (purple) and CO₂ concentrations

(black line is 5-point weighted average) across the abrupt CO₂ transition at 16.3 (a), 14.8 (b) and 11.7 kyr ago (c). Purple bands represent the durations of the abrupt CO₂ transition and orange lines show their magnitudes (durations and magnitudes also in parentheses; Extended Data Table 1).

atmosphere. Alternatively, the plateaux in CO₂ could be associated with the slow equilibration of atmospheric CO₂ with the ocean following the input of relatively large pulses of carbon during the CO₂ jumps (Extended Data Fig. 5), or uptake of atmospheric CO₂ by the slow regrowth of the terrestrial biosphere.

Our results provide evidence that the observed CO₂ rise during the last deglaciation was divided into two modes: a slow, millennial-scale mode that is closely linked to Antarctic temperature and associated with δ¹³C_{CO₂ decreases (Fig. 2), suggestive of an oceanic origin; and a fast, centennial-scale mode that is closely linked to CH₄ and Northern Hemisphere climate, potentially controlled by AMOC strength and the associated climate teleconnections. Given that a large portion of the direct radiative CO₂ forcing during the glacial–interglacial rise occurred in approximately four centuries, future modelling and isotopic investigations should focus on explaining these abrupt CO₂ variations to further our understanding of the sensitivity of the carbon cycle under natural forcing.}

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 March; accepted 27 August 2014.

- Petit, J. R. *et al.* Climate and atmospheric history of the last 420,000 years from the Vostok ice core, Antarctica. *Nature* **399**, 429–436 (1999).
- Shakun, J. D. *et al.* Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **484**, 49–54 (2012).
- Sigman, D. M. & Boyle, E. A. Glacial/interglacial variations in atmospheric carbon dioxide. *Nature* **407**, 859–869 (2000).
- Schmitt, J. *et al.* Carbon isotope constraints on the deglacial CO₂ rise from ice cores. *Science* **336**, 711–714 (2012).
- Parrenin, F. *et al.* Synchronous change of atmospheric CO₂ and Antarctic temperature during the last deglacial warming. *Science* **339**, 1060–1063 (2013).
- North Greenland Ice Core Project members. High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **431**, 147–151 (2004).
- Groote, P. M., Stuiver, M., White, J. W. C., Johnsen, S. & Jouzel, J. Comparison of oxygen isotope records from the GISP2 and GRIP Greenland ice cores. *Nature* **366**, 552–554 (1993).
- Smith, H. J., Wahlen, M., Mastroianni, D., Taylor, K. & Mayewski, P. The CO₂ concentration of air trapped in Greenland Ice Sheet Project 2 ice formed during periods of rapid climate change. *J. Geophys. Res.* **102**, 26577–26582 (1997).
- Monnin, E. *et al.* Atmospheric CO₂ concentrations over the last glacial termination. *Science* **291**, 112–114 (2001).
- Ahn, J. *et al.* A record of atmospheric CO₂ during the last 40,000 years from the Siple Dome, Antarctica ice core. *J. Geophys. Res.* **109**, D13305 (2004).
- WAIS Divide Project Members. Onset of deglacial warming in West Antarctica driven by local orbital forcing. *Nature* **500**, 440–444 (2013).
- Brook, E. J., Harder, S., Severinghaus, J., Steig, E. J. & Sucher, C. M. On the origin and timing of rapid changes in atmospheric methane during the last glacial period. *Glob. Biogeochem. Cycles* **14**, 559–572 (2000).
- Khalil, M. A. K. & Rasmussen, R. A. Sources, sinks, and seasonal cycles of atmospheric methane. *J. Geophys. Res.* **88**, 5131–5144 (1983).
- Lourantou, A., Chappellaz, J., Barnola, J. M., Masson-Delmotte, V. & Raynaud, D. Changes in atmospheric CO₂ and its carbon isotopic ratio during the penultimate deglaciation. *Quat. Sci. Rev.* **29**, 1983–1992 (2010).
- Broecker, W. S. Glacial to interglacial changes in ocean chemistry. *Prog. Oceanogr.* **11**, 151–197 (1982).
- Sigman, D. M., de Boer, A. M. & Haug, G. H. in *Ocean Circulation: Mechanisms and Impacts* (eds Schmittner, A., Chiang, J. C. H. & Hemming, S. R.) 335–349 (American Geophysical Union, 2007).
- Chiang, J. C. H. & Bitz, C. M. Influence of high latitude ice cover on the marine intertropical convergence zone. *Clim. Dyn.* **25**, 477–496 (2005).
- Anderson, R. F. *et al.* Wind-driven upwelling in the Southern Ocean and the deglacial rise in atmospheric CO₂. *Science* **323**, 1443–1448 (2009).
- McManus, J. F., Francois, R., Gherardi, J.-M., Keigwin, L. D. & Brown-Leger, S. Collapse and rapid resumption of Atlantic meridional circulation linked to deglacial climate changes. *Nature* **428**, 834–837 (2004).
- Schmittner, A. & Galbraith, E. D. Glacial greenhouse-gas fluctuations controlled by ocean circulation changes. *Nature* **456**, 373–376 (2008).
- Köhler, P., Joos, F., Gerber, S. & Knutti, R. Simulated changes in vegetation distribution, land carbon storage, and atmospheric CO₂ in response to a collapse of the North Atlantic thermohaline circulation. *Clim. Dyn.* **25**, 689–708 (2005).
- Köhler, P., Knorr, G., Buiron, D., Lourantou, A. & Chappellaz, J. Abrupt rise in atmospheric CO₂ at the onset of the Bølling/Allerød: in-situ ice core data versus true atmospheric signals. *Clim. Past* **7**, 473–486 (2011).
- Stern, J. V. & Lisiecki, L. E. North Atlantic circulation and reservoir age changes over the past 41,000 years: North Atlantic reservoir age history. *Geophys. Res. Lett.* **40**, 3693–3697 (2013).
- Toggweiler, J. R., Russell, J. L. & Carson, S. R. Midlatitude westerlies, atmospheric CO₂, and climate change during the ice ages. *Paleoceanography* **21**, PA2005 (2006).
- Sachs, J. P. & Anderson, R. F. Increased productivity in the subantarctic ocean during Heinrich events. *Nature* **434**, 1118–1121 (2005).
- Partin, J. W., Cobb, K. M., Adkins, J. F., Clark, B. & Fernandez, D. P. Millennial-scale trends in west Pacific warm pool hydrology since the Last Glacial Maximum. *Nature* **449**, 452–455 (2007).
- Wu, J., Wang, Y., Cheng, H. & Edwards, L. R. An exceptionally strengthened East Asian summer monsoon event between 19.9 and 17.1 ka BP recorded in a Hulu stalagmite. *Sci. China Ser. Earth Sci.* **52**, 360–368 (2009).
- Stager, J. C., Ryves, D. B., Chase, B. M. & Pausata, F. S. R. Catastrophic drought in the Afro-Asian monsoon region during Heinrich event 1. *Science* **331**, 1299–1302 (2011).
- Scholze, M., Knorr, W. & Heimann, M. Modelling terrestrial vegetation dynamics and carbon cycling for an abrupt climatic change event. *Holocene* **13**, 327–333 (2003).
- Weber, M. E. *et al.* Millennial-scale variability in Antarctic ice-sheet discharge during the last deglaciation. *Nature* **510**, 134–138 (2014).
- Schilt, A. *et al.* Atmospheric nitrous oxide during the last 140,000 years. *Earth Planet. Sci. Lett.* **300**, 33–43 (2010).
- Ramaswamy, V. *et al.* *Climate Change 2001: The Scientific Basis* (eds Houghton, J. T. *et al.*) 349–416 (Cambridge Univ. Press, 2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work is supported by the US National Science Foundation (NSF) (grants 0739766-ANT, 1043518-ANT, 1043092-ANT, 0839093-ANT and 1142166-ANT). We appreciate the support of the WAIS Divide Science Coordination Office at the Desert Research Institute (DRI) of Reno, Nevada, and the University of New Hampshire for the collection and distribution of the WAIS Divide ice core and related tasks (NSF grants 0230396, 0440817, 0944348 and 0944266). Additional support for this research came from the NSF Office of Polar Programs through their support of the Ice Drilling Program Office and the Ice Drilling Design and Operations group; the US National Ice Core Laboratory, for curation of the core; Raytheon Polar Services, for logistics support in Antarctica; the 109th New York Air National Guard, for airlift to Antarctica; and the Korea Meteorological Administration Research and Development Program (CATER 2012-7030). We thank T. Alig, J. Edwards and J. Lee for assisting with CO₂ and CH₄ measurements; the DRI ultracore ice-core lab, including D. Pasteris, M. Sigl and O. Maselli for their contribution to the aerosol records; and I. Fung for discussions and providing software for carbon uptake calculations.

Author Contributions S.A.M. and E.J.B. oversaw and contributed to all aspects of the research, and with T.K.B. designed the project and led the writing of the paper. J.A., M.L.K., J.P.S. and T.S. assisted with and contributed WDC gas measurements. E.J.S. contributed the WDC water isotope data. J.R.M. contributed calcium and hydrogen peroxide concentration measurements. C.B. developed the gas chronology. J.L.R. performed the firm modelling experiments and interpretation. K.C.T. led the field effort that collected the samples. K.M.C., T.J.F., J.R.M., E.J.S., K.C.T. and J.W.C.W. developed the ice chronology and interpretation. All authors discussed the results and contributed input to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to S.A.M. (smarcott@wisc.edu).

METHODS

Measurements of CO₂ were made at Oregon State University, and CH₄ measurements were made at Oregon State and Penn State University. CO₂ measurements were made with a mechanical crushing system using methods described by Ahn *et al.*³³. Approximately 1,030 measurements were made on ~320 separate depths spanning the time frame of 23,000–9,000 years BP with a median sampling resolution of 25 years. All samples were measured at least in duplicate and one to three times more for some depths. Nitrogen isotope data were used to correct the CO₂ data by 1.0–1.4 p.p.m. for gravitational fractionation. The mean standard error of all measurements was 1.0 p.p.m. A consistent 4 p.p.m. offset between WDC and EPICA Dome C CO₂ concentrations is observed. A similar offset is observed between WAIS Divide and other CO₂ reconstructions³⁴. Methane concentrations were measured for approximately 1,100 ice samples from ~580 depths covering 23,000–11,500 years BP (20-year median resolution) using methods described by Mitchell *et al.*³⁵ at Oregon State University and combined with data measured at Penn State University¹¹. Duplicate methane measurements were made on all but fifty samples, and the mean standard error of all measurements made at Oregon State University was 0.8 p.p.b. CO₂ and CH₄ measurements were calibrated using dry standard air with known mole fractions calibrated at the NOAA Earth System Research Laboratory (WMOX2007 scale for CO₂; NOAA04 scale for CH₄).

WAIS Divide timescale. The gas chronology used in this study (WDC06A-7) is based on the layer-counted WDC06A-7 chronology¹¹. Gas ages were obtained by subtracting a modelled ice-age/gas-age difference (Δ age) from the ice-age timescale³⁶. Δ age was calculated using a dynamical firn-densification model with heat diffusion. The modelling is constrained by measurements of $\delta^{15}\text{N}$ of N₂, a proxy for past firn column thickness³⁷. $\delta^{15}\text{N}$ was measured on the WAIS Divide ice core (WDC) at the Scripps Institution of Oceanography with a 300-year average resolution (Extended Data Fig. 1b) following procedures outlined elsewhere^{38,39}. The air was extracted from ~13 g ice samples using a melt–refreeze technique, and collected on a cold finger at liquid-He temperatures. The samples were analysed on a Thermo Finnigan Delta V dual-inlet mass spectrometer to determine $^{15}\text{N}/^{14}\text{N}$ of N₂ ($\delta^{15}\text{N}$), $^{18}\text{O}/^{16}\text{O}$ of O₂ ($\delta^{18}\text{O}$), $^{32}\text{O}_2/^{28}\text{N}_2$ ($\delta\text{O}_2/\text{N}_2$) and $^{40}\text{Ar}/^{28}\text{N}_2$ ($\delta\text{Ar}/\text{N}_2$). Routine analytical corrections were made to the data for pressure imbalance and chemical slope, and results were normalized to the La Jolla present-day atmosphere.

We use a dynamical adaptation of the Herron and Langway firn densification model⁴⁰ coupled to a heat diffusion–advection model⁴¹. We use the present-day convective zone thickness of 3.5 m and a parameterization for surface snow density^{42,43}. The forward model requires past temperatures and accumulation rates as inputs, and generates Δ age and $\delta^{15}\text{N}$ as outputs. Surface temperatures (Extended Data Fig. 1a) are obtained from the $\delta^{18}\text{O}$ isotopic composition of the ice¹¹, with the $\delta^{18}\text{O}$ –temperature relationship calibrated^{11,44,45} to fit observed borehole temperatures. Accumulation rates were obtained in two ways. First, accumulation can be reconstructed from the observed annual layer thickness, corrected for thinning due to ice flow (Extended Data Fig. 1c, black curve). However, the thinning function becomes increasingly uncertain with depth (the CO₂ data presented here were measured over the depth range of 1,700–2,600 m within the 3,405 m-long core). Second, accumulation can be reconstructed from the $\delta^{15}\text{N}$ data using a firn-densification inverse model⁴¹. Using the temperature history, the model determines the accumulation history that optimizes the fit to the $\delta^{15}\text{N}$ data by making small adjustments to the thinning function (Extended Data Fig. 1c, red curve). The model fit to the $\delta^{15}\text{N}$ data is shown in Extended Data Fig. 1b. We find a good agreement between both accumulation rate reconstructions.

The modelled Δ age is shown in Extended Data Fig. 1d (orange curve), where we have used the accumulation rates from the inverse $\delta^{15}\text{N}$ method. A sensitivity study in which we varied past convective zone thicknesses, temperature history, dust sensitivity⁴⁶ and firn-densification physics⁴⁷ shows a 1σ uncertainty in Δ age of ~15%. We also calculated Δ age using the depth-difference method developed by Parrenin *et al.*⁵ (Extended Data Fig. 1d, black curve). This method also suffers from uncertainty in the layer thinning function. Both methods agree within the 1σ uncertainty band. We note that at around 12 kyr BP the WDC site experienced a very sudden and short-lived increase in accumulation rates as inferred from the annual layer thicknesses (Extended Data Fig. 1c). This event is also recorded in the $\delta^{15}\text{N}$ as a thickening of the firn column (Extended Data Fig. 1b).

Potential for *in situ* production of CO₂. Prior work in Greenland⁸ has demonstrated that CO₂ records can be compromised through *in situ* production when abundant concentrations of calcium carbonate are present. Calcium is delivered to the ice in the form of carbonate dust that blows onto the ice-sheet surface and is buried by subsequent accumulation. Although problems with *in situ* production have never been detected in CO₂ records from Antarctic ice cores, probably because Antarctic sites experience an order of magnitude less calcium carbonate deposition than Greenland⁴⁸, each ice core from Antarctica is both geographically unique and affected by distinct atmospheric processes that can deliver various chemical species

to the sites. To examine whether abrupt changes in CO₂ in WDC are related to *in situ* production, we compared the non-seasalt component of calcium, which we take as a proxy for terrestrial dust, with the carbon dioxide concentration at times of abrupt change in CO₂ (Extended Data Fig. 2). We find no relationships and note that the non-seasalt Ca also is an order of magnitude lower at WDC than in Greenland ice cores, similar to other ice cores from Antarctica^{8,48}. In addition to calcium carbonate, the oxidation of organic compounds can also affect the CO₂ concentration from ice cores⁴⁹. At WDC the concentration of the oxidant H₂O₂ in the upper section of the core (WDC-05A) averaged less than 30 p.p.b.^{50,51}, which is similar to other values from Antarctic ice cores⁵², and does not covary with CO₂ concentrations over the interval from 1,940 to 2,000 m (average 20 p.p.b.) in the deeper section of WDC from where we have data.

Smoothing of the CO₂ record by firn processes. Although it may appear that the WDC and EDC CO₂ records exhibit fundamentally different patterns over short timescales (Extended Data Fig. 3), these differences are expected as a result of the dramatically different conditions at the ice-core sites, which dictate how much smoothing atmospheric gas records experience before being locked into the ice-core record. Temperature and accumulation rates at WDC vary between 233 and 243 K and, respectively, between 0.10 and 0.30 m ice equivalent per year over the deglacial transition. At EDC, in contrast, conditions are colder and dryer, averaging 209 K and 0.015 m yr⁻¹. Because of this, firn models predict that Δ age and the age distribution (a measure of smoothing) should be an order of magnitude smaller at WDC than at EDC (Extended Data Fig. 4).

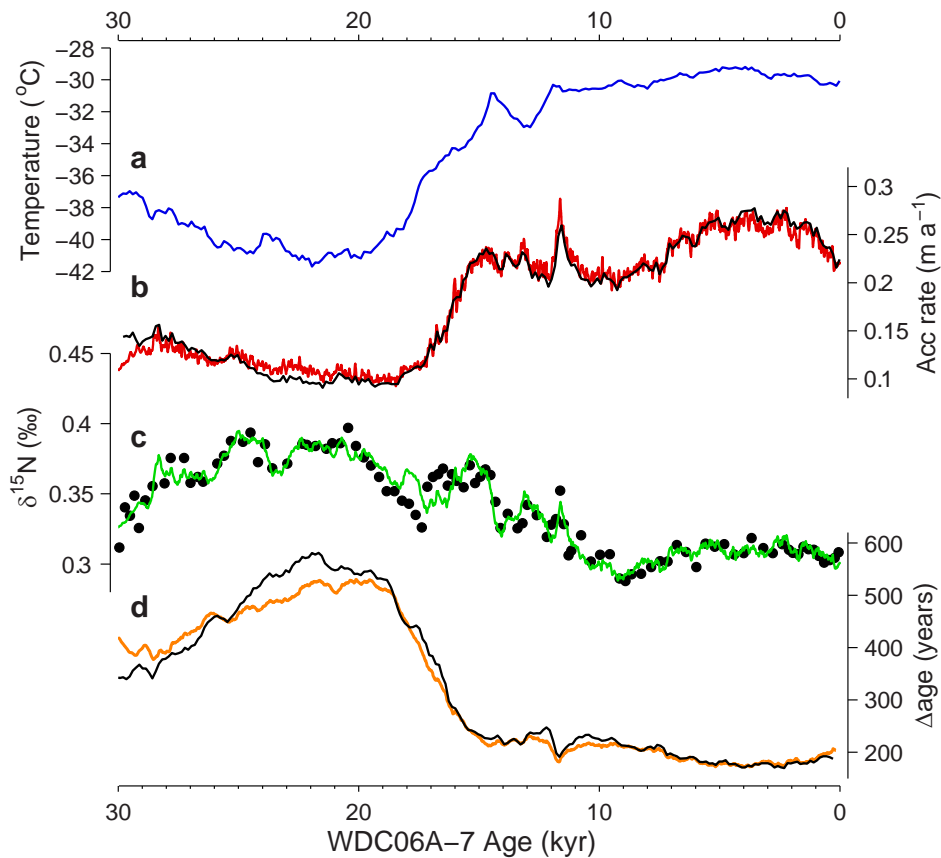
To demonstrate the consistency of the two CO₂ records, we generate a smoothing function for EDC using the OSU firn air model⁵³ (Extended Data Fig. 5a). The function describes the age distribution of the air in the closed porosity of the ice core, and accounts for gradual bubble closure as ice is advected through the lock-in zone. We then apply this filter to the WDC gas record, using it as a proxy for the true atmospheric history. Because WDC gases experience very little smoothing during bubble trapping, we argue that this is a reasonable assumption (smoothing WDC with a WDC filter does not change the record). The resulting curve agrees well with the EDC CO₂ record^{9,14,54} (Extended Data Fig. 5b), supporting the idea that differences between the two records arise because of differences in site conditions and bubble-trapping processes, and reflect no disagreement about the nature of the deglacial rise in atmospheric CO₂.

A consistent 4 p.p.m. offset between WDC and EPICA Dome C CO₂ concentrations is observed in our new deglacial reconstruction after consideration for firn smoothing processes (Extended Data Figs 3 and 5). This offset represents 1–2% of the total concentration measurements and is outside the stated uncertainties for both records. A similar offset of 2–4 p.p.m. has also been observed between WDC and CO₂ reconstructions from Law Dome and EPICA Dronning Maud Land³⁴, along with a 2 p.p.m. offset between Dronning Maud Land and the South Pole⁵². On the basis of interlaboratory comparisons between Law Dome and WDC³⁴, the offset between the ice cores is unlikely to result from laboratory methods and represents a real difference between the two ice cores, though the exact nature of the offset still remains to be determined.

Modelling CO₂ pulses and plateaux. Using the WDC CO₂ record as a proxy for the true atmospheric history, we employ a box model of the ocean carbon cycle⁵⁵ to explore how a pulse of CO₂ emissions is attenuated and potentially creates a plateau in the concentration record. The model is divided into three boxes: a well-mixed atmosphere, a well-mixed surface ocean mixed layer and a diffusive thermocline/deep ocean. The model does not account for sedimentation of CaCO₃ in the ocean. To simulate an abrupt increase in CO₂, we apply a pulse input of 0.5 Pg C yr⁻¹ that lasts for 100 years. Following the 100-year pulse, the model simulates a slow ocean uptake over the next 1,500 years (Extended Data Fig. 6). Although this model is not designed to simulate the entire CO₂ variability in the WDC record, this simple simulation suggests (but does not require) that the three abrupt CO₂ rises at 16.3, 14.8 and 11.7 kyr ago could represent brief injections of CO₂ into the atmosphere that are superimposed on a longer-term increase, with the plateau that follows each of the abrupt rises explained by the drawdown of the additional carbon in the atmosphere into the ocean.

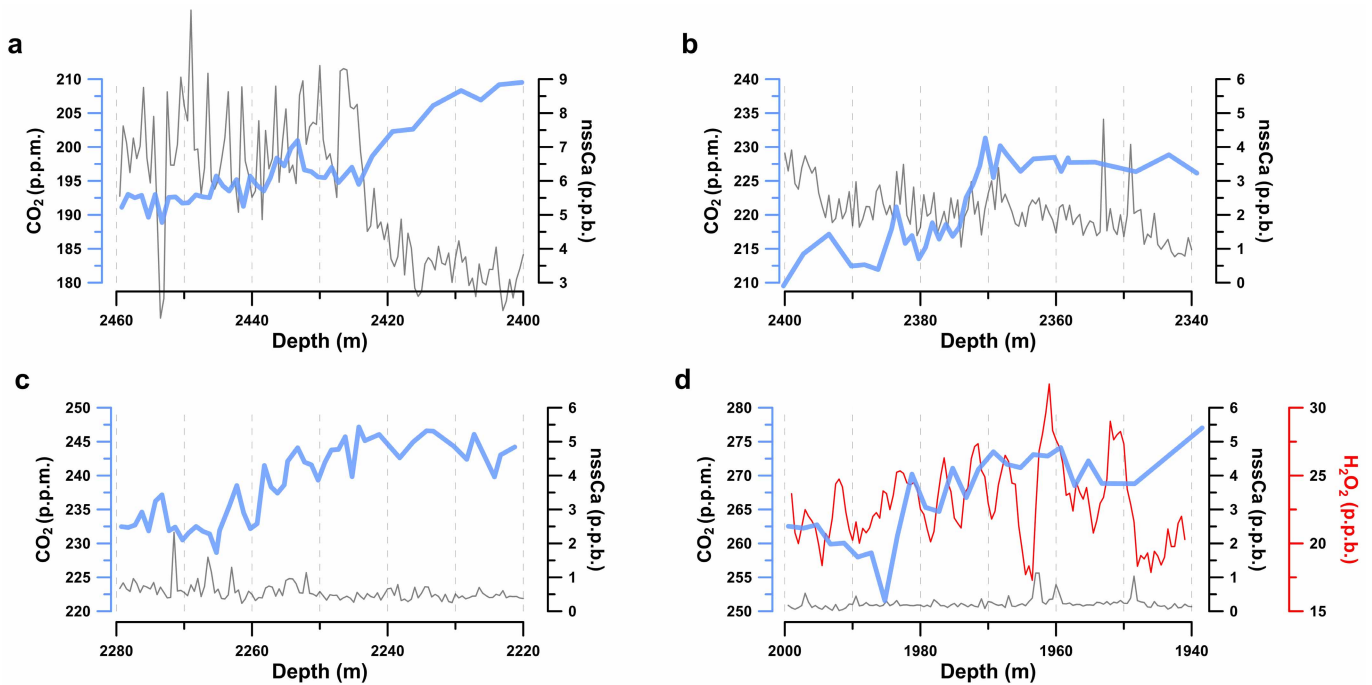
Köhler *et al.*²¹ performed a similar experiment using a more comprehensive carbon cycle box model to describe an abrupt CO₂ change at the onset of the Bølling. Using similar modelled pulses of CO₂, they also reproduced the abrupt CO₂ features. They concluded that the magnitude of the CO₂ change could have been as high as 30 p.p.m. for a 2.5 Pg yr⁻¹ injection of carbon into the atmosphere after considering smoothing of the atmospheric signal in the firn at EDC. Unlike the low-accumulation site at EDC, however, the high-accumulation WDC site provides a CO₂ record that has experienced very little firn smoothing (see Methods section on the WAIS Divide timescale), and provides a much closer approximation of the true atmospheric history. Therefore, we conclude that if there were a pulse input of CO₂ its magnitude and duration would be of the order of 0.5 Pg C yr⁻¹ and would last for 100 years.

- Breakpoint determination.** To provide an objective means for determining the timing and duration of the CO₂ and CH₄ transitions across the last deglaciation, we employed the functions Breakfit⁵⁶ and Rampfit⁵⁷. For the three abrupt jumps in CO₂, both the timing and duration of the changes were quantified (Extended Data Table 1) using the Rampfit⁵⁷ function with a 500-year search window. For each, the same seed generator number (200) was used, and 200 bootstrap (wild) iterations were performed to determine the uncertainties. For the CO₂ and CH₄ transitions at the start of the deglaciation and beginning of the Younger Dryas, the Breakfit⁵⁶ function was used to determine the timing of the transitions. The window length for each was 1,000 years and the same seed number (200) and bootstrap number (200) were used to quantify the timing uncertainty.
33. Ahn, J., Brook, E. J. & Howell, K. A high-precision method for measurement of paleoatmospheric CO₂ in small polar ice samples. *J. Glaciol.* **55**, 499–506 (2009).
 34. Ahn, J. *et al.* Atmospheric CO₂ over the last 1000 years: a high-resolution record from the West Antarctic Ice Sheet (WAIS) Divide ice core. *Glob. Biogeochem. Cycles* **26**, GB2027 (2012).
 35. Mitchell, L. E., Brook, E. J., Sowers, T., McConnell, J. R. & Taylor, K. Mulidecadal variability of atmospheric methane, 1000–1800 C.E. *J. Geophys. Res.* **116**, G02007 (2011).
 36. Schwander, J. & Stauffer, B. Age difference between polar ice and the air trapped in its bubbles. *Nature* **311**, 45–47 (1984).
 37. Sowers, T., Bender, M., Raynaud, D. & Korotkevich, Y. S. $\delta^{15}\text{N}$ of N₂ in air trapped in polar ice: a tracer of gas transport in the firn and a possible constraint on ice age-gas age difference. *J. Geophys. Res.* **97**, 15683–15697 (1992).
 38. Sowers, T., Bender, M. & Raynaud, D. Elemental and isotopic composition of occluded O₂ and N₂ in polar ice. *J. Geophys. Res.* **94**, 5137–5150 (1989).
 39. Petrenko, V. V., Severinghaus, J. P., Brook, E. J., Reeh, N. & Schaefer, H. Gas records from the West Greenland ice margin covering the Last Glacial Termination: a horizontal ice core. *Quat. Sci. Rev.* **25**, 865–875 (2006).
 40. Herron, M. M. & Langway, C. C. Firn densification: an empirical model. *J. Glaciol.* **93**, 373–383 (1980).
 41. Rasmussen, S. O. *et al.* A first chronology for the North Greenland Eemian Ice Drilling (NEEM) ice core. *Clim. Past* **9**, 2713–2730 (2013).
 42. Kaspers, K. A. *et al.* Model calculations of the age of firn air across the Antarctic continent. *Atmos. Chem. Phys.* **4**, 1365–1380 (2004).
 43. Battle, M. O. *et al.* Controls on the movement and composition of firn air at the West Antarctic Ice Sheet Divide. *Atmos. Chem. Phys.* **11**, 18633–18675 (2011).
 44. Cuffey, K. M. & Glow, G. D. Temperature, accumulation, and ice sheet elevation in central Greenland through the last deglacial transition. *J. Geophys. Res.* **102**, 26383–26396 (1997).
 45. Steig, E. J. *et al.* Recent climate and ice-sheet change in West Antarctica compared to the past 2000 years. *Nature Geosci.* **6**, 372–375 (2013).
 46. Hörhold, M. W. *et al.* On the impact of impurities on the densification of polar firn. *Earth Planet. Sci. Lett.* **325–326**, 93–99 (2012).
 47. Barnola, J. M., Pimienta, P., Raynaud, D. & Korotkevich, Y. S. CO₂–climate relationship as deduced from the Vostok ice core: a reexamination based on new measurements and on a reevaluation of the air dating. *Tellus B* **43**, 83–90 (1991).
 48. Legrand, M. R. & Delmas, R. J. Soluble impurities in four Antarctic ice cores over the last 30,000 years. *Ann. Glaciol.* **10**, 116–120 (1988).
 49. Tschumi, J. & Stauffer, B. Reconstructing past atmospheric CO₂ concentration based on ice-core analyses: open questions due to in situ production of CO₂ in the ice. *J. Glaciol.* **46**, 45–53 (2000).
 50. Sofen, E. D. *et al.* WAIS Divide ice core suggests sustained changes in the atmospheric formation pathways of sulfate and nitrate since the 19th century in the extratropical Southern Hemisphere. *Atmos. Chem. Phys. Discuss.* **13**, 23089–23138 (2013).
 51. Lamarque, J.-F., McConnell, J. R., Shindell, D. T., Orlando, J. J. & Tyndall, G. S. Understanding the drivers for the 20th century change of hydrogen peroxide in Antarctic ice-cores. *Geophys. Res. Lett.* **38**, L04810 (2011).
 52. Siegenthaler, U. *et al.* Supporting evidence from the EPICA Dronning Maud Land ice core for atmospheric CO₂ change during the past millenium. *Tellus B* **57**, 51–57 (2005).
 53. Buizert, C. *et al.* Gas transport in firn: multiple-tracer characterisation and model intercomparison for NEEM, Northern Greenland. *Atmos. Chem. Phys. Discuss.* **11**, 15975–16021 (2011).
 54. Lemieux-Dudon, B. *et al.* Consistent dating of Antarctica and Greenland ice cores. *Quat. Sci. Rev.* **29**, 8–20 (2010).
 55. Oeschger, H., Siegenthaler, U., Schotterer, U. & Gugelmann, A. A box diffusion model to study the carbon dioxide exchange in nature. *Tellus* **27**, 168–192 (1975).
 56. Mudelsee, M. Break function regression: a tool for quantifying trend changes in climate time series. *Eur. Phys. J. Spec. Top.* **174**, 49–63 (2009).
 57. Mudelsee, M. Ramp function regression: a tool for quantifying climate transitions. *Comput. Geosci.* **26**, 293–307 (2000).
 58. Veres, D. *et al.* Antarctic ice core chronology (AICC2012): an optimized multi-parameter and multi-site dating approach for the last 120 thousand years. *Clim. Past* **9**, 1733–1748 (2013).
 59. Bazin, L. *et al.* An optimized multi-proxies, multi-site Antarctic ice and gas orbital chronology (AICC2012): 120–800 ka. *Clim. Past* **9**, 1715–1731 (2013).



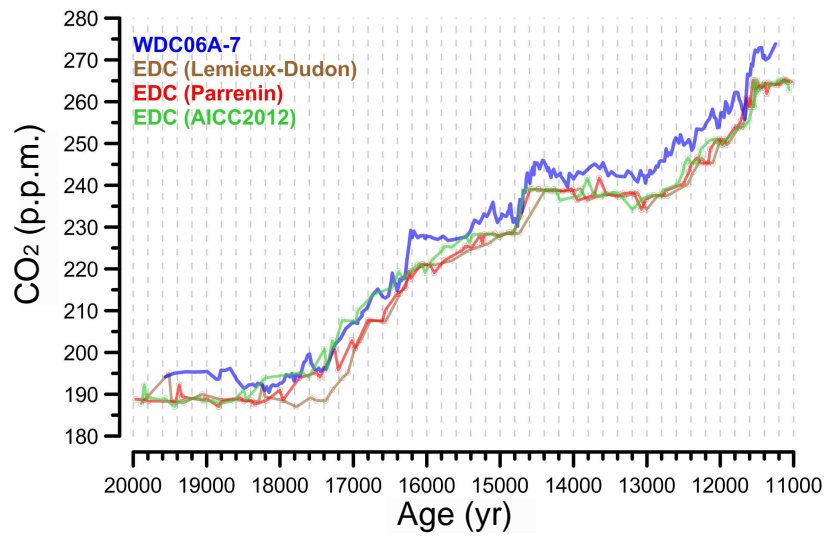
Extended Data Figure 1 | $\delta^{15}\text{N}$ and the ice-age/gas-age difference for the WDC. **a**, Borehole calibrated surface temperature reconstruction derived from $\delta^{18}\text{O}$ measurements from the ice¹. **b**, Accumulation rates reconstructed with the firn-densification inverse model (red curve) and from layer thickness

observations (black curve). **c**, $\delta^{15}\text{N}-\text{N}_2$ data for the upper 2,800 m (black dots) with model fit (green curve). **d**, Modelled age using firn-densification model (orange curve) and Δage estimate using the depth-difference technique from Parrenin *et al.*⁵ (black curve).



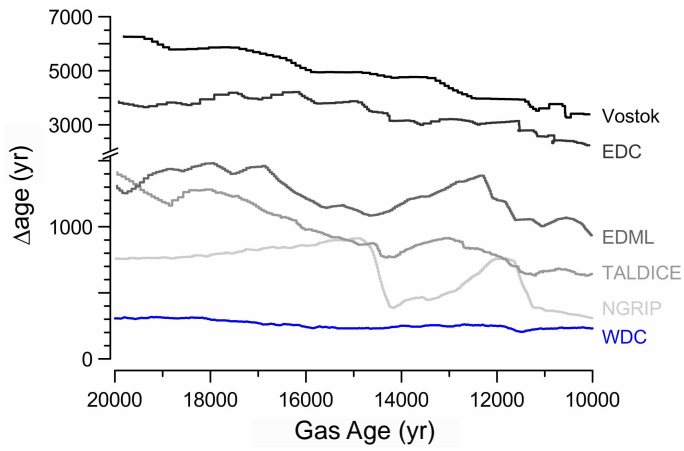
Extended Data Figure 2 | CO₂ concentrations and elemental data for WDC. WDC CO₂ concentrations (blue) plotted against non-seasalt calcium (nssCa) concentrations (black) and hydrogen peroxide (H₂O₂, red) at multiple depths

in the core where we observe abrupt changes in carbon dioxide. Hydrogen peroxide concentrations have been smoothed (2 m centred average) from original data to improve clarity.

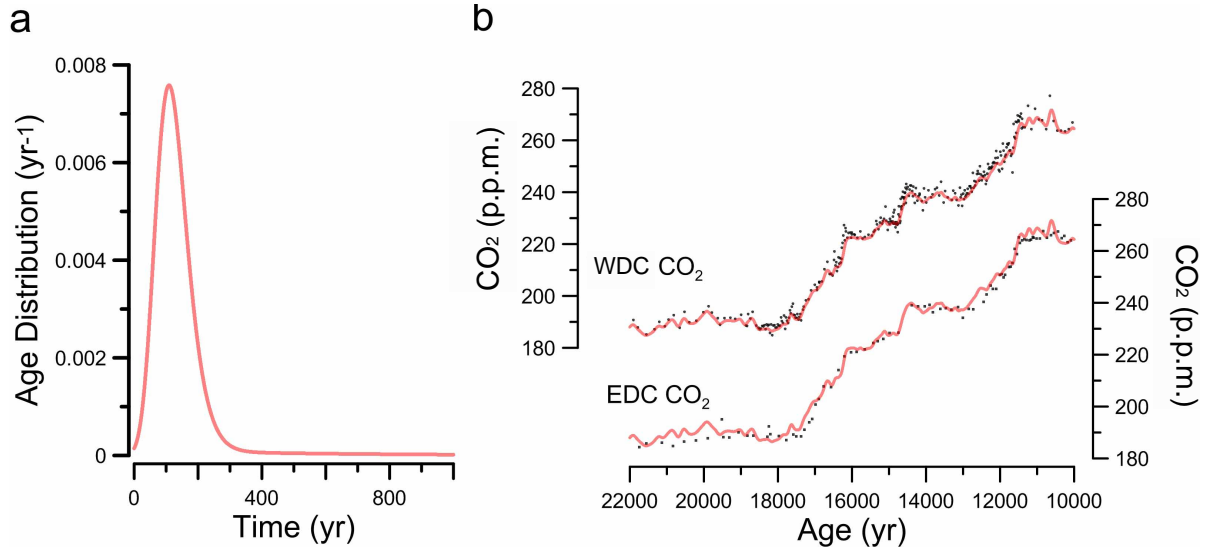


Extended Data Figure 3 | CO₂ concentrations for WDC and EDC. WDC CO₂ concentrations on layer-counted (blue; 5-point weighted average) timescale and EPICA Dome C (EDC) CO₂ concentrations on the

Lemieux-Dudon *et al.*^{9,14,54} (brown), Parrenin *et al.*⁵ (red) and Antarctic ice-core chronology^{58,59} (AICC2012; green) timescales.

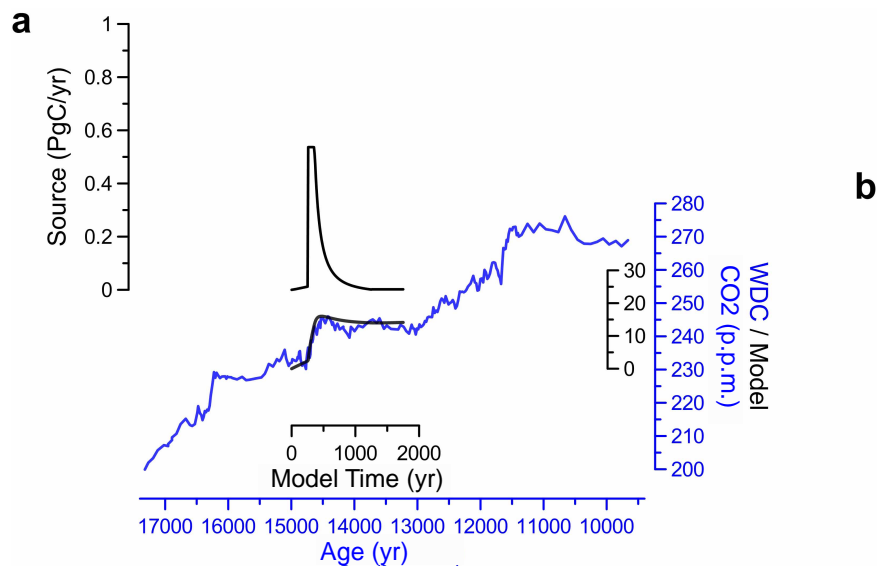


Extended Data Figure 4 | Calculated Δ age offsets across the last deglacial termination for five ice cores from Antarctica and Greenland, compared with WDC. EDML, EPICA Dronning Maud Land; TALDICE, Talos Dome Ice; NGRIP, North Greenland Ice Project. Ice-core data from refs 58, 59.



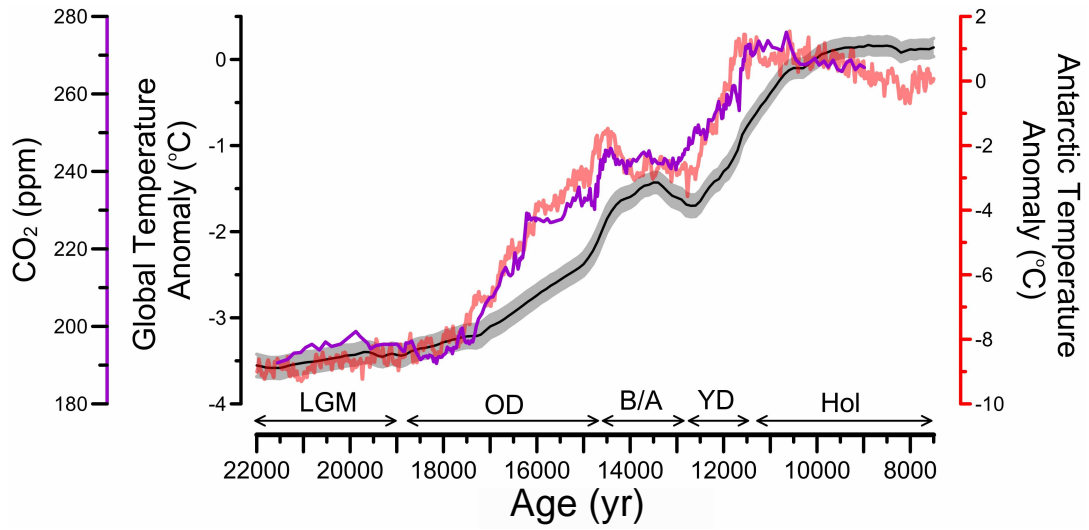
Extended Data Figure 5 | Firm smoothing functions applied to CO₂ data from WDC and EDC. **a**, The red line is the Green's function (smoothing function) produced by a firm model using an assumed EDC accumulation rate of 0.015 m yr^{-1} and a temperature of 209 K. **b**, CO₂ data from WDC (dots) and

EDC (dots) plotted against artificially smoothed CO₂ data from WDC using the EDC firm smoothing function (red line in both plots). WDC data have been systematically lowered by 4 p.p.m. for direct comparison with EDC.



Extended Data Figure 6 | Simple box model source history and atmospheric CO₂ response compared to measured data from WDC. a, Applied source history used in the modelling experiment. b, Atmospheric CO₂ record from

WDC (5-point weighted average; blue) and the model derived atmospheric history (black). Box model from ref. 55.



Extended Data Figure 7 | CO₂ concentrations and temperature reconstructions for the last deglaciation. WDC CO₂ concentrations (purple; 5-point weighted average), a global temperature reconstruction² (black; grey

band is 1 σ uncertainty envelope), and an Antarctic temperature stack based on stable isotopes from East Antarctic ice cores⁵ (red).

Extended Data Table 1 | Timing of five abrupt transitions in CO₂ and CH₄ during the last termination

Rampfit (CO ₂)	Search Window (yrs BP)	Breakpoint 1 (yrs BP)	Breakpoint 2 (yrs BP)	Duration (yrs BP)
"16.3ka" rise	16,500 - 16,000	16,200±20	16,300±20	100±30
"14.8ka" rise	14,900 - 14,400	14,560±30	14,760±30	200±40
"11.7ka" rise	11,900 - 11,400	11,550±40	11,650±30	100±50
Breakfit (CO ₂ & CH ₄)	Search Window (yrs BP)	Breakpoint CO ₂ (yrs BP)	Breakpoint CH ₄ (yrs BP)	
"18.1ka"	18,500 - 17,500	18,140±80	17,820±60	
"13.0ka"	13,500 - 12,500	13,020±40	13,010±170	

Uncertainties are 1 σ .

Calcisponges have a ParaHox gene and dynamic expression of dispersed NK homeobox genes

Sofia A. V. Fortunato^{1,2}, Marcin Adamski¹, Olivia Mendivil Ramos^{3†}, Sven Leininger^{1†}, Jing Liu¹, David E. K. Ferrier³ & Maja Adamska¹

Sponges are simple animals with few cell types, but their genomes paradoxically contain a wide variety of developmental transcription factors^{1–4}, including homeobox genes belonging to the Antennapedia (ANTP) class^{5,6}, which in bilaterians encompass Hox, ParaHox and NK genes. In the genome of the demosponge *Amphimedon queenslandica*, no Hox or ParaHox genes are present, but NK genes are linked in a tight cluster similar to the NK clusters of bilaterians⁵. It has been proposed that Hox and ParaHox genes originated from NK cluster genes after divergence of sponges from the lineage leading to cnidarians and bilaterians^{5,7}. On the other hand, synteny analysis lends support to the notion that the absence of Hox and ParaHox genes in *Amphimedon* is a result of secondary loss (the ghost locus hypothesis)⁸. Here we analysed complete suites of ANTP-class homeoboxes in two calcareous sponges, *Sycon ciliatum* and *Leucosolenia complicata*. Our phylogenetic analyses demonstrate that these calcisponges possess orthologues of bilaterian NK genes (*Hex*, *Hmx* and *Msx*), a varying number of additional NK genes and one ParaHox gene, *Cdx*. Despite the generation of scaffolds spanning multiple genes, we find no evidence of clustering of *Sycon* NK genes. All *Sycon* ANTP-class genes are developmentally expressed, with patterns suggesting their involvement in cell type specification in embryos and adults, metamorphosis and body plan patterning. These results demonstrate that ParaHox genes predate the origin of sponges, thus confirming the ghost locus hypothesis⁸, and highlight the need to analyse the genomes of multiple sponge lineages to obtain a complete picture of the ancestral composition of the first animal genome.

Sponges (Porifera) are strong candidates for being the earliest extant lineage(s) of animals⁹. The genome sequence of the demosponge *A. queenslandica* has provided rich material for comparative studies looking at the origins of metazoan developmental genes, cell types and body plans¹. Among others, it has fuelled hypotheses about the origin of one of the most widely studied groups of developmental genes: the ANTP-class homeoboxes, including Hox, ParaHox and NK genes^{5,7,8,10,11}. ANTP-class genes have been found in all animals and are involved in multiple developmental processes, including body plan patterning and neurogenesis¹². Hox, ParaHox and NK genes are often found in clusters^{5,13}, and in some animals their expression is temporally or spatially correlated to their position within the cluster (this being known as colinearity)¹². The *Amphimedon* genome contains eight NK genes, but neither Hox nor ParaHox genes are present⁵. Six NK genes are linked in a tight cluster, and their simple embryonic and larval expression patterns are not consistent with colinearity^{5,6}. The lack of Hox and ParaHox genes in *Amphimedon*, and also in the ctenophore *Mnemiopsis leidyi*⁷, has previously been interpreted as reflecting the ancestral condition, and gave rise to the ParaHoxozoa hypothesis, in which all animal lineages apart from poriferans and ctenophores are collectively known as the ParaHoxozoa. Others¹⁰ have interpreted the phylogenetic evidence differently, suggesting that both Hox and ParaHox genes were originally present in sponges, but have subsequently been lost. This view has been recently revived by the

identification of Hox and ParaHox ‘ghost loci’ (regions that display conserved synteny with bilaterian Hox and ParaHox loci, but lack the Hox/ParaHox genes themselves) in the genome of *Amphimedon*⁸.

We expected that expanding the range of sequenced sponge genomes would provide new information about the evolutionary history of genes important for the origin and evolution of the animal kingdom. Calcisponges form a poriferan lineage that has been separated from demosponges for at least 600 million years⁹. We recently started analysis of the developmental toolkits of two calcisponges, *S. ciliatum* and *L. complicata*^{2,4,14}. Here, we searched for ANTP-class homeobox genes in the transcriptomic and genomic assemblies of these species.

We retrieved ten ANTP-class homeodomains in *Sycon*, constituting nine transcripts (one with two homeoboxes), and twelve ANTP-class homeodomains in *Leucosolenia*, constituting nine transcripts (one with four homeoboxes) (Supplementary Data 1). Our phylogenetic analyses demonstrate that the repertoire of ANTP-class genes is similar between the two calcisponges, but strikingly different than in the demosponge *Amphimedon*. Calcisponges and demosponges have clear orthologues of the bilaterian genes *Hex* and *Msx*; calcisponges also have a clear *NK5* (also known as *Hmx*) orthologue, which seems to be lacking in *Amphimedon*. By contrast, this demosponge has possible *Bsh*, *BarH* and *Tlx* genes, which are not recognizable in calcisponges. While in *Amphimedon* there is a single gene associated with the bilaterian NK2/3/4 clade¹⁵, several paralogues are present in the two calcisponges. They contain multi-homeobox genes with non-orthologous relationships between *Sycon* and *Leucosolenia*, and other genes containing single homeoboxes in the NK2/3/4 clade. Affiliation of *Sycon* and *Leucosolenia* *NKB* and *NKG* and the *Leucosolenia* (*Lco*)*NKF* genes with a particular bilaterian NK family is not clear. No Hox genes were found; however, a pair of the calcisponge genes showed affinities with the ParaHox *Cdx* family, given the concordance of the results of neighbour-joining and maximum-likelihood analyses (Fig. 1 and Extended Data Fig. 1).

Given the importance of this potential assignment, we performed further phylogenetic analyses of these putative *Cdx* orthologues. In addition to the ELEKEF motif, which is shared by many Hox and ParaHox, but not NK-type homeodomains, the *Cdx* family has some distinctive residues in its homeodomain, most notably the YIT motif, which is present only in a small number of other ANTP-class homeodomains (Supplementary Note 1 and Extended Data Fig. 2). Phylogenetic analyses focused on these few families, in addition to families represented in sponges—on the basis of greater taxon sampling than in the overall classification—produced a significantly supported clustering of *Sycon* (*Sci*)*Cdx* and *LcoCdx* with *Cdx* genes from other species in neighbour-joining, maximum-likelihood and Bayesian analyses (Extended Data Figs 3–5).

We also investigated the genomic neighbourhood of *SciCdx* to help resolve the identity of this homeobox gene (Fig. 2, Supplementary Note 2 and Supplementary Table 1). From the 14 genes on the *SciCdx* scaffold that have clear human orthologues (Supplementary Table 1), four are

¹Sars International Centre for Marine Molecular Biology, University of Bergen, Thormøhlensgate 55, 5008 Bergen, Norway. ²Department of Biology, University of Bergen, Thormøhlensgate 55, 5008 Bergen, Norway. ³The Scottish Oceans Institute, Gatty Marine Laboratory, School of Biology, University of St Andrews, East Sands, St Andrews, Fife KY16 8LB, UK. †Present addresses: Stanley Institute for Cognitive Genomics, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA (O.M.R.); Institute of Marine Research, Nordnesgaten 50, 5005 Bergen, Norway (S.L.).

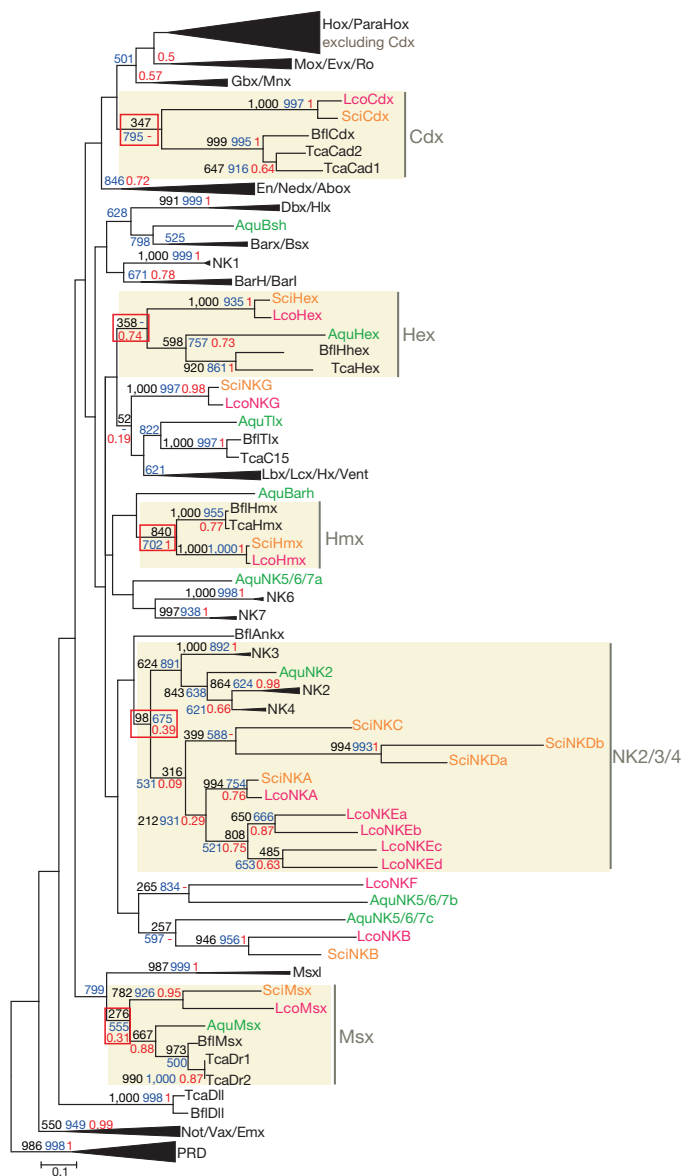


Figure 1 | Phylogenetic tree of the ANTP-class homeodomains. A neighbour-joining tree is displayed. Three support values are shown: the black value is the neighbour-joining bootstrap support, the blue value is the maximum-likelihood bootstrap support and the red value is the posterior probability from Bayesian analysis. Bootstrap values below 10% and posterior probability values below 0.5 are not shown except for associations of calcisponge sequences. The root was determined by using selected PRD-class homeodomains as an outgroup. Aqu, *A. queenslandica*; Bfl, *Branchiostoma floridae*; Tca, *Tribolium castaneum*. Scale bar indicates the number of amino acid substitutions per site.

orthologues of genes linked to ParaHox loci in humans. One of these, *SciSar1A/B*, also has a conserved neighbouring relationship with the ParaHox cluster in the cnidarian *Nematostella vectensis* (Fig. 2a, b). Although these gene numbers are insufficient to reach statistical significance, the neighbour relationships are consistent with the identification of *SciCdx* as a ParaHox gene. Furthermore, as one would expect from the ghost locus hypothesis and the identification of *SciCdx* as a bona fide ParaHox gene, we also find clustering of *Sycon* orthologues of ParaHox and Hox neighbour genes into two distinct groups in the *Sycon* genome to statistically significant levels (Fig. 2c–e). Altogether, the evidence is consistent with the identification of *SciCdx* and *LcoCdx* as the first examples of sponge ParaHox genes.

All of the *Sycon* NK genes are found on separate scaffolds (Extended Data Fig. 6), with multiple additional genes surrounding the homeobox genes. We interpret this as the ancient NK cluster having been broken apart in the *Sycon* genome. Alternatively, our current assembly is not sufficient to provide evidence of a cluster with multiple genes inserted between the NK genes. It has been previously shown that arrangements of NK genes are variable between different species, ranging from intact and conserved NK clusters^{5,16,17} to clusters that are partially broken^{15,18}.

We studied the expression of ANTP-class genes in *Sycon* using a combination of *in situ* hybridization with quantitative transcriptome analysis (Fig. 3, Supplementary Note 3 and Extended Data Figs 7–9). For all ANTP-class genes, except *SciHex*, expression can be detected in oocytes and during cleavage (Fig. 3a and Extended Data Fig. 7a–g). During embryogenesis, the most striking expression domain of the majority of the identified genes is in the cruciform cells, which are putative larval sensory cells^{2,14}. Beginning at the four-cell stage, stronger expression of *SciNKA* marks the cytoplasm destined to become partitioned into the cruciform cells (Fig. 3d and Extended Data Fig. 7h–q), and expression of *SciHmx* is also markedly elevated in these cells (Fig. 3e). *SciNKC* and *SciNKD* are uniquely and strongly expressed in the cruciform cells of more advanced (pre-inversion stage) embryos (Fig. 3f, g). *SciNKA* is

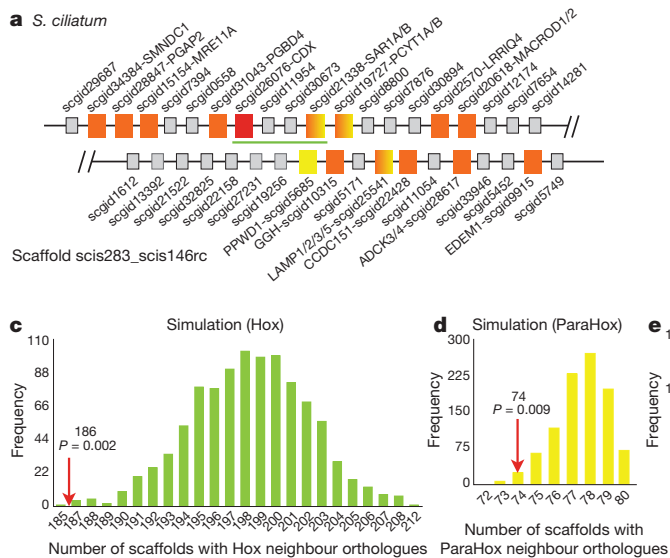


Figure 2 | *SciCdx* synteny and ghost loci simulations. a, b, Genomic neighbourhoods of the *SciCdx* gene and *N. vectensis* ParaHox cluster¹⁷. Colours indicate orthologous relationships with human genes, with chromosomal location represented as follows: yellow, ParaHox neighbours; orange, not linked to Hox/ParaHox loci; yellow–orange, mix of ParaHox and non-Hox/ParaHox neighbours; grey, no orthology. Green lines highlight the conserved *Sar1–Cdx* linkage. c–e, Monte Carlo simulations of human HOX and PARAHOX neighbour orthologue distributions and their overlap across *S. ciliatum* scaffolds. Arrows indicate numbers of scaffolds with Hox and ParaHox neighbour orthologues and their co-localization in *S. ciliatum*; the observed distributions indicate distinct Hox and ParaHox loci.

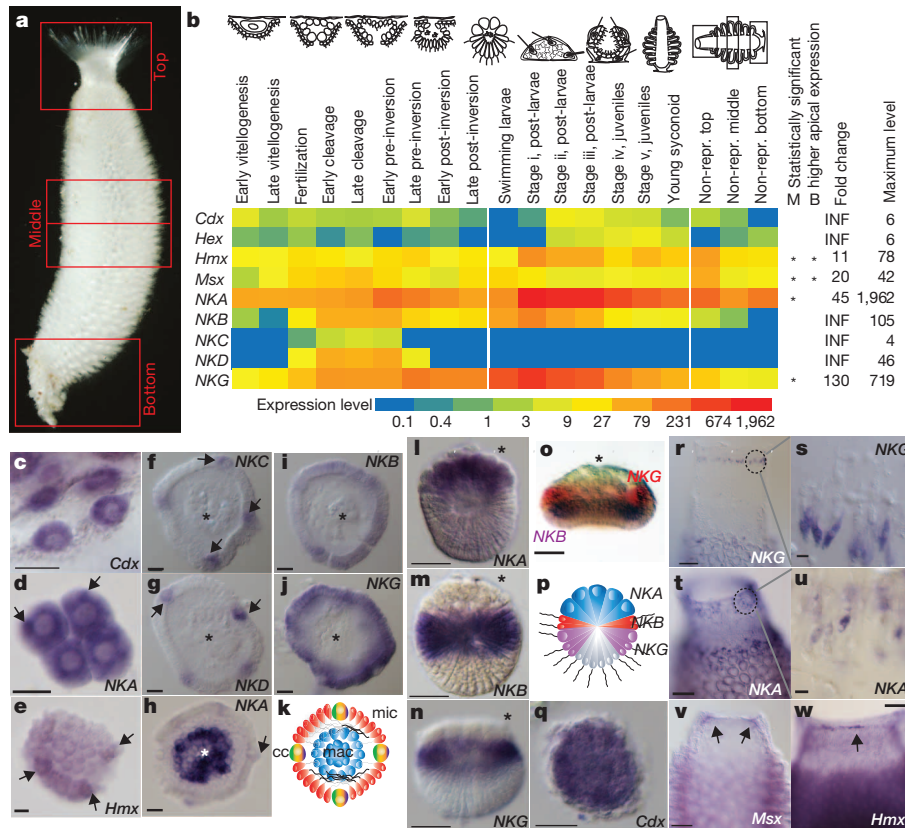


Figure 3 | Expression of *S. ciliatum* ANTP-class genes. **a**, Adult *S. ciliatum*. **b**, Expression heat map. Asterisks indicate statistically higher^{22,23} expression in the apical/top region in comparison to middle (M) or basal (B) parts. Drawings reproduced from ref. 14. INF, positive infinity; repr., reproductive. **c–k**, *S. ciliatum* oocytes (**c**), cleavage (**d**, **e**) and pre-inversion embryos (**f–k**). The formation of cross cells (cc) expressing multiple ANTP-class genes is indicated by arrows (**d–h**) or rainbow colouring (**k**). Macromeres (mac)

expressing SciNKA are indicated by asterisks (**f–o**) or blue colouring (**k**, **p**). Equatorial micromeres (mic) expressing SciNKG and SciNKB are indicated in red (**k**). **l–w**, Post-inversion embryos and larvae (**l–p**), post-larva (**q**), and top parts of sponges (**r–w**) are shown. SciNKG and SciNKA in sclerocytes, and SciMsx and SciHmx in cells of the oscular sphincter are indicated with arrows. Scale bars: 10 μm (**c–h**, **s**, **u**); 25 μm (**l–o**); 50 μm (**q**, **r**, **t**, **v**, **w**).

additionally detected in macromeres of embryos and larvae, and along with SciNKG and SciNKB domains, forms a set of adjacent stripes along the larval anterior–posterior axis (Fig. 3h–p). This pattern is reminiscent of ‘striped’ patterns reported for NK genes in bilaterians¹⁹, and might be indicative of roles for the calcisponge NK genes in axial patterning of the larval body plan or in cell type determination, with cells destined for specific fates distributed along the larval axis. For example, the macromeres give rise to the pinacocytes of the outer cell layer²⁰, and the SciNKG-positive micromeres are good candidates for future sclerocytes (spicule-producing cells), given the co-expression of SciNKG and sclerocyte-specific carbonic anhydrases²¹. All of the ANTP-class genes except SciNKC and SciNKD are expressed during metamorphosis (Fig. 3b and Extended Data Fig. 9) in subpopulations of cells in all three cell layers (Extended Data Fig. 7). The clear expression of SciCdx in the inner cell mass during formation of the choanocyte chamber (Fig. 3q) is particularly striking in light of the recently revived notion of homology of the sponge choanoderm with bilaterian endoderm¹⁴, as ParaHox expression in bilaterians is often associated with the developing gut. In adults, most of the ANTP-class genes display differential expression along the body axis (Fig. 3b and Extended Data Table 1). SciNKG and SciNKA are strongly expressed in sclerocytes, while SciMsx and SciHmx transcripts are predominantly detected within and around the oscular sphincter (Fig. 3r–w).

In summary, analysis of ANTP-class genes in a previously understudied lineage of sponges allowed us to demonstrate pre-poriferan ancestry of ParaHox genes, thus confirming the ghost locus hypothesis and rejecting the ParaHoxzoa hypothesis of Hox/ParaHox gene origins. Expression patterns of the identified genes indicate that developmental functions

of ANTP-class genes also predate poriferans, with probable involvement in the specification of potentially homologous structures (choanoderm/endoderm and cross cells/sensory cells), as well as morphological novelties (calcareous spicules). Differences in ANTP-class gene repertoires between the demosponge *Amphimedon* and the two calcisponges *Sycon* and *Leucosolenia* are striking, and the fact that both classes of sponges share a subset of genes with bilaterians indicates independent gene loss events in the two poriferan lineages.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 27 June; accepted 22 September 2014.

1. Srivastava, M. *et al.* The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature* **466**, 720–726 (2010).
2. Fortunato, S. *et al.* Genome-wide analysis of the sox family in the calcareous sponge *Sycon ciliatum*: multiple genes with unique expression patterns. *EvoDevo* **3**, 14 (2012).
3. Larroux, C. *et al.* Genesis and expansion of metazoan transcription factor gene classes. *Mol. Biol. Evol.* **25**, 980–996 (2008).
4. Fortunato, S. A., Leininger, S. & Adamska, M. Evolution of the Pax-Six-Eya-Dach network: the calcisponge case study. *EvoDevo* **5**, 23 (2014).
5. Larroux, C. *et al.* The NK homeobox gene cluster predates the origin of Hox genes. *Curr. Biol.* **17**, 706–710 (2007).
6. Fahey, B., Larroux, C., Woodcroft, B. J. & Degnan, B. M. Does the high gene density in the sponge NK homeobox gene cluster reflect limited regulatory capacity? *Biol. Bull.* **214**, 205–217 (2008).
7. Ryan, J. F. *et al.* The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to the ParaHoxzoa. *EvoDevo* **1**, 9 (2010).
8. Mendivil Ramos, O., Barker, D. & Ferrier, D. E. K. Ghost loci imply Hox and ParaHox existence in the last common ancestor of animals. *Curr. Biol.* **22**, 1951–1956 (2012).

9. Wörheide, G. *et al.* in *Advances in Marine Biology* Vol. 61 (eds Uriz, M. J., Maldonado, M., Becerro, M. A. & Xavier, T.) 1–78 (Academic, 2012).
10. Peterson, K. J. & Sperling, E. A. Poriferan ANTP genes: primitively simple or secondarily reduced? *Evol. Dev.* **9**, 405–408 (2007).
11. Ferrier, D. E. K. in *Hox Genes: Studies from the 20th to the 21st Century* Vol. 689 (ed. Deutsch, J. S.) 91–100 (Springer, 2010).
12. Garcia-Fernández, J. The genesis and evolution of homeobox gene clusters. *Nature Rev. Genet.* **6**, 881–892 (2005).
13. Chourrout, D. *et al.* Minimal ProtoHox cluster inferred from bilaterian and cnidarian Hox complements. *Nature* **442**, 684–687 (2006).
14. Leininger, S. *et al.* Developmental gene expression provides clues to relationships between sponge and eumetazoan body plans. *Nature Commun.* **5**, 3905 (2014).
15. Luke, G. N. *et al.* Dispersal of NK homeobox gene clusters in amphioxus and humans. *Proc. Natl Acad. Sci. USA* **100**, 5292–5295 (2003).
16. Hui, J. H. L. *et al.* Extensive chordate and annelid macrosynteny reveals ancestral homeobox gene organization. *Mol. Biol. Evol.* **29**, 157–165 (2012).
17. Hui, J. H. L., Holland, P. W. H. & Ferrier, D. E. K. Do cnidarians have a ParaHox cluster? Analysis of synteny around a *Nematostella* homeobox gene cluster. *Evol. Dev.* **10**, 725–730 (2008).
18. Seo, H.-C. *et al.* Hox cluster disintegration with persistent anteroposterior order of expression in *Okopleura dioica*. *Nature* **431**, 67–71 (2004).
19. Saudemont, A. *et al.* Complementary striped expression patterns of NK homeobox genes during segment formation in the annelid *Platynereis*. *Dev. Biol.* **317**, 430–443 (2008).
20. Amano, S. & Hori, I. Metamorphosis of calcareous sponges. 2. Cell rearrangement and differentiation in metamorphosis. *Invertebr. Reprod. Dev.* **24**, 13–26 (1993).
21. Voigt, C., Adamski, M., Sluzek, M. & Adamska, M. Calcareous sponge genomes reveal complex evolution of α -carbonic anhydrases and two key biomineralization enzymes. *BMC Evol. Biol.* (in the press).
22. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).
23. Li, B. & Dewey, C. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements This study was funded by the Sars Centre core budget to M. Adamska. Sequencing was performed at the Norwegian High Throughput Sequencing Centre funded by the Norwegian Research Council. O.M.R. and D.E.K.F. acknowledge support from the BBSRC and the School of Biology, University of St Andrews. We thank B. Bergum for collecting samples in the 2011 season and R. Holdhus from the Genomics Core Facility at the University of Bergen for help with DNA shearing.

Author Contributions S.A.V.F. carried out gene identification and cloning, analysed gene expression by *in situ* hybridization, and participated in phylogenetic analyses and manuscript writing. M. Adamski performed sequence assembly, annotation, quantification of gene expression and participated in sample collection, phylogenetic analyses and manuscript writing. O.M.R. performed the synteny analyses, participated in phylogenetic analyses, manuscript writing and design of the research approach for synteny and phylogenetic analyses. S.L. isolated samples for sequencing of genomes, generated Mate Pair libraries, and participated in sample collection. J.L. generated samples for sequencing of *S. ciliatum* metamorphosis stages. D.E.K.F. participated in design of the research approach for synteny and phylogenetic analyses and writing of the manuscript. M. Adamska conceived the study, and participated in data analysis, sample collection and writing of the manuscript.

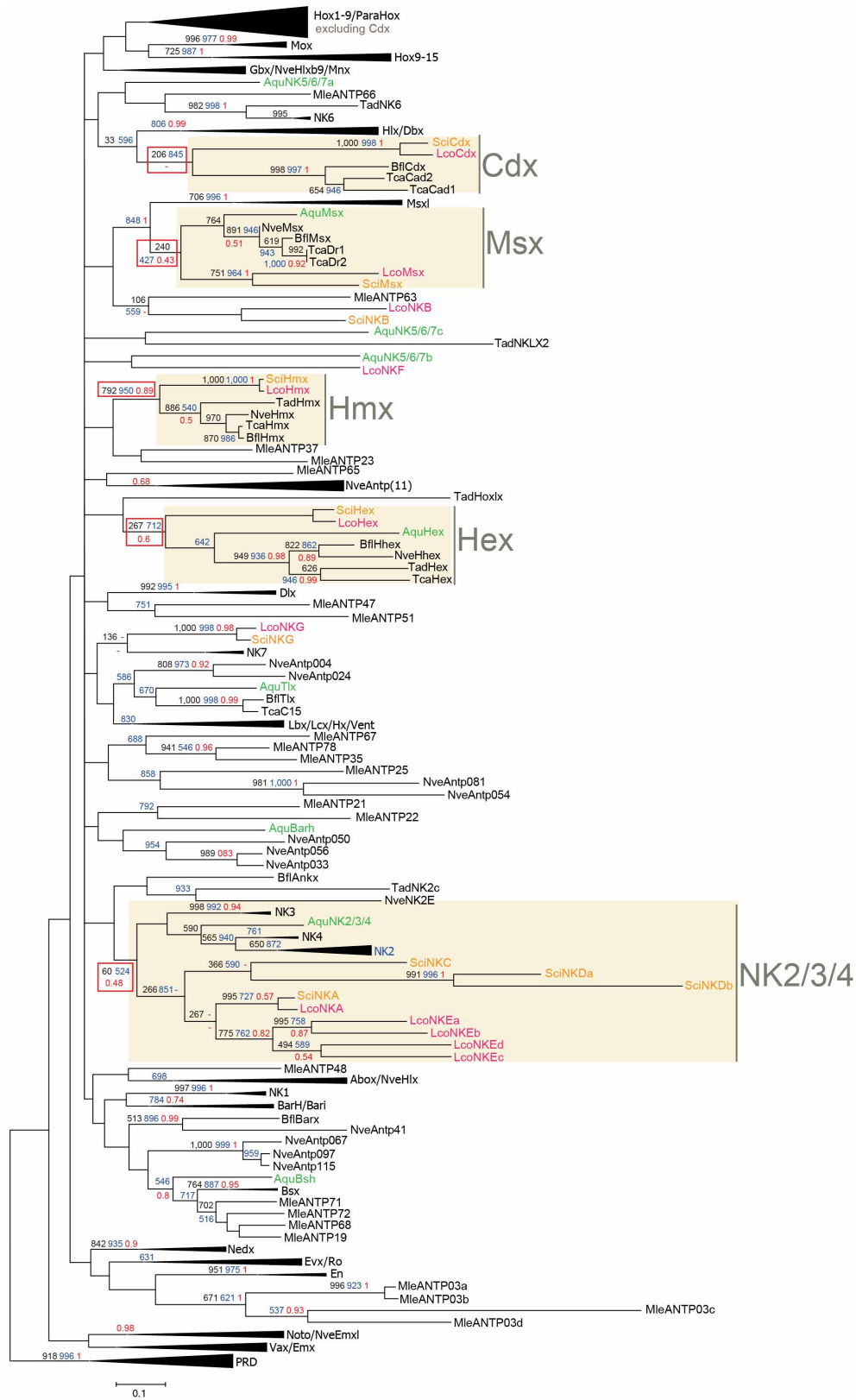
Author Information Genome assembly of *S. ciliatum* and the coding sequences and their translations from transcriptome assemblies of *S. ciliatum* and *L. complicata* used in this study can be accessed through <http://compagen.zoologie.uni-kiel.de/> and are also deposited at <http://dx.doi.org/10.5061/dryad.tn0f3>. RNA-seq data have been deposited in ArrayExpress under accession numbers E-MTAB-2430, E-MTAB-2431 and E-MTAB-2890, and the cloned coding sequences of *S. ciliatum* ANTP-class genes have been deposited in the European Nucleotide Archive under accession numbers LN609546 to LN609553. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M. Adamska (maja.adamska@sars.uib.no).

METHODS

Genome and transcriptome assemblies will be described in detail elsewhere (M. Adamski, S.L. and M. Adamska, unpublished observations). Briefly, the high-quality draft genome assembly of *S. ciliatum* was generated using two (360 bp and 530 bp) paired-end libraries and several mate-pair libraries ranging from 2.0 to 9.0 kb, and the preliminary draft assembly of *L. complicata* was generated from a single 295 bp paired-end library, all prepared and sequenced by Illumina technology. Assembly was performed using SOAPdenovo2 (ref. 24) and scaffolding was performed using SSPACE v.2.2 (ref. 25), and resulted in the length of the shortest scaffold for which scaffolds of equal or greater length contain at least half of the nucleotides in the genome (N_{50}) = 150 kb and 450 bp for *S. ciliatum* and *L. complicata*, respectively. Transcriptomes were assembled using Trinity²⁶. For *S. ciliatum*, genomic scaffolds and transcripts of sponge origin (as opposed to those derived from associated organisms) were identified by aligning the resulting assembly to reads from an Illumina-sequenced library obtained from laboratory-grown, eukaryotic-contamination-free juveniles. The calcisponge ANTP-class sequences were retrieved from these assemblies using TBLASTN with representative query homeodomain sequences from *A. queenslandica*, *Mus musculus*, *T. castaneum* and *B. floridae*. For phylogenetic analysis, we selected *B. floridae* and *T. castaneum* to provide a framework for the classification of the sponge sequences, as these species have been shown to collectively contain homologues of all major bilaterian ANTP-class genes²⁷. Their homeodomain sequences were extracted from HomeoDB²⁸. Prottest3.0 (ref. 29) and Modelgenerator v.0.85 (ref. 30) were used to determine the best suitable model of sequence evolution (LG + G). Phylogenetic analyses were based on neighbour-joining (Phylip v.3.69), maximum-likelihood (PhyML v.3.0) and Bayesian inference (MrBayes v.3.1.2) methods. Gene expression was studied using available packages^{22,23}. *S. ciliatum* ANTP-class gene amplification, cloning, sequencing, probe production and single *in situ* hybridization were performed as described previously². In the double *in situ* experiment, samples were hybridized simultaneously with digoxigenin-labelled SciNKB

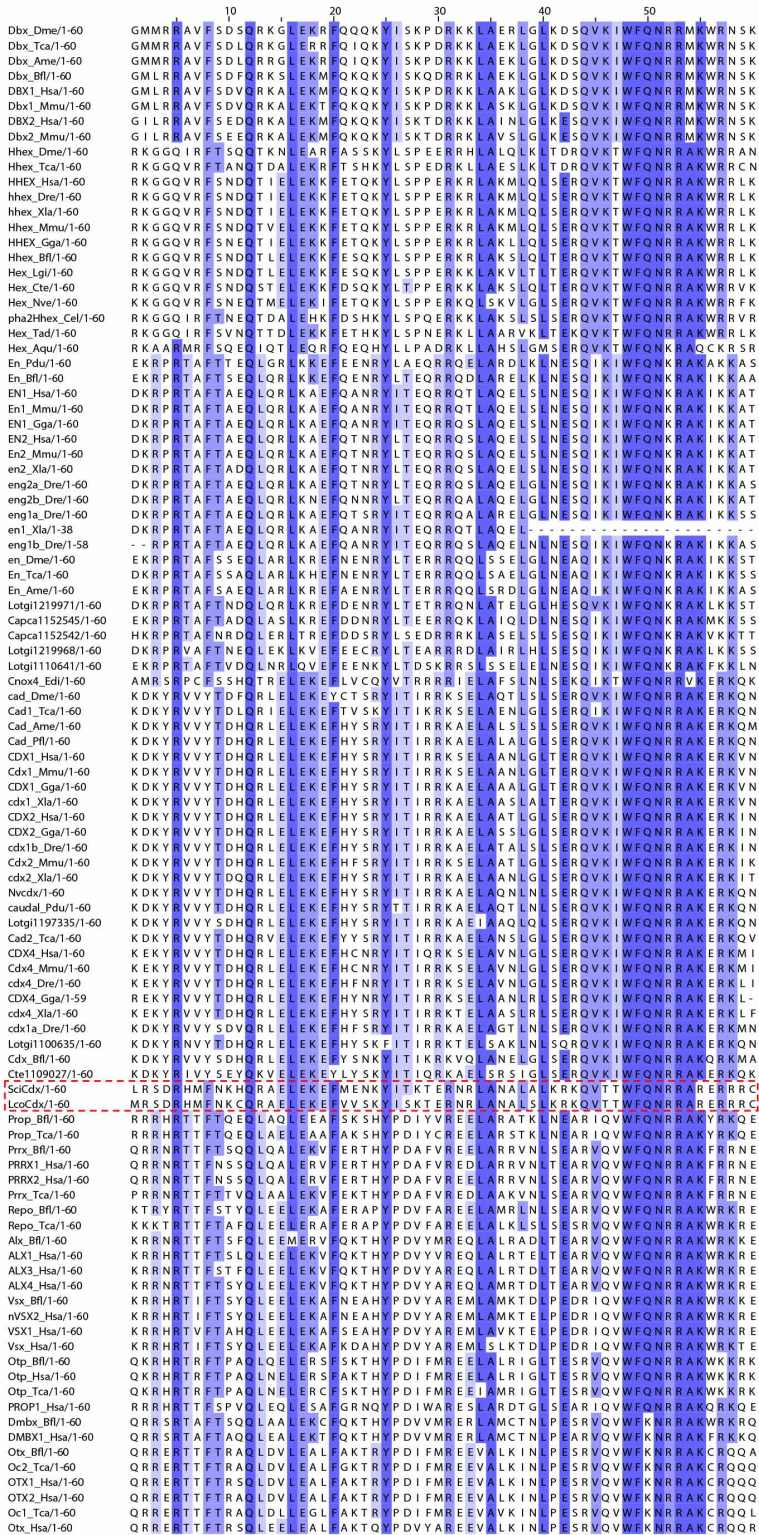
probe and fluorescein-labelled SciNKG probe. After detection of the digoxigenin-labelled probe with NBT/BCIP substrate, the anti-digoxigenin antibody was removed by two 5 min washes in 0.1 M glycine/HCl, pH 2.2, 0.1% Tween 20 followed by three additional maleic acid buffer washes. A second round of pre-blocking, antibody incubation and post-antibody washes were as in the single probe protocol with the exception that anti-Fluo-AP antibody was used and the colour was developed using Fast Red tablets (Roche) according to the manufacturer's instructions. Photographs demonstrating gene expression are representative of multiple individual specimens, with the following replicates. Oocytes and embryos: 3–4 small pieces of adult sponge, each containing tens to hundreds of oocytes or embryos of a given developmental stage; young syconoid sponges: at least 5 individual specimens; juveniles: small Petri dishes or wells of multi-well plates containing at least 10 juveniles. At least two independent experiments were carried out for each probe.

24. Luo, R. *et al.* SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* **1**, 18 (2012).
25. Boetzer, M., Henkel, C., Jansen, H., Butler, D. & Pirovano, W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579 (2011).
26. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnol.* **29**, 644–652 (2011).
27. Takatori, N. *et al.* Comprehensive survey and classification of homeobox genes in the genome of amphioxus, *Branchiostoma floridae*. *Dev. Genes Evol.* **218**, 579–590 (2008).
28. Zhong, Y.-F. & Holland, P. HomeoDB2: functional expansion of a comparative homeobox gene database for evolutionary developmental biology. *Evol. Dev.* **13**, 567–568 (2011).
29. Darriba, D., Taboada, G., Doallo, R. & Posada, D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165 (2011).
30. Keane, T. M., Creevey, C., Pentony, M., Naughton, T. & McInerney, J. Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified. *BMC Evol. Biol.* **6**, 29 (2006).



Extended Data Figure 1 | Phylogenetic tree of the ANTP-class homeodomains including representative bilaterian and non-bilaterian sequences. A neighbour-joining tree using the JTT+G (0.5) (1,000 bootstraps) model of protein evolution is displayed. A combination of three support values obtained for three phylogenetic methods is shown: black value is the bootstrap support from the neighbour-joining method, blue value is the bootstrap support from the maximum-likelihood method (LG+G 0.5), and red value is posterior probability from Bayesian analysis (LG+G 0.5). Bootstrap

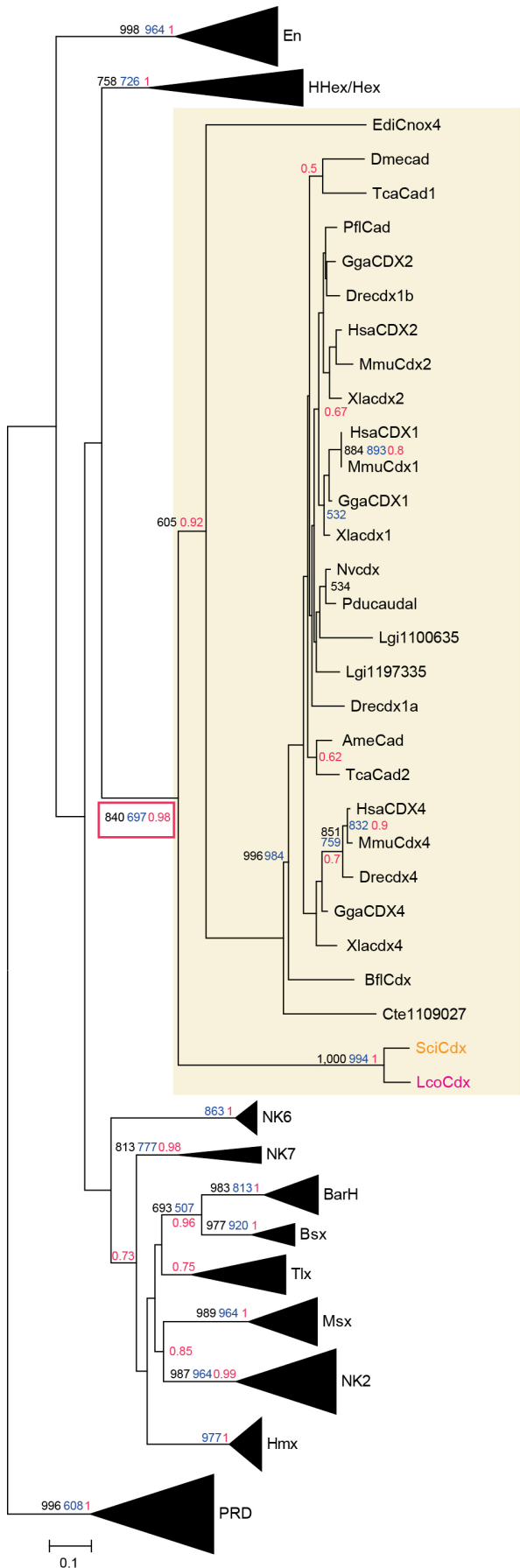
support values below 10% and posterior probability values below 0.5 are not shown except for associations of calcisponge sequences. The root was determined by using selected PRD-class homeodomains as an outgroup. Aqu, *A. queenslandica* (Porifera/demosponges); Bfl, *B. floridae* (Chordata); Lco, *L. complicata*; Mle, *M. leidyi* (Ctenophora); Nve, *N. vectensis* (Cnidaria); Sci, *S. ciliatum* (Porifera/calcsponges); Tad, *Trichoplax adhaerens* (Placozoa); Tca, *T. castaneum* (Arthropoda). Scale bar indicates the number of amino acid substitutions per site.



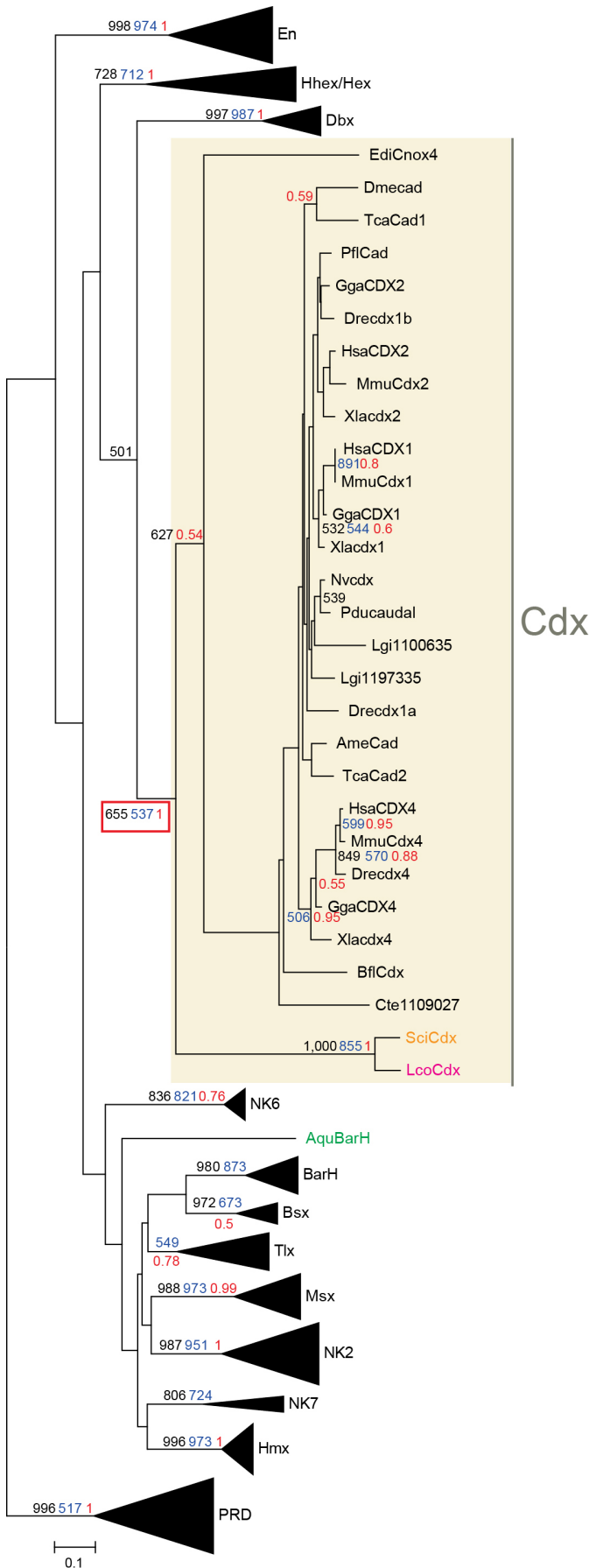
Extended Data Figure 2 | Variability of the YIT/YIS homeodomain motif within the Cdx/Cad, En and Dbx families in bilaterians, cnidarians, a placozoan and sponges. Ame, *Apis mellifera*; Aqu, *A. queenslandica*; Bfl, *B. floridae*; Cte, *Capitella teleta*; Dme, *Drosophila melanogaster*; Dre, *Danio*

erio; Edi, *Eleutheria dichotoma*; Gga, *Gallus gallus*; Hsa, *Homo sapiens*; Lgi, *Lottia gigantea*; Mmu, *Mus musculus*; Nv, *Nereis virens*; Nve, *N. vectensis*; Pdu, *Platynereis dumerilii*; Pfl, *Ptychodera flava*; Tad, *T. adhaerens*; Tca, *T. castaneum*; Xla, *Xenopus laevis*.

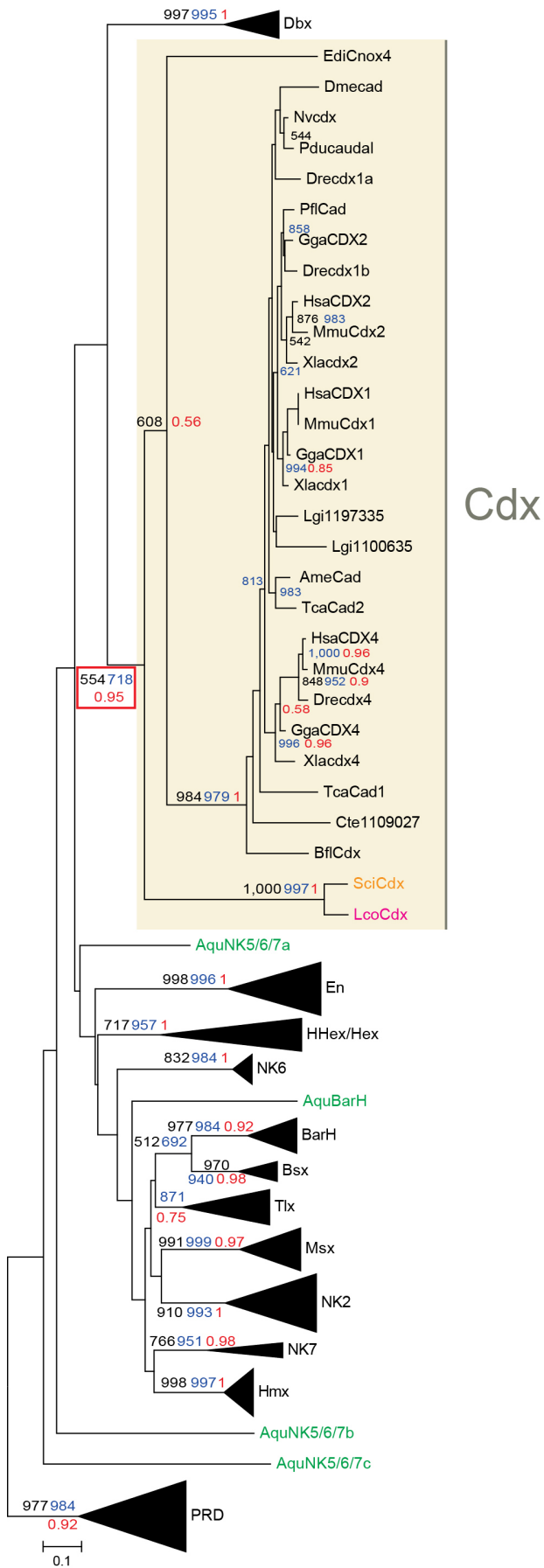
Extended Data Figure 3 | Phylogenetic tree including ANTP-class homeodomain families represented in sponges and two additional families characterized by the presence of the YIT motif but excluding divergent *A. queenslandica* sequences. Neighbour-joining (JTT, 1,000) bootstrap support values are in black, maximum-likelihood (LG+G 0.4, 1,000 replicates) bootstrap support values are in blue and Bayesian (LG+G 0.4) posterior probability values are in red. Only bootstrap support values equal to or above 500 are shown. All families except Cdx are collapsed for clarity. Ame, *A. mellifera*; Aqu, *A. queenslandica*; Bfl, *B. floridae*; Cte, *C. teleta*; Dme, *D. melanogaster*; Dre, *D. rerio*; Edi, *E. dichotoma*; Gga, *G. gallus*; Hsa, *H. sapiens*; Lgi, *L. gigantea*; Mmu, *M. musculus*; Nv, *N. virens*; Nve, *N. vectensis*; Pdu, *P. dumerilii*; Pfl, *P. flava*; Tad, *T. adhaerens*; Tca, *T. castaneum*; Xla, *X. laevis*. Scale bar indicates the number of amino acid substitutions per site.



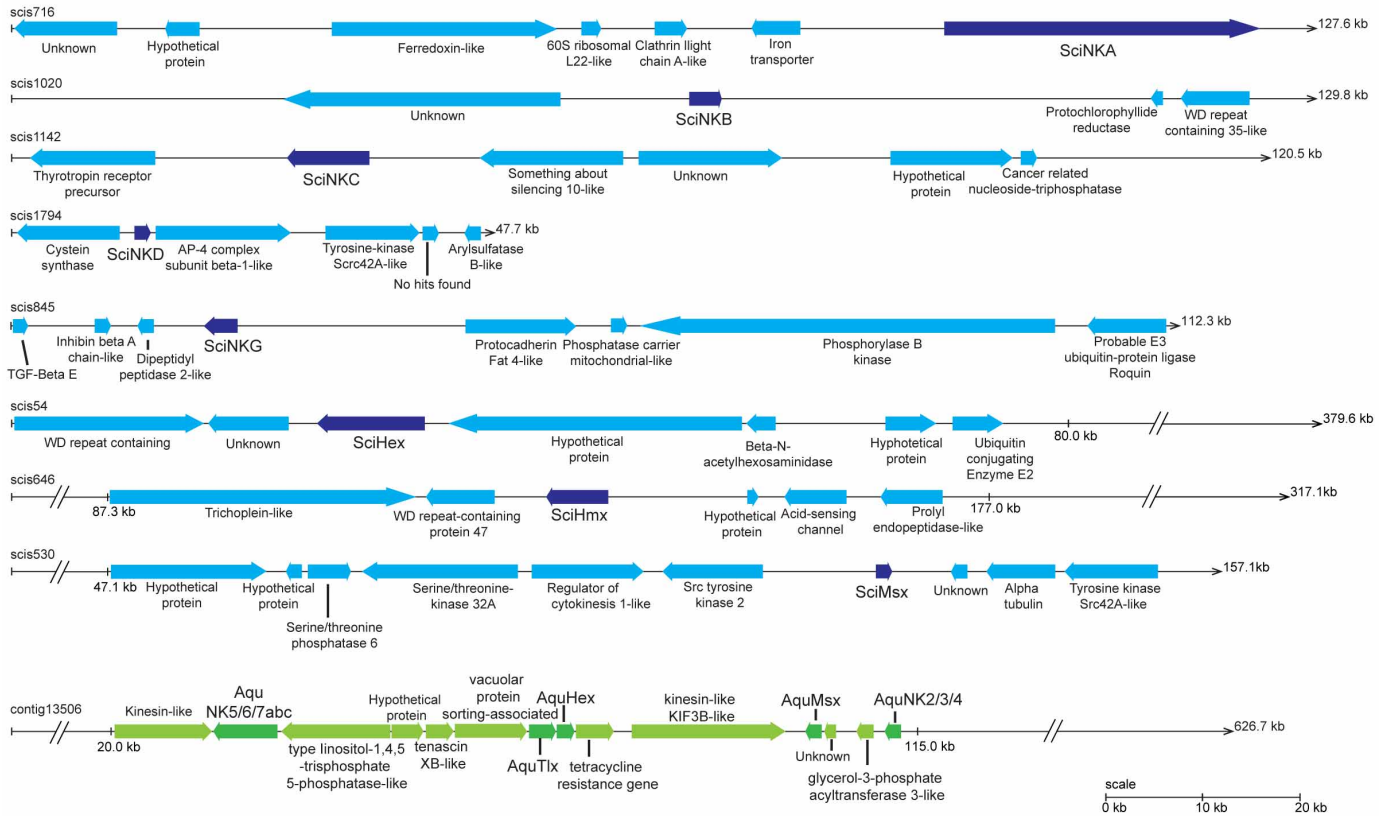
Cdx



Extended Data Figure 4 | Phylogenetic tree including ANTP-class homeodomain families represented in sponges and three additional families characterized by the presence of YIT/YIS motifs but excluding some divergent *A. queenslandica* sequences. Neighbour-joining (JTT, 1,000 replicates) bootstrap support values are in black, maximum-likelihood (LG+G 0.4, 1,000 replicates) bootstrap support values are in blue and Bayesian (LG+G 0.4) posterior probability values are in red. Only bootstrap support values equal to or above 500 are shown. All subfamilies except Cdx are collapsed for clarity. Ame, *A. mellifera*; Aqu, *A. queenslandica*; Bfl, *B. floridae*; Cte, *C. teleta*; Dme, *D. melanogaster*; Dre, *D. rerio*; Edi, *E. dichotoma*; Gga, *G. gallus*; Hsa, *H. sapiens*; Lgi, *L. gigantea*; Mmu, *M. musculus*; Nv, *N. virens*; Nve, *N. vectensis*; Pdu, *P. dumerilii*; Pfl, *P. flava*; Tad, *T. adhaerens*; Tca, *T. castaneum*; Xla, *X. laevis*. Scale bar indicates the number of amino acid substitutions per site.

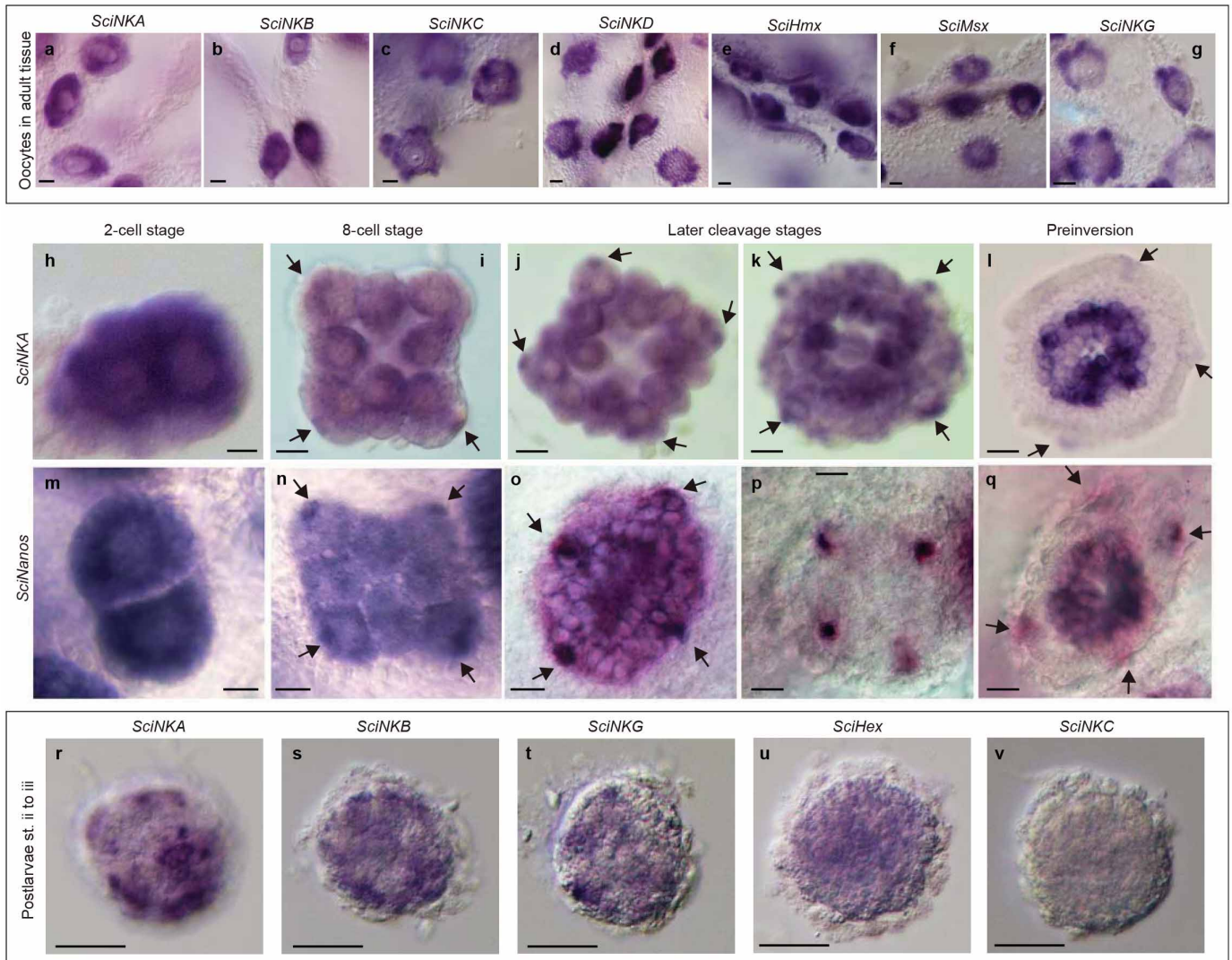


Extended Data Figure 5 | Phylogenetic tree including ANTP-class homeodomain families represented in sponges and three additional families characterized by the presence of YIT/YIS motifs. Neighbour-joining (JTT, 1,000 replicates) bootstrap support values are in black, maximum-likelihood (LG+G 0.4, 1,000 replicates) bootstrap support values are in blue and Bayesian (LG+G 0.4) posterior probability values are in red. Only bootstrap support values equal to or above 500 are shown. All families except Cdx are collapsed for clarity. Ame, *A. mellifera*; Aqu, *A. queenslandica*; Bfl, *B. floridae*; Cte, *C. teleta*; Dme, *D. melanogaster*; Dre, *D. rerio*; Edi, *E. dichotoma*; Gga, *G. gallus*; Hsa, *H. sapiens*; Lgi, *L. gigantea*; Mmu, *M. musculus*; Nv, *N. virens*; Nve, *N. vectensis*; Pdu, *P. dumerilii*; Pfl, *P. flava*; Tad, *T. adhaerens*; Tca, *T. castaneum*; Xla, *X. laevis*. Scale bar indicates the number of amino acid substitutions per site.



Extended Data Figure 6 | *S. ciliatum* scaffolds containing NK genes and *A. queenslandica* scaffold containing cluster of NK genes. *S. ciliatum* NK genes are indicated in blue and *A. queenslandica* NK genes are indicated in

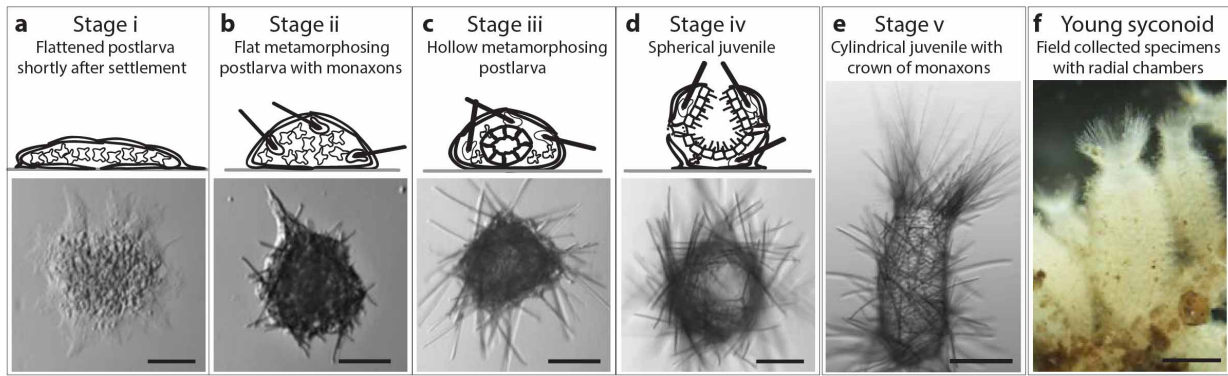
green (modified with permission from ref. 6). Annotation of the neighbouring genes (genes within 50 kb from the NK gene) in *S. ciliatum* was performed using BLASTp searches against the RefSeq database.



Extended Data Figure 7 | Additional expression patterns of ANTP-class homeobox genes in embryonic development and during metamorphosis.

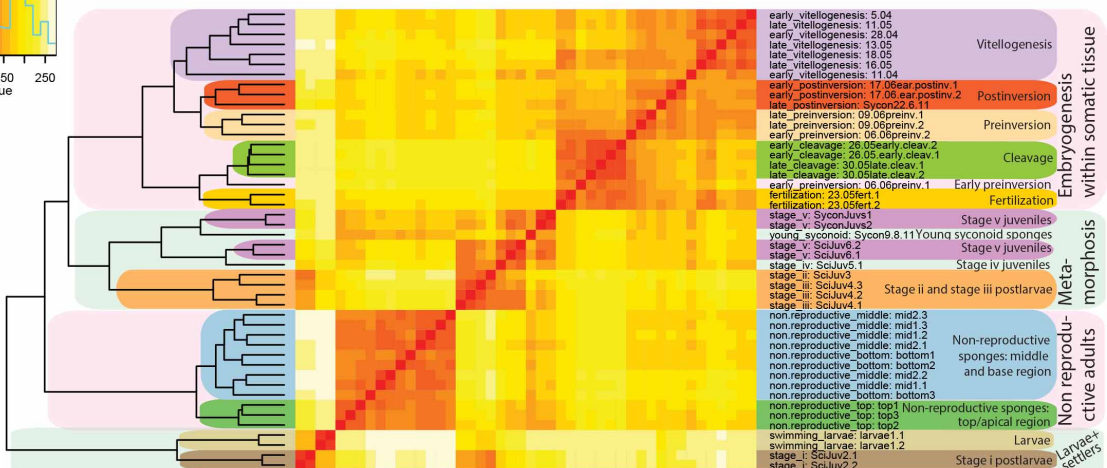
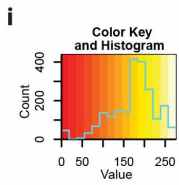
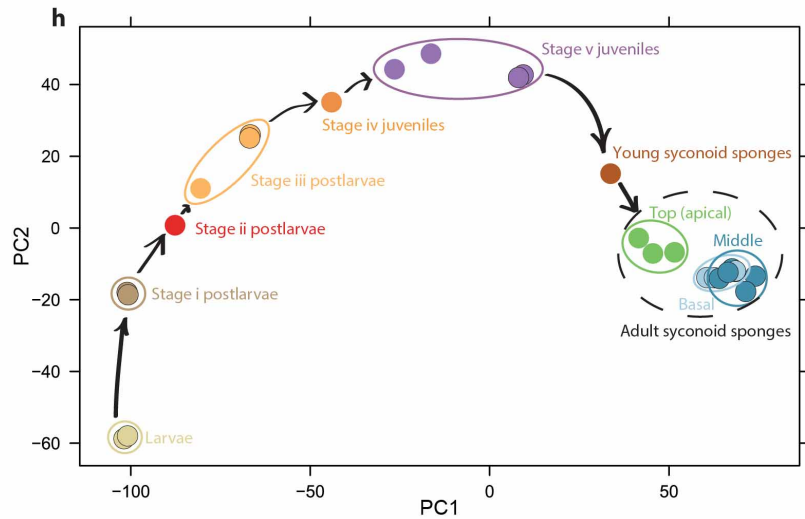
a–g. All of the investigated genes (except *Hex*, data not shown) are expressed in oocytes. The expression of *SciNKA* is detectable in all blastomeres of the cleavage-stage embryos, but the transcripts are concentrated in the corner-most cytoplasm, which becomes gradually partitioned to the cross cells (arrows). **l–q.** This subcellular localization of cross-cell-enriched transcripts is also

observed for *SciNanos*, expression of which, similarly to *SciNKA*, becomes ultimately restricted to cross cells and macromeres in pre-inversion-stage embryos. **r–v.** In metamorphosing post-larvae, *SciNKA* is expressed in the cells of the outer layer (**r**), *SciNKB* and *SciNKG* in (possibly non-overlapping) fractions of cells in the inner cell mass (**s**, **t**); *SciHex* is weakly expressed throughout the inner cell mass (**u**), and *SciNKC* (**v**) and *NKD* (data not shown) are not detectable in the juveniles. Scale bars: 10 μm (**a–q**); 25 μm (**r–v**).



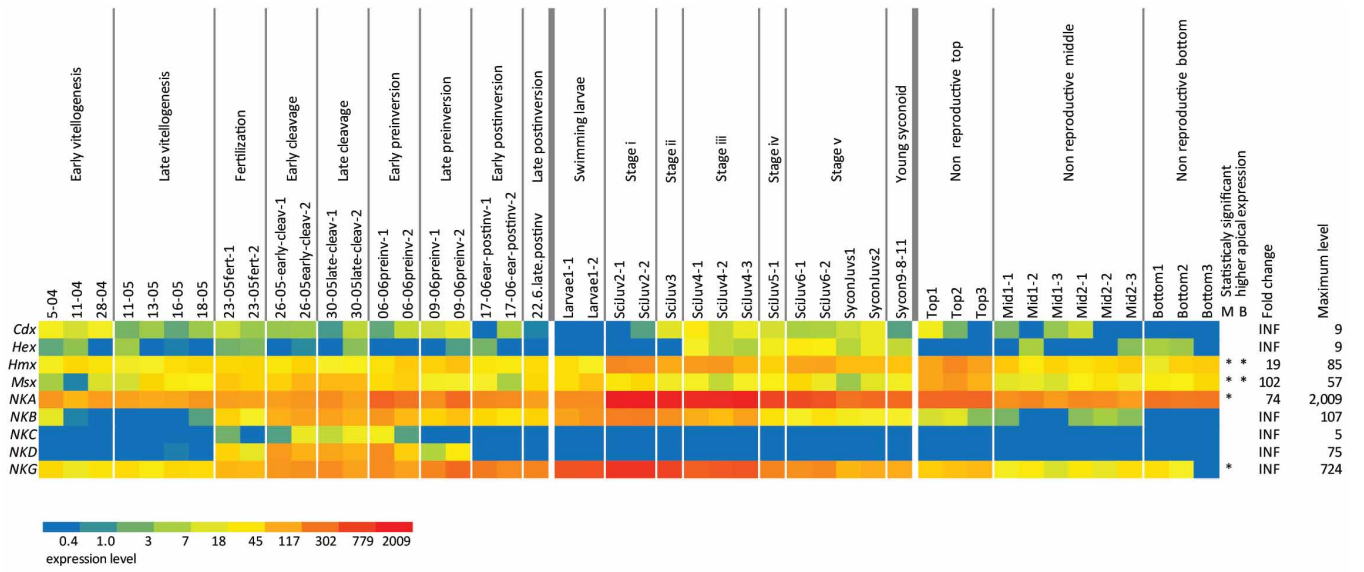
g

Sample	Replica	Collected
Stage i	1	29/06/12
	2	29/06/12
Stage ii	1	09/07/12
	1	13/07/12
Stage iii	2	13/07/12
	3	13/07/12
	1	05/07/12
Stage iv	1	16/07/12
	2	20/07/12
Stage v	3	11/08/11
	4	11/08/11
	1	09/08/11
	1	09/08/11



Extended Data Figure 8 | Samples used for quantification of expression. **a–f**, Metamorphosis in *S. ciliatum*, with stages based on those described previously²⁰ with modifications. Stage I, approximately 12 h after settlement: large flat cells derived from larval macromeres envelop the inner cell mass composed of former micromeres (**a**). Stage II, approximately 24 h after settlement: single-axis spicules (monaxons) are produced by sclerocytes, which have differentiated from the inner cell mass cells (**b**). Stage III, 2–3 days after settlement: choanocytes that have differentiated from the inner cells mass cells form a single internal chamber (**c**). Stage IV, approximately 4 days after settlement: osculum (exhalant opening) forms at the apical end of the spherical juvenile; first tri-radial spicules become apparent (**d**). Stage V, approximately 10 days after settlement: the juvenile is elongated along the apical–basal axis, long straight spicules form a crown around the osculum (**e**). Young syconoid sponges, approximately 8 weeks after settlement (**f**). **a–e**, Photographs of live specimens in culture; photographs **a–d** are top (apical) views, cartoon representations of sections and photograph **e** are side views. Scale bars: 100 μ m (**a–e**); 1 mm (**f**). **g**, Details of replicates used for the analysis.

Several hundred juveniles were used in each sample. **h**, Plot demonstrating the results of principal component analysis of the metamorphosis series and axial dissection series of non-reproductive adults calculated according to the previously described method²² and using information about the top 500 differentially expressed genes as default parameters. Metamorphosis stages and parts of sponges are colour coded, with the ovals added manually for easier visualization of similarities and differences between the samples. Progress of development, starting from freshly released larvae and until the emergence of adult, but not yet reproductive sponges, is indicated by arrows. Note similarities of samples within replicates and with neighbouring stages of the metamorphosis series, and distinctiveness of the top (apical) samples from the basal and middle samples of the adults. **i**, Heat-map representation of sample-to-sample distances among all samples used in this study, calculated according to the previously described method²² and based on expression of all coding genes in *S. ciliatum* (approximately 18,000 sequences). Note that replicates and neighbouring stages group together, as indicated by highlighting.



Extended Data Figure 9 | Heat-map representation of expression profiles. As demonstrated in Fig. 3b but with data from individual libraries presented separately.

Extended Data Table 1 | Quantification of differences in expression levels between top, middle and bottom parts of non-reproductive adult specimens of *S. ciliatum*

Gene	Expression level			Diff. expression middle-top		Diff. expression bottom-top	
	Top	Middle	Bottom	Fold change	Adj. <i>P</i> value	Fold change	Adj. <i>P</i> value
<i>Cdx</i>	1	1	0	2.1	9.5E-01	INF	9.0E-01
<i>Hex</i>	0	0	1	0.0	1.0E-00	0.0	9.6E-01
<i>Hmx</i>	32	8	7	4.3	1.2E-03*	4.6	3.0E-02*
<i>Msx</i>	24	3	5	7.6	1.7E-06*	4.5	2.8E-02*
<i>NKA</i>	106	38	71	2.8	3.6E-07*	1.5	3.8E-01
<i>NKB</i>	2	1	0	2.8	7.9E-01	INF	7.8E-01
<i>NKC</i>	0	0	0	NA	NA	NA	NA
<i>NKD</i>	0	0	0	NA	NA	NA	NA
<i>NKG</i>	12	4	3	3.2	7.0E-02*	3.9	2.5E-01

'Expression level' was calculated as the sum of the posterior probability of each read coming from a given gene over all reads²³ scaled using size factors of the libraries²²; 'fold change' was calculated between expression levels in the middle (middle-top) or bottom (bottom-top) and top part of the sponge. adj. *P* value, *P* value adjusted for multiple testing with the Benjamini-Hochberg procedure²². diff., differential; Inf, positive infinity; NA, not applicable.

* Values with statistical significance not less than 90% (adjusted *P* value \leq 0.1) and apical expression levels higher than in the middle or bottom part of the sponge.

Non-equivalent contributions of maternal and paternal genomes to early plant embryogenesis

Gerardo Del Toro-De León¹, Marcelina García-Aguilar¹ & C. Stewart Gillmor¹

Zygotic genome activation in metazoans typically occurs several hours to a day after fertilization, and thus maternal RNAs and proteins drive early animal embryo development¹. In plants, despite several molecular studies of post-fertilization transcriptional activation, the timing of zygotic genome activation remains a matter of debate. For example, two recent reports that used different hybrid ecotype combinations for RNA sequence profiling of early *Arabidopsis* embryo transcriptomes came to divergent conclusions. One identified paternal contributions that varied by gene, but with overall maternal dominance², while the other found that the maternal and paternal genomes are transcriptionally equivalent³. Here we assess paternal gene activation functionally in an isogenic background, by performing a large-scale genetic analysis of 49 *EMBRYO DEFECTIVE* genes and testing the ability of wild-type paternal alleles to complement phenotypes conditioned by mutant maternal alleles. Our results demonstrate that wild-type paternal alleles for nine of these genes are completely functional 2 days after pollination, with the remaining 40 genes showing partial activity beginning at 2, 3 or 5 days after pollination. Using our functional assay, we also demonstrate that different hybrid combinations exhibit significant variation in paternal allele activation, reconciling the apparently contradictory results of previous transcriptional studies^{2,3}. The variation in timing of gene function that we observe confirms that paternal genome activation does not occur in one early discrete step, provides large-scale functional evidence that maternal and paternal genomes make non-equivalent contributions to early plant embryogenesis, and uncovers an unexpectedly profound effect of hybrid genetic backgrounds on paternal gene activity.

In animals, maternal RNAs and proteins drive early embryo development. The transition to zygotic control requires both clearing of maternal transcripts and zygotic genome activation, a phenomenon collectively known as the maternal-to-zygotic transition¹. Research in plants has focused almost exclusively on zygotic genome activation, and has led to differing conclusions on its timing. Studies in maize and in tobacco suggested that large-scale zygotic transcription does not begin until 3–4 days after fertilization^{4,5}. In agreement with this, an *Arabidopsis* RNA sequence profiling and reporter line study² and three smaller-scale molecular studies^{6–8} also came to the conclusion that many paternal alleles are either quiescent or transcribed at lower levels than corresponding maternal alleles in the first few days after fertilization.

Other work has suggested a contrary view, namely that zygotic transcription begins immediately or soon after fertilization. Significant differences have been found between the transcript populations of egg cells and two-celled embryos in wheat⁹, and equivalent maternal and paternal expression for 25 different genes was observed in maize zygotes at 1 day after pollination¹⁰. In *Arabidopsis*, three genes have been shown by genetic means to be expressed from the paternal allele in early embryogenesis^{11–13}, and a recent genome-wide transcriptional study found biparental transcription for nearly all of the ~8,000 genes that could be assayed at the one- to two-cell zygote stage³. Even taking into account that timing of zygotic genome activation may differ among plant species, it has been difficult to reconcile these conflicting results. However, when

evaluating the two RNA profiling experiments arguing for³ or against² immediate zygotic genome activation in *Arabidopsis*, it is important to note that each study used a unique combination of polymorphic strains to distinguish maternal and paternal transcripts, such that the results are based on gene expression in embryos hybrid for different combinations of ecotypes.

To study paternal allele activation in an isogenic background, we used a genetic assay with strict functional criteria: a test of the ability of wild-type (WT) paternal alleles to complement early embryo mutant phenotypes conditioned by mutant maternal alleles. While this assay is necessarily limited in the number of genes that can be tested (and in the corresponding conclusions that can be drawn), it presents two significant advantages over RNA sequencing approaches: by design, we only evaluate genes with a proven role in embryo development, leading to biologically relevant results. Second, we avoid the complex effects on quantitative traits—such as gene expression values—that necessarily result from crossing polymorphic, distantly related ecotypes (the basis for allele specific profiling by RNA sequencing)^{14–17}.

We screened a collection of mutants in several hundred *EMBRYO DEFECTIVE* (*EMB*) genes required for early embryo development in *Arabidopsis*^{18,19}, and selected genes that were associated with clearly recognizable anatomical defects segregating in *emb/+* plants at pre-globular stages (Supplementary Table 1 and Supplementary Data Set 1). For 49 different *EMB* genes, we then crossed heterozygous mutant plants with WT pollen, and scored the resulting embryos for mutant phenotypes at 2, 3, 5 and ~14 days after pollination (Fig. 1; complete data shown in Supplementary Table 2). We identified nine genes where the WT paternal allele complemented the mutant maternal allele as soon as there was a requirement for the gene product, resulting in no observed mutant phenotypes at 2 days after pollination or later (*AtSWI3A* shown as an example in Fig. 1d–f, j). For the remaining 40 genes, the WT paternal allele gradually complemented the mutant phenotype conferred by the mutant maternal allele (*EMB2804* shown as an example in Fig. 1g–i, j). To determine more precisely the time point of the onset of paternal allele activation, we studied a subset of 13 genes in more detail (Supplementary Table 3). Comparison of mutant ratios resulting from hand-pollinated *emb/+* × *emb/+* and *emb/+* × Col crosses identified genes whose activation began at either 2, 3 or 5 days after pollination. By contrast, Col × *emb/+* control crosses resulted in no mutant phenotypes, showing that the delayed complementation we observed does not reflect haploinsufficiency expressed only in the first few days of embryogenesis, and that the maternal allele for these genes is absolutely required in early embryogenesis, while the paternal allele is not (Supplementary Table 3).

We next sought to correlate function and expression, selecting the *GCT* and *CCT* genes as examples²⁰. Figure 2a–g, i–o shows that maternally derived *gGCT-GUS* and *pCCT::GUS* reporter constructs drove embryo expression at 2 days after pollination, whereas paternally derived constructs drove embryo expression beginning at 3 days after pollination, and reached full expression at 4 days after pollination. Phenotypic complementation of *gct* and *cct* mutant maternal alleles by the WT paternal allele began at 5 days after pollination (Fig. 2h, p), 1 day after

¹Laboratorio Nacional de Genómica para la Biodiversidad (Langebio), Unidad de Genómica Avanzada, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), Irapuato, Guanajuato 36821, México.

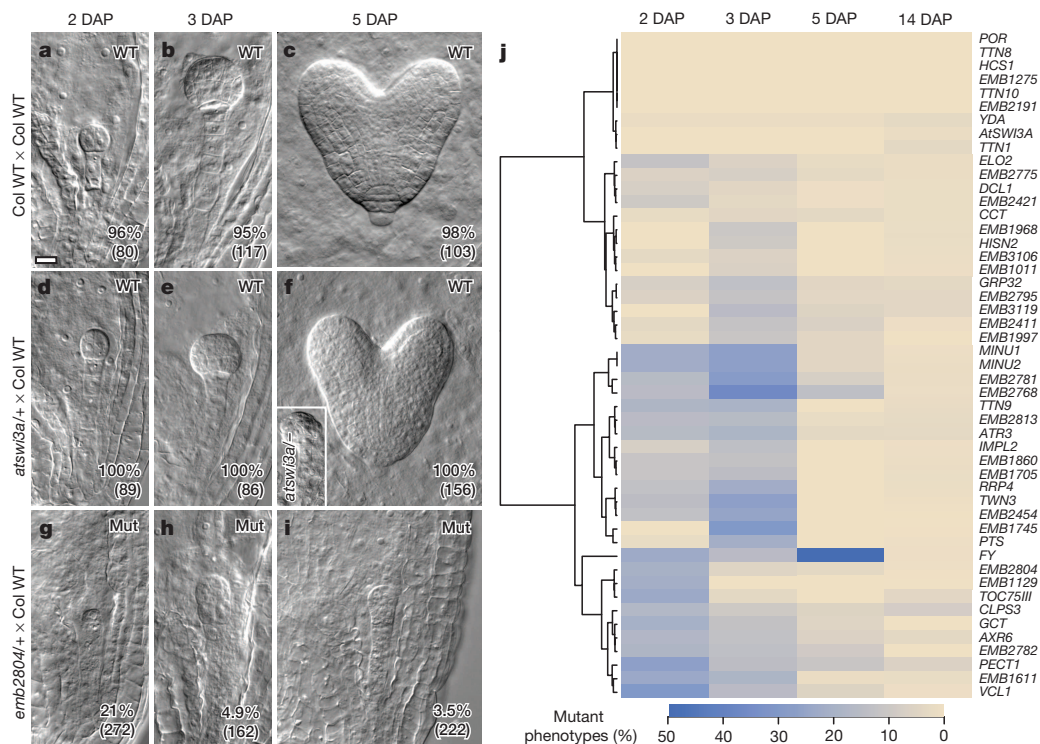


Figure 1 | Paternal allele activation in early *Arabidopsis* embryogenesis varies between genes. **a–c**, Normal phenotypes resulting from hand self-pollinated WT plants. DAP, days after pollination. **d–f**, Normal phenotypes resulting from *atswi3a/+* × Col WT. For comparison purposes, the inset in **f** shows the phenotype of a heart stage *atswi3a/–* embryo from a selfed *atswi3a/+* plant. **g–i**, Mutant (Mut) phenotypes resulting from *emb2804/+* × Col WT. In **a–i**, the percentage of embryos with WT or mutant phenotypes at 2, 3 and 5 days after pollination is shown, with the total number of embryos examined in parentheses. In crosses, the maternal genotype is listed first. Embryos shown

at equal magnification; scale bar, 10 μm. **j**, Heat map summarizing the percentage of mutant phenotypes observed at 2, 3, 5 and ~14 days after pollination in embryos resulting from *emb/+* × Col WT crosses, for 49 genes tested. Full paternal allele activity corresponds to 0% observed mutants, while complete lack of paternal allele activity corresponds to 50% observed mutants. Genes are clustered based on the similarity of their complementation profiles. For each time point, the combined number of embryos examined in at least two biological replicates is listed in Supplementary Table 2. See Methods for additional detail on crosses.

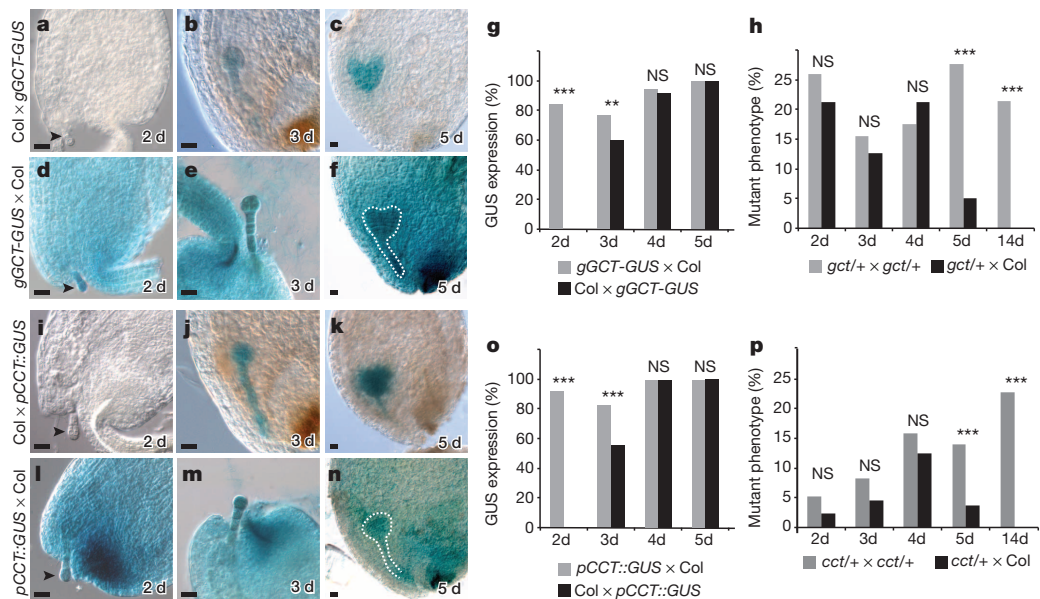


Figure 2 | Delayed molecular and functional activity for *GCT* and *CCT* paternal alleles in early embryogenesis. **a–g**, Expression of paternally and maternally derived *gGCT-GUS* at 2, 3, 4 and 5 days (d) after pollination. **h**, Quantification of phenotypically mutant embryos derived from hand-pollinated selfed *gct/+* plants, or crosses with WT pollen ($n > 90$ for all time points). **i–o**, Expression of paternally and maternally derived *pCCT-GUS* at 2, 3, 4 and 5 days after pollination. **p**, Quantification of phenotypically mutant embryos derived from hand-pollinated selfed *cct/+* plants, or from crosses with WT pollen. For all crosses, the maternal genotype is listed first. Embryos in

a, d, i and **l** are indicated with arrowheads. Embryos in **f** and **n** indicated with a dotted line. Scale bar, 20 μm. Results of a two-tailed Fisher's exact test, performed to determine whether the two samples were statistically different at each time point, are shown above the columns in **g, h, o** and **p**: *** $P < 0.001$; ** $P < 0.01$; NS, no significant difference between the two samples. For each time point, the combined number of embryos examined in at least two biological replicates, as well as exact P values for results of two-tailed Fisher's exact test, are listed in Supplementary Table 4. See Methods for additional details on crosses.

the corresponding β -glucuronidase (GUS) reporter lines reached full expression (Fig. 2g, o and Supplementary Table 4). Thus, molecular and functional data indicate that the WT paternal alleles for *GCT* and *CCT* show a gradual increase in activity in early embryogenesis.

Our data provide clear genetic evidence that most paternal alleles investigated in this study show gradual activation during the first few days of embryogenesis, whereas others are fully functional by 2 days after pollination. These findings are entirely consistent with studies of reporter gene expression in early isogenic embryos that found a gradual increase in transcription of most paternal alleles^{2,6,7}, as well as a smaller number of genes with equal maternal and paternal expression^{3,11}. Our results are also generally consistent with the overall conclusions derived from transcriptome profiling of embryos hybrid between the Landsberg *erecta* (*Ler*) and Columbia (*Col*) ecotypes, which predicted 65% of genes to have some paternal transcription at the two- to four-cell stage (with overall maternal bias)². However, our findings contrast with the equal maternal and paternal transcription discovered in early embryos hybrid between the *Col* and Cape Verde Islands (*Cvi*) ecotypes³. If early, equivalent zygotic transcription were the case in isogenic *Col*, we would expect to find no phenotypically mutant embryos resulting from *emb/ +* \times *Col* crosses for all of the 49 *EMB* genes we tested. In fact we found mutant phenotypes segregating for 40 genes, 19 of which were absent in egg and 29 in sperm cell transcriptomes²¹, demonstrating that loss of maternal contribution from the egg is not an overall explanation for lack of immediate complementation in *emb/ +* \times *Col* crosses (Supplementary Table 2). Forty-four of the 49 genes in our assay are also expressed in early endosperm (2–5 days after pollination)²² (Supplementary Table 5), although none of these genes show imprinted expression²³, and no allele-specific expression data exist, precluding any conclusion on the possibility of endosperm–embryo messenger RNA (mRNA) mobility having an effect on our complementation experiments²⁴.

On a gene-by-gene basis, comparison of our functional data for isogenic *Col* with transcriptional results for *Ler/Col*² and *Cvi/Col*³ hybrids revealed numerous discrepancies (Supplementary Table 6). One simple explanation for this divergence is that gene activity is indeed different between isogenic and hybrid embryos, and varies between hybrid combinations, as previously seen when comparing hybrid seedlings with their parental lines^{14–17}. To test this hypothesis, we crossed 11 different *emb/ +* lines (representing varying degrees of delayed paternal activity) with pollen from *Ler*, *Cvi*, the common laboratory strain Wassilewskija (*Ws*) and *Tsu* (an ecotype distantly related to *Col*)²⁵, and compared the resulting segregation ratios with those for *emb/ +* \times *Col* (Fig. 3a). For all *EMB* genes, we found significant differences between isogenic *Col* embryos and at least one hybrid combination, and for six genes we found significant differences between isogenic *Col* and three or more hybrids; in almost all instances, hybridization caused earlier gene activation compared with isogenic *Col* (Supplementary Table 3). In general, crosses with *Cvi* showed the most differences with isogenic *Col*, crosses with *Tsu* were least different, and crosses with *Ler* and *Ws* were intermediate. Importantly, directly comparing *Col/Ler* and *Col/Cvi* hybrids resulted in significant differences for 6 out of 11 genes. These trends were especially clear after principal component analysis of our data, in which PC1 separates *Col*, *Tsu* and *Ws* from *Ler* and *Cvi*, and PC2 further separates *Ler* and *Cvi* from *Col*, and from each other (Fig. 3b). Thus, significant variation in early paternal allele activity exists between isogenic *Col* embryos and embryos hybrid for *Col* and a second ecotype. In particular, direct comparison of *Col/Ler* and *Col/Cvi* hybrids revealed large differences, consistent with transcriptional profiling^{2,3}.

In summary, our data conclusively demonstrate that discrete, genome-wide zygotic activation does not occur in isogenic Columbia, the most commonly used ecotype of *Arabidopsis*. Instead, we show that gene activation varies, and individual contributions must be determined on a case-by-case basis. We have also uncovered an unexpectedly large effect of hybridization on gene activation in early embryogenesis, providing a simple explanation for different conclusions regarding zygotic genome activation derived from transcriptional profiling of different

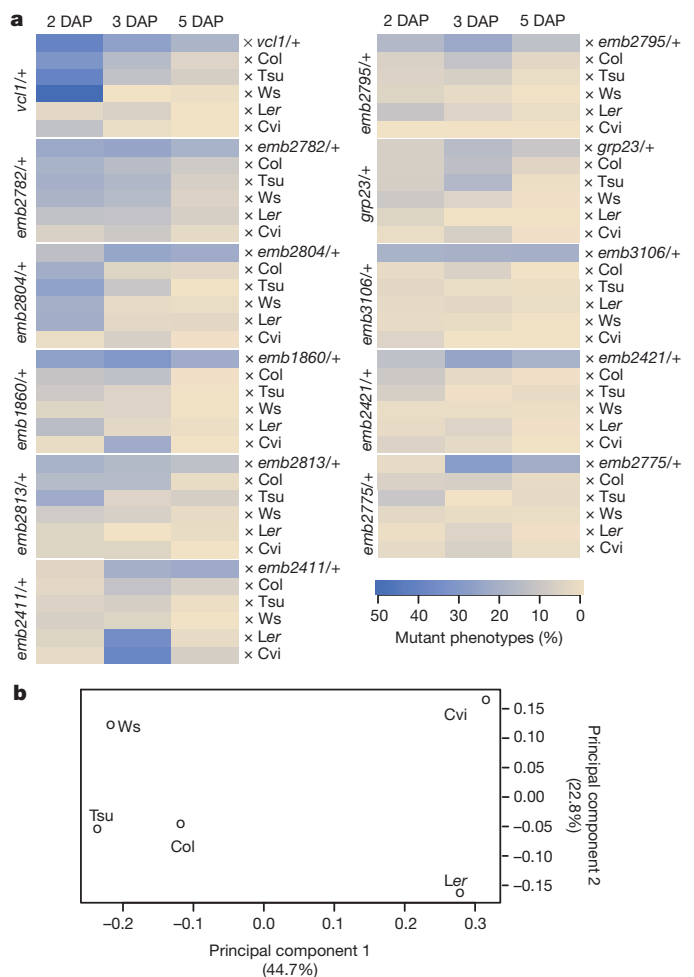


Figure 3 | Gene activation kinetics differ between isogenic and hybrid embryos. **a**, Heat map representations of mutant phenotype percentages observed for 11 different *emb/ +* mutants for self crosses ($\times emb/+$), and crosses to *Col* ($\times Col$), *Tsu* ($\times Tsu$), *Ws* ($\times Ws$), *Ler* ($\times Ler$) and *Cvi* ($\times Cvi$). **b**, Principal component analysis of mutant phenotype segregation ratios observed in *emb/ +* crosses to all WT ecotypes. For each time point, the combined number of embryos examined in at least two biological replicates is listed in Supplementary Table 3. See Methods for additional details on crosses. All emasculations and pollinations were done by hand. The *emb* mutants were in the *Col* ecotype. See Supplementary Table 3 for complete data and *P* values for the results of a two-tailed Fisher's exact test, performed to determine whether two samples were statistically different at each time point.

hybrid combinations^{2,3}. The varying effects on gene activation resulting from different combinations of inbred strains of *Arabidopsis* may be instructive for future research on the molecular basis of heterosis in seed crops.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 June 2013; accepted 27 June 2014.

Published online 7 September 2014.

1. Tadros, W. & Lipshitz, H. D. The maternal-to-zygotic transition: a play in two acts. *Development* **136**, 3033–3042 (2009).
2. Autran, D. *et al.* Maternal epigenetic pathways control parental contributions to *Arabidopsis* early embryogenesis. *Cell* **145**, 707–719 (2011).
3. Nodine, M. D. & Bartel, D. P. Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. *Nature* **482**, 94–97 (2012).
4. Grimanelli, D., Perotti, E., Ramirez, J. & Leblanc, O. Timing of the maternal-to-zygotic transition during early seed development in maize. *Plant Cell* **17**, 1061–1072 (2005).

5. Zhao, J. *et al.* Dynamic changes of transcript profiles after fertilization are associated with maternal elimination in tobacco zygote, and mark the onset of the maternal-to-zygotic transition. *Plant J.* **65**, 131–145 (2011).
6. Vielle-Calzada, J. P., Baskar, R. & Grossniklaus, U. Delayed activation of the paternal genome during seed development. *Nature* **404**, 91–94 (2000).
7. Baroux, C., Blanvillain, R. & Gallois, P. Paternally inherited transgenes are down-regulated but retain low activity during early embryogenesis in *Arabidopsis*. *FEBS Lett.* **509**, 11–16 (2001).
8. Pillot, M. *et al.* Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in *Arabidopsis*. *Plant Cell* **22**, 307–320 (2010).
9. Sprunck, S. *et al.* The transcript composition of egg cells changes significantly following fertilization in wheat (*Triticum aestivum* L.). *Plant J.* **41**, 660–672 (2005).
10. Meyer, S. & Scholten, S. Equivalent parental contribution to early plant zygotic development. *Curr. Biol.* **17**, 1686–1691 (2007).
11. Weijers, D., Geldner, N., Offringa, R. & Jürgens, G. Seed development: early paternal gene activity in *Arabidopsis*. *Nature* **414**, 709–710 (2001).
12. Lukowitz, W., Roeder, A., Parmenter, D. & Somerville, C. A MAPKK kinase gene regulates extra-embryonic cell fate in *Arabidopsis*. *Cell* **116**, 109–119 (2004).
13. Ueda, M., Zhang, Z. & Laux, T. Transcriptional activation of *Arabidopsis* axis patterning genes *WOX8/9* links zygote polarity to embryo development. *Dev. Cell* **20**, 264–270 (2011).
14. Barber, W. T. *et al.* Repeat associated small RNAs vary among parents and following hybridization in maize. *Proc. Natl Acad. Sci. USA* **109**, 10444–10449 (2012).
15. Greaves, I. K. *et al.* Trans chromosomal methylation in *Arabidopsis* hybrids. *Proc. Natl Acad. Sci. USA* **109**, 3570–3575 (2012).
16. Groszmann, M. *et al.* Changes in 24-nt siRNA levels in *Arabidopsis* hybrids suggest an epigenetic contribution to hybrid vigor. *Proc. Natl Acad. Sci. USA* **108**, 2617–2622 (2011).
17. Shen, H. *et al.* Genome-wide analysis of DNA methylation and gene expression changes in two *Arabidopsis* ecotypes and their reciprocal hybrids. *Plant Cell* **24**, 875–892 (2012).
18. Tzafirir, I. *et al.* Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiol.* **135**, 1206–1220 (2004).
19. Muralla, R., Lloyd, J. & Meinke, D. Molecular foundations of reproductive lethality in *Arabidopsis thaliana*. *PLoS ONE* **6**, e28398 (2011).
20. Gillmor, C. S. *et al.* The MED12–MED13 module of Mediator regulates the timing of embryo patterning in *Arabidopsis*. *Development* **137**, 113–122 (2010).
21. Wuest, S. E. *et al.* *Arabidopsis* female gametophyte gene expression map reveals similarities between plant and animal gametes. *Curr. Biol.* **20**, 506–512 (2010).
22. Belmonte, M. F. *et al.* Comprehensive developmental profiles of gene activity in regions and subregions of the *Arabidopsis* seed. *Proc. Natl Acad. Sci. USA* **110**, E435–E444 (2013).
23. Hsieh, T.-F. *et al.* Regulation of imprinted gene expression in *Arabidopsis* endosperm. *Proc. Natl Acad. Sci. USA* **108**, 1755–1762 (2011).
24. Costa, L. M. *et al.* Central cell-derived peptides regulate early embryo patterning in flowering plants. *Science* **344**, 168–172 (2014).
25. Nordborg, M. *et al.* The pattern of polymorphism in *Arabidopsis thaliana*. *PLoS Biol.* **3**, e196 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was facilitated by the efforts of D. Meinke and collaborators in creating and maintaining the SeedGenes Project (www.seedgenes.org). Seeds were obtained from the Arabidopsis Biological Resource Center (<http://abrc.osu.edu>). We thank S. Poethig, in whose laboratory C.S.G. generated GCT and CCT GUS reporter constructs, K. Gallagher for reagents, C. Silva-Ortega for preliminary experiments, A. de Luna, D. Lepe-Soltero and C. Abreu-Goodger for suggestions on data analysis and presentation, R. Datla for suggesting Col/Tsu hybrid, D. Bergmann for advice, and L. Herrera-Estrella, D. Autran, C. Baroux, J.-P. Vielle-Calzada and D. Grimaneli for reading the manuscript. Research was supported by a CONACyT graduate fellowship to G.D.T.-D.L. (No. 300802), a CONACyT postdoctoral fellowship to M.G.-A., a CONACyT research grant (Ciencia Básica No. 152333) to C.S.G., and by CINVESTAV institutional funds.

Author Contributions G.D.T.-D.L. performed phenotypic characterization of *emb* phenotypes and embryo complementation experiments with assistance from M.G.-A. and C.S.G.; C.S.G. made GUS reporter constructs; M.G.-A. performed GUS reporter experiments with assistance from G.D.T.-D.L.; G.D.T.-D.L., M.G.-A. and C.S.G. analysed data; G.D.T.-D.L. and C.S.G. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to C.S.G. (sgillmor@langebio.cinvestav.mx).

METHODS

Plant growth conditions. Plants were grown in growth chambers at $22 \pm 2^\circ\text{C}$ in 16/8-h light/dark cycles at constant humidity rate and $100 \mu\text{mol m}^{-2}$ fluence. Plants were grown in a mixture of vermiculite (GRACE MAN-FIN), perlite (AGROL125) and sunshine mix (PREMEZ FWSS3) (1:1:3 v/v/v). Seeds were sown directly on the soil, and given a 3-day cold treatment to promote germination, before being moved to growth chambers.

Crosses, histology and microscopy. Heterozygous plants were identified based on segregation of *emb* phenotypes. Crosses were performed using an *emb/+* plant as female parent and WT plants as the male parent. Heterozygous plants were emasculated, and pollinated 2 days later. The same ecotype was always used for both parents; with the exception of *yda-1* (*Le* ecotype), all mutants were in the Columbia ecotype. The seeds were cleared with a modified Hoyer's solution consisting of 7.5 g gum arabic, 100 g chloral hydrate, 5 ml glycerol and 30 ml water, diluted 1:1 with a dilution solution consisting of 7.5 g gum arabic, 5 ml glycerol and 30 ml water. Siliques were collected at 2, 3 and 5 days after pollination, and the ovules were cleared and observed as mentioned above. Embryos were observed on a Leica DM6000 B Nomarski microscope. For the time point at ~ 14 days after pollination, seed abortion was quantified macroscopically with a dissecting microscope. For each cross, we performed at least two biological replicates, meaning crosses were performed with at least two different *emb/+* plants as mothers. Our goal was to score at least 100 embryos for each genotype for each time point, although owing to occasional variability in seed set, for a small number of time points we scored fewer than 100 embryos. Genotypes for which there are large numbers of embryos scored (for example, \times Col reported in Fig. 3 and Supplementary Tables 2 and 3) are due to performing control crosses for multiple hybrid crosses.

GUS staining analysis. *pCCT::GUS* and *gGCT-GUS* plants were emasculated, and pollinated 2 days later using WT Col pollen. The reciprocal crosses were conducted the same way. Developing siliques at 2, 3, 4 and 5 days after pollination were dissected to expose the ovules. Ovules were then punctured with forceps and incubated in permissive GUS staining buffer (0.2% Triton X-100, 2 mM potassium ferrocyanide, 2 mM potassium ferricyanide, 2 mM X-Gluc in *N,N*-dimethylformamide (DMF) in 50 mM sodium phosphate buffer pH 7.2) overnight in a humid chamber at 37°C . The tissue was cleared in 50% glycerol and observed on a Leica DM6000 B microscope under Nomarski optics.

Construction of GCT and CCT GUS reporters. A pCAMBIA 3301 plasmid was modified by replacing the *GUS* gene with the *GUS*⁺ gene from pCAMBIA 1305.1, producing the 3301 *GUS*⁺ vector. The *gGCT-GUS* translational fusion of *GUS*⁺ to the

GCT protein was constructed by cloning the 900 base pair *GCT* 5' upstream region (from the *GCT* translational start up to the previous gene) into the 3301 *GUS*⁺ vector, in front of *GUS*⁺. Then, the entire *GCT* genomic region, as well as the 3' untranslated region corresponding to the At1g55325.2 transcript, was cloned in-frame downstream of *GUS*⁺, with a NAAIRS linker in between *GUS*⁺ and the *GCT* protein. Thus, this construct contained all immediate upstream, downstream and intron elements of the *GCT* gene. The *gGCT-GUS* construct was transformed into the Col ecotype, and more than 100 T1 plants expressing GUS were examined before choosing a representative line to use for detailed characterization. The expression pattern conferred by this construct exactly matches the native mRNA expression pattern observed for *GCT* during embryogenesis²⁰.

The 3301 *GUS*⁺ vector was modified by replacing *GUS*⁺ with a *GUS*⁺ that included the nuclear localization signal LQPKKRKRK from ref. 26, placed at the carboxy terminus of *GUS*⁺, to produce the 3301 *GUS*⁺NLS vector. The *pCCT::GUS* transcriptional reporter was constructed by cloning the 4.9 kilobase *CCT* upstream region (from the *CCT* translational start up to the previous gene) into the 3301 *GUS*⁺NLS vector, in front of *GUS*⁺NLS. Then, the *CCT* 3' untranslated region corresponding to the At4g00450.1 transcript was cloned after the *GUS*⁺NLS gene. Thus, this *CCT* transcriptional reporter contained all immediate 5' and 3' upstream and downstream regions corresponding to the *CCT* gene. The *pCCT::GUS* construct was transformed into the Col ecotype, and more than 100 *GUS*-positive T1 plants were examined before choosing a representative line to use for detailed characterization. The expression pattern conferred by this construct exactly matched the native mRNA expression pattern observed for *CCT* during embryogenesis²⁰.

Statistical analysis. To determine whether mutant phenotype segregation ratios differed significantly from 25% for selfed *emb/+* data (as presented in Supplementary Table 1), the Microsoft Excel CHITEST function was used. The two-tailed Fisher's exact test (www.langsrud.com/fisher.htm) was used to determine statistically significant differences for data in Supplementary Tables 2–4. Heat maps in Figs 1 and 3 were made in R with the heatmap.2 function of the data visualization package gplots 2.11.0; hierarchical clustering of rows was performed using the Ward method. Principal component analysis used the R package with the prcomp function. To illustrate the relationships among the variables (ecotypes) quantitatively, standard deviation percentages for each principal component were calculated as shown in Fig. 3b.

- Matsushita, T., Mochizuki, M. & Nagatani, A. Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. *Nature* **424**, 571–574 (2003).

Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function

Andrea Viale^{1,2*}, Piergiorgio Pettazoni^{1,2*}, Costas A. Lyssiotis³, Haoqiang Ying¹, Nora Sánchez^{1,2}, Matteo Marchesini^{1,2}, Alessandro Carugo^{1,2,4}, Tessa Green^{1,2}, Sahil Seth⁵, Virginia Giuliani⁵, Maria Kost-Alimova⁵, Florian Muller¹, Simona Colla¹, Luigi Nezi^{1,2}, Giannicola Genovese¹, Angela K. Deem¹, Avnish Kapoor¹, Wantong Yao^{1,2}, Emanuela Brunetto⁶, Ya'an Kang⁷, Min Yuan⁸, John M. Asara⁸, Y. Alan Wang¹, Timothy P. Heffernan⁵, Alec C. Kimmelman⁹, Huamin Wang¹⁰, Jason B. Fleming⁷, Lewis C. Cantley³, Ronald A. DePinho¹¹ & Giulio F. Draetta^{1,2}

Pancreatic ductal adenocarcinoma (PDAC) is one of the deadliest cancers in western countries, with a median survival of 6 months and an extremely low percentage of long-term surviving patients. KRAS mutations are known to be a driver event of PDAC¹, but targeting mutant KRAS has proved challenging². Targeting oncogene-driven signalling pathways is a clinically validated approach for several devastating diseases^{3,4}. Still, despite marked tumour shrinkage, the frequency of relapse indicates that a fraction of tumour cells survives shut down of oncogenic signalling^{5,6}. Here we explore the role of mutant KRAS in PDAC maintenance using a recently developed inducible mouse model of mutated *Kras*¹ (*Kras*^{G12D}, herein KRas) in a p53^{LoxP/WT} background. We demonstrate that a subpopulation of dormant tumour cells surviving oncogene ablation (surviving cells) and responsible for tumour relapse has features of cancer stem cells and relies on oxidative phosphorylation for survival. Transcriptomic and metabolic analyses of surviving cells reveal prominent expression of genes governing mitochondrial function, autophagy and lysosome activity, as well as a strong reliance on mitochondrial respiration and a decreased dependence on glycolysis for cellular energetics. Accordingly, surviving cells show high sensitivity to oxidative phosphorylation inhibitors, which can inhibit tumour recurrence. Our integrated analyses illuminate a therapeutic strategy of combined targeting of the KRAS pathway and mitochondrial respiration to manage pancreatic cancer.

Using our tetracycline-inducible KRas allele¹, we controlled the expression of KRas in a temporal and pancreas-specific manner. Upon doxycycline withdrawal, we observed regression of pancreatic tumours within 2–3 weeks followed by relapse after 4–5 months, suggesting that a fraction of tumour cells survived oncogene ablation. To investigate the impact of KRas ablation in detail, we transplanted cells from primary tumours subcutaneously into recipient mice fed with doxycycline. When tumours reached a diameter of 1 cm, doxycycline was withdrawn and the lesions rapidly and apparently completely regressed (Fig. 1a and Extended Data Fig. 1a). However, analysis of residual scars detected epithelial remnants embedded in fibrotic tissue (Fig. 1b and Extended Data Fig. 1b, c). This phenotype was confirmed *in vitro* using a three-dimensional culture system in which cells from primary lesions were grown as spheres in semi-solid medium. After doxycycline withdrawal (Extended Data Fig. 1d, e), tumour spheres underwent regression due to apoptosis (Extended Data Fig. 1f), and only a small population of dormant cells survived (Extended Data Fig. 1d, g). Notably, upon KRas re-activation, surviving cells (SCs) massively re-entered the cell cycle both *in vitro* and *in vivo* (Fig. 1c and Extended Data Fig. 1g, h) and rapidly reconstituted spheres and tumours,

suggesting that subpopulations of cells differently addicted to KRas co-exist in pancreatic tumours.

To assess the tumorigenic potential of SCs, we isolated KRas-expressing cells and SCs from tumour spheres (*ex vitro*) and tumours (*ex vivo*) and transplanted them in limiting dilution into recipient mice. Surprisingly, SCs were highly enriched in tumour-initiating cells (TICs). Virtually all SCs *ex vitro* initiated tumours in mice (TIC frequency $\gg 1:5$ versus 1:31 in KRas-expressing cells; $P < 0.001$) (Fig. 1d and Extended Data Fig. 2a), and TIC frequency was similarly enriched in SCs *ex vivo* (1:10 versus 1:100 in KRas-expressing cells) (Fig. 1d and Extended Data Fig. 2b). Then, to assess whether pharmacological ablation of oncogenic pathways could mimic the genetic suppression of KRas, we treated tumour spheres derived from a KRas constitutive mouse model⁷ with a combination of Mek1 (AZD8330) and dual PI3K/mTOR (BEZ235) inhibitors (Extended Data Fig. 2c). The treatment resulted in an enrichment of tumorigenic cells (TIC frequency 1:7 versus 1:47 for treated versus non-treated cells, respectively) (Fig. 1d and Extended Data Fig. 2d). Collectively, our data demonstrate that PDAC tumours are heterogeneous, and that a population of spherogenic and tumorigenic cells survives genetic and pharmacological ablation of oncogenic pathways.

To exclude the possibility that SCs represent a more aggressive subclone of tumour cells, we performed exome sequencing of tumour cells during cycles of KRas activation–inactivation–re-activation (ON–OFF–ON cycles) and evaluated changes in the allelic frequency of single nucleotide variants (SNVs), a hallmark of clonal selection. Mutational profiles did not show any significant modification in allelic frequencies before versus after ON–OFF–ON cycles (Fig. 1e and Extended Data Fig. 2e), demonstrating that tumours after KRas reactivation are genetically identical to their primary counterparts. While these data formally exclude genetic clonal selection among SCs, epigenetically driven clonal selection of a more aggressive subclone remains possible.

To characterize SCs further, we examined the expression of markers used to isolate cancer stem cells in human tumours^{8–10}. We found that different subpopulations of tumour cells were differentially sensitive to KRas ablation; specifically, only CD133⁺ CD44^{high} cells avoided undergoing massive apoptosis (Fig. 1f and Extended Data Fig. 1i). Consequently, tumour remnants *in vivo* are strongly positive for stem cell markers (Fig. 1g, h and Extended Data Fig. 2f, g). Together, the tumorigenicity and immunophenotypic similarity between SCs and previously identified human pancreatic cancer stem cells^{8–10} suggests that SCs may possess cancer stem cell characteristics.

We next performed a transcriptomic analysis of cells isolated from tumour spheres. Gene set enrichment analysis (GSEA) revealed significant

¹Department of Genomic Medicine, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ²Department of Molecular and Cellular Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ³Department of Medicine, Weill Cornell Medical College, New York, New York 10065, USA. ⁴Department of Experimental Oncology, European Institute of Oncology, Milan 20139, Italy. ⁵Institute for Applied Cancer Science, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ⁶Pathology Unit, San Raffaele Scientific Institute, Milan 20132, Italy. ⁷Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ⁸Department of Medicine, Division of Signal Transduction, Beth Israel Deaconess Medical Center, Boston, Massachusetts 02115, USA. ⁹Department of Radiation Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts 02215, USA. ¹⁰Department of Pathology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA. ¹¹Department of Cancer Biology, The University of Texas MD Anderson Cancer Center, Houston, Texas 77030, USA.

*These authors contributed equally to this work.

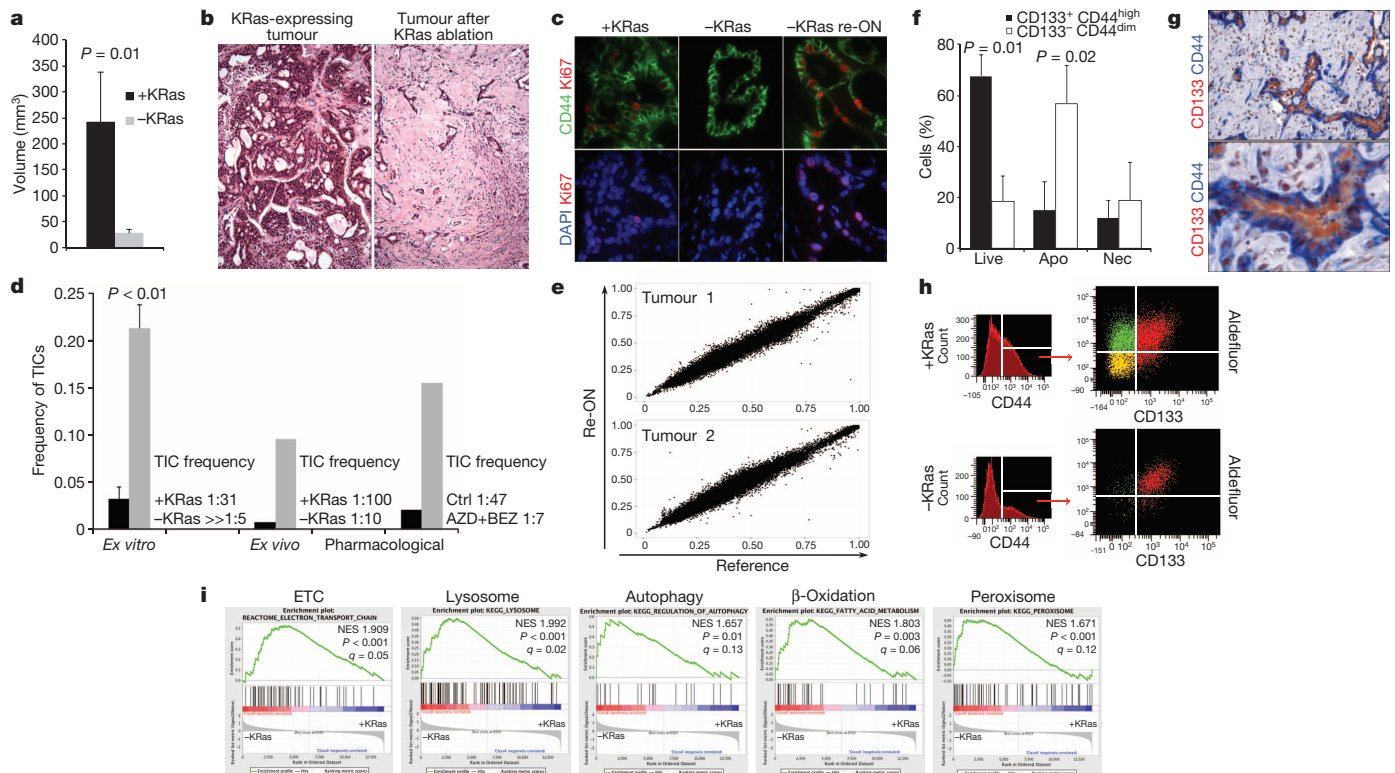


Figure 1 | Cells surviving oncogene ablation are enriched in tumorigenic cells. **a**, Tumour volume before/after KRas ablation (+/−KRas) ($n = 6$). **b**, Histology depicting tumour remnants ($\times 10$). **c**, Immunofluorescence of KRas-expressing tumour (+KRas), regressed tumour (−KRas) and regressed tumours 48 h after KRas re-activation (−KRas re-ON) for Ki67 (red), CD44 (green) and 4',6-diamidino-2-phenylindole (DAPI; blue) ($\times 40$). **d**, Limiting dilution transplantation, TIC frequency. Genetic model: +KRas (black) versus −KRas (grey) *ex vitro* ($n = 4$) or *ex vivo* ($n = 2$). Pharmacological downregulation: control (black) versus treated spheres (grey; AZD8330 plus BEZ235 (AZD+BEZ)) ($n = 2$). Ctrl, control. **e**, Exome sequencing: allele

frequencies after KRas re-activation in SCs (Re-ON) versus KRas-expressing cells (Reference) at 40,383 and 44,182 SNVs for two independent tumours. **f**, Annexin V in spheres +/−KRas with respect to CD44/CD133 expression. Percentage of live, apoptotic (Apo) and necrotic (Nec) cells are reported ($n = 3$). **g**, Immunohistochemistry of −KRas tumours for CD44 (blue) and CD133 (red) ($\times 20$ –40). **h**, Immunophenotyping of +/−KRas tumours for CD44/CD133/aldefluor. **i**, GSEA of pathways enriched in −KRas versus +KRas cells. Normalized enriched scores (NES) are reported. Data are mean \pm standard deviation (s.d.).

enrichment of genes involved in several metabolic pathways (for example, mitochondrial electron transport chain (ETC), lysosome activity, autophagy, mitochondrial and peroxisomal β -oxidation) (Fig. 1i and Extended Data Fig. 3a–e), which suggested that SCs might have increased mitochondrial activity. Indeed, *Ppargc1a* (PGC1a), a key regulator of mitochondrial biogenesis¹¹, was increased at the messenger RNA and protein levels in SCs (Fig. 2a and Extended Data Fig. 4a), and we detected PGC1a accumulation in the nuclei of SCs *in vivo* (Fig. 2c). Furthermore, SCs from tumour spheres stained intensely for MitoTracker Green, a marker of mitochondrial mass (Extended Data Fig. 4b). These data were corroborated by increased expression of the mitochondrial marker VDAC1 in SCs *in vitro* and *in vivo* (Fig. 2b, d).

We functionally validated our findings by measuring respiratory capacity. SCs had a fourfold increase in oxygen consumption rate (OCR) compared with KRas-expressing cells (118 versus 33 pmol min^{−1}, respectively; $P = 0.001$) (Fig. 2e). Likewise, both *in vivo* and *in vitro*, mitochondria of SCs that were either genetically or pharmacologically selected were more hyperpolarized (Fig. 2f and Extended Data Fig. 4c–f) and generated more reactive oxygen species (ROS) (Extended Data Fig. 4g, h), indicating a more active ETC. We also detected morphological differences in mitochondria using transmission electron microscopy (TEM) (Fig. 2g). Because transmembrane mitochondrial potential regulates the mitochondrial permeability transition pore (lower potential = lower threshold for pore opening) and because cells positive for stem cell markers in KRas-expressing tumours have hyperpolarized mitochondria (Extended Data Fig. 4i, j), the higher mitochondrial potential *per se* could explain why cells positive for stem cell markers are less prone to KRas ablation-mediated

apoptosis. Altogether, our data suggest that altered metabolic and mitochondrial functions are critical features of SCs.

SCs and KRas-expressing cells were next exposed to mitochondrial stressors. Treatment with oligomycin, a Fo-ATPase inhibitor of Complex V, significantly reduced mitochondrial respiration in both populations (Extended Data Fig. 5a). Conversely, normalization to basal respiration revealed different effects of the uncoupler carbonyl cyanide-*p*-trifluoromethoxyphenylhydrazone (FCCP) (Extended Data Fig. 5b), indicating that the mitochondria of SCs function near their maximal rate and have minimal spare respiratory capacity. Despite similar overall responses to Complex V inhibition, ATP levels of SCs were dramatically reduced upon oligomycin treatment compared to KRas-expressing cells (Fig. 2h), suggesting a deficit in energy compensatory mechanisms. Indeed, KRas-expressing cells exposed to oligomycin strongly upregulated their extracellular acidification rate (ECAR) and lactate production, a compensation that did not occur in SCs (Fig. 2i and Extended Data Fig. 5c, e), confirming that SCs failed to increase glycolysis after oxidative phosphorylation (OXPHOS) inhibition in a manner sufficient to rescue the defects in ATP production.

To assess comprehensively differences in metabolism, we performed a metabolomic analysis using a liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS)-based platform^{11,12}, revealing that several metabolic pathways were deregulated in SCs (Extended Data Fig. 5d). Consistent with the above findings, glycolytic intermediates downstream of phospho-fructose kinase (PFK) were significantly less abundant in SCs versus KRas-expressing cells (Fig. 2j). The impaired glycolysis of SCs was independently validated by measuring glucose

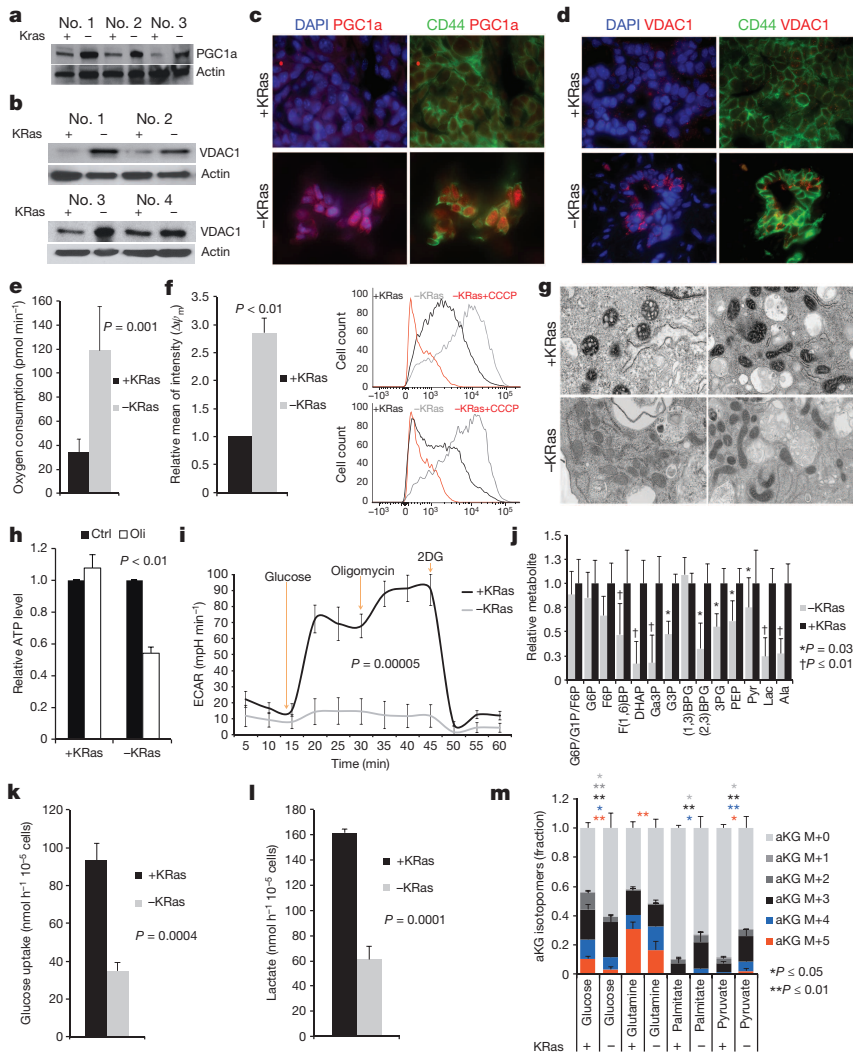


Figure 2 | Surviving cells have more active mitochondria and impaired glycolysis.

a, b, Immunoblot of +/- KRas cells derived from independent tumours probed with PGC1a (**a**) and VDAC1 (**b**) antibodies. Numbers indicate different independent tumours. **c, d**, *In vivo* immunofluorescence for CD44 (green) and PGC1a (red) (**c**), and CD44 and VDAC1 (red) (**d**) in +/- KRas tumours ($\times 60$). **e**, Oxygen consumption of +/- KRas cells ($n = 7$). **f**, *In vivo* mitochondrial potential of +/- KRas tumours ($n = 3$); representative flow cytometry of two tumours, as control uncoupler CCCP was added to acquired -KRas. **g**, TEM micrographs representing mitochondrial morphology ($\times 25,000$). **h**, ATP levels of +/- KRas cells in response to oligomycin (Oli) or dimethylsulphoxide (DMSO; Ctrl) ($n = 4$). **i**, ECAR response of +/- KRas cells to glucose, oligomycin and 2-deoxy-D-glucose (2DG) ($n = 4$). **j**, Fold change of glycolytic intermediates in +/- KRas cells ($n = 4$). Ala, alanine; BPG, bisphosphoglycerate; DHAP, dihydroxyacetone phosphate; F6P, fructose-6-phosphate; F(1,6)BP, fructose-1,6-bisphosphate; G6P, glucose-6-phosphate; G1P, glucose-1-phosphate; Ga3P, glyceraldehyde-3-phosphate; G3P, glycerol-3-phosphate; Lac, lactate; PEP, phosphoenolpyruvate; 3PG, 3-phosphoglycerate; Pyr, pyruvate. **k, l**, Glucose uptake (**k**) and lactate production (**l**) of +/- KRas cells ($n = 3$). **m**, Isotopomer distribution for α -ketoglutarate (aKG) following steady-state tracing with uniformly carbon-13-labelled substrates ($n = 3$). Data are mean \pm s.d.

uptake and lactate production *in vitro* (Fig. 2k, l and Extended Data Fig. 6a) and *in vivo* (Extended Data Fig. 6b). The ratio between lactate and glucose in spent media supports this difference even more strongly (concentration of lactate/concentration of glucose: KRas-expressing cells, 16.9; SCs, 0.9). Importantly, SCs surviving pharmacological ablation of KRas showed a similar phenotype (Extended Data Fig. 6c, d).

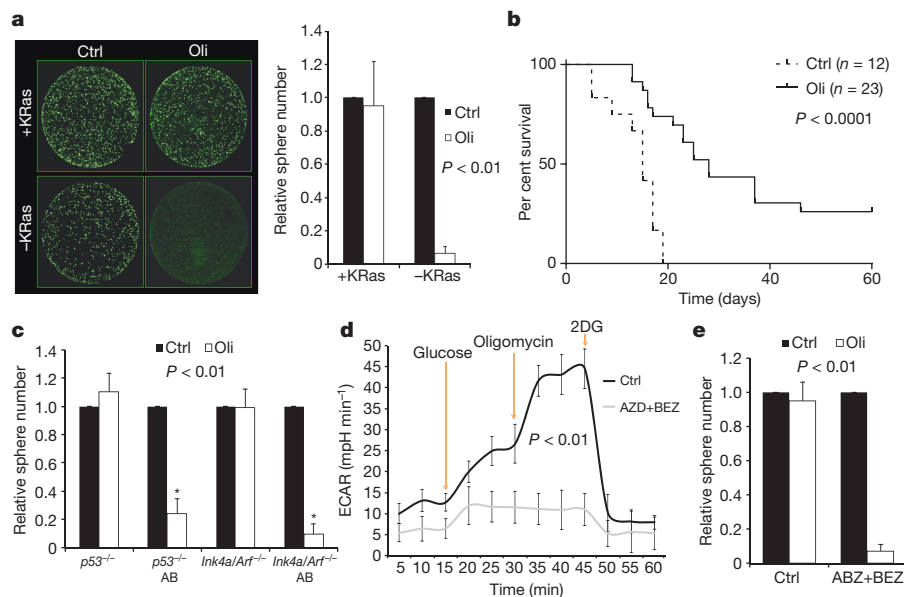
We also detected a decreased abundance of metabolic intermediates specific to the tricarboxylic acid (TCA) cycle in SCs (Extended Data Fig. 5f). We used carbon-13-labelled primary metabolic substrates to trace their contribution to central carbon metabolism. After 36 h labelling, SCs relied less on glucose and glutamine and more on pyruvate and palmitate to generate TCA cycle intermediates and branching metabolites (Fig. 2m and Extended Data Fig. 7e–h). This is consistent with reports describing the mutant KRas-mediated activation of anabolic glucose and glutamine metabolism in PDAC^{1,13–15}. It is also worth noting that SCs had lower levels of high-energy metabolites (Extended Data Fig. 5g) (compatible with less anabolic, dormant cells) and increased total glutathione (Extended Data Figs 5g and 8h). Importantly, any effect of cell cycle fluctuations on metabolic processes was excluded by our experimental design, as comparisons were made between quiescent SCs and quiescent fully formed KRas-expressing spheres (Extended Data Fig. 1g). In fact, we demonstrated that sphere formation is a dynamic and regulated process in which tumour cells expressing KRas exit the cell cycle when tumour spheres are fully formed. Thus, observed metabolic alterations can appropriately be attributed to an autonomous metabolic program.

The lower energy levels and a lack of energetic compensatory mechanisms in response to mitochondrial stressors in SCs suggested that

treatment with OXPHOS inhibitors might impact their survival. As expected, even a short, 24 h exposure to oligomycin completely abrogated the ability of SCs to reform tumour spheres upon KRas re-expression, whereas KRas-expressing cells were unaffected (Fig. 3a and Extended Data Fig. 8a). Similar effects were observed with other OXPHOS inhibitors, although with less dramatic effects relative to Complex V inhibitors (Extended Data Fig. 8b). To determine whether OXPHOS inhibition synergized with KRas ablation *in vivo*, we induced tumour regression for 2 weeks via doxycycline withdrawal, treated mice with oligomycin (0.5 mg kg⁻¹ daily, intraperitoneal (i.p.)) or vehicle, and subsequently re-induced KRas and tumour relapse (Extended Data Fig. 8c–e). While mice treated with vehicle relapsed immediately (median survival 15 days, maximal survival 19 days), 25% of oligomycin-treated mice survived >60 days ($P < 0.0001$) (Fig. 3b and Extended Data Fig. 8f). Oligomycin treatment also strongly abrogated the spherogenic potential of SCs treated with AZD8330 plus BEZ235 independently of their p53 status (Fig. 3c). Because OXPHOS inhibitors induce ROS¹⁶ and SCs have elevated basal mitochondrial ROS levels (Extended Data Fig. 4g and Extended Data Fig. 8g), we excluded the possibility that the observed synthetic lethality of oligomycin was caused by oxidative stress (Extended Data Fig. 8h–k).

To explore the therapeutic potential of inhibiting OXPHOS in human tumours, we derived spheres from early passage patient-derived xenografts¹⁷. Similar to the mouse model, combination treatment with AZD8330 and BEZ235 for 1 week strongly decreased sphere formation (Extended Data Fig. 9a–c), sparing cells endowed with spherogenic potential, increased mitochondrial potential (Extended Data Fig. 9d) and, most importantly, unable to upregulate ECAR in response to OXPHOS inhibition (Fig. 3d). As with mouse cells, short exposure of human SCs to oligomycin

Figure 3 | OXPPOS inhibition specifically targets surviving cells. **a**, Effect of oligomycin (Oli) and DMSO (Ctrl) on spherogenic potential of +/-KRas cells ($n = 8$); representative calcein staining. **b**, Kaplan–Meier overall survival after KRas reactivation in mice bearing regressed tumours treated for 2 weeks with oligomycin or vehicle. **c**, Effects of oligomycin and DMSO on spherogenic potential of cells treated (AB) or not with AZD8330 plus BEZ235 in $p53^{-/-}$ and $Ink4a/Arf^{-/-}$ backgrounds ($n = 3$). **d**, ECAR response of human tumour spheres treated (AZD+BEZ) and untreated (Ctrl) to glucose, oligomycin and 2DG ($n = 3$). **e**, Effect of oligomycin and DMSO on spherogenic potential of human treated and untreated tumour cells ($n = 3$). Data are mean \pm s.d.



dramatically decreased their spherogenic potential (Fig. 3e). Similarly, transient genetic downregulation of TFAM (a key regulator of mitochondrial transcription¹⁸) and TUFM (a GTPase that delivers aminoacylated transfer RNAs to mitochondrial ribosomes¹⁹) decreased the spherogenic potential of human cells surviving combined AZD8330 plus BEZ235 treatment (Extended Data Fig. 9e–h). Altogether, these data highlight mitochondrial respiration as an attractive, druggable target that may effectively eradicate SCs in PDAC.

The low energy state of SCs suggested that these cells may be experiencing metabolic stress. We thus measured AMPK phosphorylation at

Thr 172, an established indicator of metabolic stress. Surprisingly, AMPK phosphorylation was lower in SCs versus KRas-expressing cells (Fig. 4a); however, marked induction of AMPK phosphorylation was observed specifically in SCs upon oligomycin treatment, stressing again the importance of ETC activity for their energy production (Fig. 4b and Extended Data Fig. 9i).

To elucidate further the metabolic mechanisms active in SCs, we investigated other targets identified by transcriptomic analysis, including autophagy and β -oxidation. Autophagy and lysosome activity, two of the most activated pathways on the basis of GSEA (Fig. 1i), have an important

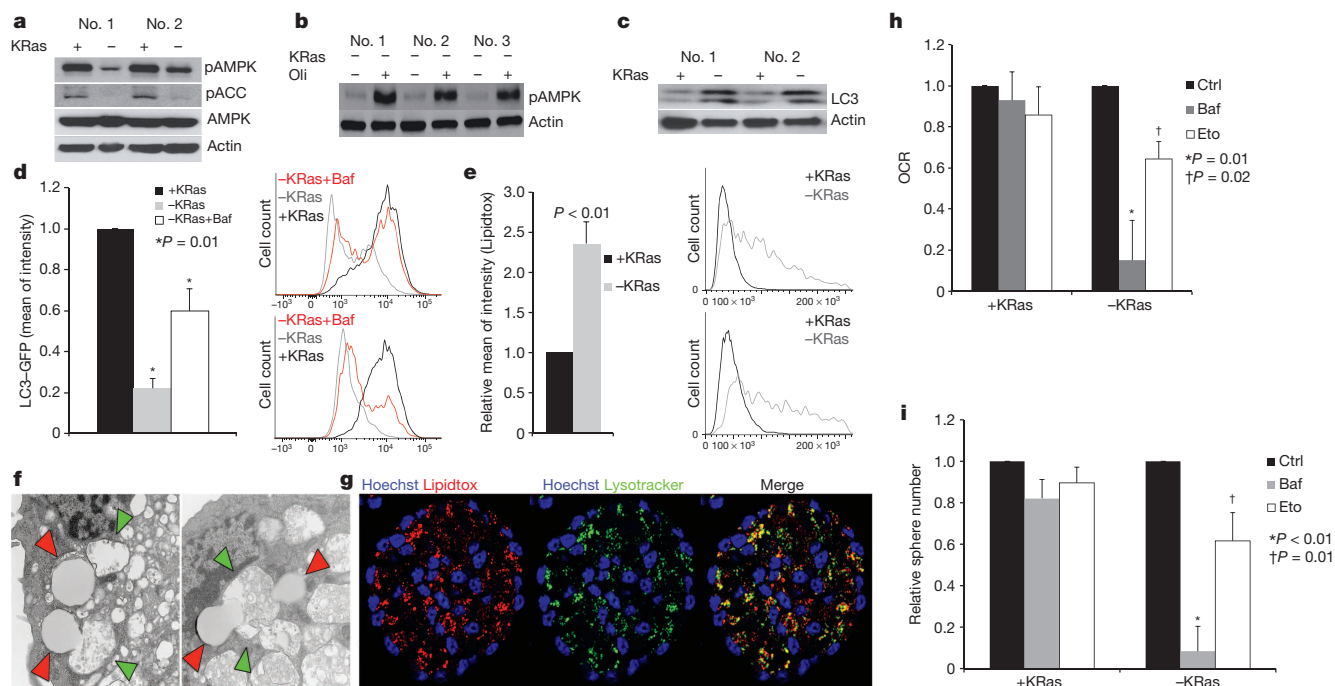


Figure 4 | Surviving cells are not in metabolic stress and activate autophagy. **a**, Immunoblot of Thr-172-phosphorylated and total AMPK, Ser-79-phospho-acetyl-coA-carboxylase (pACC) and β -actin in +/-KRas cells for two independent tumours. **b**, Immunoblot of -KRas cells treated or not with oligomycin (Oli) probed with anti-Thr-172-phospho-AMPK and β -actin antibodies for three independent tumours. **c**, Immunoblot of LC3 and β -actin in +/-KRas cells for two independent tumours. **d**, Autophagic flux of +/-KRas cells stably expressing GFP–LC3. Bafilomycin treatment rescues GFP (–KRas+Baf) ($n = 3$); representative flow cytometry of two tumours.

e, Lipid droplet quantification in +/-KRas cells ($n = 4$); representative flow cytometry of two tumours. **f**, Fusion between lipid droplets (red arrowheads) and autophagosomes (green arrowheads) in -KRas cells (TEM; $\times 25,000$). **g**, Confocal microscopy for lipid droplets (Lipidtoxi; red) and lysosomes (Lysotracker; green) in -KRas spheres. Hoechst, blue. **h**, Oxygen consumption of +/-KRas cells pre-treated with bafilomycin (Baf), etomoxir (Eto) or vehicle (Ctrl) ($n = 4$). **i**, Effect of bafilomycin and etomoxir on spherogenic potential of +/-KRas cells ($n = 6$). Data are mean \pm s.d.

role in the metabolism of SCs. The autophagic marker microtubule-associated protein light-chain 3 (LC3) was abundantly expressed in SCs (Fig. 4c), and lysosomes and autophagosomes at various stages of evolution were abundant (Extended Data Fig. 10a). To investigate the activation of autophagic flux, we transduced KRas-expressing tumour spheres with LC3–green fluorescent protein (GFP) fusion protein. Cells stably expressing LC3–GFP were maintained in the presence or absence of doxycycline and subsequently analysed by flow cytometry. As shown in Fig. 4d, SCs had a dramatically reduced GFP signal (80%) compared with KRas-expressing cells, indicating that GFP was quenched upon autolysosome formation, as also confirmed by the treatment with bafilomycin, a vacuolar H⁺-ATPase inhibitor (Fig. 4d).

Upon TEM analysis, we observed increased content of cytoplasmic lipid droplets in SCs (Extended Data Fig. 10b) that we quantified with neutral lipid-specific dyes (Fig. 4e). Importantly, and unique to SCs, we observed unambiguous evidence of microlipophagy²⁰—the fusion of lipid droplets with autophagosomes (Fig. 4f)—which we confirmed by demonstrating strong co-localization of lipid droplets and lysosomes upon staining in SCs (Fig. 4g). Consistent with observations that SCs catabolize more endogenous protein and fatty acid, tracing metabolic experiments illustrate that more carbon in central metabolism is unaccounted for from labelled substrates in SCs versus KRas-expressing cells (Extended Data Fig. 7e–h). To determine whether autophagy and microlipophagy may represent a strategy for SCs to maintain their energetic balance^{21,22}, we explored their contribution in maintaining mitochondrial activity by measuring OCR in cells pre-treated with autophagic and β -oxidation inhibitors. Neither inhibitor affected OCR in KRas-expressing cells, whereas treatment with either bafilomycin or etomoxir both markedly reduced OCR in SCs (85% and 35%, respectively) (Fig. 4h). Importantly, both inhibitors also induced metabolic stress specifically in SCs (Extended Data Fig. 9j, k) and dramatically affected both their spherogenic potential (decrease of >90% and 39%, respectively; $P \leq 0.01$) (Fig. 4i) and survival (Extended Data Fig. 9l). We therefore conclude that autophagy and microlipophagy are as critical as ETC activity for the survival of SCs.

We isolated and characterized tumour cells surviving the genetic and pharmacological ablation of oncogenic pathways, demonstrating that they have features of TICs and a distinct metabolic profile. Unlike highly glycolytic tumour cells, which are dependent on lactate production to regenerate NAD⁺ to support continued glycolysis for their anabolic metabolism (Warburg effect), SCs are less glycolytic and more dependent on mitochondrial respiration. Their dependence on OXPHOS for energy production is consistent with recent reports showing that both normal and leukaemic stem cells rely on mitochondrial respiration^{19,23–25}. Importantly, the strong dependence of ETC activity of SCs on autophagic and catabolic processes makes them more resistant to nutrient deprivation and environmental stressors. However, their inability to increase compensatory fluxes (namely, glycolysis) following OXPHOS inhibition confers extreme sensitivity to inhibition of mitochondrial function, which dramatically impacts their spherogenic and tumorigenic potential. Our findings hold important implications for tumour treatment, paving the way for targeting OXPHOS in association with oncogenic pathway inhibitors to eradicate SCs and prevent tumour relapse in pancreatic cancer.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 3 May 2013; accepted 19 June 2014.

Published online 10 August 2014.

1. Ying, H. *et al.* Oncogenic Kras maintains pancreatic tumors through regulation of anabolic glucose metabolism. *Cell* **149**, 656–670 (2012).
2. Karmoub, A. E. & Weinberg, R. A. Ras oncogenes: split personalities. *Nature Rev. Mol. Cell Biol.* **9**, 517–531 (2008).
3. Kantarjian, H. *et al.* Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. *N. Engl. J. Med.* **346**, 645–652 (2002).
4. Flaherty, K. T. *et al.* Inhibition of mutated, activated BRAF in metastatic melanoma. *N. Engl. J. Med.* **363**, 809–819 (2010).
5. Quintas-Cardama, A., Kantarjian, H. & Cortes, J. Imatinib and beyond—exploring the full potential of targeted therapy for CML. *Nature Rev. Clin. Oncol.* **6**, 535–543 (2009).
6. Jang, S. & Atkins, M. B. Which drug, and when, for patients with BRAF-mutant melanoma? *Lancet Oncol.* **14**, e60–e69 (2013).
7. Bardeesy, N. *et al.* Both p16^{Ink4a} and the p19^{Arf}-p53 pathway constrain progression of pancreatic adenocarcinoma in the mouse. *Proc. Natl Acad. Sci. USA* **103**, 5947–5952 (2006).
8. Hermann, P. C. *et al.* Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* **1**, 313–323 (2007).
9. Li, C. *et al.* Identification of pancreatic cancer stem cells. *Cancer Res.* **67**, 1030–1037 (2007).
10. Kim, M. P. *et al.* ALDH activity selectively defines an enhanced tumor-initiating cell population relative to CD133 expression in human pancreatic adenocarcinoma. *PLoS ONE* **6**, e20636 (2011).
11. Fernandez-Marcos, P. J. & Auwerx, J. Regulation of PGC-1 α , a nodal regulator of mitochondrial biogenesis. *Am. J. Clin. Nutr.* **93**, 884S–890S (2011).
12. Yuan, M., Breitkopf, S. B., Yang, X. & Asara, J. M. A positive/negative ion-switching, targeted mass spectrometry-based metabolomics platform for bodily fluids, cells, and fresh and fixed tissue. *Nature Protocols* **7**, 872–881 (2012).
13. Gaglio, D. *et al.* Oncogenic K-Ras decouples glucose and glutamine metabolism to support cancer cell growth. *Mol. Syst. Biol.* **7**, 523 (2011).
14. Son, J. *et al.* Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* **496**, 101–105 (2013).
15. Vizán, P. *et al.* K-ras codon-specific mutations produce distinctive metabolic phenotypes in NIH3T3 mice fibroblasts. *Cancer Res.* **65**, 5512–5515 (2005).
16. Liu, Y. & Schubert, D. R. The specificity of neuroprotection by antioxidants. *J. Biomed. Sci.* **16**, 98 (2009).
17. Kim, M. P. *et al.* Generation of orthotopic and heterotopic human pancreatic cancer xenografts in immunodeficient mice. *Nature Protocols* **4**, 1670–1680 (2009).
18. Shi, Y. *et al.* Mammalian transcription factor A is a core component of the mitochondrial transcription machinery. *Proc. Natl Acad. Sci. USA* **109**, 16510–16515 (2012).
19. Škrtić, M. *et al.* Inhibition of mitochondrial translation as a therapeutic strategy for human acute myeloid leukemia. *Cancer Cell* **20**, 674–688 (2011).
20. Singh, R. *et al.* Autophagy regulates lipid metabolism. *Nature* **458**, 1131–1135 (2009).
21. Guo, J. Y. *et al.* Activated Ras requires autophagy to maintain oxidative metabolism and tumorigenesis. *Genes Dev.* **25**, 460–470 (2011).
22. Yang, S. *et al.* Pancreatic cancers require autophagy for tumor growth. *Genes Dev.* **25**, 717–729 (2011).
23. Nakada, D., Saunders, T. L. & Morrison, S. J. Lkb1 regulates cell cycle and energy metabolism in haematopoietic stem cells. *Nature* **468**, 653–658 (2010).
24. Lagadinou, E. D. *et al.* BCL-2 inhibition targets oxidative phosphorylation and selectively eradicates quiescent human leukemia stem cells. *Cell Stem Cell* **12**, 329–341 (2013).
25. Samudio, I. *et al.* Pharmacologic inhibition of fatty acid oxidation sensitizes human leukemia cells to apoptosis induction. *J. Clin. Invest.* **120**, 142–156 (2010).

Acknowledgements We thank A. Divakaruni, J. Dunn, C. Smith, K. McGirr and D. Ferrick for their support with the Seahorse Bioscience XF96 Analyser; T. Tieu for vector cloning and J. Kovacs for support with the YSI analyser; J. D. Lechleiter for the protocol to measure mitochondrial potential *in vivo*; H. Sandoval, C. Tacchetti, M. E. Di Francesco, J. Marszalek and P. Jones for discussions and suggestions; K. Dunner Jr and the High Resolution Electron Microscopy Facility at the MD Anderson Cancer Center (MDACC) for TEM (Cancer Center Core Grant CA16672); W. N. Hittelman and the Center for Targeted Therapy for confocal microscopy; the Dana-Farber Cancer Institute Microarray Core Facility for Affymetrix expression profiling and the MDACC Sequencing and Microarray Facility (SMF) funded by National Cancer Institute (NCI) grant CA016672 (SMF) for exome sequencing; the MDACC Flow Cytometry and Cellular Imaging Core Facility supported by grant NCI#P30CA16672 for flow cytometers and FACS; D. Jayanta for providing GFP–LC3 constructs; B. Perrazzona, U. Varadarajan and R. Dewan for lab management; and S. Jiang for assistance in maintenance of mouse colonies. A.V. is thankful to A. Fantino, S. Rapi, V. Giuliani and P. Viale for their continuous support. This study was supported by grants from the Hirshberg Foundation for Pancreatic Cancer Research to A.V., Harvard Stem Cell Institute to R.A.D. and A.V., Sheikh Ahmed Center for Pancreatic Cancer Research to G.F.D., T.P.H. and A.V., American Italian Cancer Foundation to G.F.D., National Institutes of Health (NIH) P01CA117969 to R.A.D., NIH/NCI P01CA120964 to J.M.A., The Viragh Family Foundation to J.B.F.; C.A.L. is a Pancreatic Cancer Action Network-AACR Pathway to Leadership Fellow.

Author Contributions A.V., P.G., R.A.D. and G.F.D. designed the studies, interpreted the data and wrote the manuscript; A.V., P.G., H.Y., N.S., M.M., A.C., T.G. and V.G. performed the experiments; C.A.L. was responsible for metabolomics and carbon-13 tracing experiments; S.S. was responsible for CNV and bioinformatics analysis; M.K.-A., F.M., S.C., L.N., G.G., A.K.D., A.K., W.Y., E.B., Y.K., T.P.H., A.C.K., H.W. and J.B.F. contributed essential reagents and resources; M.Y. and J.M.A. helped with the metabolomics analysis; F.M., Y.A.W. and L.C.C. assisted with data interpretation; A.K.D. edited the manuscript.

Author Information Complete transcriptomic profiles are available in the Gene Expression Omnibus under accession number GSE58307. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.V. (aviale@mdanderson.org) or G.F.D. (gdraetta@mdanderson.org).

METHODS

Mice. The inducible tetO-LSL-KRas^{G12D}, ROSA26-LSL-rtTA, p53^{LoxP/WT}, p48-Cre mouse model was generated as described¹. After weaning, mice were fed doxycycline (2 g l⁻¹) in drinking water supplemented with sucrose (20 g l⁻¹) and monitored for tumour development. Constitutive LSL-KRas^{G12D}, p53^{LoxP/LoxP}, Pdx1-Cre and LSL-KRas^{G12D}, Ink4a/Arf^{LoxP/LoxP}, Pdx1-Cre mice were generated as previously reported⁷. All animals were kept in FVB/C57Bl6 mix or pure FVB backgrounds in a pathogen-free facility at the University of Texas MD Anderson Cancer Center (MDACC). All manipulations were performed under Institutional Animal Care and Use Committee (IACUC)-approved protocols.

Antibodies, plasmids and chemical reagents. Primary antibodies used for flow cytometry, immunofluorescence, immunohistochemistry and immunoblotting were: CD44 (eBioscience, IM7), CD133 (eBioscience, 13A4; Abcam, 165118), phospho-p44/42 (Cell Signaling, D10), VDACL1 (Abcam, ab15895), PGC1 α (Abcam, ab54481), BrdU (Abcam, ab82421), Ras (05-516, Millipore), LC3A/B (Cell Signaling, 4108), phospho-AMPK α (Thr 172) (Cell Signaling, 40H9), AMPK α (Cell Signaling, D63G4), TFAM (Cell Signaling, D5C8), TUFM (Abcam, ab175199), β -actin (sc-1615, Santa Cruz). Plasmids: pBABE-puro GFP-LC3 (ref. 26) (Addgene plasmid 22405), pLKO TFAM shRNAs (TRCN000016093, 5'-CCGGCAGTGTATATTGATCCAGAACTCGAGTTCTGGATCAATATACTCACGTTTTT-3'; TRCN0000016095, 5'-CCGGTAAGTTCTTACCTTCGATTTCTCGAAGAACTCAAGTAAGAAGTACTTTTT-3') and pLKO TUFM shRNAs (TRCN0000280863, 5'-CCGGGCTTACCCGAGTTGGCTATACTCGAGTTATAGCCAACTCGTTCGAGTTG-3'; TRCN0000280864, 5'-CCGGGAGGACCTGAAGTTCAACTACTCGAGTGTGAACTTCAGGTCCTCTTTTTG-3') (Sigma), Tet-inducible pLKO was provided by the Institute for Applied Cancer Science at MDACC. Chemical reagents: oligomycin A (Sigma), buthioninesulphoximine (Sigma), 4-hydroxynonenal (Cayman Chemical), α -tocopherol (Sigma), N-acetylcysteine (Sigma), tetrakis (Calbiochem), bafilomycin (Sigma), etomoxir (Sigma), dicyclohexylcarbodiimide (Sigma), venturicidin A (Sigma), rotenone (Seahorse), antimycin (Seahorse), doxycycline (Research Product International), AZD8330-AZD6244-BE235 (Selleckchem).

Tumour culture. Patient-derived xenograft tumours were generated by transplanting subcutaneously in mice small tumour fragments isolated directly from surgical specimens according to Kim *et al.*¹⁷. Patient-derived samples were obtained from consented patients under an Institutional Review Board (IRB)-approved protocol LAB07-0854 chaired by J.B.F. Cells derived from early passage patient-derived xenograft tumours (F1–F2) and primary mouse tumours were kept in culture as spheres in semi-solid media for <15–16 passages. Briefly, after explant, tumours were digested at 37 °C for 1 h (Collagenase IV-Dispase 4 mg ml⁻¹; Invitrogen). Single-cell suspensions were plated in stem cell medium (SCM) MEBM (Lonza) supplemented with 2 mM glutamine (Invitrogen), B27 (Invitrogen), 20 ng ml⁻¹ hEGF (PeproTech), 20 ng ml⁻¹ hFGF (PeproTech), 5 μ g ml⁻¹ h-Insulin (Roche), 0.5 μ M hydrocortisone (Sigma), 100 μ M β -mercaptoethanol (Sigma), 4 μ g ml⁻¹ heparin (Sigma). Methocult M3134 (StemCell Technologies) was added to SCM (final concentration 0.8%) to keep tumour cells growing as clonal spheres versus aggregates. Fully formed tumour spheres were collected and digested with 0.05% trypsin (Gibco) to single cells and re-plated in culture. KRas inducible system: for maintenance, tumour cells were kept in doxycycline (2 μ g ml⁻¹). For experiments, cells were plated in the presence or absence of doxycycline to select SCs. After 8 days in culture, tumour spheres or SCs were collected, washed and processed for experiments.

In vitro treatments. For drug treatments, KRas-expressing spheres or surviving tumour cells after 8 days of doxycycline withdrawal were collected, washed, digested with trypsin and repeatedly counted (Countless, Invitrogen). Similar procedures were used for spheres derived from human tumours and from the KRas constitutive mouse model except that surviving cells were selected by treating tumour spheres with AZD8330 (10 nM) and BEZ235 (100 nM) for 1 week. Equal numbers of live KRas-expressing cells and surviving tumour cells (AZD8330 plus BEZ235 treated cells for the constitutive KRas and human tumours) were treated with oligomycin (200 nM for 24 h), venturicidin (500 nM for 24 h), DCCD (1 μ M for 24 h), bafilomycin (50 nM for 48 h) or etomoxir (100 μ M for 48 h). For OCR measurement and western-blot experiments, cells were treated with bafilomycin and etomoxir for 24 and 6 h, respectively. To test the effects of ROS on spherogenic potential, cells were treated with buthionine-sulphoximine (100 μ M) to deplete GSH or pre-treated with α -tocopherol (100 μ M), N-acetylcysteine (1 mM) or Mn-tetrakis (50 μ M) before oligomycin treatment. After drug wash out, treated cells were re-plated for 7 days (on doxycycline when using the KRas-inducible system). The number of tumour spheres (spherogenic potential) was quantified using ImageXpress Velos Laser Scanning Cytometer (Molecular Devices) upon calcein staining (Molecular Probes).

Tumour transplantation, limiting dilution and in vivo treatments. Tumour cells isolated from *in vivo* tumours or from *in vitro* spheres were digested to single cells (see earlier). Usually, 10⁴–10⁵ tumour cells were used for routine transplantation, for transplantation in limiting dilution 10³, 10² or 10 tumour cells were used. Tumour cells were suspended in SCM (see earlier) and Matrigel (BD Biosciences,

356231) (1:1 dilution) and injected subcutaneously into the flank of 6- to 8-week-old female immunodeficient mice (NCR-NU, Taconic). If cells were derived from KRas-inducible tumours, mice were injected with doxycycline (3 mg kg⁻¹, i.p.) at the time of transplantation and then fed with doxycycline in drinking water. Tumour-initiating cell (TIC) frequencies were determined by Poisson statistical analysis using L-Calc software (StemCell Technologies). For BrdU incorporation experiments, mice were injected i.p. with 1 mg of BrdU 3 times (every 8 h) after 24 h or 48 h of KRas reactivation before being killed. For *in vivo* oligomycin treatment, mice were transplanted with tumour cells and fed with doxycycline in drinking water (+Dox) until tumours reached 1 cm in diameter. At that time, doxycycline was withdrawn (-Dox) and, after 2 weeks, when tumours were almost completely regressed, mice were injected with oligomycin (0.5 mg kg⁻¹, i.p.; Sigma) or vehicle, 5 days a week, for 2 weeks. After treatment, doxycycline was added back to drinking water and mice were monitored for tumour relapse. Genetically identical and age matched recipient mice were used for transplantation experiments and were randomly allocated for treatment with oligomycin upon tumour regression. Experimenters were not blinded to the experimental groups in evaluating treatment outcome.

For pharmacological inhibition of KRas pathways using small molecule inhibitors, mice bearing tumours derived from the KRas constitutive system were treated with a combination of AZD6244 and BEZ235 (100 mg kg⁻¹ and 40 mg kg⁻¹ respectively, per oral gavage daily) for at least 7 days. Tumour volume was calculated using the formula: $V = l^2 \times L/2$ (where l is length and L is width). All manipulations were performed under IACUC-approved protocols.

Isolation of tumour cells. To isolate pure populations of tumour cells from transplanted tumours, we took advantage of dim-high ubiquitous expression of CD44 in pancreatic tumour cells (Extended Data Fig. 2f). Single-cell suspensions of digested tumours were stained with anti-CD44 biotinylated antibody (eBioscience, IM7). Positive cells were then purified using Easy Sep Biotin Selection Kit (StemCell Technologies) according to manufacturer's instructions. Counterstaining of isolated cells with anti-CD45 and CD31 antibodies and fluorochrome-conjugated streptavidin were used to check the purity by flow cytometry.

shRNA expression and gene downregulation. shRNAs against TFAM (TRCN000016093; TRCN0000016095) and TUFM (TRCN0000280863; TRCN0000280864) (Sigma) were cloned in a Tet-inducible pLKO vector provided by the Institute for Applied Cancer Science at MDACC (for sequences see also earlier). Human tumour spheres were transduced with viral particles and selected with puromycin. Upon selection, tumour spheres were treated or not with a combination of Mek/Pi3K inhibitors (10 nM AZD8330 plus 100 nM BEZ235) for 1 week. On the second day after beginning the combination drug treatment, doxycycline was added to the culture to induce shRNA expression and was maintained for 5 days. Then tumour cells were collected, washed, and replated to evaluate their spherogenic potential. Downregulation of the target was evaluated by western blot at 72 h after shRNA induction.

Flow cytometry and cell sorting. Single cells isolated from tumours or spheres were stained with primary antibodies after blocking with 10% BSA and 5% rat serum. Aldefluor (StemCell Technologies) and annexin V (eBioscience) staining were performed according to the manufacturer's instructions. To study the cell cycle of tumour spheres, BrdU Flow Kit (BD Pharmingen) was used according to datasheet specifications. Mitochondrial potential *in vitro* was measured using MitoProbe DilC₁(5) Assay Kit for Flow Cytometry (Molecular Probes) according to specifications, and CCCP treatment was used for controls. *In vivo* mitochondrial potential was evaluated according to Zheng *et al.*²⁷. Briefly, mice bearing KRas-expressing or regressed tumours (inducible model or pharmacologically treated with AZD6244 plus BEZ235) were injected with 25 nmol kg⁻¹ of TMRE (Molecular Probes) as a tail vein bolus. After 1 h, mice were killed and tumours explanted, digested (as described earlier, adding 10% FBS to the digestion mix) and stained for CD44. Samples were kept on ice and immediately acquired gating CD44-positive DAPI-negative cells. The same approach, timeline and staining were used to evaluate glucose uptake *in vivo*, injecting 25 μ mol kg⁻¹ 2NBDG (Molecular Probes) as a tail vein bolus according to Ytoh *et al.*²⁸. For *in vitro* experiments, 2NBDG was used at a concentration of 10 μ M in complete stem cell medium containing 2 mM glucose²⁸. Cells were incubated for 6 h then washed and analysed by flow cytometry. MitoTracker Green and Deep Red (Molecular Probes) were used to measure mitochondrial mass. ROS were evaluated using MitoSOX and CellROX (Molecular Probes) and induced by 4-hydroxynonenal treatment (10 μ M) for positive controls. LipidTox deep red (Molecular Probes) was used for quantifying lipid droplet content. All staining procedures were performed according to the manufacturer's protocols. DAPI (Invitrogen) was used to stain DNA content or to exclude dead cells depending on the experiment. For measuring autophagic flux, KRas-expressing cells were transduced with pBABE-puro GFP-LC3 (ref. 26) and, upon selection, doxycycline was withdrawn for 8 days. Mean of fluorescence of surviving and matched KRas-expressing cells was quantified and surviving tumour cells treated for 24 h with bafilomycin (50 nM) were used as a control. Gating strategies to exclude doublets and dead cells (DAPI) were always employed. After staining,

samples were acquired using a BD FACSCantoII flow cytometer or sorted using BD Influx cell sorter. Data were analysed by BD FACSDiva or FlowJo (Tree Star). **Immunohistochemistry and immunofluorescence.** Tumour samples were fixed in 4% formaldehyde for 2 to 4 h on ice, moved in 70% ethanol for 12 h, and then embedded in paraffin (Leica ASP300S). After cutting (Leica RM2235), baking and deparaffinization, slides were treated with Citra-Plus Solution (BioGenex) according to specifications. For immunohistochemistry (IHC) staining, endogenous peroxidases were inactivated by 3% hydrogen peroxide. Non-specific signals were blocked using 3% BSA, 10% goat serum and 0.1% Triton. Tumour samples were stained with primary antibodies. For BrdU detection, samples were digested on slides for 1 h at 37 °C with DNase I (300 µg ml⁻¹) before staining. For IHC, ImmPress and ImmPress-AP (Vector Lab) were used as secondary antibodies and Nova RED, Vector BLUE and DAB were used for detection (Vector Lab). Images were captured with a Nikon DS-Fi1 digital camera using a wide-field Nikon EclipseCi microscope. For immunofluorescence, secondary antibodies conjugated with Alexa488 and 555 (Molecular Probes) were used. Images were captured with a Hamamatsu C11440 digital camera, using a wide-field Nikon EclipseNi microscope. LipidTox, Lysotracker, MitoTracker, CellRox and Hoechst 33342 (Molecular Probes) were used on live spheres and cells at the concentrations suggested by manufacturer's protocols and images were acquired using a Nikon high-speed multiphoton confocal microscope A1 R MP.

Immunoblotting. Protein lysates were resolved on 5–15% gradient polyacrylamide SDS gels and transferred onto PVDF membranes according to standard procedures. Membranes were incubated with indicated primary antibodies, washed, and probed with HRP-conjugated secondary antibodies. The detection of bands was carried out upon chemiluminescence reaction followed by film exposure. For the Ras pull-down assay, the amount of active Ras was evaluated by detecting the fraction of Ras protein that co-precipitated with RAF kinase. Cell lysates from KRas-expressing cells and cells surviving KRas ablation were co-incubated with RAF-linked agarose beads for 2 h. After incubation, beads were collected, washed and boiled for 5 min in the presence of laemmli loading buffer supplemented with 10% 2-mercaptoethanol and ultimately loaded onto SDS-PAGE gels. The detection of the active fraction of RAS was carried out using standard western blot procedures with anti-Ras antibody.

Transmission electron microscopy. TEM was performed at the MDACC High Resolution Electron Microscopy Facility. Samples were fixed with a solution containing 3% glutaraldehyde plus 2% paraformaldehyde in 0.1 M cacodylate buffer, pH 7.3, for 1 h. After fixation, the samples were washed and treated with 0.1% Millipore-filtered cacodylate buffered tannic acid, post-fixed with 1% buffered osmium tetroxide for 30 min, and stained en bloc with 1% Millipore-filtered uranyl acetate. The samples were dehydrated in increasing concentrations of ethanol, infiltrated and embedded in LX-112 medium. The samples were polymerized in a 60 °C oven for 2 days. Ultrathin sections were cut in a Leica Ultracut microtome, stained with uranyl acetate and lead citrate in a Leica EM Stainer and examined in a JEM 1010 transmission electron microscope (JEOL) at an accelerating voltage of 80 kV. Digital images were obtained using an AMT Imaging System (Advanced Microscopy Techniques Corp).

DNA, RNA, cDNA and qPCR. DNA and RNA were extracted using DNeasy Blood and Tissue Kit (Qiagen) and RNeasy Mini Kit (Qiagen) according to technical specifications. A mix of random hexamers and oligo(dT) were applied for cDNA synthesis using SuperScript III First-Strand-Synthesis System (Invitrogen). For qPCR, 10 ng of DNA or cDNA was amplified with EXPRESS SYBR GreenER qPCR Super-Mix (Invitrogen) using a Startagene Mx3005p thermal-cycler. Primers used for mitochondrial and lipid metabolic gene amplification are shown in Extended Data Fig. 10c. β -Actin (F-GACGGCCAGGTCATCACTAATTG, R-AGGAAGGCTG GAAAAGAGCC), 28S (F-TCATCAGACCCAGAAAAGG, R-GATTCGGCA GGTGAGTTGTT) and β -microglobulin (F-ATTCACCCCACTGAGACTG, R-TGCTATTTCTTCTGCGTGC) were used as house-keeping genes for normalization. Expression of genes involved in the ETC, mitochondria and autophagy was evaluated using Qiagen commercial arrays: RT2 Profiler PCR Array Mouse Mitochondrial Energy Metabolism, Mouse Mitochondria and Mouse Autophagy.

Expression profiling and data analysis. Gene expression profiling was performed at the Dana-Farber Cancer Institute Microarray Core Facility. RNA isolated from KRas-expressing and surviving tumour cells was hybridized on a Gene Chip Mouse Genome 430 2.0 Array (Affymetrix). Raw data (CEL files) were pre-processed using a robust multi-array analysis (RMA) and analysed with GSEA²⁹ using Signalling Pathways c2.cp.v3.0 and TFT c3.tft.v3.0 symbols gene sets.

Exome sequencing. Two independent tumour spheres expressing KRas were collected, digested to single cells, and replated on two plates in the presence of doxycycline (ON Dox, +KRas) for 1 week. Spheres from one plate were collected and snap frozen for DNA extraction (Reference), while spheres from the other plate were plated OFF doxycycline (–KRas) for 1 week before re-addition of doxycycline to re-express KRas. After 1 week, reformed spheres were collected and snap frozen for

DNA extraction (Re-ON). Genomic DNA was extracted by phenol-chloroform and further purified from contamination on columns (Qiagen). The exonic DNA regions were captured using Nimblegen SeqCap EZ Mouse Exome kit. The DNA sequences recovered were processed through a standard SNP calling pipeline. Reads were aligned using BWA following removal of duplicates, realignment and recalibration (using Broad's Genome Analysis Toolkit or GATK). Further SNPs were called using GATK's Unified Genotyper and annotated using Annovar. For the analysis we considered SNVs which were (1) called in both samples, and (2) had a minimum coverage of 200 in both, providing enough confidence to compare allelic frequencies between the sample pairs. Roughly 12% (40,383/335,076) and 13% (44,182/339,398) of the common SNVs for tumour #2 and tumour #1, respectively, satisfied the coverage threshold of 200.

Metabolomics. For metabolomic analyses, surviving and KRas-expressing spheres (grown as described earlier), were collected after 8 days of culture when both KRas-expressing spheres and surviving tumour cells were confirmed to have exited the cell cycle (to avoid the confounding effects of proliferation on metabolism) (Extended Data Fig. 1g). The day before collection, medium was changed and spheres re-plated in fresh medium. After 24 h, cells were collected by centrifugation, washed three times and samples were then immediately lysed in methanol:water (80:20) at dry-ice temperature. The quantity of the metabolite fraction analysed was adjusted to the corresponding protein concentration calculated upon processing a parallel sample. Metabolite fractions were processed and analysed by targeted LC-MS/MS via selected reaction monitoring (SRM), as described^{11,12,14}. Processed data were analysed in Cluster 3.0 and TreeViewer. The analysis was performed on four independent tumours in biological triplicate.

Carbon-13 metabolic tracing. Cells were prepared in a manner identical to that for routine metabolomic profiling (see earlier), except that media formulations were adjusted to account for isotopically labelled substrates. Specifically, surviving and KRas-expressing cells were plated in RPMI medium (reconstituted with growth factors) devoid of glucose, glutamine, pyruvate, palmitate and supplemented with one of the four carbon-13-labelled substrates where the remaining substrates were unlabelled (glucose 10 mM, glutamine 2 mM, pyruvate 1 mM and palmitate 75 µM). After 24 h of incubation with labelled substrates, cells were collected and replated again in fresh media supplemented with carbon-13-labelled substrates for another 12 h to minimize the effects of cellular uptake on the concentration of different substrates. After 36 total hours of labelling, samples were collected by centrifugation and immediately lysed in methanol:water (80:20) at dry-ice temperature. As for metabolomics, the quantity of the metabolite fraction analysed was adjusted to the corresponding protein concentration from a sample processed in parallel. The analysis was performed for each of the four substrates on three independent tumours in biological triplicate. Data were collected and processed as above, as described^{11,12,14}.

Oxygen consumption and glycolytic capacity. Tumour spheres and surviving cells were digested to single cells and spun into XF96 Cell Culture Microplates (Seahorse Bioscience) previously treated with Cell-Tak (BD Biosciences) immediately before the experiment. To measure OCR and the response to OXPHOS inhibition, we plated cells in complete stem cell medium (see earlier) prepared with MEMB lacking NaHCO₃ (Lonza) and supplemented with 5 mM pyruvate (Sigma). Oligomycin, FCCP, antimycin and rotenone (XF Cell Mito Stress Test Kit, Seahorse Bioscience) were sequentially injected at a final concentration of 1 µM. To measure ECAR, we plated tumour cells in base assay medium (D5030, Sigma) adjusted to pH 7.4. Glucose (10 mM final concentration, Sigma), oligomycin (1 µM final concentration) and 2-deoxy-glucose (100 mM final concentration) were sequentially injected (XF Glycolysis Stress Test Kit, Seahorse Bioscience). Specifically for the experiment in Extended Data Fig. 5c, ECAR was measured in complete stem cell media to evaluate the glycolytic reserve of tumour cells in a nutrient-rich environment. Oligomycin and 2DG were sequentially injected at a final concentration of 1 µM and 100 mM, respectively. Experiments were run using a XF96 analyser and raw data were normalized to metabolically active cells, evaluated as Hoechst 33342 positive/propidium iodide negative, by an Operetta High-Content Imaging System (Perkin-Elmer) immediately after each experiment.

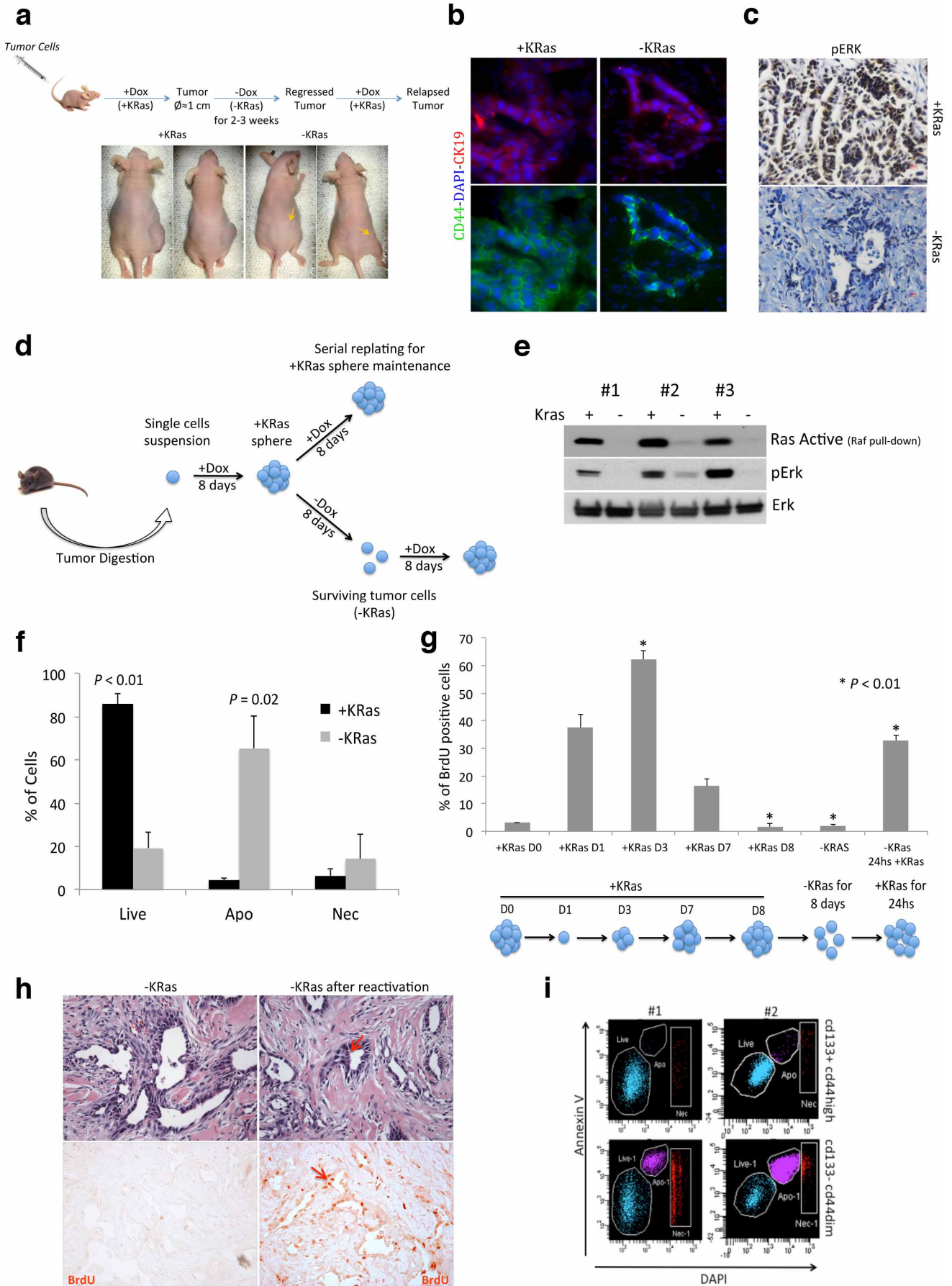
Glucose uptake and lactate production. To measure glucose uptake, we used two different approaches: by flow cytometry using the fluorescent glucose analogue 2NBDG (see earlier) and by YSI 2950 Biochemistry Analyzer (YSI Life Sciences). For YSI, the same number of surviving cells and KRas-expressing cells was plated in triplicate in a 96-well plate in 200 µl of complete stem cell medium. After 24 h, the medium was collected and glucose and lactate concentrations were evaluated by YSI analyser, using unconditioned fresh medium as a reference. Lactate production in response to oligomycin treatment (200 nM) was independently measured by Lactate Kit (Trinity Biotech) to directly evaluate the concentration of lactic acid in stem cell medium 6 to 12 h after tumour cell plating.

ATP and glutathione assays. ATP production of tumour cells in response to 6 to 12 h of oligomycin treatment (200 nM) was measured using Cell Titer Glo (Promega)

following the manufacturer's instructions. For glutathione, cells were lysed by sonication in 1× GSH MES buffer (Cayman Chemical) and debris was pelleted by high-speed centrifugation. Samples were deproteinized by adding vol:vol of meta-phosphoric acid (1 mg ml⁻¹) followed by centrifugation and pH equilibration by the addition of triethanolamine. Glutathione content was evaluated through an enzymatic recycling method by using a commercially available kit (Cayman Chemical). PHERA-Star plus microplate reader (BMG Labtech) was used to measure luminescence and absorbance.

Statistical analysis. *In vitro* and *in vivo* data are presented as the mean ± s.d. Statistical analyses were performed using a two-tailed Student's *t*-test after the evaluation of variance. Results from survival experiments were analysed with a Gehan–Breslow–Wilcoxon test and expressed as Kaplan–Meier survival curves.

26. Fung, C., Lock, R., Gao, S., Salas, E. & Debnath, J. Induction of autophagy during extracellular matrix detachment promotes cell survival. *Mol. Biol. Cell* **19**, 797–806 (2008).
27. Zheng, W., Talley Watts, L., Holstein, D. M., Wewer, J. & Lechleiter, J. D. P2Y1R-initiated, IP3R-dependent stimulation of astrocyte mitochondrial metabolism reduces and partially reverses ischemic neuronal damage in mouse. *J. Cereb. Blood Flow Metab.* **33**, 600–611 (2013).
28. Itoh, Y., Abe, T., Takaoka, R. & Tanahashi, N. Fluorometric determination of glucose utilization in neurons *in vitro* and *in vivo*. *J. Cereb. Blood Flow Metab.* **24**, 993–1003 (2004).
29. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).



Extended Data Figure 1 | Oncogene ablation leads to tumour regression *in vitro* and *in vivo*. **a**, *In vivo* experimental scheme. Tumour cells isolated from primary tumours or tumour spheres were injected in nude mice fed with doxycycline in drinking water (+Dox). When mice developed tumours, doxycycline was withdrawn (-Dox) and tumours underwent a complete macroscopic regression after 2–3 weeks (arrows indicate regressed tumours). In residual lesions few tumour cells remain quiescent for months and they can quickly reform tumours upon KRas reactivation (+Dox). **b**, Tumours expressing KRas (+KRas) and tumour remnants after regression (-KRas) are positive for ductal epithelial marker CK19 ($\times 40$). **c**, Tumours expressing KRas (+KRas) and epithelial remnants after tumour regression (-KRas) were stained for phosphorylated-p42/44 (pErk). No signal is detected in surviving cells ($\times 20$). **d**, *In vitro* experimental scheme. After digestion to a single-cell suspension, tumour cells isolated from primary tumours were plated in stem cell medium in presence of doxycycline (+Dox, +KRas). Spherogenic cells form tumour spheres (+KRas) that can be maintained by serial replating in presence of doxycycline. Upon doxycycline withdrawal (-Dox) tumour spheres undergo involution and only a minority of cells survive the ablation of KRas (SCs, -KRas). SCs readily reform tumour spheres upon re-activation of KRas (+Dox). **e**, The amount of active Ras in KRas-expressing cells (+KRas) and SCs (-KRas) was evaluated in three independent tumour spheres by detecting the fraction of Ras protein that co-precipitates with Raf kinase. Total

lysates were probed with anti-phospho-p42/44 (pErk), total p42/44 (Erk) antibodies. **f**, Annexin V staining in tumour spheres after 3 days +/-KRas ($n = 3$). **g**, Sphere formation is a regulated process and tumour cells enter and exit cell cycle. BrdU incorporation (pulse of 3 h) was evaluated at different time points during sphere formation and regression. KRas-expressing fully formed spheres (day (D)0 and 8) are quiescent. Upon sphere dissociation and replating (D0), spherogenic cells enter cell cycle (D1) and tumour cells continue to grow until day 3–4, when spheres reach their maximal S phase. Then tumour cells gradually exit the cell cycle and become quiescent (D8). After doxycycline withdrawal (-KRas), tumour spheres undergo involution and surviving cells remain quiescent until KRas is re-expressed (-KRas 24 h +KRas) and spheres are reformed. Ruling out the effect of the cell cycle, transcriptomic and metabolomic characterizations were done, matching quiescent surviving tumour cells to quiescent fully formed KRas-expressing spheres at D8 ($n = 3$). **h**, Haematoxylin and eosin staining and immunohistochemistry of regressed tumours after three 8 h pulses of BrdU show that epithelial remnants in regressed tumours after KRas ablation (-KRas) are completely quiescent (left panels). Forty-eight hours after KRas reactivation (doxycycline i.p. injection) tumour cells re-enter massively the cell cycle (right panels). Red arrows indicate mitotic cells ($\times 20$). **i**, Representative annexin V staining with respect to CD133 and CD44 after 3 days of KRas ablation, two independent tumours are represented. Data are mean \pm s.d.

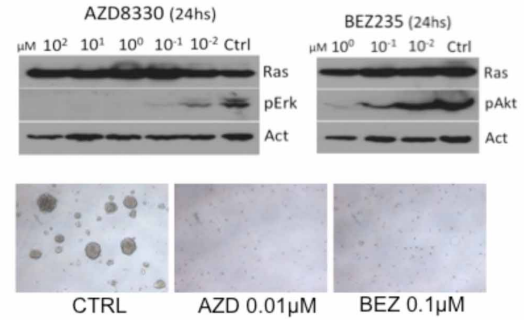
a

Transplantation in Limiting Dilution: Genetic Model <i>Ex Vivo</i>				
Experiment	Cells	Dose	Transplanted Mice	Developed Tumors
I	+KRas	100	2	2 (100%)
		10	6	1 (16.7%)
	-KRas	100	2	2 (100%)
		10	6	6 (100%)
II	+KRas	100	2	2 (100%)
		10	8	2 (25%)
	-KRas	100	2	2 (100%)
		10	7	7 (100%)
III	+KRas	100	2	2 (100%)
		10	8	3 (37.5%)
	-KRas	100	2	2 (100%)
		10	8	8 (100%)
IV	+KRas	100	2	2 (100%)
		10	10	3 (30%)
	-KRas	100	2	2 (100%)
		10	10	10 (100%)

b

Transplantation in Limiting Dilution: Genetic Model <i>Ex Vivo</i>				
Experiment	Cells	Dose	Transplanted Mice	Developed Tumors
I	+KRas	100	3	1 (33.3%)
		10	8	0 (0%)
	-KRas	100	3	3 (100%)
		10	8	5 (62.5%)
II	+KRas	100	2	1 (50%)
		10	10	1 (10%)
	-KRas	100	2	2 (100%)
		10	10	6 (60%)

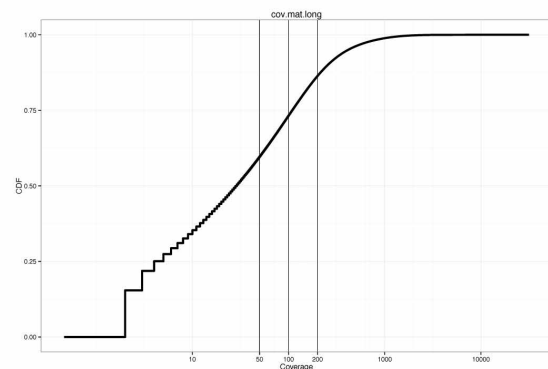
c



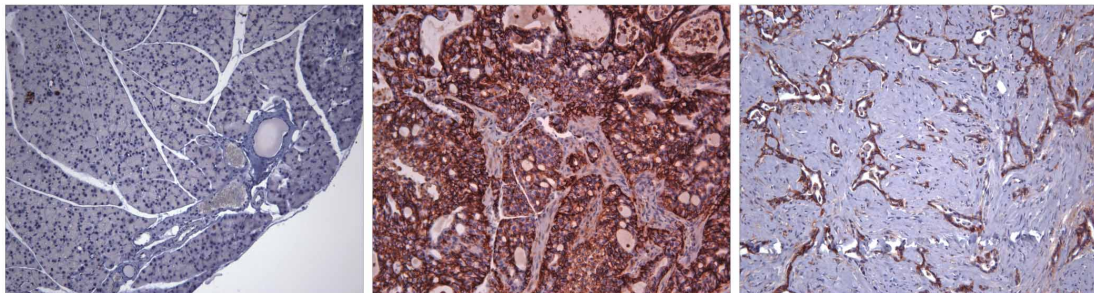
d

Transplantation in Limiting Dilution: Pharmacological Downregulation				
Experiment	Cells	Dose	Transplanted Mice	Developed Tumors
I	CTRL	100	2	2 (100%)
		10	8	1 (12.5%)
	AZD+BEZ	100	2	2 (100%)
		10	6	5 (83.3%)
II	CTRL	100	2	2 (100%)
		10	8	2 (25%)
	AZD+BEZ	100	2	2 (100%)
		10	8	6 (75%)

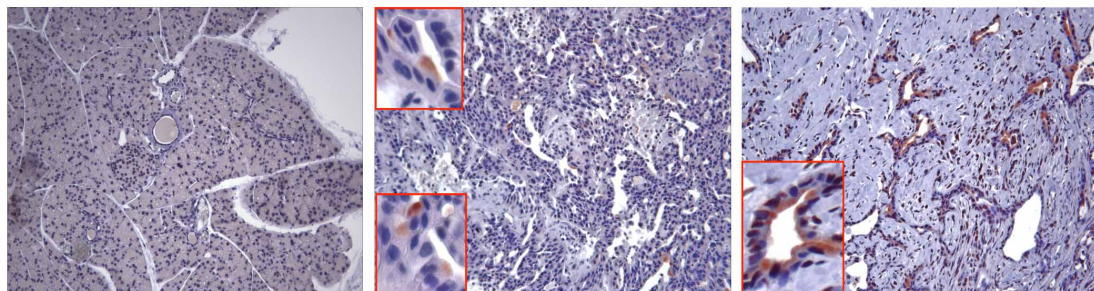
e



f



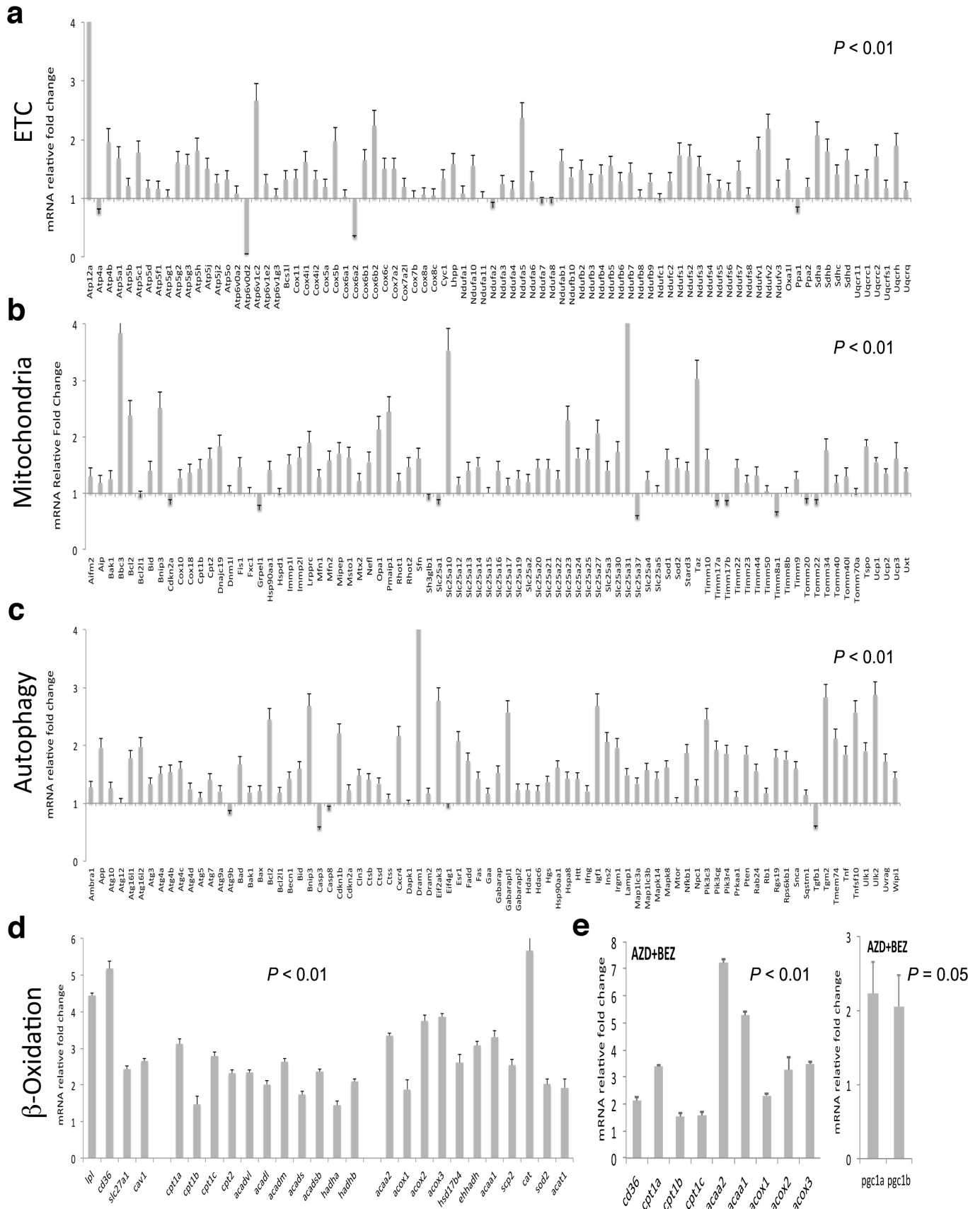
g



Extended Data Figure 2 | Transplantation in limiting dilution and characterization of epithelial remnants.

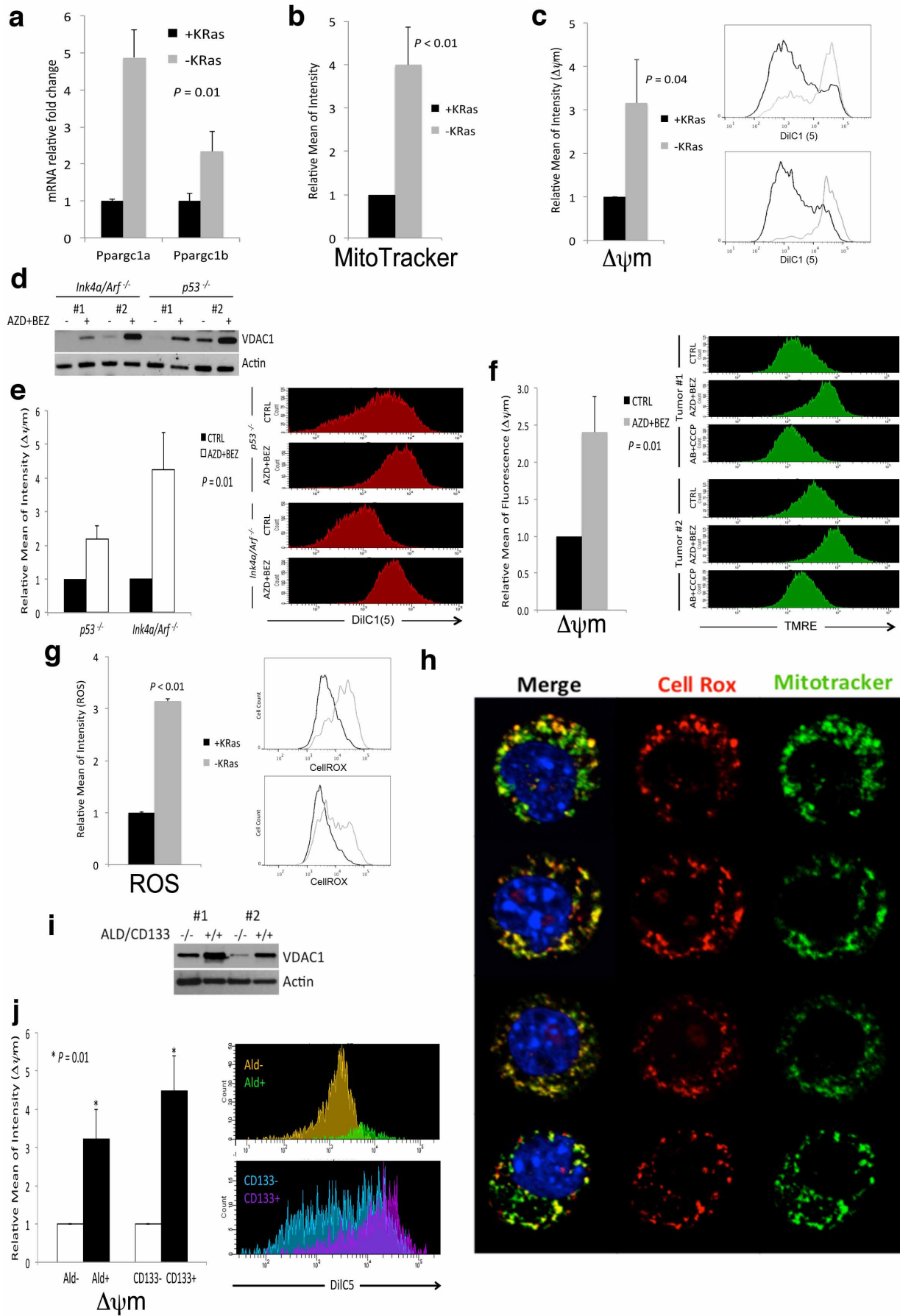
a–d, Transplantation in limiting dilution. Experiments, number of transplanted mice and percentage of developed tumour are shown. **a**, Limiting dilution experiments using tumour spheres (+KRas) and surviving cells (–KRas) (genetic model *ex vitro*). **b**, Limiting dilution experiments using cells isolated from KRas-expressing (+KRas) and regressed tumours (–KRas) (genetic model *ex vivo*). **c**, Top panels, immunoblots of tumour spheres treated with different concentrations of Mek1 (AZD8330) and dual PI3K/mTOR (BEZ235) inhibitors probed with anti-phospho-p42/44 (pErk), phospho-Akt (pAkt), pan-Ras (Ras) and β -actin (Act) antibodies. Bottom panels, effects of AZD8330 (AZD 0.01 μ M) and BEZ235 (BEZ 0.1 μ M) treatment for 1 week on tumour sphere formation; some cells, as single cells or in small clusters, are able to survive the treatment ($\times 5$). **d**, Limiting dilution experiments using cells surviving pharmacological

downregulation of oncogenic pathways (AZD+BEZ, combination of AZD8330 and BEZ235) and control cells (CTRL). **e**, The plot shows the cumulative distribution of coverage at all the SNVs called by Unified Genotyper (across samples). **f**, CD44 is expressed during tumorigenesis in mice: no positive cells are detected in normal pancreas (left panel), KRas-expressing tumours express high levels of CD44 (middle panel), epithelial remnants in regressed tumours maintain their positivity for CD44 (right panel) ($\times 10$). **g**, Validation of CD133 (ab16518) in immunohistochemistry: this antibody does not recognize cells and ductal structures in normal pancreas (left panel), a small population of cells is stained by ab16518 in KRas-expressing tumours (middle panel), epithelial remnants in regressed tumours are strongly positive for CD133 (right panel) ($\times 10$). At higher magnification (red boxes), it is possible to appreciate the classical polarized pattern of CD133.



Extended Data Figure 3 | qPCR validation of pathways enriched in surviving cells. a–e, mRNA fold change in surviving cells normalized to KRas-expressing cells. **a**, Genes involved in ETC ($n = 5$). **b**, Genes involved in the biogenesis and function of mitochondria (Mitochondria) ($n = 5$).

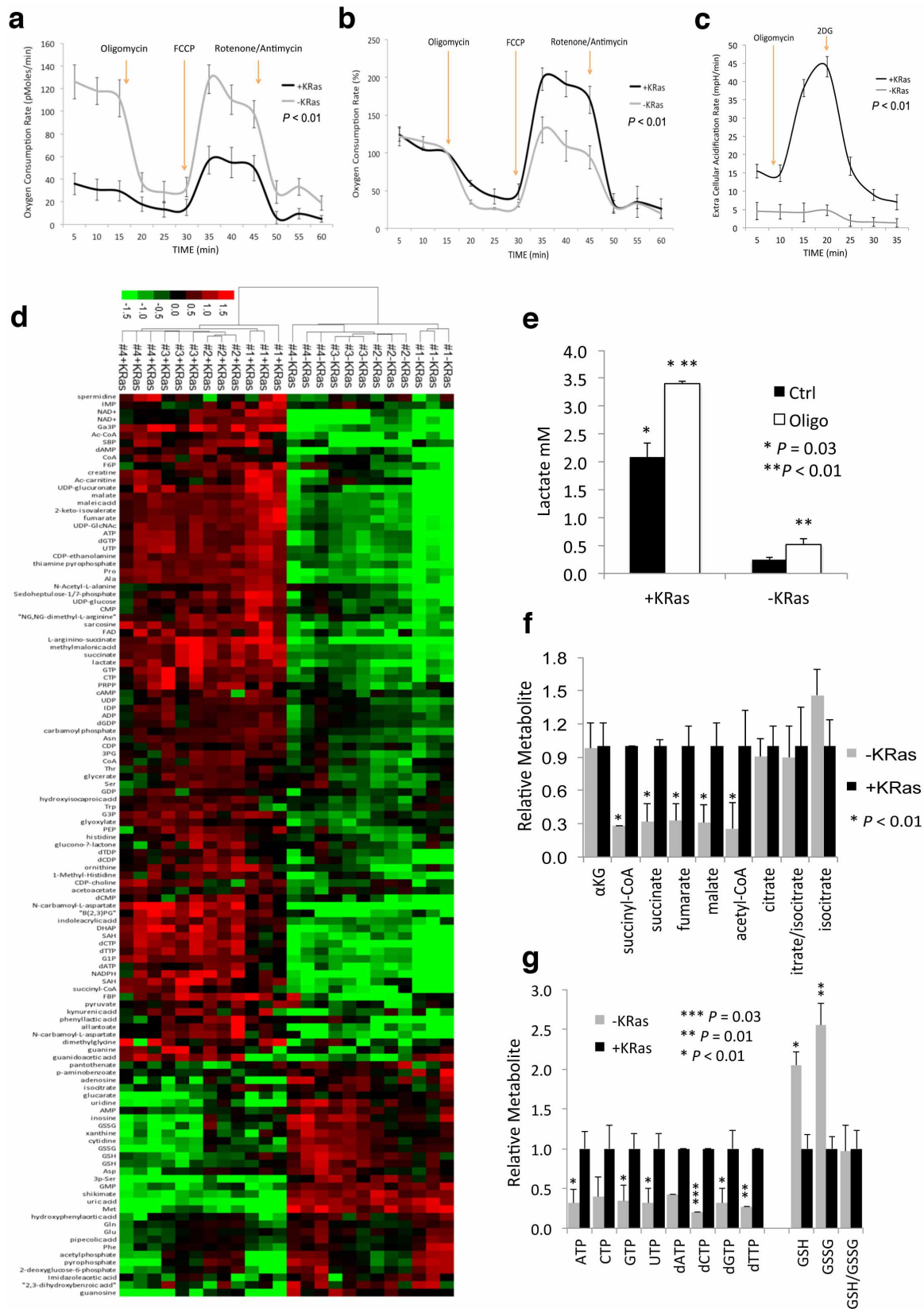
c, d, Genes encoding proteins of the autophagic molecular machinery and its key regulators ($n = 5$) (**c**) and β -oxidation ($n = 5$) (**d**). **e**, mRNA fold change in cells surviving AZD8330 plus BEZ235 treatment (AZD+BEZ) versus controls ($n = 3$). Data are mean \pm s.d.



Extended Data Figure 4 | Surviving cells have more active mitochondria.

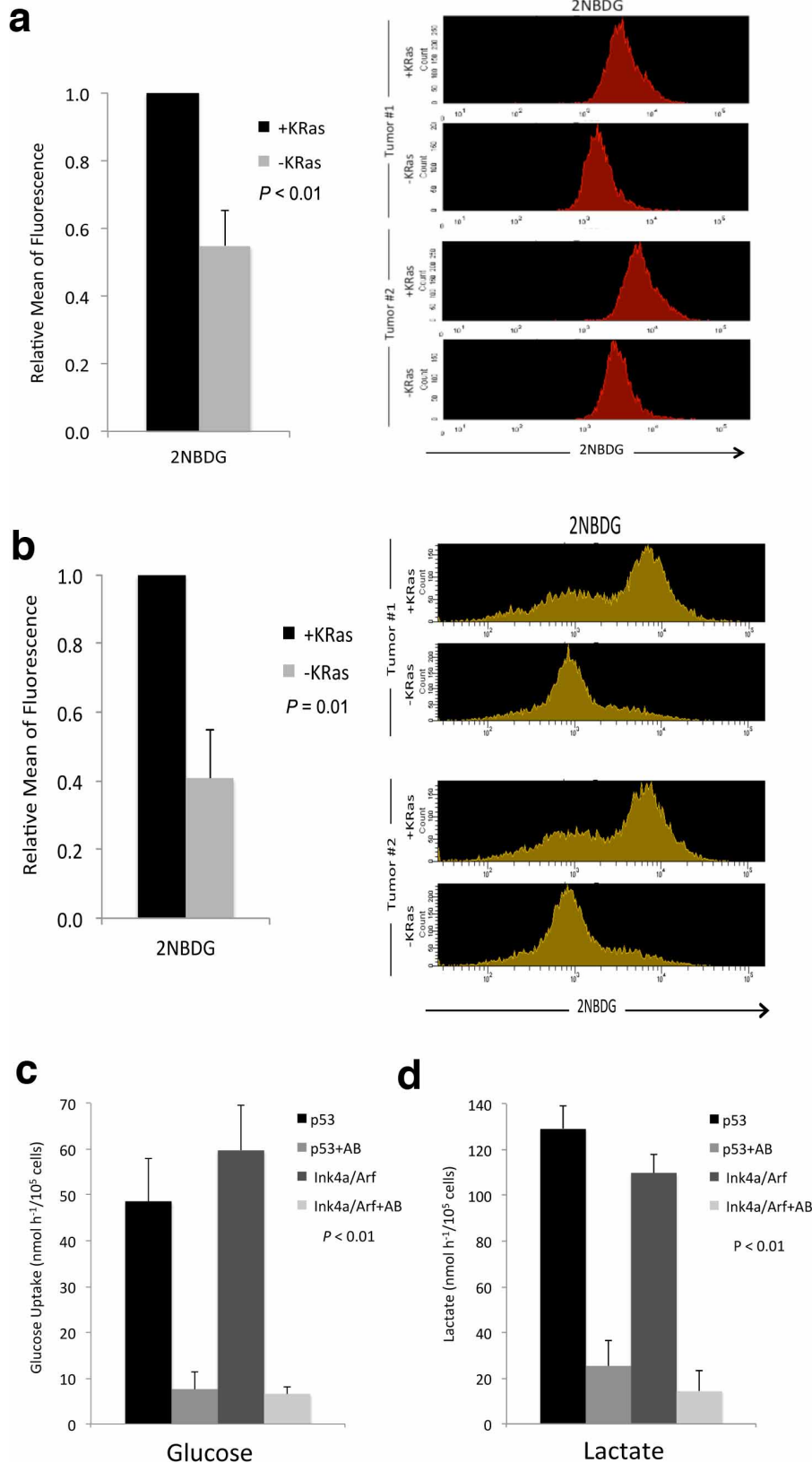
a, mRNA fold change of *Pparg1* genes in -KRas versus +KRas cells ($n = 5$). **b**, Quantification of MitoTracker Green staining in +KRas and -KRas cells ($n = 3$). **c**, Mitochondrial membrane potential ($\Delta\psi_m$) of +KRas and -KRas cells ($n = 4$); representative flow-cytometry analysis of two tumours. **d**, Immunoblot of two independent tumour spheres derived from different genetic backgrounds (*Ink4a/Arf*^{-/-} and *p53*^{-/-}) treated or not with AZD8330 and BEZ235 (AZD+BEZ) for 7 days and probed with anti-VDAC1 (VDAC1) and β -actin (Actin) antibodies. **e**, Cells surviving AZD+BEZ treatment have higher mitochondrial transmembrane potential ($\Delta\psi_m$) than untreated cells (CTRL) ($n = 3$); representative flow-cytometry analysis is reported. **f**, Mice bearing tumours have been treated (AZD+BEZ) or not (CTRL) with a combination of AZD6244 and BEZ235 for 1 week. Upon tail vein injection of a bolus of TMRE, tumours were explanted and analysed by flow cytometry for

their mitochondrial potential ($\Delta\psi_m$) upon gating on CD44⁺ DAPI⁻ cells ($n = 3$). A representative flow-cytometry analysis of two different tumours is reported, AB+CCCP samples represent reacquisition of AZD+BEZ samples after incubation with CCCP for 5 min. **g**, ROS production in +KRas and -KRas cells ($n = 3$); representative flow-cytometry analysis of two tumours. **h**, Live confocal imaging of SCs stained for mitochondria (MitoTracker Green), ROS (CellRox-Red) and DNA (Hoechst). The vast majority of signal generated by ROS colocalizes with mitochondria. **i**, Immunoblot of Aldefluor/CD133 double-positive and double-negative cells sorted from two independent tumours probed with anti-VDAC1 (VDAC1) and β -actin (Actin) antibodies. **j**, KRas-expressing cells positive for aldefluor (Ald⁺) and CD133 (CD133⁺) have higher mitochondrial transmembrane potential ($\Delta\psi_m$) than tumour cells that do not express the same markers (Ald⁻ and CD133⁻) ($n = 3$); a representative flow-cytometry analysis is reported. Data are mean \pm s.d.



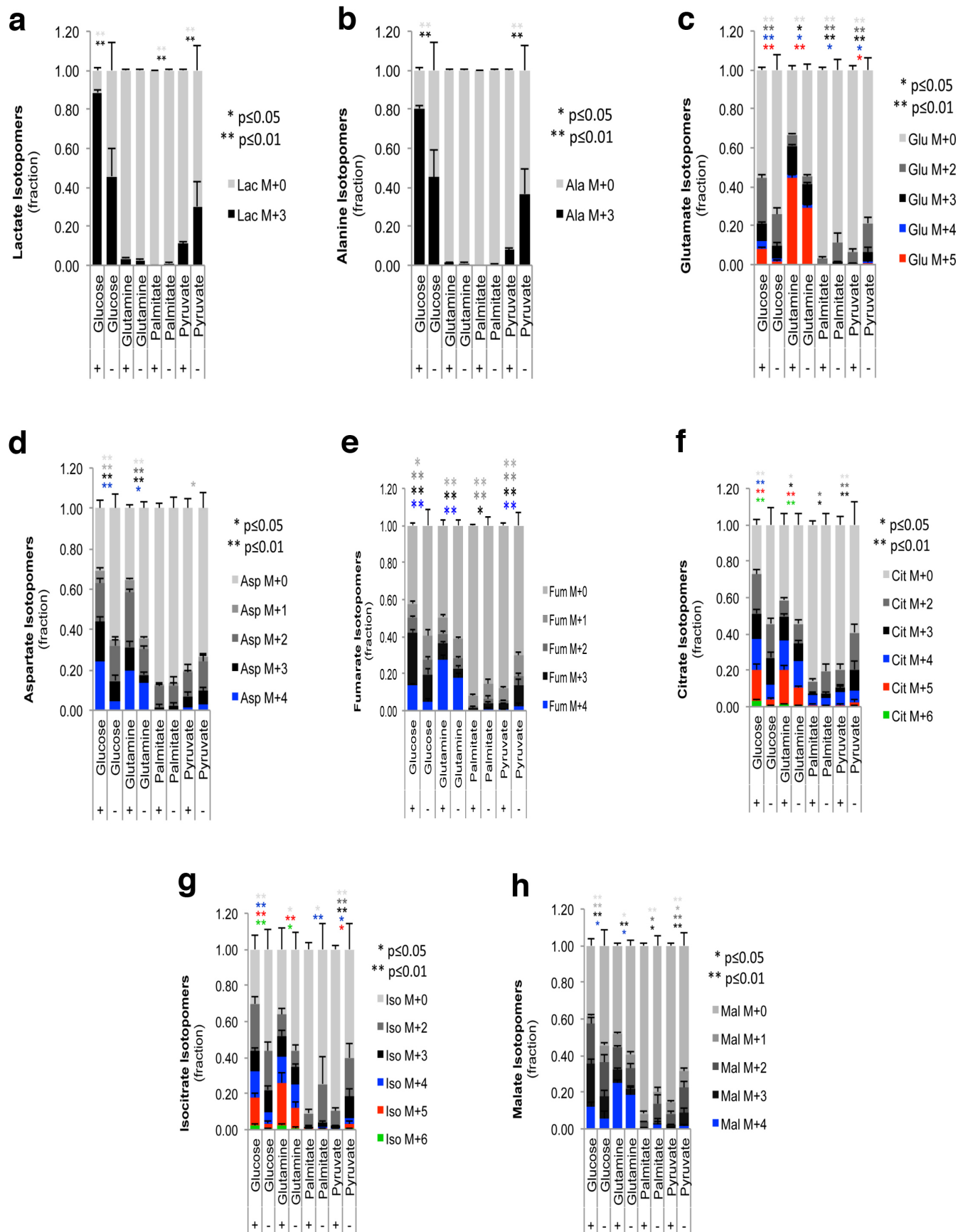
Extended Data Figure 5 | OCR, ECAR and metabolomics. **a**, OCR of KRas-expressing cells (+KRas) and SCs (−KRas) in response to oligomycin, FCCP and rotenone/antimycin ($n = 4$). **b**, Same as in **a** but normalized to basal respiration of +KRas and −KRas cells. **c**, ECAR response of +KRas and −KRas cells to oligomycin and 2DG. The experiment has been carried out in complete stem cell media to evaluate the glycolytic reserve of tumour cells in a nutrient rich environment. **d**, Metabolome analysis for +/−KRas cells; unsupervised

hierarchical clustering and heat map of significantly ($P < 0.05$) deregulated metabolites ($n = 4$). **e**, Lactate production of +KRas and −KRas cells in response to oligomycin (Oligo) or DMSO (Ctrl) treatment ($n = 3$). **f**, Fold change of TCA cycle intermediates in +KRas versus −KRas cells (α KG, α -ketoglutarate) ($n = 4$). **g**, Fold change of nucleotide triphosphates and deoxynucleotide triphosphates, glutathione (GSH) and glutathione disulphide (GSSG) in −KRas versus +KRas cells ($n = 4$). Data are mean \pm s.d.



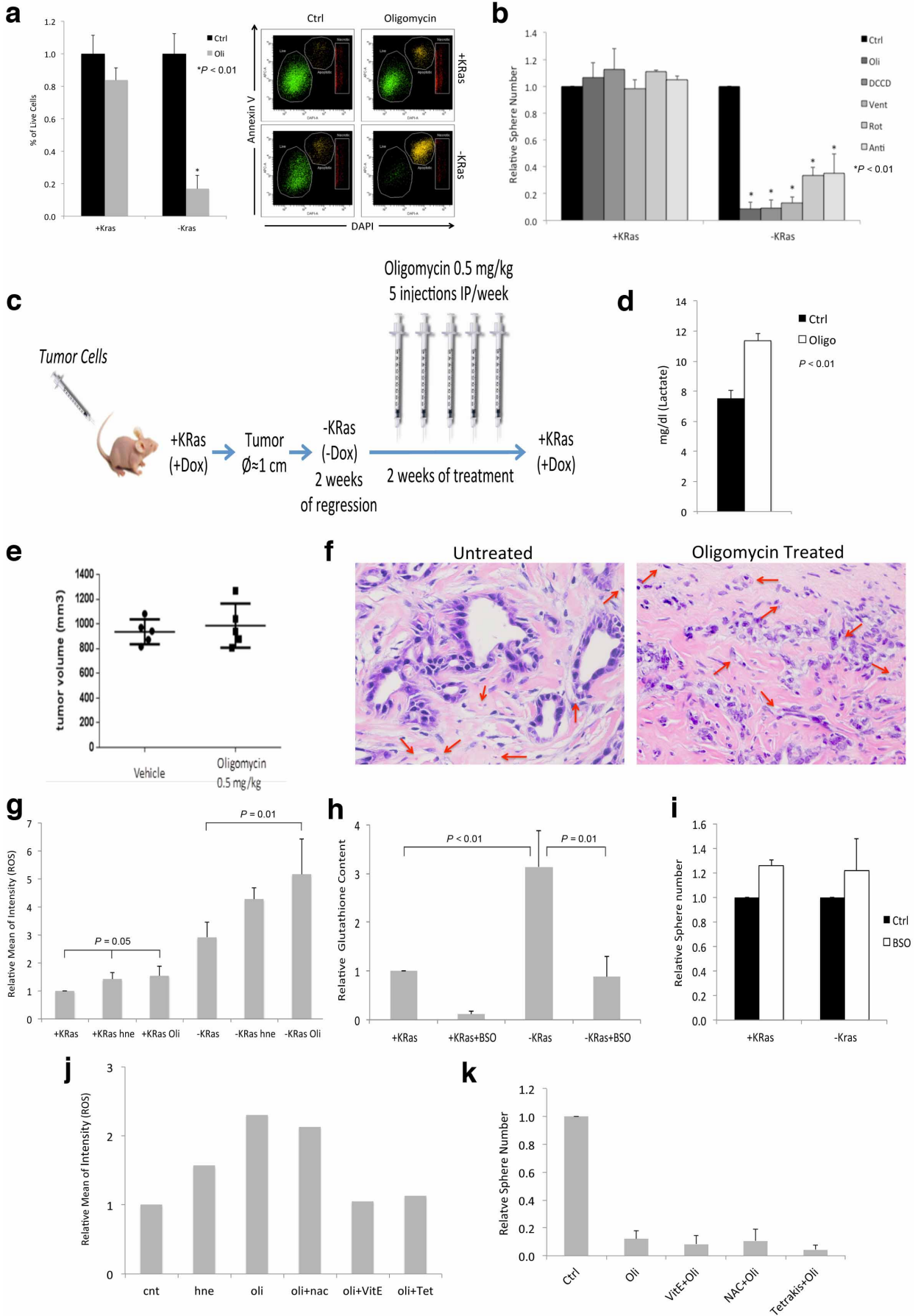
Extended Data Figure 6 | Surviving cells *in vitro* and *in vivo* have an impaired glucose uptake. **a**, KRas-expressing cells (+KRas) and SCs (-KRas) were incubated with 2NBDG (2-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose) for 6 h then analysed by flow cytometry ($n = 3$); a representative flow-cytometry analysis of spheres derived from two different tumours is reported. **b**, Mice bearing KRas-expressing tumours (+KRas) and 3-week regressed tumours (-KRas) were injected with a tail vein bolus of

2NBDG. After 1 h tumours were explanted and analysed by flow cytometry upon gating on $\text{CD44}^+ \text{DAPI}^-$ cells ($n = 3$); a representative flow-cytometry analysis of two different tumours is reported. **c, d**, Tumour spheres derived from different genetic backgrounds ($\text{Ink4a/Arf}^{-/-}$ and $\text{p53}^{-/-}$) were treated (+AB) or not with AZD8330 and BEZ235 for 7 days then plated in fresh stem cell medium. After 24 h medium was collected and analysed by YSI analyser for glucose uptake (**c**) and lactate production (**d**) ($n = 3$). Data are mean \pm s.d.



Extended Data Figure 7 | Fuel carbon contribution to TCA cycle and TCA branch metabolites. a–h, Isotopomer distribution for lactate (a), alanine (b), glutamate (c), aspartate (d), fumarate (e), citrate (f), isocitrate (g) and

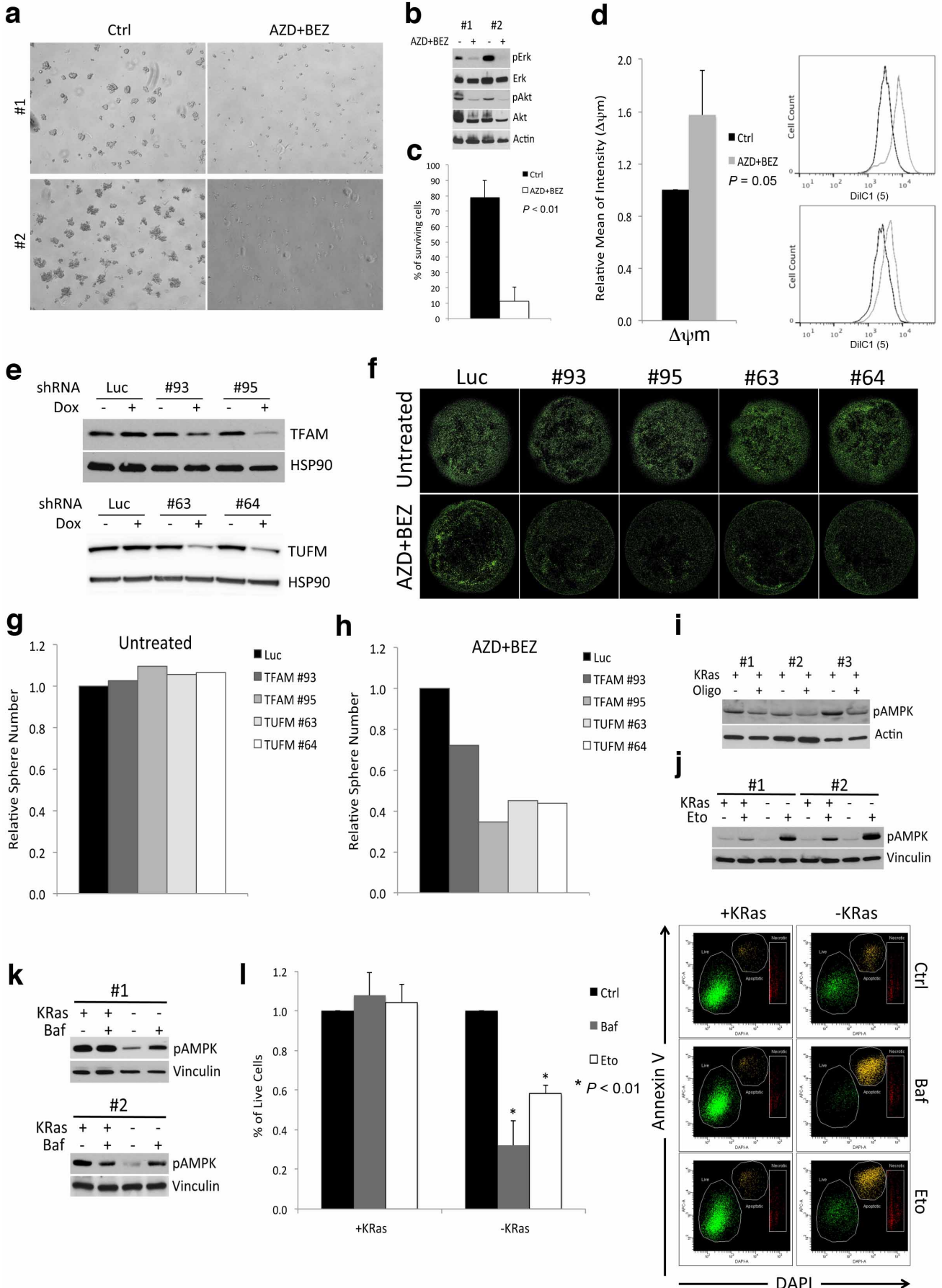
malate (h) in KRas-expressing (+) and SCs (-) following steady-state tracing (36 h labelling) with uniformly carbon-13-labelled substrates: glucose, glutamine, palmitate and pyruvate ($n = 3$). Data are mean \pm s.d.



Extended Data Figure 8 | Differential sensitivity of tumour cells to

OXPPOS inhibition. **a**, Annexin V staining of cells treated with oligomycin 200 nM (Oli) for 24 h shows a significant decrease in viability in surviving cells (−KRas). By contrast, control cells expressing KRas (+KRas) are minimally affected ($n = 3$); a representative flow-cytometry analysis is reported. **b**, Effect of oligomycin (Oli), dicyclohexylcarbodiimide (DCCD), veturicidin (Vent), rotenone (Rot), antimycin (Anti) and DMSO (Ctrl) on spherogenic potential of KRas-expressing (+KRas) and SCs (−KRas) ($n = 4$). **c**, *In vivo* treatment experimental scheme: mice were transplanted with tumour cells and fed with doxycycline in drinking water (+KRas, +Dox) until they developed tumours of 1 cm in diameter. Then doxycycline was withdrawn (−KRas, −Dox) and after 2 weeks, when tumours were regressed, mice were treated with oligomycin (0.5 mg kg^{-1} , i.p.) or vehicle for 5 days a week, for 2 weeks. After treatment, KRas was re-induced (+Dox) and mice were monitored for tumour relapse. **d**, One dose of oligomycin (0.5 mg kg^{-1} , i.p.) is sufficient to increase lactate concentration in plasma of treated mice after 4 h from injection. Oligo, oligomycin; Ctrl, vehicle. $n = 4$. **e**, Tumour volume of KRas-expressing

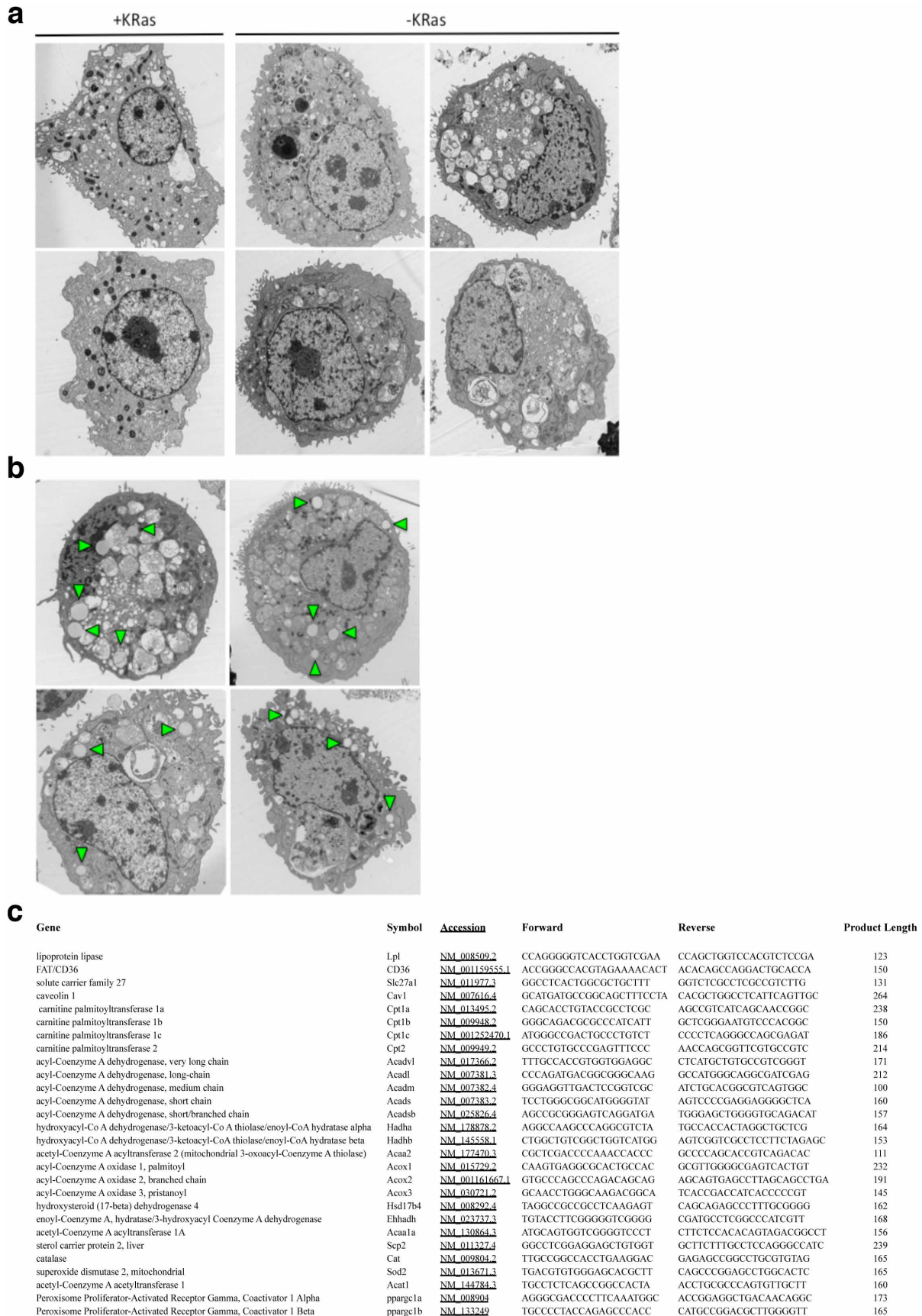
tumours treated with either vehicle or oligomycin 0.5 mg kg^{-1} , 5 days a week, for 2 weeks. Treatment was started when tumours reached 5 mm of diameter (5 mice per group). **f**, SCs after treatment with oligomycin show signs of degeneration and epithelial remnants change their morphology. Red arrows indicate the presence of capillaries (red blood cells) indicating that regressed tumours are vascularized ($\times 40$). **g**, Oligomycin (Oli) induces ROS production in KRas-expressing cells (+KRas) and SCs (−KRas). Its effect is even stronger than that of positive control 4-hydroxynonenal (hne) ($n = 3$). **h**, Glutathione levels in KRas-expressing cells (+KRas) and SCs (−KRas) before and after buthionine sulphoximine (BSO) treatment. Glutathione is increased in SCs and BSO treatment is effective in reducing its level ($n = 3$). **i**, Effect of glutathione depletion on spherogenic potential of KRas-expressing cells (+KRas) and SCs (−KRas) ($n = 3$). **j**, ROS production in SCs after treatment with 4-hydroxynonenal (hne) and oligomycin (oli) in the presence or absence of antioxidants: α -tocopherol (vitE), *N*-acetylcysteine (nac) and tetrakis (Tet) ($n = 2$). **k**, Effect of oligomycin on spherogenic potential of surviving cells pre-treated with antioxidants ($n = 4$). Data are mean \pm s.d.



Extended Data Figure 9 | Effect of mitochondrial downregulation in human tumour spheres and metabolic stress mediated by inhibition of autophagy.

a, Effects of the combination of AZD8330 and BEZ235 (AZD+BEZ) on human tumour spheres. Some cells, usually doublets, are able to survive the treatment ($\times 5$). **b**, Immunoblots of human tumour spheres treated or not with AZD+BEZ probed with anti-phospho-p42/44 (pErk), total-Erk (Erk), phospho-Akt (pAkt), Akt and β -actin (Actin) antibodies; two independent tumours are reported. **c**, Annexin V staining of treated (AZD+BEZ) and control (Ctrl) cells after 4 days of treatment ($n = 3$). **d**, Mitochondrial transmembrane potential ($\Delta\psi_m$) of untreated (Ctrl) and treated (AZD+BEZ) human spheres with AZD8330 and BEZ235 for 7 days ($n = 3$); representative flow-cytometry analysis of two tumours. **e–h**, TFAM and TUFM were downregulated using two inducible short hairpin RNAs (shRNAs) each (TFAM: #93, #95; TUFM: #63, #64) in human spheres expressing KRas (untreated) and cells surviving 1 week of treatment with AZD8330 and BEZ235 (AZD+BEZ); after 5 days of shRNA induction cells

were replated for evaluating their spherogenic capacity. **e**, Immunoblots of tumour spheres after 72 h of shRNA induction (+Dox) probed with anti-TFAM, TUFM and HSP90 antibodies. **f**, Representative calcein staining after sphere replating. **g, h**, Effects of downregulation of TFAM and TUFM on spherogenic potential of untreated (**g**) and treated (**h**) cells; data represent the average of two independent human tumours. **i**, Immunoblot of KRas-expressing cells treated or not with oligomycin 200 nM (Oligo, +/-) probed with anti-Thr-172-phospho-AMPK and actin antibodies. **j, k**, Immunoblots of +KRas and -KRas cells treated with etomoxir (Eto, 100 μ M for 6 h) (**j**) and bafilomycin (Baf, 50 nM for 24 h) (**k**) probed with anti-Thr-172-phospho-AMPK and vinculin antibodies. **l**, Annexin V staining of cells treated for 48 h with bafilomycin 50 nM (Baf) and etomoxir 100 μ M (Eto) clearly shows a significant decrease in viability of SCs (-KRas). Control cells expressing KRas (+KRas) are not affected ($n = 3$); representative dot-plots are reported. Data are mean \pm s.d.



Extended Data Figure 10 | Cells surviving oncogene ablation are engorged with autophagosomes and lysosomes and contain lipid droplets. **a**, SCs (-KRas) have a cytoplasm full of phagosomes and autophagosomes, a feature absent in KRas-expressing cells (+KRas) (TEM; $\times 7,500$). **b**, SCs are

characterized by the presence of several lipid droplets (arrowheads) in the cytoplasm (TEM; $\times 7,500$). **c**, Primers used for amplification of mitochondrial and lipid metabolic genes.

Conditional tolerance of temperate phages via transcription-dependent CRISPR-Cas targeting

Gregory W. Goldberg¹, Wenyan Jiang¹, David Bikard^{1†} & Luciano A. Marraffini¹

A fundamental feature of immune systems is the ability to distinguish pathogenic from self and commensal elements, and to attack the former but tolerate the latter¹. Prokaryotic CRISPR-Cas immune systems defend against phage infection by using Cas nucleases and small RNA guides that specify one or more target sites for cleavage of the viral genome^{2,3}. Temperate phages include viruses that can integrate into the bacterial chromosome, and they can carry genes that provide a fitness advantage to the lysogenic host^{4,5}. However, CRISPR-Cas targeting that relies strictly on DNA sequence recognition provides indiscriminate immunity both to lytic and lysogenic infection by temperate phages⁶—compromising the genetic stability of these potentially beneficial elements altogether. Here we show that the *Staphylococcus epidermidis* CRISPR-Cas system can prevent lytic infection but tolerate lysogenization by temperate phages. Conditional tolerance is achieved through transcription-dependent DNA targeting, and ensures that targeting is resumed upon induction of the prophage lytic cycle. Our results provide evidence for the functional divergence of CRISPR-Cas systems and highlight the importance of targeting mechanism diversity. In addition, they extend the concept of ‘tolerance to non-self’ to the prokaryotic branch of adaptive immunity.

Clustered, regularly interspaced, short palindromic repeat (CRISPR) loci and their CRISPR-associated (*cas*) genes function together as a prokaryotic adaptive immune system which can protect bacteria and archaea from invading genetic elements such as viruses (phages)^{2,3}. Repeat elements of the CRISPR locus are intercalated with short ‘spacer’ sequences which typically match phage or plasmid genomes and dictate the targets for immunity on the basis of sequence identity. Active immunity requires transcription of the CRISPR locus, followed by cleavage of the transcript within repeat sequences by Cas endoribonucleases^{7,8}. This liberates small CRISPR RNAs (crRNAs) which specify the target for RNA-guided Cas nucleases that defend the cell from infection by degrading invading genomes^{2,3}. Furthermore, CRISPR-Cas systems can behave in an adaptive fashion through what appears to be direct acquisition of new spacer sequences from invading elements, thereby conferring sequence-specific, heritable immunity. On the basis of *cas* gene content and organization, CRISPR-Cas systems have been classified into three main types (I–III) and at least 12 subtypes (A–F)⁹. A growing body of work indicates a common role for these systems in antiviral defence, but the diversity of crRNA processing and targeting mechanisms^{2,3} that have been described suggest the potential for differences in their precise functions *in vivo*.

It is well established that CRISPR-Cas systems can tolerate ‘self’ spacer elements within the CRISPR locus DNA via sequence discrimination at the flanking repeats. For type I and type II systems, this requires that short sequences which license targeting, known as protospacer adjacent motifs, are absent from the repeat sequences flanking each spacer^{10,11}. For type III systems, targeting is prevented by excessive base pairing between the repeat-derived crRNA tag and its corresponding DNA sequence¹². Tolerance to ‘non-self’ DNA elements, on the other hand, has yet to be described. Previous reports indicate that active CRISPR-Cas systems and their targets cannot co-exist in the same cell^{6,13,14}. Thus, CRISPR-Cas targeting that relies strictly on DNA sequence recognition does not offer the flexibility to accommodate genetic elements with ambiguous fitness

costs, such as temperate phages. Upon infection, temperate phages can kill the host cell by initiating a lytic cycle, but they may also spare the cell from lysis and establish a lysogenic cycle, typically via repression of lytic genes and integration into the host chromosome as a so-called prophage¹⁵. In addition to preventing lysis, lysogenization can result in a variety of phenotypic outcomes which can improve host fitness, for example via expression of non-viral ‘moron’ genes carried on temperate phage genomes^{4,5}. The lysis/lysogeny decision is generally governed by a central promoter region which responds to stochastic and environmental factors to control transcription in divergent directions, thereby promoting one or the other infection cycle¹⁵. Under certain conditions, the prophage can re-initiate a lytic cycle and excise from the chromosome—a process referred to as prophage induction. Commitment to either the lytic or lysogenic cycle does not involve changes in the viral genome sequence. Hence, it is generally accepted that CRISPR-Cas targeting of temperate phages should exclude both infection outcomes; in addition to preventing lysis, CRISPR attack of an integrated prophage target precludes stable lysogenization. Although this appears to be the case for type I–E (ref. 6) and II–A (ref. 16) CRISPR-Cas systems, the potential for tolerance during type III immunity had not been explored.

To investigate the behaviour of type III CRISPR immunity during temperate phage infection, we introduced pGG3, a plasmid carrying the type III–A CRISPR-Cas system of *S. epidermidis* RP62a (ref. 17), into *Staphylococcus aureus* RN4220 (ref. 18). This strain is sensitive to the lambda-like temperate phages of *S. aureus* Newman, a clinical isolate harbouring four heteroimmune prophages (ΦNM1–4) which carry genes that enhance the pathogenicity of their host¹⁹. We also identified a spacer in one of the CRISPR loci of *S. aureus* MSHR1132 (ref. 20) with near-perfect identity to a conserved target sequence present in ΦNM1 (Fig. 1a), ΦNM2 and ΦNM4. This spacer, referred to as 32T (Supplementary Table 1), was added to the CRISPR locus of pGG3. Using ΦNM1, we first established that this spacer prevents lytic infection by showing that plaquing efficiency is reduced approximately seven orders of magnitude compared with a strain carrying the pGG3 plasmid without the ΦNM1-targeting spacer (Fig. 1b). We then introduced an erythromycin resistance gene (*ermC*) into ΦNM1 to facilitate quantification of lysogens that have stably integrated a chromosomal prophage (creating ΦNM1-Erm^R). Using this system, we expected to find results consistent with a report describing CRISPR-mediated immunity to lysogenization by phage lambda in *E. coli*⁶. Surprisingly, we obtained the same efficiency of lysogenization compared with the control strain lacking spacer 32T (Fig. 1c). To test whether the presence of mismatches between the 32T crRNA and its target was influencing this phenomenon, we engineered spacer 32T* with a perfect match to its target, but obtained the same results (Fig. 1b, c). We next sought to determine whether genetic CRISPR-Cas inactivation is responsible for the apparent tolerance of these lysogens by testing them for sensitivity to ΦNM2. All 14 clones maintained resistance to ΦNM2 mediated by spacer 32T (Extended Data Fig. 1a, d). Finally, we demonstrated that spacer 32T tolerance does not result from genetic alteration of the target phage (Extended Data Fig. 1b, c). Tolerance was also observed for ΦNM4 (Extended Data Fig. 2), demonstrating that the tolerance phenomenon is not specific for the ΦNM1-Erm^R phage or its integration

¹Laboratory of Bacteriology, The Rockefeller University, New York, New York 10065, USA. †Present address: Synthetic Biology Group, Institut Pasteur, 28 Rue du Dr. Roux, 75015 Paris, France.

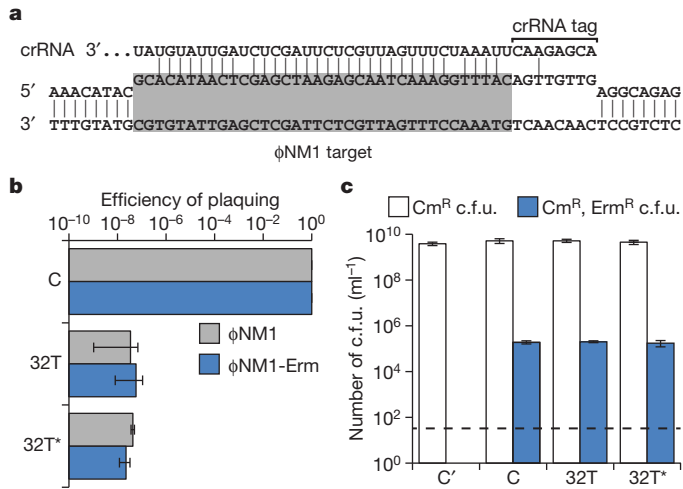


Figure 1 | Type III-A CRISPR immunity can block lytic infection but tolerate lysogenization. **a**, Base pairing interaction between crRNA 32T and its target in the Φ NM1 genome (highlighted in grey). The crRNA tag is a sequence transcribed from the CRISPR repeat that needs to be unpaired with the flanking region of the target to license immunity. The target gene is transcribed from left to right. **b**, CRISPR immunity against Φ NM1 infection provided by spacers 32T and 32T* (similar to 32T but without mismatches), measured as a decrease in the number of plaque-forming units (p.f.u.) with respect to the non-targeting control pGG3 (C). **c**, Lysogenization with Φ NM1-Erm^R in the presence of spacers 32T and 32T* or the pGG3 control (C), measured as the number of chloramphenicol- and erythromycin-resistant colony-forming units (c.f.u.) per milliliter obtained after infection. Control cells lysogenized with Φ NM1 (C') lack the *ermC* insertion and do not yield erythromycin-resistant colony-forming units. Error bars, mean \pm s.d. ($n = 3$).

locus. These results demonstrate that type III-A CRISPR immunity can block lytic infection but tolerate lysogenization without concomitant genetic CRISPR-Cas inactivation or alteration of the phage genome.

To determine whether prophage tolerance is a spacer-specific phenomenon, we designed a variety of spacers with 100% target identity, targeting different regions of the Φ NM1 genome on both strands (Fig. 2a). We first tested the ability of each spacer to prevent lytic infection (Fig. 2b). Surprisingly, spacer functionality varied with the predicted transcriptional context of each target sequence. Spacers matching putative lytic genes to the right of the central promoter which are predicted to be unidirectionally transcribed were only effective when they targeted the predicted non-template strand (top strand according to our spacer nomenclature). Meanwhile, transcription is predicted to be bi-directional to the left of the central promoter¹⁹. Spacers targeting this region prevented plaque formation regardless of the strand targeted. This resembled the activity reported for the type III-B CRISPR-Cas system of the archaeon *Sulfolobus islandicus* REY15A, where immunity to plasmid transformation depended on the presence of promoters flanking a target sequence²¹. We thus reasoned that transcription-dependent targeting could explain the discrepancies in spacer functionality. Indeed, Φ NM1 transcription profiles assessed by RNA sequencing of RN4220 cultures 6, 15, 30 and 45 min after infection revealed predominantly unidirectional transcription to the right of the central promoter, while bi-directional transcription was detected to the left of the central promoter (Extended Data Fig. 3).

Further evidence for the transcription-dependence of type III-A CRISPR-Cas targeting was obtained via the characterization of a spacer 2B CRISPR-escape mutant phage, Φ NM1 γ 6, exhibiting a clear plaque phenotype characteristic of phages that cannot establish lysogeny (Fig. 2c, inset). Sanger sequencing of the spacer 2B target sequence did not reveal any mutations in the target or flanking sequences (data not shown); thus, we measured the Φ NM1 γ 6 plaquing efficiency with other spacers to determine whether it possessed a sequence-independent, general CRISPR-escape phenotype (Fig. 2c). Although most spacers provided immunity

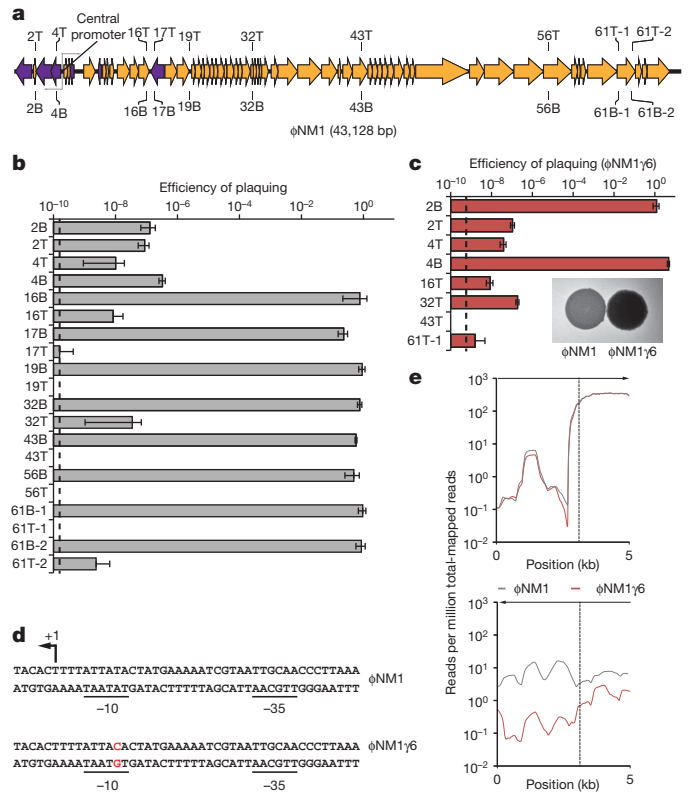


Figure 2 | Transcription of target sequences is required for type III-A CRISPR immunity. **a**, The Φ NM1 genome and the position of targets used in this study. T, crRNA anneals to the top strand; B, bottom strand. Grey arrows represent the Φ NM1 central promoter driving divergent transcription. **b**, Immunity against Φ NM1 infection provided by spacers targeting the phage regions shown in **a**. Dotted line indicates the limit of detection for the assay. **c**, Immunity against Φ NM1 γ 6 infection. Inset, comparison of lysis phenotypes for Φ NM1 (turbid) and Φ NM1 γ 6 (clear), representative of four technical replicates. **d**, Leftward promoter consensus sequences at the Φ NM1 and Φ NM1 γ 6 central promoter. The Φ NM1 γ 6 mutation in the -10 element is shown in red. The putative transcription start site is noted (+1). **e**, Comparison of phage transcription profiles from cells infected with Φ NM1 (grey line) or Φ NM1 γ 6 (red line), 15 min post-infection. Phage-derived transcripts are plotted in reads per million total-mapped reads relative to their position on the genome; arrows indicate the direction of transcription plotted in each graph; the vertical dotted line marks the position of the central promoter. Error bars, mean \pm s.d. ($n = 3$).

against Φ NM1 γ 6, we identified one additional spacer, 4B, which was escaped by the mutant phage. Both the 2B and 4B spacers target the same strand in the lysogenization operon to the left of the central promoter. Importantly, the two complementary spacers (2T and 4T) targeting the opposite strand of spacers 2B and 4B were not escaped by Φ NM1 γ 6, indicating that the 2B/4B escape phenotype did not result from changes to the target DNA per se. Consistent with this, we did not observe differences in the Φ NM1 and Φ NM1 γ 6 plaquing efficiency when targeting the 4B region via Cas9-mediated type II-A CRISPR immunity (Extended Data Fig. 4a), which was shown to cleave double-stranded DNA even in the absence of target transcription^{22,23}. We thus reasoned that the Φ NM1 γ 6 type III-A CRISPR-escape and clear-plaque phenotypes could result from a localized, unidirectional reduction in transcription, for example, leftward from the central promoters. Indeed, *de novo* sequencing of Φ NM1 γ 6 revealed a single nucleotide polymorphism in a crucial residue of the leftward -10 element (Fig. 2d) of the central promoters, immediately upstream of the SAPPV1_g4 *cl*-like repressor gene required for lysogenic establishment, and ~ 1700 base pairs (bp) away from the 2B target sequence. Encouraged by this result, we directly assessed Φ NM1 γ 6 transcription profiles using RNA sequencing, 6 and 15 min after infection

(Extended Data Fig. 5a). Consistent with our hypothesis, leftward transcription (Fig. 2e, lower panel) of the lysogenization operon 15 min after infection was strongly reduced, while rightward transcription (Fig. 2e, upper panel) in this region was relatively unchanged. Taken together, these findings suggest that transcription across target sequences is a requirement for type III-A CRISPR immunity. Previously reported²⁴ strand-independent immunity against plasmids in *S. epidermidis* may also follow this rule, as bi-directional transcription was detected across targets (Extended Data Fig. 6).

Given that temperate phages silence transcription of their lytic genes during lysogeny¹⁵, we hypothesized that transcription-dependent targeting would allow a variety of otherwise effective spacers to tolerate prophage target sequences. The corollary to this prediction is that targets which are constitutively transcribed during lysogeny (for example, leftward from the central promoter; Extended Data Fig. 5b) should not be tolerated. Five additional spacers were tested for their ability to tolerate lysogenization by Φ NM1-Erm^R (Fig. 3a). As expected, lysogenization was tolerated by spacers 17T, 43T and 61T-1 targeting lytic genes to the right of the central promoter, but not by spacers 2B and 17B targeting genes constitutively expressed leftwards during lysogeny. Consistent with the notion that type III-A tolerance results from differences in transcription at the target, we did not observe tolerance with the transcription-independent type II-A CRISPR-Cas system, even when targeting a lytic gene (Extended Data Fig. 4b–h). To rule out the possibility that type III-A tolerance is influenced by processes that occur during phage infection, we corroborated these results using a ‘reverse’ CRISPR immunity assay by electroporating CRISPR plasmids into pre-existing lysogens. In this experiment, transformation with a targeting CRISPR-Cas system results in attack of chromosomal prophage sequences and an inability to form

colonies¹⁴. Again, results varied with the transcriptional context of target sequences in a manner consistent with our previous findings (Fig. 3b). Collectively, these results demonstrate that prophages are not intrinsically tolerated during CRISPR immunity in staphylococci, and suggest rather that type III-A tolerance is only achieved during lysogeny under the condition that transcription is silenced at target sequences.

To demonstrate definitively that transcription-dependent targeting offers a biological mechanism for conditional tolerance, we integrated the Φ NM1 target sequence for spacers 43T and 43B into the chromosome of *S. aureus* RN4220 under the control of a tightly regulated tetracycline-inducible promoter, thus emulating target lysogenization. The target was placed in both orientations with respect to the inducible promoter (Fig. 3c) and with respect to the chromosomal origin of replication (Extended Data Fig. 7a). The resulting strains were then transformed with the spacer 43T or 43B plasmids in a reverse CRISPR immunity assay, and plated in the absence or presence of the inducer. CRISPR immunity was only achieved when transcription across the target was induced with anhydrotetracycline (ATc) in the presence of an antisense crRNA, regardless of the target’s orientation (Fig. 3d and Extended Data Fig. 7a). Once again, we confirmed this finding to be a type-III-specific phenomenon by transforming the strains from Fig. 3c with the spacer 43B-III type II-A CRISPR plasmid targeting the same region (Extended Data Fig. 7b). We corroborated this result by following the growth of spacer 43T transformants in liquid media (Fig. 3e). Upon addition of the inducer, growth was only inhibited for cells with the target in the forward orientation for which spacer 43T produces an antisense crRNA. Importantly, tolerance achieved in the absence of the inducer did not appear to affect growth (Fig. 3e, dotted lines). Finally, having established that type III-A CRISPR-Cas systems can block lytic infection but tolerate lysogenization, we examined the effect of tolerant spacers on prophage induction of Φ NM1 lysogens in culture. Compared with a spacerless lysogen control, the phage titre resulting from spontaneous induction of overnight cultures was significantly lower for lysogens harbouring a tolerant spacer (Fig. 4a). We next followed the growth of cultures induced directly with the DNA-damaging agent mitomycin C (Fig. 4b, solid lines). While the spacerless lysogen control cultures succumbed to prophage induction, the presence of a tolerant spacer prevented lysis.

We report here that type III-A immunity can offer conditional tolerance to ‘non-self’ genetic elements, in this case, temperate phages. This has several important implications for the CRISPR-Cas system and its host population. Tolerance helps ensure the genetic stability of the CRISPR-Cas system, since selective pressure to integrate prophages in the presence of intolerant spacers can drive genetic CRISPR-Cas inactivation (Extended Data Fig. 4c–e)—similar to what occurs during plasmid uptake²⁵.

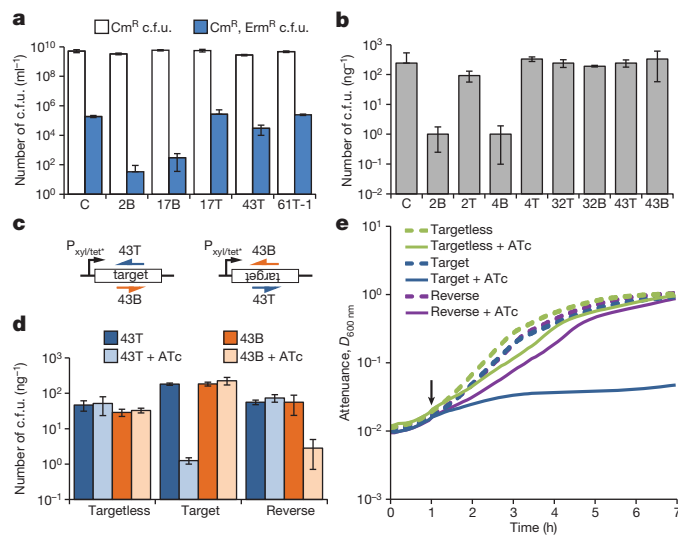


Figure 3 | Conditional tolerance is achieved via transcription-dependent CRISPR-Cas targeting. **a**, Φ NM1-Erm^R lysogenization for additional spacers. C, pGG3 non-targeting control. **b**, Transformation of Φ NM1-lysogenic competent cells with CRISPR-Cas plasmids containing different spacers (transformation efficiency is measured as colony-forming units per nanogram of plasmid DNA). C, pGG3 non-targeting control. **c**, Integration of the 43T/B Φ NM1 target region into the chromosome of *S. aureus*. Target sequences (inserted in both forward and reverse orientations) are under the control of the tetracycline-inducible promoter P_{xyl/tet^*} . The 43T/B crRNAs are shown annealing to either the top or bottom strands. **d**, Transformation of both strains shown in **c**, as well as an isogenic control strain lacking the target insertion, with CRISPR-Cas plasmids containing spacers 43T or 43B. Transformants were plated on selective plates with or without ATc for induction of the P_{xyl/tet^*} promoter. **e**, Growth curve of strains shown in **d** expressing the spacer 43T CRISPR-Cas system, in the presence or absence of ATc addition at the indicated time point (black arrow). Error bars, mean \pm s.d. ($n = 3$).

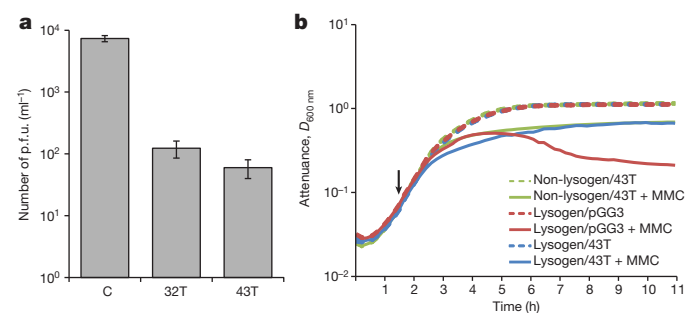


Figure 4 | Prophage induction is limited by type III-A CRISPR immunity in tolerant lysogens. **a**, Plaque-forming potential (measured in plaque-forming units per millilitre) of supernatants from overnight cultures of Φ NM1 lysogens carrying the tolerant spacer 32T or 43T CRISPR plasmids, or the pGG3 non-targeting control (C). **b**, Growth curve of Φ NM1-Erm^R lysogens or a non-lysogen control harbouring the pGG3 or spacer 43T CRISPR plasmids as indicated, with or without the addition of the prophage-inducing agent mitomycin C (MMC) at the indicated time point (black arrow). Error bars, mean \pm s.d. ($n = 3$).

In other words, tolerant spacers ensure that a population can sample potentially beneficial phenotypes that result from prophage integration without compromising their CRISPR-mediated immunity. Tolerance may also be particularly vital for type III systems, which were recently shown to provide immunity despite up to 15 mismatches with their spacer²⁶. Thus, without the potential for phages to evade targeting readily via point mutation that is seen for type I and type II systems^{10,11}, the transcriptional dependence of type III-A targeting offers temperate phages an alternative route to lysogenization that need not provide selection for mutants. Furthermore, tolerant lysogens had the added potential of resisting lysis via prophage induction. Although a few spacers targeting the lysogenization functions did not provide tolerance, it is important to note that these genes only constitute a small portion of the phage genome. Hence, spacers targeting this region should be acquired less frequently, even if spacer acquisition occurs randomly without an additional mechanism for distinguishing tolerant from intolerant spacers during acquisition.

The requirement for transcription across target sequences during type III-A immunity contrasts with the transcription-independent targeting reported for type I (ref. 27) and type II (refs 22, 23) CRISPR-Cas systems. Given the temporal pattern of target transcription observed during the phage lytic cycle, it might be expected that CRISPR targeting of late genes would not provide immunity if the cell's survival is already compromised at the onset of targeting. Indeed, we observed some differences in spacer effectiveness when infecting cells in liquid culture at very high multiplicity of infection (MOI; ~100): spacers targeting late genes were less protective (Extended Data Fig. 8). However, this effect was not pronounced at an MOI of 10, suggesting that the system is generally robust to delays in target transcription, in accordance with what we observed in efficiency of plaquing assays. Consistent with our findings at large, our survey of sequenced staphylococcal type III spacers showed that naturally acquired spacers with known target sequences produced crRNAs complementary to the non-template strand of predicted ORFs in nine out of ten cases (Supplementary Table 2). This bias suggests negative selection on non-functional spacers targeting template strands. Alternatively, type III systems may utilize an unknown mechanism to discriminate template and non-template strands during spacer acquisition. The molecular details of the transcription-dependent targeting mechanism remain unclear. Preliminary experiments indicate that the presence of a transcript provided *in trans* is not sufficient to license DNA targeting (Extended Data Fig. 9). Consistent with this, induction of transcription across a plasmid-borne target results specifically in loss of the targeted plasmid (Extended Data Fig. 10). Hence, transcription *in cis* is probably required for DNA targeting. One possibility is that negative supercoiling generated in the wake of a passing transcription bubble could facilitate target DNA melting and improve crRNA recognition²⁷. This hypothesis, however, would not account for the template and non-template strand asymmetry observed in our system. Another possibility is that exposure of the target non-template strand within the transcription bubble is required for annealing of a crRNA²¹. In this scenario, the observed asymmetry might be explained by occlusion of base-pairing to the template strand by either the nascent transcript or the RNA polymerase²⁸. However, given the small size of the transcription bubble, this explanation may also be incomplete. Alternatively, transcription may be required to activate a targeting mechanism rather than facilitating target recognition or binding per se. In this case, effective targeting could require base-pairing potential between the crRNA and the nascent transcript *in cis*, which would be absent for crRNAs with complementarity to the template strand. Previous reports^{29,30} have revealed specific RNA target cleavage activity for type III(-B/C) systems that could be relevant to the transcription-dependent DNA targeting observed in staphylococci. Further experimentation will be needed to clarify these points. In summary, our work expands the repertoire of CRISPR-based immune functions to include a novel capacity for conditional tolerance of foreign elements, and establishes distinct genetic outcomes resulting from immunity to temperate phages via divergent CRISPR-Cas targeting mechanisms.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 13 January; accepted 2 July 2014.

Published online 31 August 2014.

1. Belkaid, Y. & Hand, T. W. Role of the microbiota in immunity and inflammation. *Cell* **157**, 121–141 (2014).
2. Barrangou, R. CRISPR-Cas systems and RNA-guided interference. *Wiley Interdiscip. Rev. RNA* **4**, 267–278 (2013).
3. Sorek, R., Lawrence, C. M. & Wiedenheft, B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Annu. Rev. Biochem.* **82**, 237–266 (2013).
4. Brüßow, H., Canchaya, C. & Hardt, W. D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol. Mol. Biol. Rev.* **68**, 560–602 (2004).
5. Cumby, N., Davidson, A. R. & Maxwell, K. L. The moron comes of age. *Bacteriophage* **2**, 225–228 (2012).
6. Edgar, R. & Qimron, U. The *Escherichia coli* CRISPR system protects from lambda lysogenization, lysogens, and prophage induction. *J. Bacteriol.* **192**, 6291–6294 (2010).
7. Brouns, S. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
8. Carte, J., Wang, R., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22**, 3489–3496 (2008).
9. Makarova, K. S. *et al.* Evolution and classification of the CRISPR-Cas systems. *Nature Rev. Microbiol.* **9**, 467–477 (2011).
10. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190**, 1390–1400 (2008).
11. Semenova, E. *et al.* Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl Acad. Sci. USA* **108**, 10098–10103 (2011).
12. Marraffini, L. A. & Sontheimer, E. J. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**, 568–571 (2010).
13. Bikard, D., Hatoum-Aslan, A., Mucida, D. & Marraffini, L. A. CRISPR interference can prevent natural transformation and virulence acquisition during *in vivo* bacterial infection. *Cell Host Microbe* **12**, 177–186 (2012).
14. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnol.* **31**, 233–239 (2013).
15. Johnson, A. D. *et al.* lambda Repressor and cro-components of an efficient molecular switch. *Nature* **294**, 217–223 (1981).
16. Nozawa, T. *et al.* CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. *PLoS ONE* **6**, e19543 (2011).
17. Hatoum-Aslan, A., Samai, P., Maniv, I., Jiang, W. & Marraffini, L. A. A ruler protein in a complex for antiviral defense determines the length of small interfering CRISPR RNAs. *J. Biol. Chem.* **288**, 27888–27897 (2013).
18. Kreiswirth, B. N. *et al.* The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* **305**, 709–712 (1983).
19. Bae, T., Baba, T., Hiramatsu, K. & Schneewind, O. Prophages of *Staphylococcus aureus* Newman and their contribution to virulence. *Mol. Microbiol.* **62**, 1035–1047 (2006).
20. Holt, D. C. *et al.* A very early-branching *Staphylococcus aureus* lineage lacking the carotenoid pigment staphyloxanthin. *Genome Biol. Evol.* **3**, 881–895 (2011).
21. Deng, L., Garrett, R. A., Shah, S. A., Peng, X. & She, Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* **87**, 1088–1099 (2013).
22. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl Acad. Sci. USA* **109**, E2579–E2586 (2012).
23. Jinek, M. *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
24. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**, 1843–1845 (2008).
25. Jiang, W. *et al.* Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet.* **9**, e1003844 (2013).
26. Manica, A., Zebec, Z., Steinkellner, J. & Schleper, C. Unexpectedly broad target recognition of the CRISPR-mediated virus defence system in the archaeon *Sulfolobus solfataricus*. *Nucleic Acids Res.* **41**, 10509–10517 (2013).
27. Westra, E. R. *et al.* CRISPR immunity relies on the consecutive binding and degradation of negatively supercoiled invader DNA by Cascade and Cas3. *Mol. Cell* **46**, 595–605 (2012).
28. Nudler, E. RNA polymerase active center: the molecular engine of transcription. *Annu. Rev. Biochem.* **78**, 335–361 (2009).
29. Hale, C. R. *et al.* RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945–956 (2009).
30. Zhang, J. *et al.* Structure and Mechanism of the CMR Complex for CRISPR-Mediated Antiviral Immunity. *Mol. Cell* **45**, 303–313 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We acknowledge T. Bae for providing strains; B. Levin, J. Modell and P. Samai for discussion of the paper, and A. Hatoum-Aslan, as well the Torres laboratory (New York University), for their optimization of RNA preparation protocols for *S. aureus*. We also thank J. Chen of the Novick laboratory for advice on the

construction of temperate phages with a selectable marker. Finally, we acknowledge The Rockefeller University Genomics Resource Center core facility for performing the next-generation sequencing in this work. L.A.M. is supported by the Searle Scholars Program, the Rita Allen Scholars Program, an Irma T. Hirsch Award, a Sinsheimer Foundation Award and a National Institutes of Health Director's New Innovator Award (1DP2AI04556-01).

Author Contributions G.W.G. and L.A.M. designed experiments. Research was done by G.W.G. W.J. constructed the pWJ40 and pWJ153 plasmids and performed the

plasmid-curing experiment. D.B. constructed plasmid pDB184, assisted with phage *de novo* assembly and provided the reads per million normalization script for RNA sequencing data. G.W.G. and L.A.M. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to L.A.M. (marraffini@rockefeller.edu).

METHODS

Bacterial strains and growth conditions. Cultivation of *S. aureus* RN4220 (ref. 18), TB4 (ref. 19) and derivative strains was done in TSB media (BD) at 37 °C, except when phage infections were performed, or when otherwise noted (see below). Whenever applicable, media were supplemented with chloramphenicol at 10 µg ml⁻¹ to ensure CRISPR plasmid maintenance. RN4220 strains harbouring pCL55-derived insertion vectors were grown similarly, but kanamycin was provided at 25 µg ml⁻¹ except during re-culture for competent cell preparation. *E. coli* DH5α was grown in LB Broth (BD) supplemented with kanamycin at 25–50 µg ml⁻¹ to maintain pCL55-derived plasmids. Selection for ΦNM1-Erm lysogens with resistance to erythromycin (10 µg ml⁻¹) was only applied during the lysogenization protocol as described below, and, where applicable, during the subsequent ΦNM2 sensitivity assays.

Estimation of phage lysate titres. Serial dilutions were prepared in triplicate and plated on soft agar lawns of RN4220 in HIB-agar (BD) supplemented with 5 mM CaCl₂ (technical replicates). Plates were incubated at 37 °C for 16–24 h after drying at room temperature (25 °C).

DNA preparation and cloning. Plasmid DNA was purified from 2 to 6 ml of *E. coli* DH5α or *S. aureus* RN4220 overnight cultures. For preparation from *S. aureus* cultures, cells were pelleted, re-suspended in 100 µl TSM buffer (50 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 0.5 M sucrose) then treated with 5 µl lysostaphin (2 mg ml⁻¹) at 37 °C for 1.5 h before treatment with plasmid miniprep reagents from Qiagen. Purification used Qiagen or EconoSpin columns.

Cloning used RN4220 electrocompetent cells unless otherwise stated. For most type III CRISPR plasmids, scarless addition of repeat-spacer units to the pGG3 parent vector was accomplished by ‘round-the-horn’ PCR (ref. 31) followed by blunt ligation, using common primer oGG12 and spacer-specific oligonucleotides listed in Supplementary Table 3. The pGG3 vector was itself constructed by ‘round-the-horn’ PCR using primers L55 and A10 to remove extraneous repeat-spacer elements from the pWJ30β (ref. 32) CRISPR array. For construction of the remaining type III CRISPR plasmids, a modified parent vector (pGG3-BsaI) was created by introducing a placeholder spacer harbouring two BsaI restriction sites, to facilitate scarless cloning of spacers by replacement with annealed oligonucleotide pairs possessing BsaI-compatible overhangs. Type III-A CRISPR arrays were amplified with primers L50/L6 and sequenced by Sanger using either forward or reverse primers. The BsaI cloning method was also used to construct type II CRISPR plasmids from the pDB184 parent vector, a modified version of pWJ40 with only the single placeholder spacer. The pC194-derived pWJ40 vector contained the full *S. pyogenes* MIGAS type II CRISPR-Cas system and was constructed by amplifying *S. pyogenes* genomic DNA with oligonucleotides L362/W278 and pC194 with oligonucleotides W270/W282, followed by digestion of the PCR products with BglII and BssSI and a subsequent ligation. Type II CRISPR arrays were amplified with primers L448 and W176, and sequenced by Sanger using L448. After the cloning of each spacer, plasmid sizes were verified by restriction digest with BssSI for type III plasmids or BtgI for type II plasmids. pDB184 was created via Gibson assembly of two PCR fragments: a pWJ40 backbone amplified using primers B220/B334, and a CRISPR array amplified from pCas9 (ref. 14) using primers L448/B333.

For construction of pCL55-derived inducible target vectors, cloning used chemically competent DH5α cells. Briefly, the chloramphenicol resistance cassette was first replaced with a kanamycin resistance cassette amplified from strep LAM202-3 using primers L484/L485. This was accomplished by ‘round-the-horn’ PCR on the pCL55-iTET parent vector using primers L482/L483, followed by blunt ligation with the PCR-amplified resistance cassette to create the new pKL55-iTET-B parent vector. Directionality of the insertion was verified afterwards by restriction digest with BtgI. Modification of the P_{xyI/tet} promoter in accordance with pRAB12 (ref. 33) architecture was achieved via two consecutive overlap PCR steps to introduce point mutations using oligonucleotide pairs oGG108/oGG109 and oGG110/oGG111, followed by a ‘round-the-horn’ PCR step and blunt ligation to introduce the downstream operator sequence using oligonucleotides oGG112 and oGG113. The resulting pKL55-iTET-RC12 vector harbouring the P_{xyI/tet*} modifications was used for downstream manipulations, as well as integration into the RN4220 chromosome to create the ‘Targetless’ control strain. For forward and reverse target insertions, annealed oligonucleotide pairs (oGG124/oGG125 and oGG126/oGG127, respectively) with appropriate overhangs were ligated into the multiple cloning site after digesting the vector with BglII and SacII restriction enzymes (NEB). Target insertions were verified by PCR amplification and Sanger sequencing using the primers oGG64 and oGG88. Inversion of the *attP* motif for both forward and reverse target vectors was achieved by Gibson assembly of two PCR fragments, using oligonucleotides oGG102/oGG103 for the *attP* motif and oGG104/oGG105 for the backbone. Directional integration into the RN4220 chromosome was verified by amplification of either the *attL* or *attR* junctions using primer pairs oGG50/oGG96 and oGG51/oGG96, respectively. The pWJ153-inducible target vector is a pKL55-iTET-RC12- and pE194- (ref. 34) derived plasmid constructed via multiple steps of either

‘round-the-horn’ PCR³¹ followed by blunt ligation or Gibson assembly. The full sequence is provided as Supplementary Sequence 1.

Construction of the ΦNM1-Erm^R lysogen was achieved via pKOR allelic exchange³⁵. Homology arms (~1 kb) were amplified from the chromosome of *S. aureus* RN4220: ΦNM1 using primer pairs oGG181/oGG182 and oGG185/oGG186, while the ~1.25 kb *ermC* resistance cassette was amplified from a pE194 plasmid preparation using primers oGG183 and oGG184. An ~3.25 kb fragment was assembled by SOEing PCR³⁶ using external primers oGG181 and oGG186 with clonase (QuikChange) *attB* adapters that allowed directional integration into the pKOR vector³⁵. Sequence integrity of the ~3.25 kb insertion was verified by Sanger using primers L29, oGG191, oGG192, W277 and L325.

Preparation of electrocompetent *S. aureus* cells. *S. aureus* RN4220, TB4 or derivative strains were grown overnight in TSB medium, diluted 1:100 in fresh medium without antibiotics, then allowed to grow to an attenuation (*D*_{600nm}) reading of 0.8–1.0 for RN4220 or 0.7–0.9 for TB4. Measurements were taken using a NanoDrop 2000c Spectrophotometer (Thermo Scientific) and disposable polystyrene cuvettes. After re-culture, cells were pelleted at 4 °C, and two or three washes were performed using chilled, sterile dH₂O or 10% glycerol. Cells were ultimately re-suspended in 1/100th volume of chilled, sterile 10% glycerol and 50 µl aliquots were distributed for storage at –80 °C.

Efficiency of plaquing assays. High-titre lysates (~10¹² p.f.u. ml⁻¹) of either ΦNM1, ΦNM1γ6 or ΦNM2 were serially diluted in triplicate and applied to soft agar lawns of RN4220 strains harbouring CRISPR plasmids, including pGG3 or pDB184 spacerless control lawns infected in parallel (technical replicates). Plates were incubated at 37 °C for 18 h. After incubation, plates were monitored at bench top for up to 24 h to facilitate quantification of plaque-forming units.

Quantification of erythromycin-resistant lysogens. Overnight cultures of RN4220 with respective CRISPR plasmids were inoculated in triplicate from single colonies in HIB medium supplemented with chloramphenicol (biological replicates). After chilling at 4 °C, 1:10 dilutions were prepared in 1 ml fresh HIB supplemented with chloramphenicol and 5 mM CaCl₂. Diluted cultures were infected with ΦNM1-Erm at ~MOI 10 and incubated on ice for 30 min. After incubation on ice, cultures were transferred to a 37 °C incubator for 30 min with shaking. Serial dilutions from each culture were then applied to HIB-agar plates supplemented with chloramphenicol, erythromycin and 5 mM CaCl₂ for quantification of lysogenic colony-forming units. In selected cases, type III-A CRISPR locus and target sequence integrity was verified by colony PCR after re-streaking single colonies using primer pairs L6/L50 (CRISPR array) and oGG25/oGG26 (ORF 2) or oGG38/oGG39 (ORF 32). Where applicable, Sanger sequencing of PCR products was also performed using these primers. When verifying type II lysogenization isolates, the spacer 43B-III target region was amplified using primers oGG233 and oGG234, and the type II CRISPR array was amplified using L448 and W176. The presence of integrated ΦNM1 or ΦNM1-Erm^R prophages was confirmed by colony PCR using primer pairs oGG191/W277 and oGG206/W276 to amplify the *attL* and *attR* junctions, respectively. To estimate the total number of recipient cells, serial dilutions of untreated overnight cultures were plated on TSB- or HIB-agar supplemented with chloramphenicol.

ΦNM2-sensitivity assay. High-titre lysate of ΦNM2 (~10¹² p.f.u. ml⁻¹) was applied to the surface of a pre-dried HIB-agar plate supplemented with 5 mM CaCl₂ and appropriate antibiotics, then allowed to dry for an additional ~30 min at room temperature. Single colonies isolated from Erm lysogeny experiments or CRISPR plasmid transformations were streaked through the ΦNM2-seeded region using a sterile plastic loop then incubated for ~12 h at 37 °C.

Enumeration of plaque-forming units liberated from lysogen cultures. Overnight cultures of either RN4220:ΦNM1-Erm or RN4220:ΦNM1 lysogens harbouring targeting CRISPR plasmids or non-targeting control plasmids were inoculated in triplicate from single colonies in HIB media supplemented with chloramphenicol (biological replicates). After overnight growth, cells were transferred to 4 °C then pelleted by centrifugation at 4696g for 5 min. Supernatants were filtered, and 100 µl from each lysate was mixed with 100 µl of either an indicator strain or targeting strain overnight culture for plating by the soft agar method. After drying at room temperature (25 °C), plates were incubated 18 h at 37 °C.

Screen for lipase-negative ΦNM4 lysogens. An overnight culture of *S. aureus* TB4 harbouring the spacer 32T CRISPR plasmid was re-cultured to log phase growth in HIB medium supplemented with 5 mM CaCl₂. After measurement of *D*_{600nm}, cells were treated with ΦNM4 at ~MOI 50. After incubation for 1 h, cells were plated on TSA supplemented with 5% egg yolk emulsion. After ~24 h incubation at 37 °C, approximately 1,000 colonies were inspected for lipase secretion. Two lipase-negative candidates were re-streaked to single colonies, and the presence of an integrated ΦNM4 prophage was confirmed by colony PCR using primers oGG50 and oGG96 to amplify the *attL* junction.

Phage DNA isolation and deep sequencing. Samples of high titre phage lysates (~10¹² p.f.u. ml⁻¹) were treated with DNase and RNase to a final volume of 150 µl for 1 h at 37 °C. Samples were treated with EDTA (pH 8.0) to a final concentration

of 20 mM, followed by treatment with SDS to a final concentration of 0.5% and 2 µl proteinase K. Samples were incubated for 1 h at 65 °C, then subjected to a PCR purification protocol (Qiagen). Paired-end library preparation was performed on purified phage DNA using a Nextera Tagmentation protocol (Illumina), and samples were pooled for multiplexed sequencing on a MiSeq (Illumina). *De novo* assembly of phage genomes used ABySS³⁷.

RNA preparation for reverse transcription PCR and RNA sequencing. For reverse transcription PCR (RT-PCR), overnight cultures were diluted 1:20 in 25 ml fresh media and grown for 2.5 h at 37 °C with shaking. After re-culture, cells were pelleted and washed twice in 1 ml ice cold TSM buffer, then treated with 3 µl lyso-staphin (2 mg ml⁻¹) for 20 min at 37 °C in 500 µl TSM buffer. Treated cells were pelleted then re-suspended in 750 µl cold TRIzol Reagent (Life Technologies) after discarding the supernatant. The following chloroform extraction and precipitation was performed according to the manufacturer's protocol. After resuspension in dH₂O, samples were treated with Qiagen DNase I for 45 min at 30 °C, then re-purified using RNeasy Cleanup columns (Qiagen). In some cases, it was necessary to repeat this step a second time to ensure the complete removal of DNA. After cleanup, all samples were again treated with DNase I (Sigma-Aldrich) for 30–45 min at 30 °C, before use in the reverse transcription reaction.

For RNA sequencing, overnight cultures were diluted 1:100 in fresh HIB supplemented with chloramphenicol and 500 µM CaCl₂, and grown for 1.5 h (approximately mid-log phase) at 37 °C with shaking. Cultures were removed, infected at MOI ≈ 20 then split into 10 ml portions for an additional 6, 15, 30 or 45 min of growth. Immediately following incubation, samples were mixed with 10 ml of a 1:1 acetone/ethanol solution and transferred to -80 °C. The ΦNM1 lysogen was grown similarly, except without antibiotics, and harvested immediately after the 1.5 h re-culture at 37 °C. After at least one overnight at -80 °C, samples were thawed on ice and pelleted by centrifugation at 4696g for 10 min. After two washes of 1 ml TE buffer, cells were re-suspended in 1 ml RLT buffer (Qiagen) supplemented with BME, and transferred to 2-ml tubes pre-loaded with ~0.5–1 cm³ of 0.1 mm glass beads (BioSpec). Samples were processed in a Mini-Beadbeater instrument (BioSpec) three times for 10 s at 4,200 oscillations per minute, with 40 s of chilling on ice between runs. After bead-beating, samples were spun down for 2 min at 16,100g in a refrigerated microcentrifuge. Supernatant (750 µl) was transferred to a clean tube for mixing with 500 µl of 100% ethanol, and the following RNeasy purification was done according to the manufacturer's protocol (Qiagen). After elution, samples were treated with either Qiagen or Sigma-Aldrich DNase I for 30–45 min at 30 °C, then re-purified using RNeasy cleanup columns. In some cases, it was necessary to repeat this step a second time to ensure the complete removal of DNA. rRNA-depleted samples were subsequently generated using the RiboZero Magnetic Kit for bacteria (Epicentre), according to the manufacturer's protocol.

RT-PCR. Reverse transcription used M-Mulv Reverse Transcriptase (NEB), with DNA-free total RNA isolated from RN4220 cultures harbouring either the pNes(wt-*d*) or pNes(wt-*i*) plasmids as templates for cDNA synthesis. For pNes(wt-*d*), reverse transcription used either the L8 or L86 primers in two separate 30 µl reactions, alongside mock reactions (-RT enzyme). For pNes(wt-*i*), the same was done using primers L8 or L87. After incubation, 1 µl of each reaction was used as a template for PCR, with respective primer pairs for each sample.

Phage transcriptome analysis and visualization. Reads were aligned to reference genomes using Bowtie and sorted using Samtools. Using a custom script, sorted reads were accessed via Pysam, normalized as reads per million values, and plotted in log scale as the average over consecutive windows of 500 base pairs using matplotlib tools for IPython.

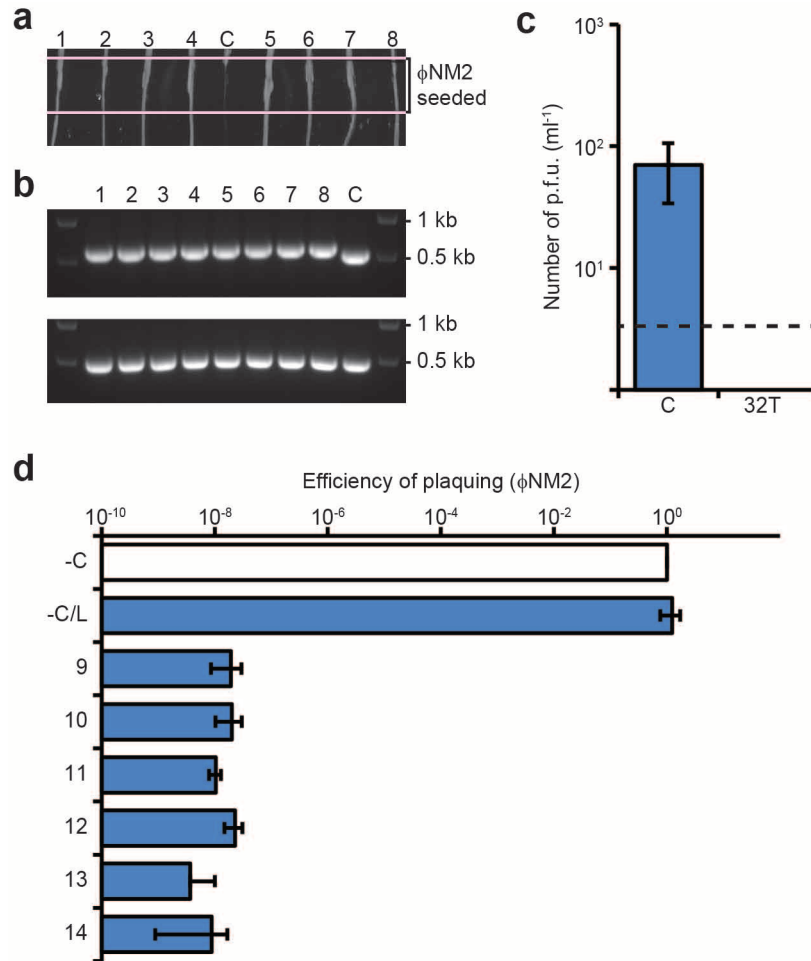
Transformation assays. *S. aureus* RN4220 plasmid preparations were dialysed on 0.025 µm nitrocellulose filters (Millipore) then quantified using a NanoDrop 2000c

Spectrophotometer (Thermo Scientific). Aliquots (50 µl) of electrocompetent cells were transformed in triplicate with 80 ng dialysed DNA per transformation using a GenePulser Xcell (BioRad) with the following parameters: 2900 V, 25 µF, 100 Ω, 2 mm (technical replicates). After electroporation, cells were immediately re-suspended in TSB to a final volume of 200 µl and recovered at 30 °C for 2 h with shaking. Serial dilutions were then prepared before plating with appropriate antibiotics. For reverse CRISPR immunity assays targeting insertion vectors, additional plating in the presence of ATc at a final concentration of 0.5 µg ml⁻¹ was performed in parallel using the same dilutions. Plates were incubated at 37 °C for 18–24 h.

Plate reader growth curves. For ATc induction experiments, overnight cultures were launched from single colonies in triplicate and diluted 1:200 in TSB broth (biological replicates). After 1 h of growth, ATc was added at a final concentration of 0.5 µg ml⁻¹ where applicable. Measurements were taken every 5 min. For mitomycin C induction experiments, overnight cultures were launched from single colonies in duplicate and diluted 1:100 in HIB broth (biological replicates). After 1.5 h of growth, mitomycin C was added at a final concentration of 0.5 µg ml⁻¹ where applicable. Measurements were taken every 10 min. For ΦNM1 infections, overnight cultures were launched from single colonies in triplicate and diluted 1:100 in HIB broth supplemented with CaCl₂ 5 mM (biological replicates). After 1 h 25 min of growth, *D*_{600 nm} was measured for three representative cultures to estimate MOI. Aliquots were then loaded into 96-well plates along with ΦNM1 at the appropriate MOI (10 or 100), where applicable. Measurements were taken every 5 min. For ΦNM1γ6 infections, overnight cultures were launched from single colonies in triplicate and diluted 1:200 in HIB supplemented with CaCl₂ 5 mM (biological replicates). An average *D*_{600 nm} was measured after 1 h of growth, and ΦNM1γ6 was added at an MOI of 10 on the basis of this value, where applicable. Measurements were taken every 5 min.

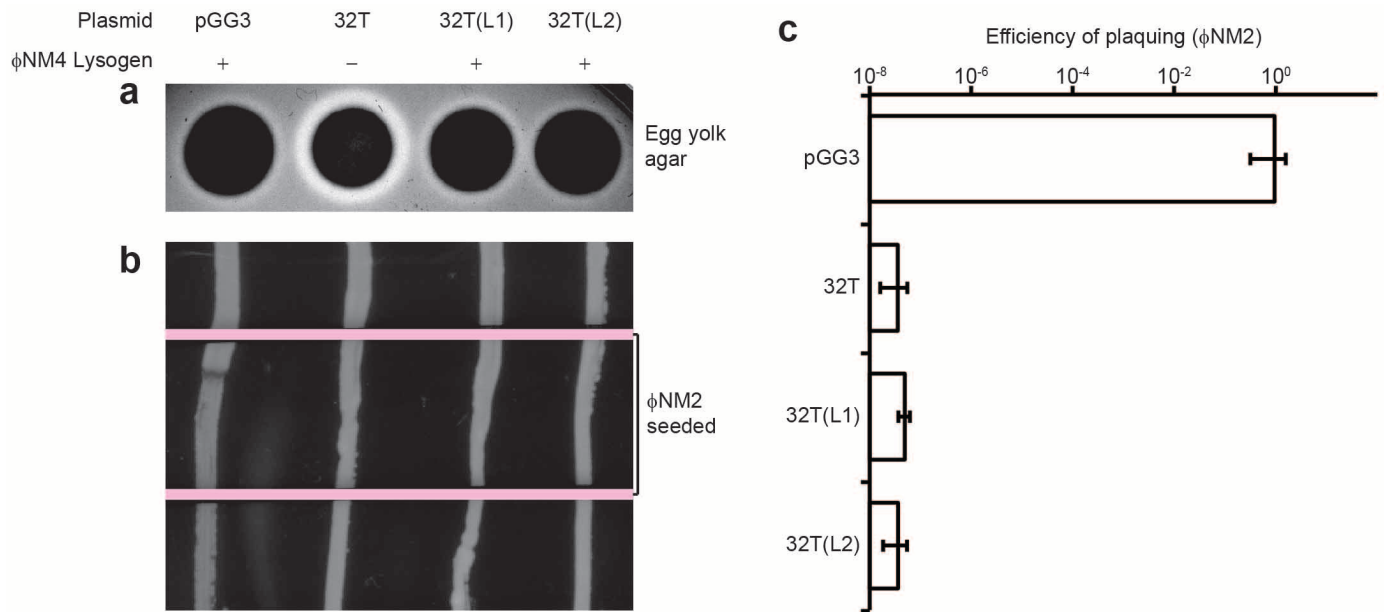
Plasmid-curing assay. RN4220 cells harbouring both the pGG3 CRISPR-Cas plasmid and the pWJ153 target plasmid were cultured in TSB supplemented with chloramphenicol (10 µg ml⁻¹) to a *D*_{600 nm} of 0.45. After splitting the culture in two, transcription across the target was induced for one of the cultures via the addition of ATc to a final concentration of 0.25 µg ml⁻¹. Aliquots of cells were harvested before (0) and after (1, 2, 3, 4, 5 and 6 h) the time of induction. After purification of DNA, plasmids were linearized with the common single cutter BamHI and subjected to agarose gel electrophoresis. In parallel, serial dilutions of both cultures were prepared in triplicate for each time point and plated on TSA plates supplemented with chloramphenicol and erythromycin or chloramphenicol alone, for quantification of antibiotic-resistant colony-forming units (technical replicates).

- Moore, S. D. & Prevelige, P. E. Jr A P22 scaffold protein mutation increases the robustness of head assembly in the presence of excess portal protein. *J. Virol.* **76**, 10245–10255 (2002).
- Hatoum-Aslan, A., Maniv, I., Samai, P. & Marraffini, L. A. Genetic Characterization of antiplasmid immunity through a type III-A CRISPR-Cas system. *J. Bacteriol.* **196**, 310–317 (2014).
- Helle, L. *et al.* Vectors for improved Tet repressor-dependent gradual gene induction or silencing in *Staphylococcus aureus*. *Microbiology* **157**, 3314–3323 (2011).
- Horinouchi, S. & Weisblum, B. Nucleotide sequence and functional map of pE194, a plasmid that specifies inducible resistance to macrolide, lincosamide, and streptogramin type B antibiotics. *J. Bacteriol.* **150**, 804–814 (1982).
- Bae, T. & Schneewind, O. Allelic replacement in *Staphylococcus aureus* with inducible counter-selection. *Plasmid* **55**, 58–63 (2006).
- Horton, R. M. In vitro recombination and mutagenesis of DNA: SOEing together tailor-made genes. *Methods Mol. Biol.* **15**, 251–261 (1993).
- Simpson, J. T. *et al.* ABySS: a parallel assembler for short read sequence data. *Genome Res.* **19**, 1117–1123 (2009).



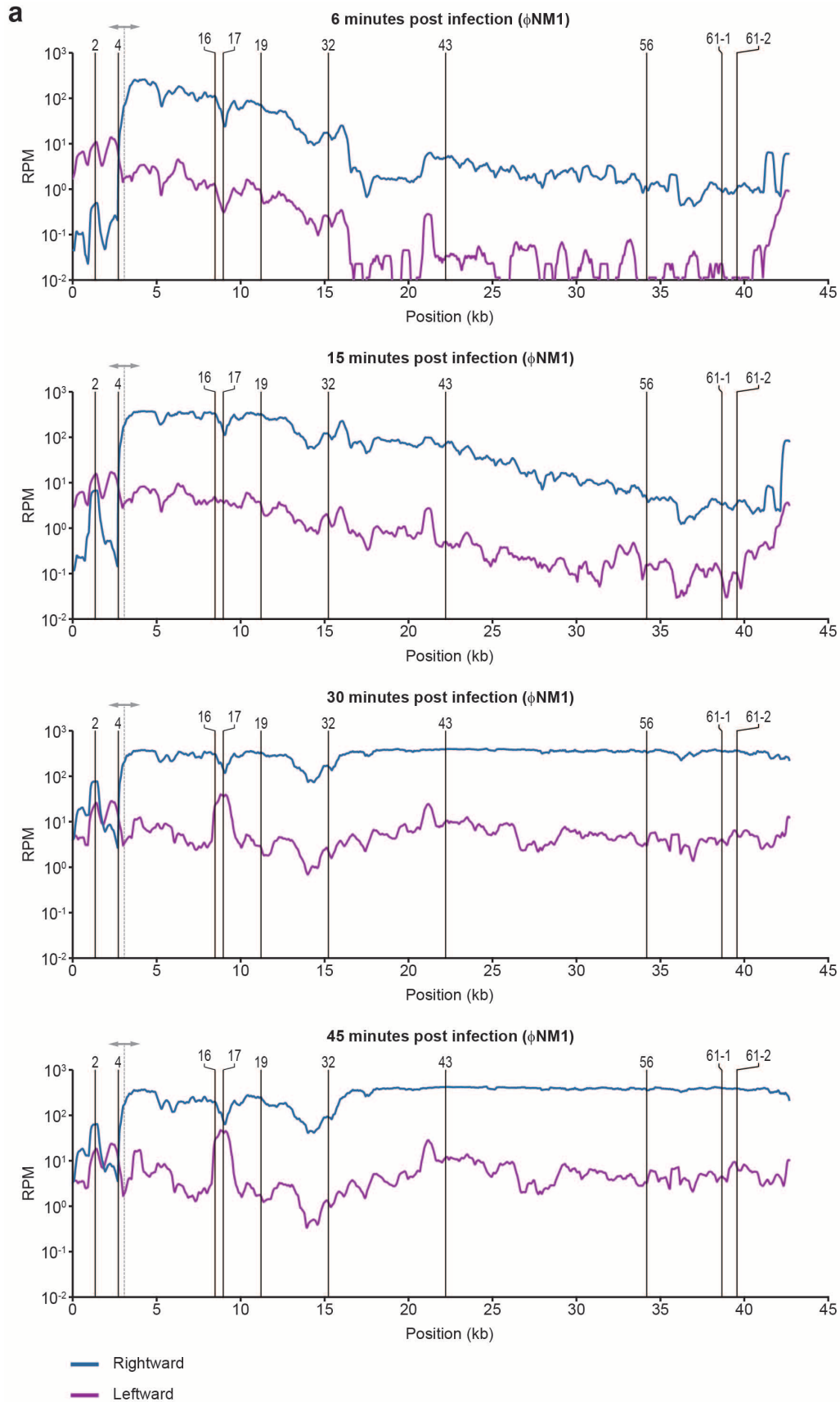
Extended Data Figure 1 | Characterization of spacer 32T isolates lysogenized with Φ NM1-Erm^R. **a**, Φ NM2 sensitivity assay. Eight randomly selected Φ NM1-Erm^R lysogen clones were re-streaked through the indicated Φ NM2-seeded region from top to bottom (1–8); C, sensitive Φ NM1-Erm^R lysogen harbouring the pGG3 control plasmid. **b**, PCR amplification of the CRISPR array (upper panel) and spacer 32T target region (lower panel) for the strains tested in **a**. The pGG3 control lysogen (C) lacks a phage-targeting spacer in its CRISPR array. Size markers of 1 kb and 0.5 kb are indicated. All eight PCR products for the target region were sequenced by the Sanger method and no mutations were found (data not shown). **c**, Plaque-forming potential of

filtered supernatants from spacer 32T lysogen overnight cultures inoculated in triplicate. Plaque-forming units were enumerated on soft agar lawns of RN4220 harbouring either the pGG3 control (C) or spacer 32T CRISPR plasmids. Dotted line represents the limit of detection for this assay. **d**, Φ NM2 plaquing efficiency on soft agar lawns of an additional six randomly selected Φ NM1-Erm^R lysogen clones isolated during infection of RN4220/spacer 32T (9–14); a Φ NM1-Erm^R lysogen harbouring the pGG3 control plasmid was also tested (-C/L). Plaquing efficiency on the non-lysogenic indicator strain harbouring pGG3 is shown for comparison (-C). Error bars, mean \pm s.d. ($n = 3$). **a**, **b**, Single experiments performed for 8 of 32 isolates.



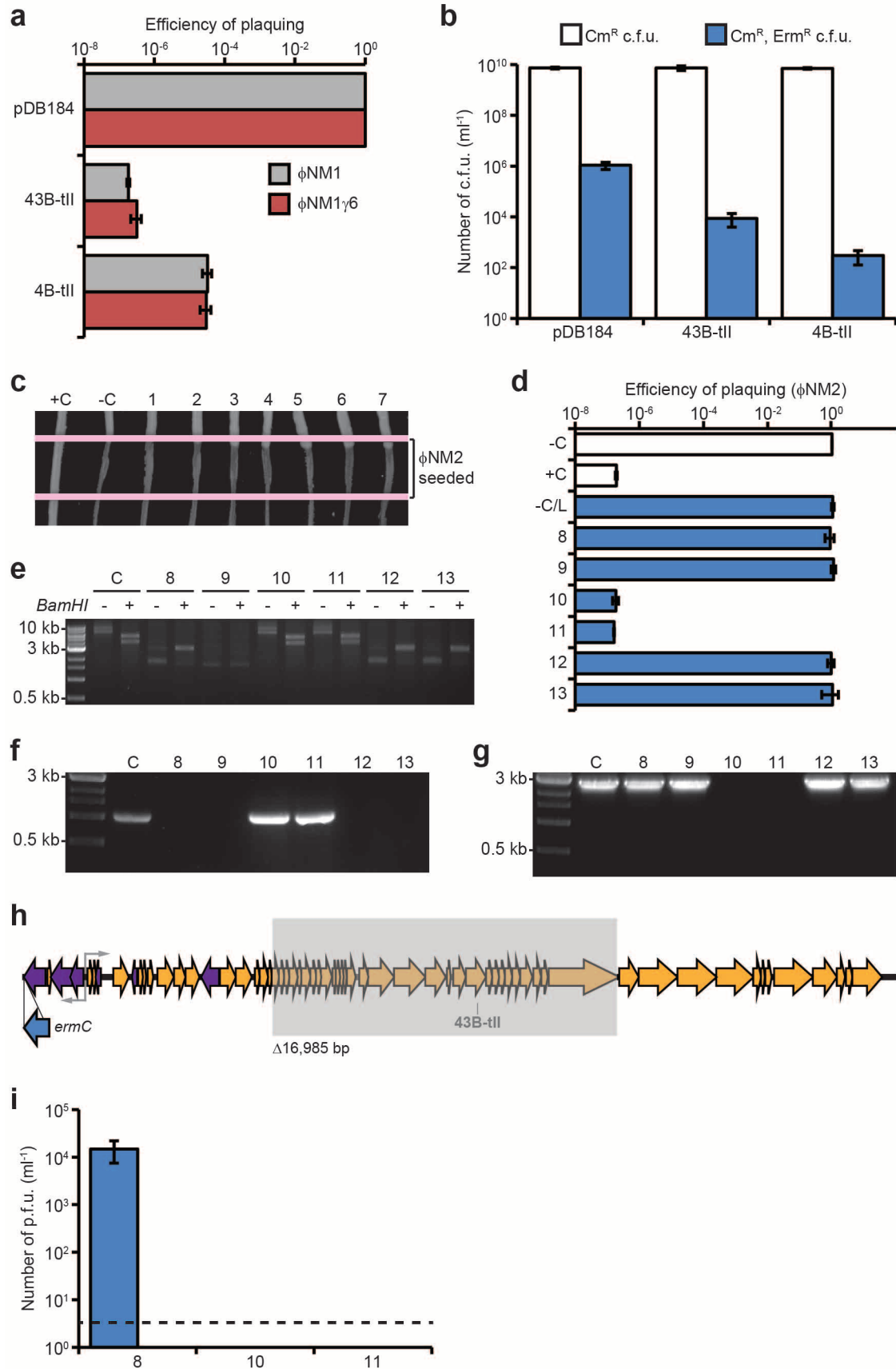
Extended Data Figure 2 | Characterization of spacer 32T isolates lysogenized with Φ NM4. **a**, Visualization of TB4-derived strains grown on egg-yolk agar. Integration of Φ NM4 within the *geh* locus of TB4 results in strongly reduced lipase secretion, enabling a screen for Φ NM4 lysogenization with spacer 32T. Right-most lanes display two lipase-negative isolates from the lysogenization screen; picture is representative of five technical replicates for each isolate. **b**, Φ NM2 sensitivity assay. Strains shown in **a** were re-streaked

through the indicated Φ NM2-seeded region from top to bottom. The pGG3 lysogen and spacer 32T non-lysogen in the two left-most lanes serve as sensitive and insensitive controls, respectively. Picture is representative of three technical replicates for each isolate. **c**, Φ NM2 plaquing efficiency on soft agar lawns of the strains analysed in **a** and **b**. 32T(L1) and 32T(L2) refer to the two Φ NM4 lysogens isolated during the spacer 32T egg-yolk screen. Error bars, mean \pm s.d. ($n = 3$).



Extended Data Figure 3 | Visualization of Φ NM1 transcription profiles 6, 15, 30 and 45 min after infection (MOI 20). Rightward and leftward expression values are plotted as blue and fuchsia lines, respectively, in reads per million (RPM). Position of relevant spacer targets are indicated with vertical solid lines. The dotted line with arrowheads marks the position of the central

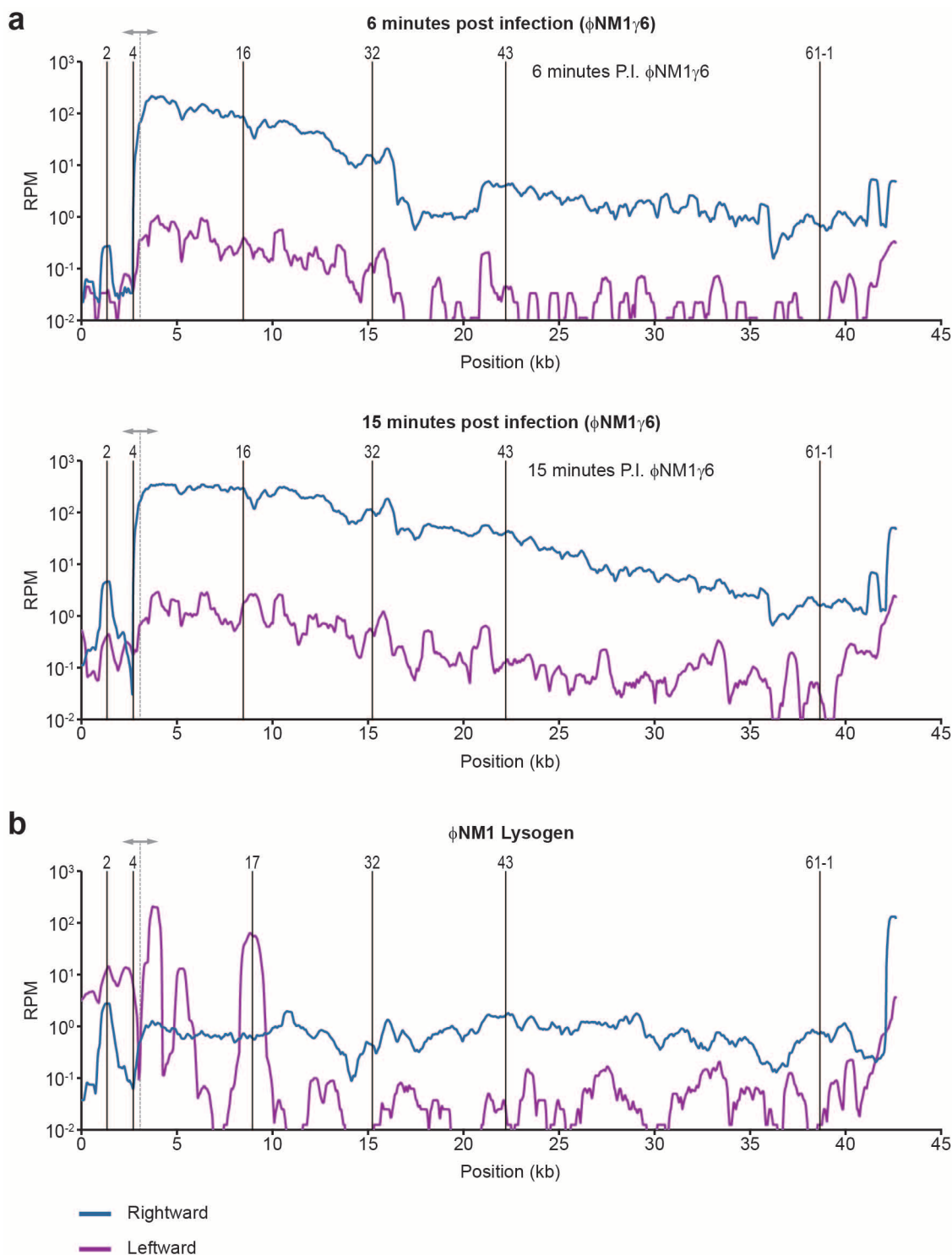
promoter. To improve readability, all curves were smoothed by plotting the average reads per million values over a 500 bp sliding-window. To the left of the central promoter, rightward expression is comparable to leftward expression by 30 min after infection, consistent with the strand-independent targeting observed for this region.



Extended Data Figure 4 | Type II CRISPR-Cas targeting in *S. aureus*

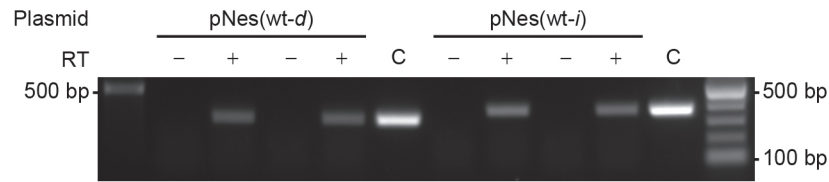
prevents both lytic and lysogenic infection. **a**, Plaquing efficiency of Φ NM1 and Φ NM1 γ 6 on lawns of RN4220 harbouring type II-A CRISPR-Cas plasmids as indicated. The parental vector, pDB184, serves as a non-targeting control. **b**, Φ NM1-Erm^R lysogenization of RN4220 harbouring either the spacer 43B-tII, 4B-tII, or non-targeting type II-A CRISPR plasmids. **c**, Φ NM2 sensitivity assay for seven randomly selected Φ NM1-Erm^R lysogen clones isolated during infection of RN4220/spacer 43B-tII (1–7). For comparison, a resistant non-lysogen harbouring the spacer 43B-tII plasmid and a sensitive lysogen harbouring the pDB184 plasmid were included as controls (respectively, C+ and C–). Picture represents a single experiment for 7 of 22 isolates. **d**, Φ NM2 plaquing efficiency on soft agar lawns for an additional six randomly selected Φ NM1-Erm^R lysogen clones isolated during infection of RN4220/spacer 43B-tII (8–13); a Φ NM1-Erm^R lysogen harbouring the pDB184 plasmid is also tested (-C/L). For comparison, plaquing efficiency of Φ NM2 on the non-lysogenic indicator strain harbouring pDB184 or the targeting spacer 43B-tII plasmid are also shown (-C and +C, respectively). **e**, Agarose gel electrophoresis of plasmid DNA purified from isolates 8–13 and the parental spacer 43B-tII strain (C). The symbols + or – indicate the presence or absence of treatment with the BamHI restriction enzyme, which produces two bands for the wild-type spacer 43B-tII plasmid: 5367 bp and 3972 bp. Size markers correspond to 10 kb, 3 kb and 0.5 kb bands of the 1 kb DNA ladder from NEB. **f**, Colony PCR spanning the type II CRISPR array for isolates 8–13. Spacer 43B-tII plasmid DNA was used as a template for the

control (C). Size markers of 3 kb and 0.5 kb are indicated. **g**, Colony PCR spanning the target region for isolates 8–13 and a Φ NM1-Erm^R lysogen harbouring the pDB184 control plasmid (C). Isolates 10 and 11 harbour identical deletions within the prophage that remove the target region (see below). Size markers of 3 kb and 0.5 kb are indicated. The presence of *attL* and *attR* prophage integration arms was also verified independently for each isolate using PCR (data not shown). **h**, Location of the 16,985 bp deletion identified within the prophage harboured by isolates 10 and 11 (shaded grey box). The location and orientation of the *ermC* insertion cassette is also shown (blue arrow). Deletion was mapped by primer walking. An ~9.1 kb product spanning the deletion was ultimately amplified using primers oGG6 and oGG241, and the deletion junction was sequenced by the Sanger method using oGG245. A perfect 14 bp direct repeat micro-homology flanks the deletion. **i**, Plaques-forming potential of overnight culture supernatants from isolates 8, 10 and 11. Supernatants were plated by the soft agar method with RN4220 cells harbouring the non-targeting pDB184 control plasmid as an indicator strain. Supernatants were also plated with spacer 43B-tII targeting lawns, yielding no detectable plaque-forming units. Isolate 8 appears to exhibit wild-type levels of spontaneous prophage induction (compare with pGG3 control in Fig. 4a). No plaque-forming units were detected from the supernatants of isolates 10 and 11 whatsoever, presumably resulting from their deletion of genes essential for prophage induction, including the ORF 43 major capsid protein. Dotted line represents the limit of detection for this assay. Error bars, mean \pm s.d. ($n = 3$). **e–g**, Single experiments for 6 of 22 isolates.



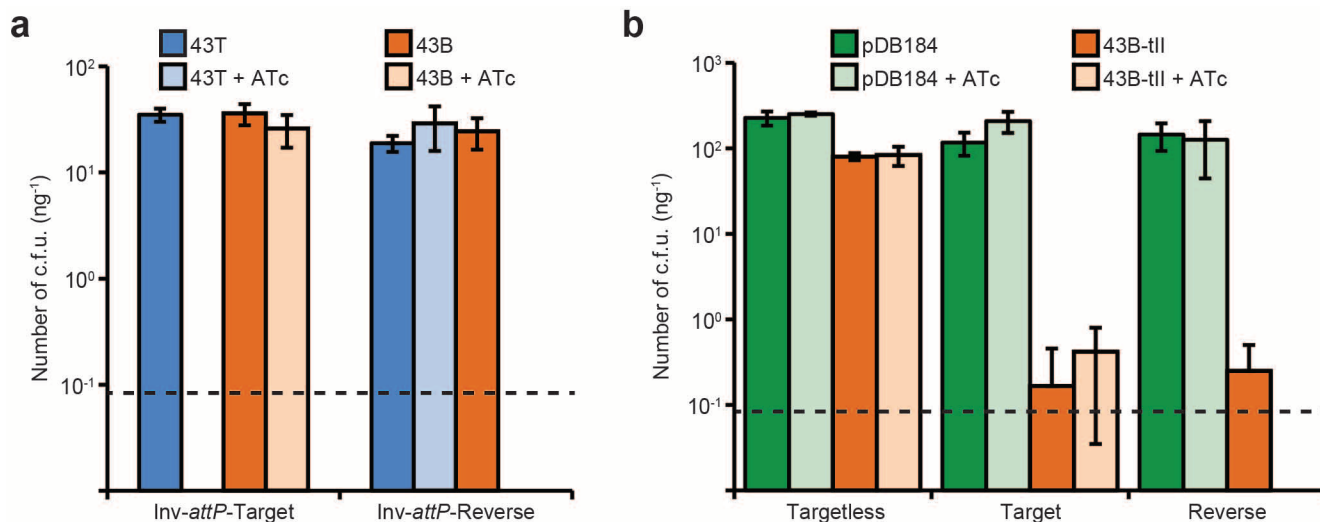
Extended Data Figure 5 | Visualization of transcription profiles for Φ NM1 γ 6 and the Φ NM1 prophage. Graphical presentation is the same as in Extended Data Figure 3. **a**, Φ NM1 γ 6 transcription profiles 6 and 15 min after infection (MOI 20). Comparison with Φ NM1 samples at equivalent time points (Extended Data Fig. 3) reveals a marked decrease in leftward transcription to the left of the central promoter region. We calculated the fold-change in reads per million between Φ NM1 and Φ NM1 γ 6 samples 15 min after infection. Leftward expression within the region bounded by the start of the genome and the central promoter was reduced 32-fold, while only a fourfold reduction in leftward expression was observed overall. Meanwhile, rightward expression

was reduced fourfold both overall and in this region. This suggests an approximately eightfold net reduction in leftward transcription originating from the central promoter. **b**, Φ NM1 prophage transcription profiles. Strong leftward transcription originates from the central promoter and a few upstream regions, which are presumed to be important for lysogenic maintenance. Rightward transcription was weaker than leftward transcription as expected, but not absent. Given the strength of rightward transcription observed during the lytic cycle (Extended Data Fig. 3), however, this transcription may originate from a subpopulation of cells undergoing prophage induction, rather than the stable lysogen majority.



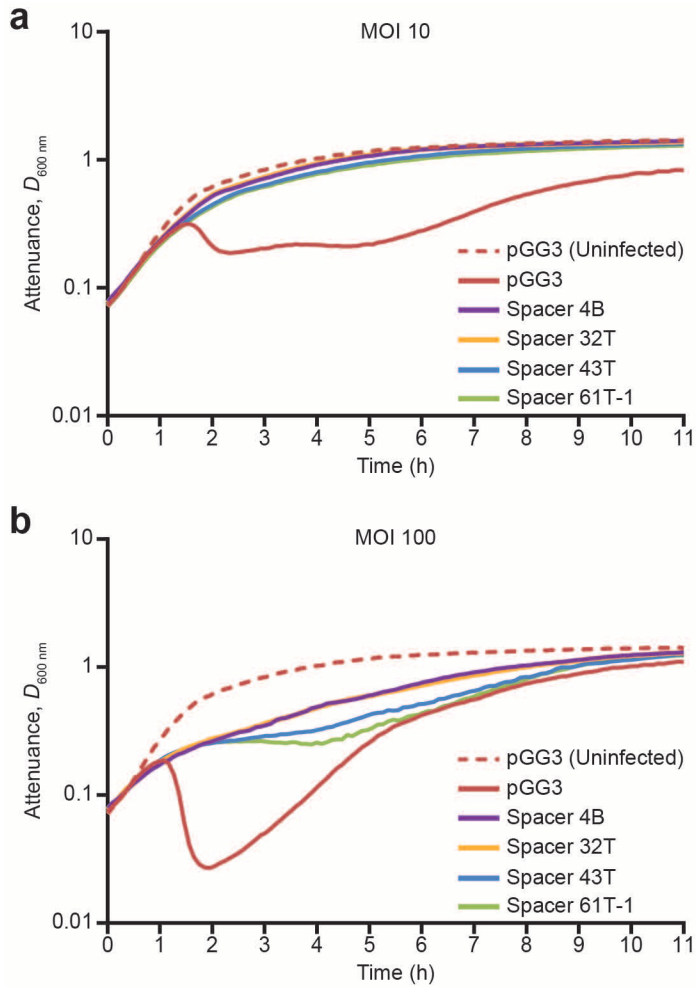
Extended Data Figure 6 | Detection of transcription across target insertions for the pNes(wt-d) and pNes(wt-i) plasmids. For each target plasmid²⁴, reverse transcription was performed in both directions with DNase-treated total RNA from RN4220 cells harbouring the indicated plasmids, using either forward or reverse primers for cDNA synthesis in two separate reactions.

PCR was performed on cDNA products, or plasmid DNA templates for control (C) lanes. The symbols + or - indicate the presence or absence of reverse transcriptase enzyme in the reverse transcription reaction mixture used for PCR. Size markers of 500 bp and 100 bp are indicated. Picture is representative of a single technical replicate.

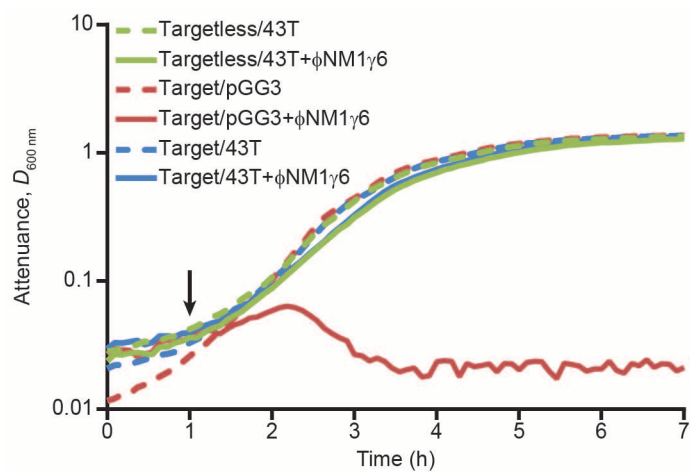


Extended Data Figure 7 | Reverse CRISPR-immunity assays using inverted chromosomal target insertions or type II CRISPR-Cas plasmids. Values represent the average transformation efficiency of three transformations in colony-forming units per microgram of plasmid DNA transformed. ATc, anhydrotetracycline at $0.5 \mu\text{g ml}^{-1}$. Dotted lines represent the limit of detection for these assays. **a**, Reverse CRISPR-immunity assays using inverted target

vector insertions and spacer 43T or 43B plasmid DNA. Inversion of the *attP* motif ('Inv-*attP*-') for forward and reverse insertion vectors causes integration in the opposite orientation relative to the chromosomal origin of replication. **b**, Reverse CRISPR-immunity assays using type II-A CRISPR plasmid DNA to transform strains from Fig. 3b. The pDB184 parent vector serves as a non-targeting control. Error bars, mean \pm s.d. ($n = 3$).

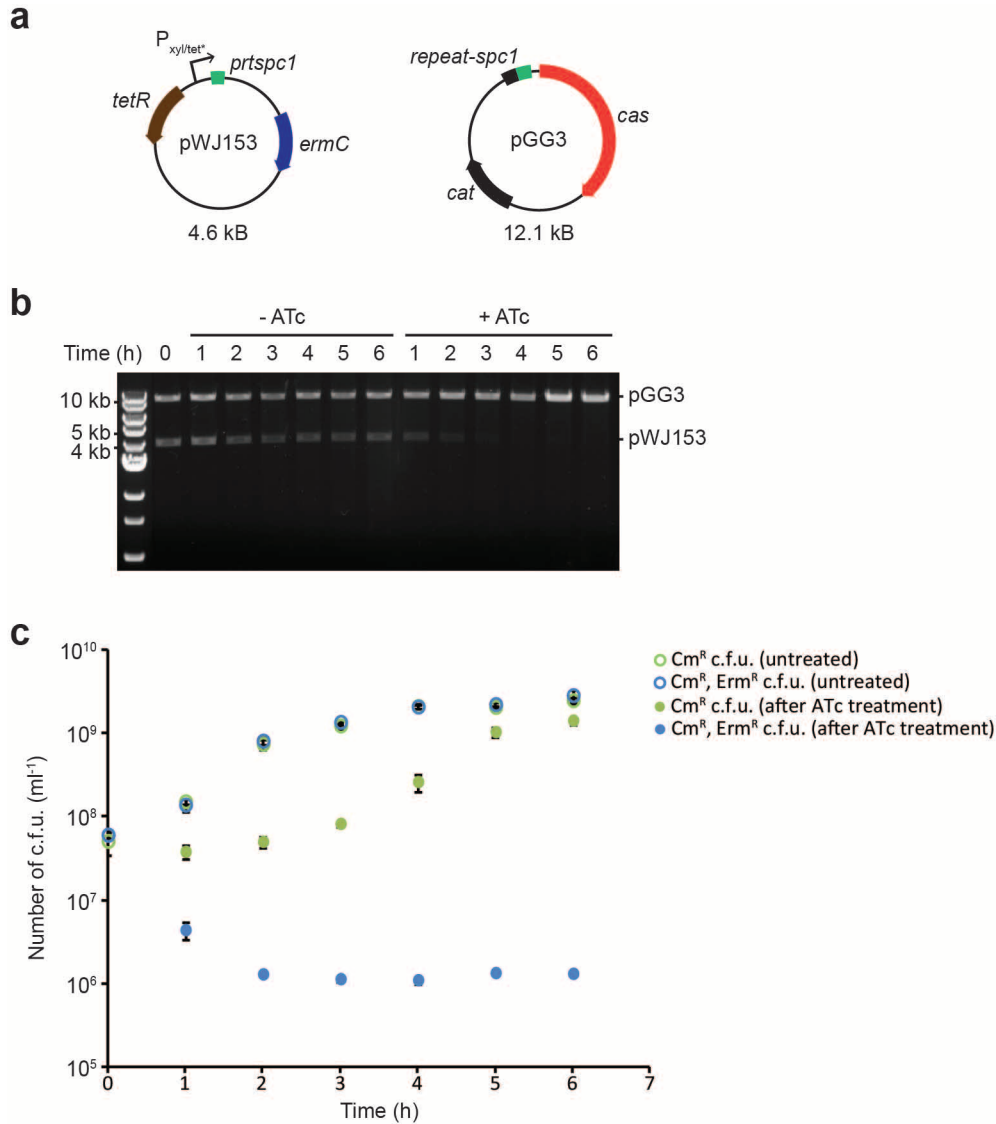


Extended Data Figure 8 | Infection with ΦNM1 in liquid culture. Growth curves of RN4220 cells harbouring the indicated CRISPR plasmids were infected at time zero with ΦNM1 at an MOI of 10 (a) or 100 (b). Growth of uninfected RN4220/pGG3 cultures is also shown (dotted red lines).



Extended Data Figure 9 | Immunity to Φ NM1 γ 6 in liquid culture is unaffected by the presence of a tolerated chromosomal target.

Growth curves of the indicated chromosomal insertion strains from Fig. 3 harbouring either spacer 43T or pGG3 CRISPR plasmids, in the absence (dotted lines) or presence (solid lines) of Φ NM1 γ 6 addition at an MOI of 10. Black arrow denotes the time of phage addition; no ATc induction is used in this assay. The presence of a chromosomal target for spacer 43T has no discernable effect on culture growth during spacer 43T-mediated immunity to Φ NM1 γ 6 (compare solid green and blue lines).



Extended Data Figure 10 | Inducible curing of a target plasmid. **a**, Diagram of plasmids used in the plasmid-curing experiment. The pGG3 CRISPR plasmid harbours a single spacer (*spc1*) targeting a sequence (*prtspc1*) inserted downstream of the P_{xyl/tet^*} -inducible promoter in pWJ153. **b**, Agarose gel electrophoresis of linearized plasmid DNA purified both from anhydrotetracycline-treated (+ATc) and untreated (-ATc) cultures at the indicated time points. Size markers of 10 kb, 5 kb and 4 kb are indicated.

Picture is representative of a single technical replicate. **c**, Colony-forming units recovered from cultures analysed in **b** at each time point. Cells were plated with selection for either Cm^R colony-forming units (green) or Cm^R, Erm^R colony-forming units (blue). Targeting of the pWJ153 plasmid via induction with ATc (filled circles) is accompanied by a severe drop in erythromycin-resistant colony-forming units relative to untreated cultures (open circles). Error bars, mean \pm s.d. ($n = 3$).

Rapid fucosylation of intestinal epithelium sustains host–commensal symbiosis in sickness

Joseph M. Pickard¹, Corinne F. Maurice², Melissa A. Kinnebrew³, Michael C. Abt³, Dominik Schenten⁴, Tatyana V. Golovkina⁵, Said R. Bogatyrev⁶, Rustem F. Ismagilov⁶, Eric G. Pamer³, Peter J. Turnbaugh² & Alexander V. Chervonsky¹

Systemic infection induces conserved physiological responses that include both resistance and ‘tolerance of infection’ mechanisms¹. Temporary anorexia associated with an infection is often beneficial^{2,3}, reallocating energy from food foraging towards resistance to infection⁴ or depriving pathogens of nutrients⁵. However, it imposes a stress on intestinal commensals, as they also experience reduced substrate availability; this affects host fitness owing to the loss of caloric intake and colonization resistance (protection from additional infections)⁶. We hypothesized that the host might utilize internal resources to support the gut microbiota during the acute phase of the disease. Here we show that systemic exposure to Toll-like receptor (TLR) ligands causes rapid $\alpha(1,2)$ -fucosylation of small intestine epithelial cells (IECs) in mice, which requires the sensing of TLR agonists, as well as the production of interleukin (IL)-23 by dendritic cells, activation of innate lymphoid cells and expression of fucosyltransferase 2 (Fut2) by IL-22-stimulated IECs. Fucosylated proteins are shed into the lumen and fucose is liberated and metabolized by the gut microbiota, as shown by reporter bacteria and community-wide analysis of microbial gene expression. Fucose affects the expression of microbial metabolic pathways and reduces the expression of bacterial virulence genes. It also improves host tolerance of the mild pathogen *Citrobacter rodentium*. Thus, rapid IEC fucosylation appears to be a protective mechanism that utilizes the host’s resources to maintain host–microbial interactions during pathogen-induced stress.

To maintain healthy gut microbiota during a systemic response induced by microbial products the host may use its internal resources. L-Fucosylation could present an example of such a resource: L-fucose attached to glycoproteins and glycolipids is accessible for microbial, but not for host energy harvest^{7,8}. Constitutive $\alpha(1,2)$ fucosylation affects the microbial community in a diet-dependent manner⁹, serves as a substrate for pathogens during antibiotic exposure¹⁰ and for microbes colonizing the ileum of newborns or of adult germ-free (GF) mice^{11–13}. Under normal conditions, however, the small intestine of specific pathogen-free (SPF) BALB/c mice is largely free of surface fucose. By contrast, a systemic injection of agonists of TLRs, such as lipopolysaccharide (LPS; TLR4 ligand) (Fig. 1), CpG DNA (TLR9 ligand), or Pam₃CSK₄ (TLR2 agonist), led to ubiquitous $\alpha(1,2)$ fucosylation of the small intestine in mice of different genetic backgrounds, which started within a few hours after LPS exposure and lasted for several days (Extended Data Fig. 1a–c). It did not result in differentiation of IECs into functional M cells¹⁴, which are permanently fucosylated and are involved in microbial sensing and translocation (Extended Data Fig. 1d). Induced fucosylation was independent of the gut microbiota (observed in GF mice), and was not induced by oral LPS (Extended Data Fig. 1e).

Global deletion of the TLR signalling adaptor molecule MyD88 prevented IEC fucosylation, and its conditional deletion from dendritic cells, but not from IECs, abrogated the process (Fig. 1). The inducible fucosylation pathway was similar to induction of antimicrobial peptides by a systemic microbial signal¹⁵: it required MyD88-expressing dendritic

cells, production of IL-23, the transcriptional regulator ROR γ t and IL-22 (Fig. 1 and Extended Data Fig. 2a), and was induced by a direct injection of IL-22 into *Myd88*^{-/-} mice (Fig. 1). IEC fucosylation in mice lacking T cells (Fig. 1) suggested that innate lymphoid cells (ILCs) were a sufficient source of IL-22. *Salmonella enterica* ssp. Typhimurium, known to spread systemically, induced small intestine IEC fucosylation (Extended Data Fig. 2b). The $\alpha(1,2)$ fucosyltransferase responsible for fucosylation of IECs in the small intestine was identified as *Fut2* (Fig. 2a), which is inducible by stress conditions^{16,17} and constitutively expressed in the stomach and large intestine¹⁸. Genetic ablation of the *Fut2* gene blocked IEC fucosylation in response to LPS (Fig. 2b, c). The overall chain of events is shown in Extended Data Fig. 3.

LPS injection caused marked sickness in BALB/c *Fut2*-deficient (*Fut2*^{-/-}) and *Fut2*-sufficient (*Fut2*^{+/+} or *Fut2*^{+/-}) littermates hours after injection: mice displayed measurable anorexia, adipsia, reduced activity, diarrhoea and weight loss. Food consumption, weight loss, and production of inflammatory cytokines and antimicrobial peptides (Fig. 2d, e and Extended Data Fig. 4a, b) were similar in both groups. Although

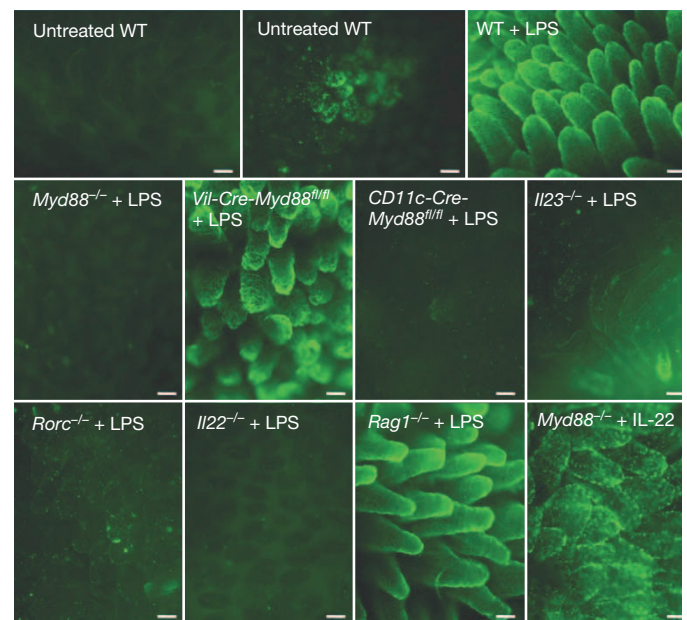


Figure 1 | MyD88-dependent fucosylation of small intestine IECs by systemic stimulation of TLRs. UEA-1, which binds $\alpha(1,2)$ -fucosylated substrates, staining in the proximal one-third of the small intestine of mice untreated or 24 h after intraperitoneal (i.p.) LPS injection, or 6 h after injection of IL-22 (*Myd88*^{-/-} mouse). Scale bars = 100 μ m. Staining of tissue from mutant mice was always accompanied by staining of wild-type (WT) controls, and is representative of at least two independent experiments for each genotype.

¹Department of Pathology and Committee on Immunology, The University of Chicago, Chicago, Illinois 60637, USA. ²FAS Center for Systems Biology, Harvard University, Cambridge, Massachusetts 02138, USA. ³Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. ⁴The University of Arizona, Tucson, Arizona 85721, USA. ⁵Department of Microbiology, The University of Chicago, Chicago, Illinois 60637, USA. ⁶California Institute of Technology, Pasadena, California 91125, USA.

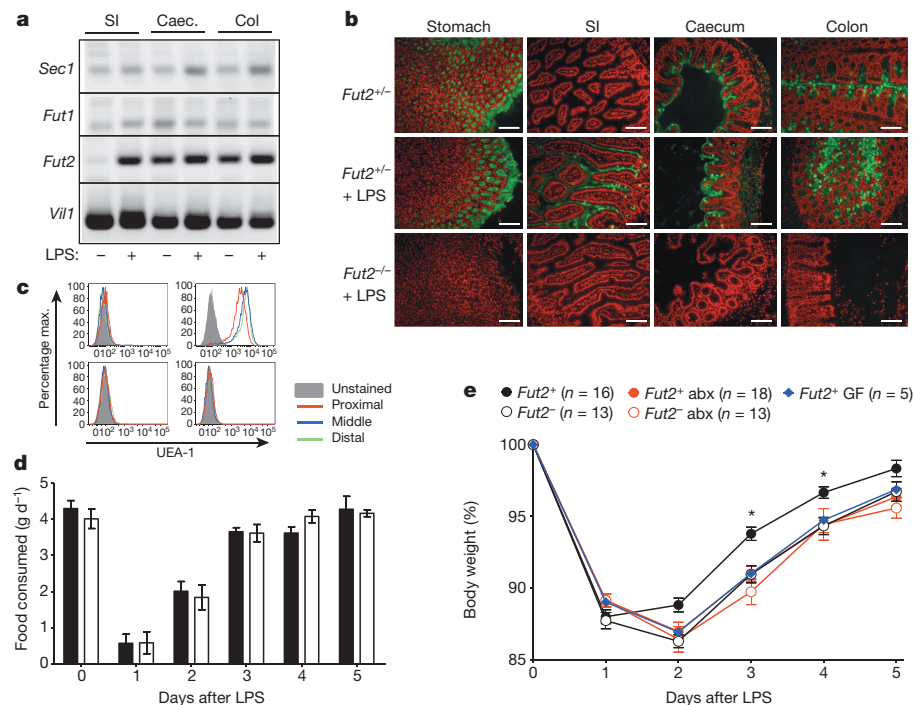


Figure 2 | Consequences of the loss of Fut-2-dependent fucosylation. **a**, Expression of mouse $\alpha(1,2)$ fucosyltransferase genes (*Fut2*, *Fut1* and *Sec1*) and control (villin-encoding gene, *Vil1*) in the gut 24 h after LPS injection (semiquantitative polymerase chain reaction with reverse transcription (RT-PCR)). Caec, caecum; Col., colon; SI, small intestine. **b**, Intestinal fucosylation (green) of *Fut2*-sufficient and *Fut2*-deficient mice. Red, propidium iodide. Scale bars = 100 μ m. **c**, Fluorescence-activated cell sorting (FACS) histograms of small intestine IECs from PBS- (left) or LPS- (right) injected, *Fut2*⁺ (top) or *Fut2*⁻ (bottom) mice. **d**, Food consumption in LPS-treated *Fut2*⁺ ($n = 5$, black bars) and *Fut2*⁻ mice ($n = 3$, open bars) (mean \pm standard error of the mean (s.e.m.)). **e**, Dependence of weight recovery after LPS challenge on the presence of microbiota and expression of *Fut2* (mean \pm s.e.m. of percentage of starting body weight, data combined from four experiments). * $P < 0.05$, one-way ANOVA. abx, antibiotics (ampicillin and vancomycin). All data are representative of at least two independent experiments.

Fut2^{-/-} mice have been shown to be healthy under normal SPF conditions¹⁹, they were significantly slower than controls in recovering their body weight (Fig. 2e) after LPS injection. Importantly, food deprivation without LPS challenge did not cause small intestine fucosylation and the weight recovery was identical in *Fut2*-sufficient and *Fut2*-deficient animals (Extended Data Figs 1a and 4c). Thus, fucosylation of the small intestine was a response to activation by microbial stimuli and not to anorexia per se.

To understand the reason for the slow weight recovery in *Fut2*^{-/-} mice, we tested whether fucosylation affected the function of host IEC proteins, which were identified by direct sequencing as secreted mucins and digestive enzymes (Fig. 3a). Fucosylation did not change the activity of several enzymes (Fig. 3b). Thus, changes in enzymatic activity are unlikely to explain the slow weight recovery in *Fut2*^{-/-} mice, although the role of less abundant fucosylated proteins cannot be excluded.

The beneficial effect of fucosylation was dependent on the microbiota: the weight recovery after LPS challenge was delayed in GF mice and in wild-type SPF mice treated with antibiotics (Fig. 2e). Antibiotics did not have a direct effect on the host's responses to LPS (Extended Data Fig. 4a, b, d–g) and did not further impair the recovery of *Fut2*^{-/-} mice (Fig. 2e). Thus, *Fut2* and an intact gut microbiota were both necessary for efficient recovery of body weight after LPS challenge.

To serve as a substrate to the bacteria residing in the large intestine²⁰, fucosylated proteins (Extended Data Fig. 5a–d) must be released into the small intestine lumen. After LPS challenge, *Fut2*-dependent fucosylation of luminal proteins was detectable at much higher levels in GF mice and in SPF mice treated with antibiotics or with a fucosidase inhibitor, than in SPF controls (Fig. 3c and Extended Data Fig. 5e). Thus, fucose is available, released and used by microbes in the large intestine.

Next, to directly show fucose sensing by intestinal commensal bacteria, we used a reporter system in which *Escherichia coli* expressed green fluorescent protein (GFP) driven by the promoter of the *E. coli* fucose metabolism operon^{21,22} (Fig. 3d, e and Extended Data Fig. 6a). Because *E. coli* lacks $\alpha(1,2)$ fucosidase that cleaves fucose off substrates, in GF mice mono-colonized with the reporter *E. coli* it did not upregulate GFP, even after LPS injection (Fig. 3d). Thus, free fucose was not readily available for reporter bacteria in the gut and required bacterial fucosidase activity, which was sensitive to antibiotics (Extended Data Fig. 6b). A commensal bacterium with $\alpha(1,2)$ fucosidase activity, *Bacteroides acidifaciens*, was isolated

from our mouse colony (Extended Data Fig. 6c–e). In LPS-injected GF mice co-colonized with the reporter *E. coli* and *B. acidifaciens*, the reporter strain expressed significantly more GFP (Fig. 3d), as well as genes for fucose import (*fucP*) and metabolism (*fucA*) (Fig. 3f). In LPS-treated SPF mice, *Fut2* was required for GFP reporter expression (Fig. 3e and Extended Data Fig. 6f). These findings made it clear that fucose can serve as a substrate for the microbiota under conditions of stress applied to the host, and underscored the interdependence between members of the gut microbial community²³.

To confirm these findings in mice with complex gut microbiota, we profiled microbial community structure, gene abundance and transcriptional activity before and after LPS treatment of *Fut2*^{+/-} and *Fut2*^{-/-} mice (Supplementary Table 1). Analyses of community structure based on 16S and shotgun DNA sequencing revealed that the gut microbiota was largely robust to host genotype and LPS exposure (Fig. 3g and Extended Data Fig. 7a, b). We did not detect: (1) significant clustering of microbial communities based on genotype or LPS treatment ($P > 0.05$ for both comparisons; permutational multivariate analysis of variance (PERMANOVA) of unweighted UniFrac distances); (2) species-level operational taxonomic units (OTUs) significantly associated with host genotype before or after LPS treatment (all were $q > 0.05$; analysis of variance (ANOVA)); or (3) significant differences in overall microbial diversity (Extended Data Fig. 7c). However, we were able to detect a significantly increased abundance of *B. acidifaciens* after LPS treatment in *Fut2*-sufficient mice ($P < 0.05$, linear discriminant analysis (LDA) score > 4 ; linear discriminant analysis effect size (LefSe) analysis of 16S profiles), consistent with its ability to utilize fucosylated glycans.

At the same time, LPS markedly altered community-wide gene expression in both *Fut2*^{+/-} and *Fut2*^{-/-} mice (Fig. 3h), with multiple orthologous groups differentially expressed upon LPS treatment: 61 in *Fut2*^{+/-} mice and 56 in *Fut2*^{-/-} animals. These changes were not due to altered community structure: only one differentially expressed orthologous group (K05351, xylulose reductase) also exhibited significant changes in gene abundance (Supplementary Table 2). As expected, we detected a significant upregulation of fucose permease (*fucP*; K02429) in *Fut2*-sufficient mice after exposure to LPS (Fig. 3i), and increased expression of metabolic modules for anaerobic respiration, protein and ATP synthesis, isoprenoid biosynthesis and amino sugar import, in addition to pathways

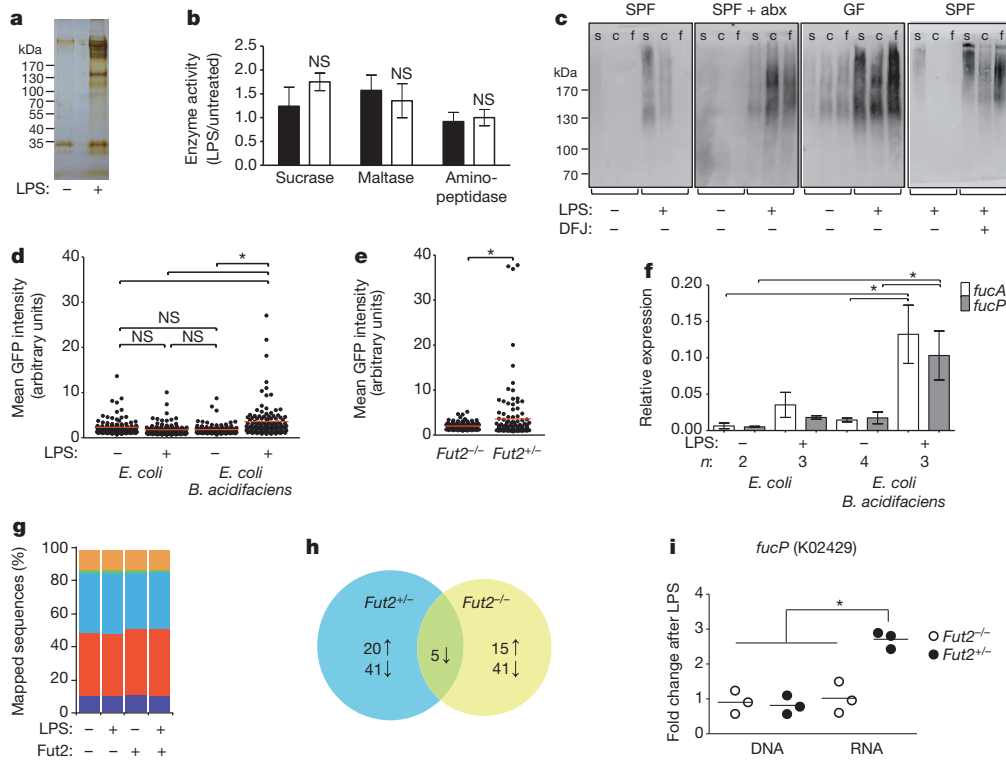


Figure 3 | Commensals utilize fucose detached from proteins fucosylated by Fut2 upon systemic challenge with LPS. **a**, Silver-stained SDS–polyacrylamide gel electrophoresis (SDS–PAGE) of UEA-1-precipitated small intestine IEC protein from control or LPS-treated mouse. **b**, Ratios of digestive enzyme activities in small intestine IECs of LPS-treated to untreated Fut2-sufficient (black bars) or Fut2-deficient (open bars) mice 2 days after LPS injection. Mean \pm s.e.m. of four combined experiments; four mice per group. NS, not significant ($P > 0.05$, two-tailed Student's *t*-test). **c**, SDS–PAGE of intestinal contents blotted on nitrocellulose and stained with UEA-1–peroxidase complexes. abx, antibiotic-treated mice; c, caecum; DFJ, deoxyfuconojirimycin; f, faeces; s, small intestine. Data are representative of at least two experiments. **d**, **e**, Fucose-sensitive GFP reporter expression in gnotobiotic mice colonized with the indicated strains (**d**) or SPF mice (**e**). Dots are values for individual bacteria, lines are means; $n = 120$. * $P < 0.05$, one-way ANOVA with Bonferroni post-test (**d**), two-tailed Student's *t*-test (**e**). Data are

representative of three independent experiments. **f**, *fucA* and *fucP* gene expression relative to housekeeping gene *rpoA* (quantitative RT–PCR) in *E. coli* tested as in **d**. * $P < 0.05$ by ANOVA with Bonferroni's post-test. Data are combined from three experiments. **g**, Stable relative abundance of bacterial phyla across treatment groups and genotypes, as indicated by shotgun sequencing of community DNA. Phyla with a mean reads per kilobase per million mapped reads (RPKM) $> 40,000$ are shown, including Actinobacteria (purple), Bacteroidetes (red), Firmicutes (blue), Fusobacteria (green), Proteobacteria (orange) and Tenericutes (yellow). 16S rRNA gene sequencing confirms these observations. Extended Data Fig. 7 shows 16S rRNA gene sequencing results. **h**, Differentially expressed KEGG orthologous groups following LPS treatment (paired glm edgeR analysis; $q < 0.05$, > 2 -fold change; see Supplementary Table 2 for complete list). **i**, Increased gut microbial expression of *fucP* (K02429) in Fut2-sufficient mice (mean \pm s.e.m.; * $P < 0.01$, Mann–Whitney test).

for aminoglycan degradation (Extended Data Fig. 7d). Thus, intact host fucosylation appears to affect gut microbial metabolism. Importantly, LPS challenge led to the significantly increased expression of microbial virulence genes in Fut2-negative but not Fut2-sufficient mice, including *rtxA* (K10953) and haemolysin III (K11068) (Supplementary Table 2). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways potentially involved in microbial pathogenesis (defined as flagella synthesis, chemotaxis, plant/pathogen interaction and *Vibrio cholera* infection) were overrepresented in Fut2-deficient mice (Extended Data Fig. 7d). We hypothesized that fucosylation induced by systemic microbial challenge might limit the effects of additional exogenous or endogenous pathogens. We tested this by infecting Fut2-sufficient and Fut2-deficient mice with a non-lethal intestinal pathogen, *C. rodentium*. Four days after infection, mice were treated with LPS. Infected Fut2-negative mice lost significantly more weight than Fut2-sufficient animals compared with respective LPS-treated non-infected controls (Fig. 4a). Thus, infection with a non-lethal pathogen further reduced the overall fitness of Fut2-deficient mice in response to LPS. *C. rodentium* did not induce small intestine IEC fucosylation and did not colonize the small intestine (Extended Data Fig. 8), indicating that systemic challenge by a microbial product was required to reveal the role of inducible fucosylation.

The fitness of infected animals can be maintained through either decreased pathogen burden (resistance), or by an increase in pathogen tolerance without a change in pathogen burden. We quantified the abundance

of *C. rodentium* in the faeces of infected mice, as well as the adherence to IECs of *C. rodentium*-expressing luciferase^{24,25} (Fig. 4b, c). No differences in pathogen loads were found between Fut2-sufficient and Fut2-deficient animals treated with LPS. Thus, fucosylation of the small intestine upon systemic treatment with LPS probably enhances tolerance of the pathogen. Fut2-negative mice infected with *C. rodentium* and injected with LPS had more pronounced colonic hyperplasia (a trademark of this infection) compared with Fut2-sufficient mice or mice that did not receive LPS (Fig. 4d, e). Thus, inducible IEC fucosylation might be viewed as an emergency measure taken by the host to support its gut commensals. Fucose used by microbes as an energy source may contribute to protection of the host from endogenous opportunistic pathogens, or it could increase tolerance of infection by regulating bacterial genes responsible for quorum sensing²⁶ or virulence²⁷. Fucose can also serve as a substrate for microbial production of the short-chain fatty acid propionate (Extended Data Fig. 9), which is primarily produced by members of the Bacteroidetes phylum²⁸. Whether this process contributes to overall fitness of the animals under infection-induced stress remains to be elucidated. Of note, around 20% of humans lack a functional *FUT2* gene, which is linked to Crohn's disease²⁹ and to lethality from sepsis in premature infants³⁰. Overall, fucosylation of the small intestine in response to systemic microbial exposure can be considered a type of 'tolerance of infection' response. It is interesting, however, that a very similar pathway regulates secretion of antimicrobial

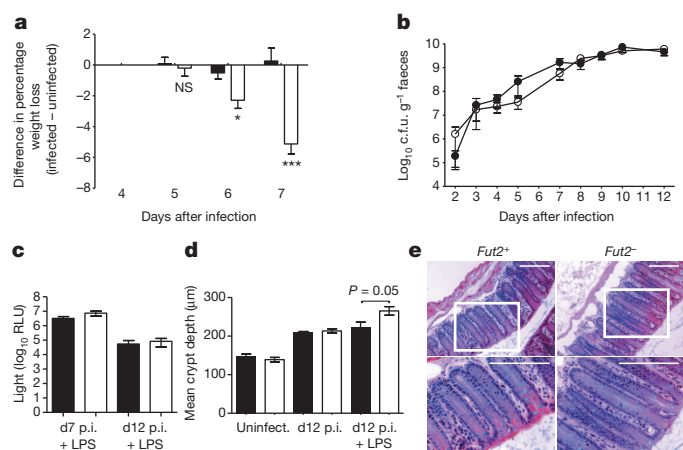


Figure 4 | Host fucosylation increases tolerance of a pathogen. **a**, Difference in percentage weight loss between LPS-injected *C. rodentium*-infected and uninfected mice. Data are shown as mean \pm s.e.m.; * $P = 0.01$ *** $P = 0.0001$, two-tailed Student's *t* test; combined from six experiments. NS, not significant. **b**, Faecal colony forming units (c.f.u.) of *C. rodentium* from Fut2⁺ or Fut2⁻ mice (mean \pm s.e.m.), representative of three experiments). **c**, Luminescence of thoroughly washed mid-colon of mice infected with *pler-lux*⁺ *C. rodentium*. $n = 4$ for day (d)12 and 8 for d7; mean \pm s.e.m. RLU, relative luminescence units; p.i., post-infection. Data combined from three experiments. **d**, Average crypt depth in uninfected ($n = 3$) or infected mice ($n = 4$) at day 12 post-infection, with or without LPS injection on day 4 post-infection. Data are shown as mean \pm s.e.m. (two-tailed Student's *t*-test). In **a–d** black bars and circles indicate Fut2-positive mice; open bars and circles indicate Fut2-negative mice. **e**, Representative haematoxylin and eosin staining of distal colon of LPS-treated mice at day 12 after infection. Scale bars = 100 μ m. Bottom row shows magnified boxed regions.

proteins—a resistance mechanism¹⁵. Thus, the mechanisms of resistance and tolerance to pathogens could be evolutionarily linked together to increase the fitness of the host.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 December 2013; accepted 3 September 2014.

Published online 1 October 2014.

1. Ayres, J. S. & Schneider, D. S. Tolerance of infections. *Annu. Rev. Immunol.* **30**, 271–294 (2012).
2. Ayres, J. S. & Schneider, D. S. The role of anorexia in resistance and tolerance to infections in *Drosophila*. *PLoS Biol.* **7**, e1000150 (2009).
3. Murray, M. J. & Murray, A. B. Anorexia of infection as a mechanism of host defense. *Am. J. Clin. Nutr.* **32**, 593–596 (1979).
4. Kyriazakis, I. I., Tolkamp, B. J. & Hutchings, M. R. Towards a functional explanation for the occurrence of anorexia during parasitic infections. *Anim. Behav.* **56**, 265–274 (1998).
5. Exton, M. S. Infection-induced anorexia: active host defence strategy. *Appetite* **29**, 369–383 (1997).
6. Stecher, B. & Hardt, W. D. Mechanisms controlling pathogen colonization of the gut. *Curr. Opin. Microbiol.* **14**, 82–91 (2011).
7. Bocci, V. & Winzler, R. J. Metabolism of L-fucose-1-¹⁴C and of fucose glycoproteins in the rat. *Am. J. Physiol.* **216**, 1337–1342 (1969).
8. Becker, D. J. & Lowe, J. B. Fucose: biosynthesis and biological function in mammals. *Glycobiology* **13**, 41R–53R (2003).
9. Kashyap, P. C. *et al.* Genetically dictated change in host mucus carbohydrate landscape exerts a diet-dependent effect on the gut microbiota. *Proc. Natl Acad. Sci. USA* **110**, 17059–17064 (2013).
10. Ng, K. M. *et al.* Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* **502**, 96–99 (2013).
11. Umesaki, Y., Tohyama, K. & Mutai, M. Appearance of fucolipid after conventionalization of germ-free mice. *J. Biochem.* **90**, 559–561 (1981).

12. Bry, L., Falk, P. G., Midtved, T. & Gordon, J. I. A model of host-microbial interactions in an open mammalian ecosystem. *Science* **273**, 1380–1383 (1996).
13. Sonnenburg, J. L. *et al.* Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. *Science* **307**, 1955–1959 (2005).
14. Clark, M. A., Jepsen, M. A., Simmons, N. L., Booth, T. A. & Hirst, B. H. Differential expression of lectin-binding sites defines mouse intestinal M-cells. *J. Histochem. Cytochem.* **41**, 1679–1687 (1993).
15. Kinnebrew, M. A. *et al.* Interleukin 23 production by intestinal CD103⁺CD11b⁺ dendritic cells in response to bacterial flagellin enhances mucosal innate immune defense. *Immunity* **36**, 276–287 (2012).
16. Thomsson, K. A. *et al.* Intestinal mucins from cystic fibrosis mice show increased fucosylation due to an induced Fucal-2 glycosyltransferase. *Biochem. J.* **367**, 609–616 (2002).
17. Holmén, J. M., Olson, F. J., Karlsson, H. & Hansson, G. C. Two glycosylation alterations of mouse intestinal mucins due to infection caused by the parasite *Nippostrongylus brasiliensis*. *Glycoconj. J.* **19**, 67–75 (2002).
18. Domino, S. E., Zhang, L. & Lowe, J. B. Molecular cloning, genomic mapping, and expression of two secretor blood group α (1,2)fucosyltransferase genes differentially regulated in mouse uterine epithelium and gastrointestinal tract. *J. Biol. Chem.* **276**, 23748–23756 (2001).
19. Domino, S. E., Zhang, L., Gillespie, P. J., Saunders, T. L. & Lowe, J. B. Deficiency of reproductive tract α (1,2)fucosylated glycans and normal fertility in mice with targeted deletions of the FUT1 or FUT2 α (1,2)fucosyltransferase locus. *Mol. Cell. Biol.* **21**, 8336–8345 (2001).
20. Vaishnava, S. *et al.* The antibacterial lectin RegIII γ promotes the spatial segregation of microbiota and host in the intestine. *Science* **334**, 255–258 (2011).
21. Hovel-Miner, G., Faucher, S. P., Charpentier, X. & Shuman, H. A. ArgR-regulated genes are derepressed in the *Legionella*-containing vacuole. *J. Bacteriol.* **192**, 4504–4516 (2010).
22. Zhang, Z., Yen, M. R. & Saier, M. H. Jr. Precise excision of IS5 from the intergenic region between the *fucPIK* and the *fucAO* operons and mutational control of *fucPIK* operon expression in *Escherichia coli*. *J. Bacteriol.* **192**, 2013–2019 (2010).
23. Fischbach, M. A. & Sonnenburg, J. L. Eating for two: how metabolism establishes interspecies interactions in the gut. *Cell Host Microbe* **10**, 336–347 (2011).
24. Kamada, N. *et al.* Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science* **336**, 1325–1329 (2012).
25. Barba, J. *et al.* A positive regulatory loop controls expression of the locus of enterocyte effacement-encoded regulators Ler and GrlA. *J. Bacteriol.* **187**, 7918–7930 (2005).
26. Scott, K. P., Martin, J. C., Campbell, G., Mayer, C. D. & Flint, H. J. Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium “*Roseburia inulinivorans*”. *J. Bacteriol.* **188**, 4340–4349 (2006).
27. Pacheco, A. R. *et al.* Fucose sensing regulates bacterial intestinal colonization. *Nature* **492**, 113–117 (2012).
28. De Vadder, F. *et al.* Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits. *Cell* **156**, 84–96 (2014).
29. McGovern, D. P. *et al.* Fucosyltransferase 2 (FUT2) non-secretor status is associated with Crohn's disease. *Hum. Mol. Genet.* **19**, 3468–3476 (2010).
30. Morrow, A. L. *et al.* Fucosyltransferase 2 non-secretor and low secretor status predicts severe outcomes in premature infants. *J. Pediatr.* **158**, 745–751 (2011).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank C. Reardon and C. Daly for sequencing support, H. Ye for help with metabolic cage analysis, N. F. Dalleska for assistance and use of GC-MS instrumentation in the Environmental Analysis Center at the California Institute of Technology, and G. Nuñez for luciferase-expressing *C. rodentium*. This work was supported by grants from the National Institutes of Health (P50 GM068763 to P.J.T., AI96706 and AI42135 to E.G.P., T32 AI065382 to J.M.P.), the Harvard Bauer Fellows Program, National Science Foundation grant EFRI-1137089 to R.F.I. and A.V.C., Digestive Disease Research Core Center grant DK42086 and a Kenneth Rainin Foundation grant to A.V.C.

Author Contributions J.M.P., M.A.K., M.C.A. and E.G.P. performed analysis of inducible fucosylation in mice, including mutant strains; J.M.P. and C.F.M. produced DNA and RNA sequencing data and P.J.T. analysed these data; D.S. produced *Myd88*^{fl/fl} mice; T.V.G. produced GF BALB/c mice and performed cytokine ELISA analysis; S.R.B. and R.F.I. performed analysis of short-chain fatty acids; R.F.I., E.G.P. and P.J.T. contributed to writing of the manuscript; A.V.C. conceived the project, analysed the results and wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information The DNA and RNA shotgun sequencing data have been deposited in the Gene Expression Omnibus database under accession number GSE60874; 16S rRNA gene sequencing reads have been deposited in MG-RAST under accession number 10494. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.V.C. (achervon@bsd.uchicago.edu).

METHODS

Mice. BALB/c, C57BL/6J, NOD/LtJ, C3H/HeN, *Rag1*^{-/-} (B6.129S7-*Rag1*^{tm1.1Mom/J}), *Fut2*^{-/-} (B6.129X1-*Fut2*^{tm15d0/J}), Villin-Cre (B6.SJL-Tg(Vil-Cre)997Gum/J), and CD11c-Cre (C57BL/6J-Tg(Itgax-Cre, EGFP)4097Ach/J) mice were purchased from The Jackson Laboratory. *Fut2*^{-/-} mice were backcrossed greater than seven generations to BALB/c. Knockout mice produced litters of mixed genotypes kept co-housed to homogenize their gut microbiota. B6 *Myd88*^{-/-} mice were a gift from S. Akira. B6 *MyD88* floxed mice were described previously³¹. *Rorc*^{-/-} (ref. 32), and *IL23p19*^{-/-} (also known as *IL23a*^{-/-})³³ mice were provided by Y.-X. Fu. *IL22*^{-/-} mice³⁴ were maintained at the Memorial Sloan-Kettering Cancer Center. Mice were housed in a specific pathogen-free facility and used in accordance with institutional guidelines for animal welfare. Six- to twelve-week-old male and female mice were used for randomization purposes. The numbers of mice per group were chosen as the minimum needed to obtain biologically significant results, based on previous experience. Evaluations were made in a blind fashion. Functional experiments were done with *Fut2*-negative mice on the BALB/c genetic background using *Fut2*-sufficient littermates as controls.

Metabolic studies. Mice were single-housed in TSE Systems LabMaster cages to monitor physical activity (*x/y* and *z* axes by infrared beams crossed), drinking and feeding.

Fucosylation activation *in vivo*. LPS from *S. enterica* ssp. Typhimurium (#L6511, Sigma-Aldrich) was injected i.p. at 1 µg g⁻¹ body weight, or gavaged to GF mice at 1 mg in 400 µl sterile PBS. CpG (100 µg CpG-B ODN 1826, Coley Pharmaceutical Group) and Pam₃CSK₄ (100 µg, InvivoGen), were injected i.p. Carrier-free recombinant mouse IL-22 (Biolegend,) was diluted in 1% BSA/PBS and 1.5 µg was given i.p.

Antibiotic treatment. Ampicillin (1 g l⁻¹) and 200 mg l⁻¹ vancomycin (Sigma-Aldrich) were 0.22 µm filtered and added to autoclaved drinking water starting 2 days before LPS treatment (day -2) and lasting for the duration of the experiment.

Lectin staining. For whole-mount staining, small intestine was removed, a 1 cm piece from the upper third was excised, opened, cleaned of mucus in cold PBS, and incubated with *Ulex europaeus* agglutinin-1 (UEA-1) conjugated to FITC, TRITC or atto-594 (Vector Laboratories or Sigma-Aldrich) for 15 min on ice. Tissue was placed lumen side up on a slide for microscopy. For *IL22*^{-/-} experiments, whole tissue was fixed in 4% paraformaldehyde before proceeding with whole-mount staining. Fixation did not affect staining pattern or intensity.

For staining of sections, tissues were fixed in 2% paraformaldehyde overnight at 4 °C, cryoprotected in 20% sucrose/PBS overnight at 4 °C, and embedded in OCT compound (Sakura Finetek), and 10 µm sections were stained with UEA-1-FITC (1 µg ml⁻¹, Vector) for 30 min at room temperature, and incubated for 20 min at 37 °C in propidium iodide (0.5 µg ml⁻¹, Sigma-Aldrich) with RNase A (10 µg ml⁻¹, Sigma-Aldrich) to label nuclei.

For single-cell analysis by FACS, small intestine (with Peyer's patches and mesenteric fat removed) was divided into equal thirds, opened longitudinally and washed with cold PBS. Tissue was cut into 1 cm pieces and shaken in 10 mM EDTA/1 mM dithiothreitol (DTT)/PBS at 37 °C for 20 min, and filtered through nylon mesh. Single-cell suspensions were pelleted and fixed in 5 ml of 1% paraformaldehyde overnight at room temperature. Cells were then stained with UEA-1-FITC (1 µg ml⁻¹, Vector), and gated on the FSC/SSC high epithelial cell population.

Light microscopy. Fluorescence microscopy of whole-mounts and sections used a Leica DM LB microscope (Leica Camera AG) and Spot RT Slider camera and software (Diagnostic Instruments). Confocal microscopy of bacteria was performed with an Olympus DSU spinning disk microscope and Slidebook software (3I). All images in an experiment were taken using the same exposure settings.

Scanning electron microscopy. Tissues were fixed in 2% paraformaldehyde/2% glutaraldehyde/0.1 M cacodylate buffer, transferred to cacodylate buffer overnight and processed with the OTOTO procedure³⁵. To dehydrate, samples were passed through increasing concentrations of acetone in water, followed by hexamethyldisilazane (Electron Microscopy Sciences). Samples were mounted with colloidal silver paste (Electron Microscopy Sciences) and imaged with an FEI Nova NanoSEM 230 (FEI) at 5 kV.

Identification of fucosylated proteins. To visualize fucosylated proteins in the lumen contents, the intestinal contents were gently removed by squeezing and homogenized at 1 g sample per 5 ml of Tris-Triton-X-100 buffer (150 mM NaCl, 50 mM Tris pH 8, 1% Triton X-100, and protease inhibitor tablet (Roche)) on ice and spun at 17,000g for 20 min at 4 °C. Supernatant proteins were separated by 6% non-reducing SDS-PAGE, transferred to PVDF membrane (Bio-Rad), blocked with 0.5% gelatin in 0.05% Tween-20/PBS, and visualized with UEA-1-HRP (1 µg ml⁻¹, Sigma-Aldrich).

For sequencing of fucosylated epithelial cell proteins, small intestine epithelial cells were isolated as described earlier and lysed in Tris-Triton-X-100 buffer. Lysates were pre-cleared by incubating with unconjugated agarose beads (Vector) twice for 45 min at 4 °C. Cleared lysates were then incubated with washed UEA-1-conjugated

beads (Vector) for 45 min. Beads were washed five times to remove unbound protein. To elute UEA-1-bound protein, beads were incubated with 200 mM L-fucose (Sigma-Aldrich) for 30 min. Eluted proteins were separated on a 4–15% gradient SDS-PAGE gel (Bio-Rad), silver-stained (Thermo Fisher Scientific), and bands were excised for identification by mass spectrometry at the Taplin Biological Mass Spectrometry Facility, Harvard Medical School.

Isolation of fucosidase-positive bacteria. Brain heart infusion (BHIS) agar plates (Becton, Dickinson, and Co.) with added thioglycolic acid, menadione and hematin/histidine (Sigma-Aldrich) were pre-reduced in an anaerobic chamber with 2.5% hydrogen atmosphere at 37 °C, and spread with 40 µl of 5-bromo-4-chloro-3-indolyl- α -L-fucopyranoside (50 mM in DMF) (Carbosynth). Faecal pellets were homogenized in reduced 0.01% thioglycolic acid/PBS, plated, and grown for 3 days at 37 °C anaerobically, then at 4 °C aerobically to develop colour. Blue colonies were identified by sequencing of their 16S rDNA genes using primers 8F (5'-AGAGTTTGTATCC TGGCTCAG-3') and 1391R (5'-GACGGGCGGTGWGTRCA-3'), and sequencing of their *gyrB* genes using primers *gyrB* F (5'-GAAGTCATCATGACCGTTCTGC AYGCGNGGNGNAARTTYGA-3') and *gyrB* R (5'-AGCAGGGTACGGATGT GCGAGCCRTCACRTCCNGTCAT-3') for amplification and *gyrB* FS (5'-GAAGTCATCATGACCGTTCTGCA-3') and *gyrB* RS (5'-AGCAGGGTAC GGATGTGCGAGCC-3') for sequencing³⁶. *B. thetaiotaomicron* VPI-5482 and *B. uniformis* (a gift from C. Nagler) were used as controls.

Measurement and inhibition of fucosidase activity. Faecal pellets were weighed and homogenized in 10 µl PBS per 1 mg sample, and centrifuged at 17,000g for 10 min. Fifty microlitres of this supernatant was incubated for 1 h at 37 °C with or without adding 0.5 µl of 50 mM 4-methylumbelliferyl fucopyranoside (in dimethylsulphoxide (DMSO), Sigma or Gold Biotechnology). The reaction was diluted 100-fold in 0.2 M glycine-NaOH buffer, pH 10.5, and fluorescence was measured at 365 nm excitation, 445 nm emission. Fluorescence of the no-substrate control was subtracted from the substrate-containing reaction. The amount of cleaved substrate and fucosidase activity was then calculated by comparison to a standard curve of 4-methylumbelliferone (Sigma) in glycine-NaOH buffer. Deoxyfuconojirimycin (DFJ; Enzo Life Sciences) was dissolved in PBS and gavaged to mice in 100 µl (5 µmol total).

Reporter *E. coli*. The pXDC94 plasmid²¹, containing the *ptac* promoter controlling mCherry expression and a multiple cloning site upstream of the promoter-less *GFP* gene, was a gift from H. Shuman. The promoter region upstream of the *E. coli* *fucPIK* genes²² was amplified using primers that added restriction sites (bold) at the 5' and 3' ends of the amplicon, respectively: *fucproF*, 5'-TATGGTACCGGAT TCATTTTCAAATAAAA-3'; *fucproR*, 5'-TATCCGGGTAGCTACCTCTCTC TGATTC-3'. The PCR product and vector were digested with *XmaI* and *KpnI* (New England Biolabs), gel purified, ligated, and introduced into *E. coli* K-12 strain BW25113 (Yale Coli Genetic Stock Center) by electroporation. Correct expression of mCherry and GFP was verified by growth in minimal medium³⁷ with 10 mM glucose and the indicated concentrations of fucose.

E. coli K12 bacteria carrying the pXDC94-*fucPIK*-pro reporter were grown overnight in Luria-Bertani broth (LB) with shaking at 37 °C, centrifuged at 5,000g, resuspended to ~10⁹ c.f.u. ml⁻¹ in PBS, and 400 µl gavaged to mice that had received LPS i.p. 6 h earlier (SPF mice), or at the time of gavage (gnotobiotic mice). For dual colonization with *B. acidifaciens*, 100 µl of stationary-phase culture was gavaged at the same time as *E. coli*. Twenty-four hours after gavage, mice were killed and their caecum and colon contents were homogenized in PBS. Bacteria were enriched by centrifuging at 200g for 5 min, centrifuging the supernatant at 5,000g for 5 min, and resuspending the pellet in 750 µl PBS. This was underlain with 300 µl Histopaque-1119 (Sigma-Aldrich) and centrifuged for 1 min at 11,600g. The interface containing mostly bacteria was washed with PBS, fixed in 2% paraformaldehyde for 20 min at room temperature, washed and resuspended in 50–100 µl PBS. Four microlitres of bacterial suspension was placed on a slide and coverslipped. Random fields were selected using red fluorescence only, and an image was acquired in both the red and green channels. Fluorescence of individual bacteria were measured in ImageJ 1.41 software³⁸ by gating on areas of red fluorescence and measuring the mean green pixel intensity within the gated area.

Estimation of total bacterial loads. Fresh faecal pellets were placed in an anaerobic chamber and mashed in 500 µl of reduced PBS containing 0.01% thioglycolic acid. Serial dilutions were made in reduced PBS and plated on pre-reduced Brucella blood agar plates (Becton, Dickinson, and Co.; for anaerobic counts), removed from the anaerobic chamber, and plated on tryptic soy agar (TSA)/5% sheep's blood plates (Becton, Dickinson, and Co.; aerobic counts). Total colonies were counted after 2 days (aerobic) or 3 days (anaerobic) incubation at 37 °C.

For 16S copy number, faecal pellets were weighed, DNA was isolated by a bead beating and phenol/chloroform extraction method³⁹, and qPCR performed and copy number determined as described⁴⁰.

***S. enterica* Typhimurium and *C. rodentium* infection.** For *S. enterica* Typhimurium infection, mice were gavaged with 20 mg streptomycin (Sigma) in 100 µl sterile

water, 24 h before infection. *S. enterica* Typhimurium strain SL1344 was grown overnight in LB with streptomycin ($50 \mu\text{g ml}^{-1}$) and gavaged to mice at 5×10^8 c.f.u. in 100 μl volume.

C. rodentium strains DBS100 or DBS120 *pler-lux* were grown in LB overnight at 37 °C, then diluted 1:100 and grown for 2.5 h, centrifuged and resuspended in 0.01 volumes PBS, and mice were gavaged with $\sim 5 \times 10^9$ bacteria in 100 μl LB. To determine mouse colonization levels, a fresh faecal pellet or small intestine contents (gently squeezed to remove, except last 3 cm of ileum) was weighed, mashed in 500 μl of PBS, serially diluted, and plated on MacConkey (Becton, Dickinson, and Co.) or LB agar with $50 \mu\text{g ml}^{-1}$ kanamycin. For luciferase measurements, the faecal homogenate was adjusted to 10 mg in 1 ml PBS in an Eppendorf tube and light measured in a Triathler scintillation counter (Hidex), before plating dilutions in PBS on agar with kanamycin. Colon-adherent bacteria were measured in a 1 cm piece from the middle of the colon. The piece was opened longitudinally, washed in 1 mM DTT/PBS by vortexing for 10 s, then washed in PBS, and placed in an Eppendorf in 500 μl PBS for light measurement as before. The piece was then mashed and dilutions plated on agar with kanamycin.

Histology. Distal colon from *C. rodentium*-infected mice or uninfected controls was fixed in neutral formalin, then kept in 70% EtOH until being embedded in paraffin, and 5 μm sections were cut and stained with haematoxylin and eosin. Well-oriented crypts were photographed and their lengths measured in ImageJ, and the mean taken for each mouse.

Serum cytokine ELISA. Serum concentrations were measured by ELISA according to the manufacturer's instructions (IL-1 β : eBioscience; IL-6 and TNF- α : Becton, Dickinson and Co.).

Short-chain fatty acid measurements. For gavage experiments, food was removed and SPF mice were gavaged at 0, 3 and 6 h with 300 μl of 0.1 M L-fucose or 0.1 M D-galactose (Sigma-Aldrich) in autoclaved tap water, or water only. At 8 h, mice were killed and caecal contents removed and kept at -80 °C until processing.

Concentrations of the short-chain fatty acids in caecal contents were measured using the direct injection gas chromatography-mass spectrometry (GC-MS) method adapted from that described previously⁴¹⁻⁴³.

Caecal contents were extracted in two steps. Deionized water was added to the samples at the amount of 5 ml per 1 g of sample, followed by brief vortex-mixing and sonication for 15 min. In the second step, acetonitrile (#1103, BDH) containing 20 mM of tetradeutoeroacetic acid (#16621, Acros Organics) was added to the aqueous extracts at the amount of 5 ml per 1 g of initial sample, followed by the second round of brief vortex-mixing and sonication for 15 min.

Extracted samples were centrifuged for 5 min at 17,000g and 0.02 ml of clear supernatant was mixed with 0.98 ml of acetonitrile containing 20 mM of formic acid (#94318, Fluka) and 0.05 mM of 2-ethylbutyric acid (#109959, Aldrich). The mixtures were briefly vortex-mixed and centrifuged for 5 min at 17,000g and 1 μl of the obtained supernatants were analysed by direct injection GC-MS on an Agilent 6890N GC system equipped with a Mass Selective (MS) Detector 5973 (Agilent Technologies).

The instrument was used in a splitless mode with an installed double taper inlet liner (#23308, Sky by Restek) and fused-silica column with polyethylene glycol stationary phase (INNOWax #19091N-133, J&W Scientific, Agilent Technologies) 30 m \times 0.25 mm ID coated with 0.25 μm film.

Helium carrier gas was supplied at 1.00 ml min⁻¹ flow rate. The injection port temperature was 260 °C. The initial oven temperature of 60 °C was maintained for 2 min, then increased to 150 °C at 10 °C min⁻¹ and further to 250 °C at 25 °C min⁻¹ and held at that temperature for 4 min, bringing the total duration of the run to 19 min. The MS detector temperature was set to 245 °C.

GC-MS data were analysed using the MSD ChemStation D.01.02.16 software (Agilent Technologies). Tetradeutoeroacetic and 2-ethylbutyric acids served as internal standards.

Enzymatic activity of digestive enzymes. Epithelial cells were dissociated from the upper half of the small intestine as described earlier. Haematopoietic cells were removed with anti-CD45-biotin (clone 30-F11, Biolegend) and streptavidin magnetic beads with a MACS LS column (Miltenyi Biotec). Epithelial cells were diluted in 96-well plates in triplicate in appropriate substrate solutions (for sucrase, 60 mM sucrose in PBS; for maltase, 30 mM maltose in PBS; for aminopeptidase, 5 mM leu-pNA (Enzo Life Sciences) in 50 mM TRIS, pH 7.4). Cells were incubated for 30-60 min at 37 °C. For the aminopeptidase assay, absorbance was measured at 405 nm. For the sucrase and maltase assays, plates were centrifuged and 5 μl of supernatant was used for a colorimetric assay to measure liberated glucose (Cayman Chemical).

Metagenomic library preparation and sequencing. Methods for microbial community DNA and mRNA sequencing were as previously described^{44,45}. Faecal pellets collected from mice before or after LPS treatment (pooled from 2 and 3 days post-injection) were kept at -80 °C until DNA and RNA isolation using the guanidinium thiocyanate/caesium chloride gradient method⁴⁶ as described, except that crude particles were removed by centrifugation before overlaying the gradient. RNA was subjected to DNase treatment (Ambion), purification using MEGAClear columns

(Ambion), and rRNA depletion via subtractive hybridization (MICROBExpress, Ambion, in addition to custom depletion oligonucleotides). The presence of genomic DNA contamination was assessed by PCR with universal 16S rDNA primers. cDNA was synthesized using SuperScript II and random hexamers (Invitrogen), followed by second-strand synthesis with RNaseH and *E. coli* DNA polymerase (New England Biolabs). Samples were prepared for sequencing with an Illumina HiSeq instrument after enzymatic fragmentation (NEBE6040L/M0348S). Libraries were quantified by qPCR according to the Illumina protocol. qPCR assays were run using ABsoluteTM QPCR SYBR Green ROX Mix (Thermo Scientific) on an Mx3000P QPCR System instrument (Stratagene). The size distribution of each library was quantified on an Agilent HS-DNA chip (Agilent).

16S rRNA gene sequencing and analysis. Community DNA was PCR-amplified using universal bacterial primers targeting variable region 4 of the 16S rRNA gene with the following thermocycler protocol: denature at 94 °C for 3 min, 35 cycles of 94 °C for 45 s, 50 °C for 30 s, and 72 °C for 90 s, with a final extension at 72 °C for 10 min^{45,47,48}. Triplicate reactions for each sample were pooled and amplification was confirmed by 1.5% gel electrophoresis. 16S rRNA gene amplicons were cleaned with the Ampure XP kit (Agencourt) and quantified using the Quant-iT Picogreen ds DNA Assay Kit (Invitrogen). Barcoded amplicons from multiple samples were pooled and sequenced using the Illumina HiSeq platform⁴⁷. 16S rRNA gene sequences were analysed using the QIIME (Quantitative Insights Into Microbial Ecology)⁴⁴ software package along with custom Perl scripts. Data sets were randomly subsampled before clustering analyses at a depth that retained all of the individual samples (180,000 sequences per sample). All sequences were used for the comparison of the relative abundance of bacterial taxonomic groups. OTUs were picked at 97% similarity against the Greengenes database⁴⁹ (constructed by the nested_gg_workflow.py QiimeUtils script on 4 February 2011), which we trimmed to span only the 16S rRNA region flanked by our sequencing primers (positions 521-773). The LefSe package was used to identify taxonomic groups significantly associated with each treatment⁵⁰. LefSe was run on the sub-sampled data sets, after filtering out species-level OTUs with <100 sequences or present in only 1 sample. Statistical analyses were also performed using the QIIME scripts 'otu_category_significance' (ANOVA) and 'compare_categories.py' (PERMANOVA).

Reference genome database. A custom database was constructed from draft and finished reference genomes obtained from human-associated microbial isolates (538 genomes from the Human Microbiome Project Data Analysis and Coordination Center (<http://www.hmpdacc.org>)), in addition to the *Escherichia lenta* DSM2243 reference genome. All predicted proteins from the reference genome database were annotated with KEGG orthologous groups (KOs) using the KEGG database (version 52; BLASTX *e*-value < 10⁻⁵, bit score > 50, and >50% identity)⁵¹. For query genes with multiple matches, the annotated reference gene with the lowest *e*-value was used. When multiple annotated genes with an identical *e*-value were encountered after a BLAST query, we included all KOs assigned to those genes. Genes from the database with significant homology (BLASTN *e*-value < 10⁻²⁰) to non-coding transcripts from the 539 microbial genomes were excluded from subsequent analysis.

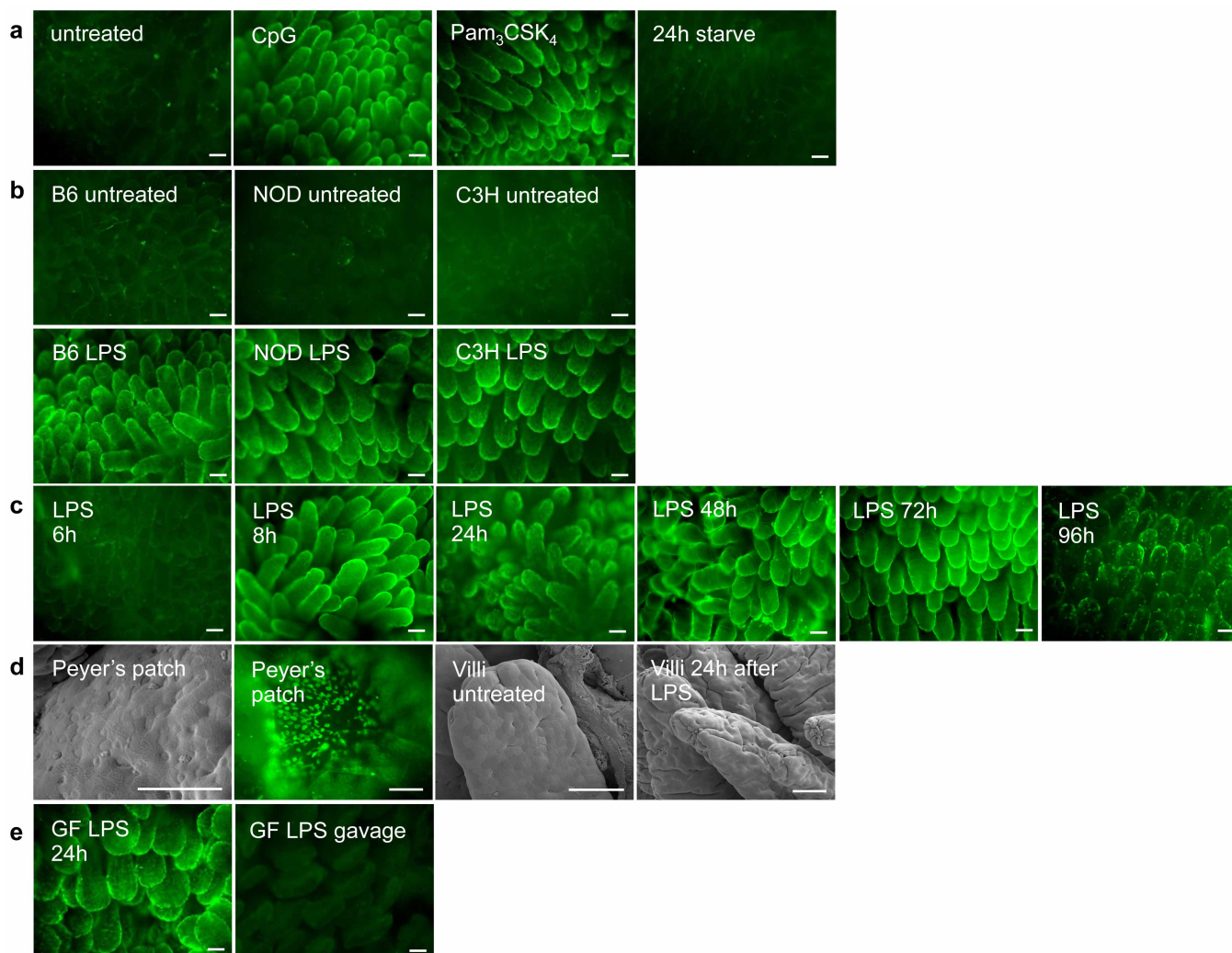
Metagenomic sequence analysis. DNA- and RNA-seq analysis was performed with our recently described pipeline⁴⁵. Briefly, high-quality reads (see Supplementary Table 1 for sequencing statistics) were mapped using SSAHA2⁵² to our custom 539-genome database and the Illumina adaptor sequences (SSAHA2 parameters: '-best 1 -score 20 -solexa'). The number of transcripts assigned to each gene was then tallied and normalized to reads per kilobase per million mapped reads (RPKM). To account for genes that were not detected owing to limited sequencing depth, a pseudocount of 0.01 was added to all samples. Samples were clustered in Matlab (version 7.10.0) based on gene expression or abundance. Genes were grouped by KOs by calculating the cumulative RPKM for each sample for genes present in at least six samples. We used HUMAnN, a recently developed software package for metabolic reconstruction from metagenomic data⁵³, followed by LefSe analysis to identify metagenomic biomarkers⁵⁰. A modified version of the "SConstruct" file was used to input KO counts into the HUMAnN pipeline for each RNA-seq data set. We then ran LefSe on the resulting KEGG module abundance file using the '-o 1000000' flag. We used the edgeR package⁵⁴ to identify orthologous groups with significantly altered abundance or expression. Prior to analysis, we calculated the cumulative number of sequencing reads assigned to each sample for each KO (without RPKM normalization). We then used a paired glm analysis to determine consistent changes within each animal following LPS treatment. Default parameters were used, with one exception: for the estimateGLMTrendedDisp step 'min.n' was set to 50/300 for the genomes and KO, respectively. Significance was accepted at a false discovery rate (FDR) < 0.05 and >2-fold change.

PCR and RT-PCR. RNA was isolated from mouse tissues and gut contents by the guanidinium thiocyanate/caesium chloride gradient method⁴⁶. RNA was DNase treated (Sigma) and reverse transcribed with Superscript III (Invitrogen). Primer sequences were as follows: fut1F, 5'-CAAGGAGCTCAGCTATGTGG-3', fut1R, 5'-GACTGCTCAGGACAGGAAGG-3'; fut2F, 5'-ACAGCCAGAAGAGCCAT

GGC-3', fut2R, 5'-TAACACCGGGAGACTGATCC-3'; sec1F, 5'-ATCCAAGC AGTGCTCCAGC-3', sec1R, 5'-CAATATTCGCCCATCTGGTTC-3'; villinF, 5'-G CTTGCCACAACCTCCTAAG-3', villinR, 5'-CTTGCTTGAAGTAGCTCCGG-3'. Quantitative RT-PCR was performed on an Applied Biosystems StepOnePlus instrument with Universal Sybr Green Universal Supermix (Bio-Rad), and the following primers²⁷: *E. coli* fucA F, 5'-GGCGCGCAAGGAATAGAA-3', *E. coli* fucA R, 5'-GATCCCCGCTATTCACATACATGA-3'; *E. coli* fucP F, 5'-CCAAATACGGTTC GTCCTTCA-3', *E. coli* fucP R, 5'-ACCCATGACCGGAGTGACAA-3'; *E. coli* rpoA F, 5'-GCGCTCATCTTCTCCGAAT-3', *E. coli* rpoA R, 5'-CGCGGTCTGGT TATGTG-3'.

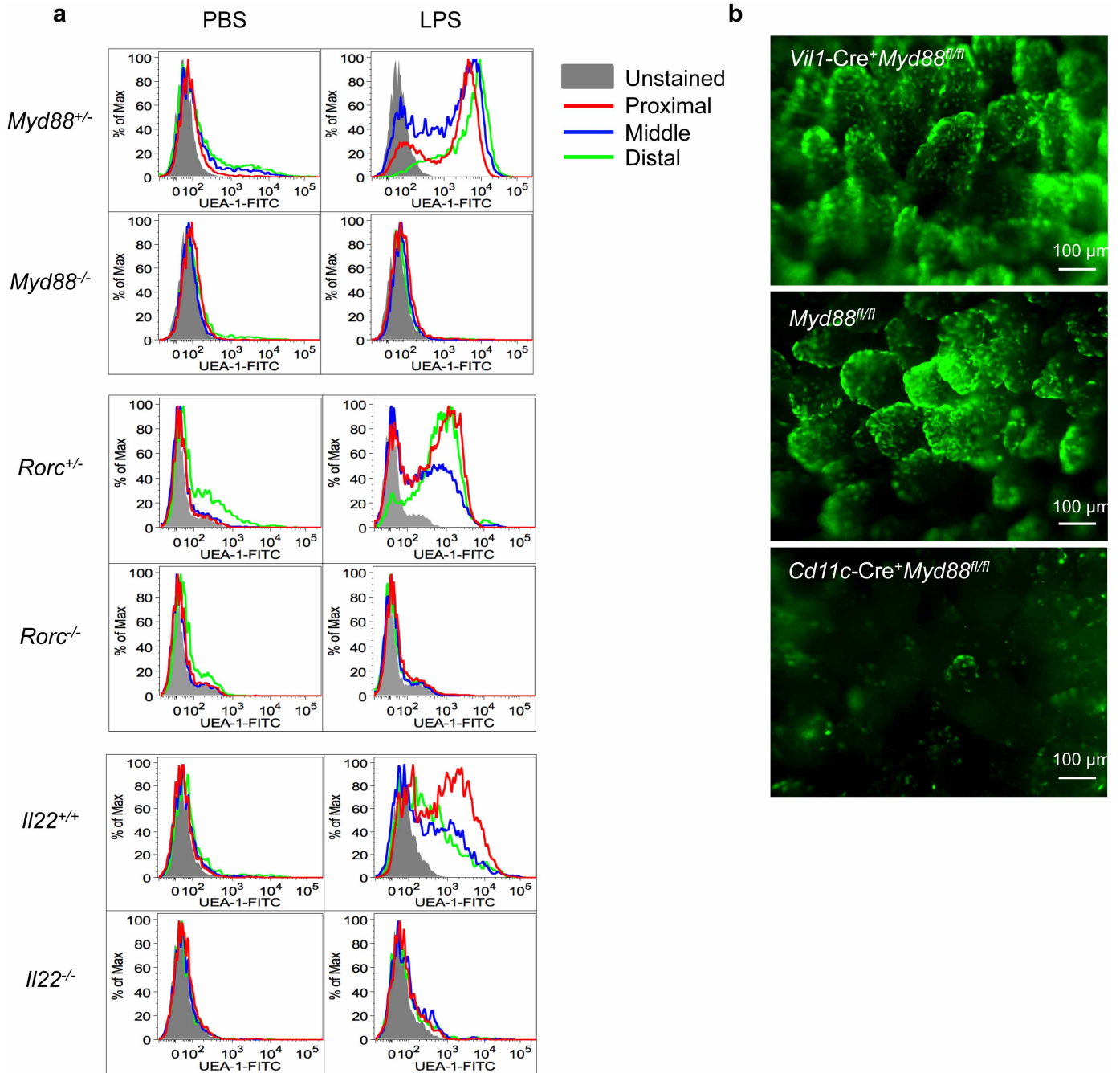
Statistics. Statistical analyses were performed with GraphPad Prism 5 software (GraphPad Software).

31. Kleinridders, A. *et al.* MyD88 signaling in the CNS is required for development of fatty acid-induced leptin resistance and diet-induced obesity. *Cell Metab.* **10**, 249–259 (2009).
32. Eberl, G. *et al.* An essential function for the nuclear receptor ROR γ t in the generation of fetal lymphoid tissue inducer cells. *Nature Immunol.* **5**, 64–73 (2004).
33. Cua, D. J. *et al.* Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* **421**, 744–748 (2003).
34. Zenewicz, L. A. *et al.* Interleukin-22 but not interleukin-17 provides protection to hepatocytes during acute liver inflammation. *Immunity* **27**, 647–659 (2007).
35. Malick, L. E. & Wilson, R. B. Modified thiocarbonylhydrazide procedure for scanning electron microscopy: routine use for normal, pathological, or experimental tissues. *Stain Technol.* **50**, 265–269 (1975).
36. Sakamoto, M. & Ohkuma, M. Identification and classification of the genus *Bacteroides* by multilocus sequence analysis. *Microbiology* **157**, 3388–3397 (2011).
37. Boronat, A. & Aguilar, J. Rhamnose-induced propanediol oxidoreductase in *Escherichia coli*: purification, properties, and comparison with the fucose-induced enzyme. *J. Bacteriol.* **140**, 320–326 (1979).
38. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).
39. Ubeda, C. *et al.* Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J. Clin. Invest.* **120**, 4332–4341 (2010).
40. Buffie, C. G. *et al.* Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to *Clostridium difficile*-induced colitis. *Infect. Immun.* **80**, 62–73 (2012).
41. Fleming, S. E., Traitler, H. & Koellreuter, B. Analysis of volatile fatty acids in biological specimens by capillary column gas chromatography. *Lipids* **22**, 195–200 (1987).
42. Tangerman, A. & Nagengast, F. M. A gas chromatographic analysis of fecal short-chain fatty acids, using the direct injection method. *Anal. Biochem.* **236**, 1–8 (1996).
43. Zhao, G., Nyman, M. & Jonsson, J. A. Rapid determination of short-chain fatty acids in colonic contents and faeces of humans and rats by acidified water-extraction and direct-injection gas chromatography. *Biomed. Chromatogr.* **20**, 674–682 (2006).
44. Caporaso, J. G. *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336 (2010).
45. Maurice, C. F., Haiser, H. J. & Turnbaugh, P. J. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* **152**, 39–50 (2013).
46. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry* **18**, 5294–5299 (1979).
47. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).
48. Caporaso, J. G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl Acad. Sci. USA* **108** (suppl. 1), 4516–4522 (2011).
49. DeSantis, T. Z. *et al.* Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **72**, 5069–5072 (2006).
50. Segata, N. *et al.* Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).
51. Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y. & Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **32**, D277–D280 (2004).
52. Ning, Z., Cox, A. J. & Mullikin, J. C. SSAHA: a fast search method for large DNA databases. *Genome Res.* **11**, 1725–1729 (2001).
53. Abubucker, S. *et al.* Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput. Biol.* **8**, e1002358 (2012).
54. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).



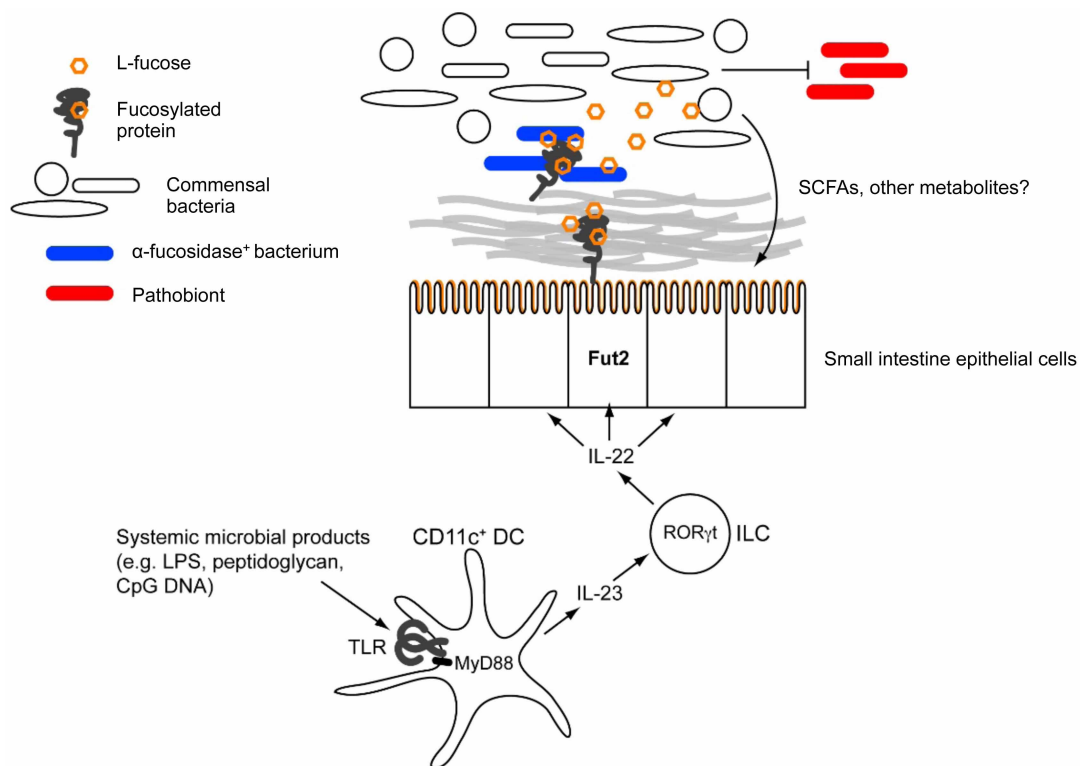
Extended Data Figure 1 | Requirements and kinetics for small intestine fucosylation induced by systemic injection of TLR ligands. **a**, Systemic injection of bacterial TLR ligands induces small intestine fucosylation, but simple starvation does not. UEA-1 staining (as in Fig. 1) after i.p. injection of CpG DNA, or Pam₃CSK₄, or food deprivation for 24 h of BALB/c SPF mouse. **b**, LPS injection causes small intestine fucosylation in various inbred mouse strains. SPF mice of the indicated strains were injected with LPS i.p. and the small intestine was stained with UEA-1 after 24 h, as in Fig. 1. **c**, Fucosylation peaks at 8 h after LPS injection and is still detectable at 96 h. **d**, M cells can be

readily detected by scanning electron microscope and UEA-1 staining of the domes of the Peyer's patches, but are rare in the villi and are not massively induced in the villi by LPS injection. UEA-1 staining and scanning electron microscopy were performed on adjacent pieces from the proximal one-third of the small intestine. Scale bars = 100 μ m for UEA-1 staining, 50 μ m for scanning electron microscope images. **e**, Small intestine fucosylation does not require the presence of endogenous microbiota (LPS injection in GF mouse) and is not induced by oral administration of LPS (1 mg). All data are representative of at least two independent experiments.



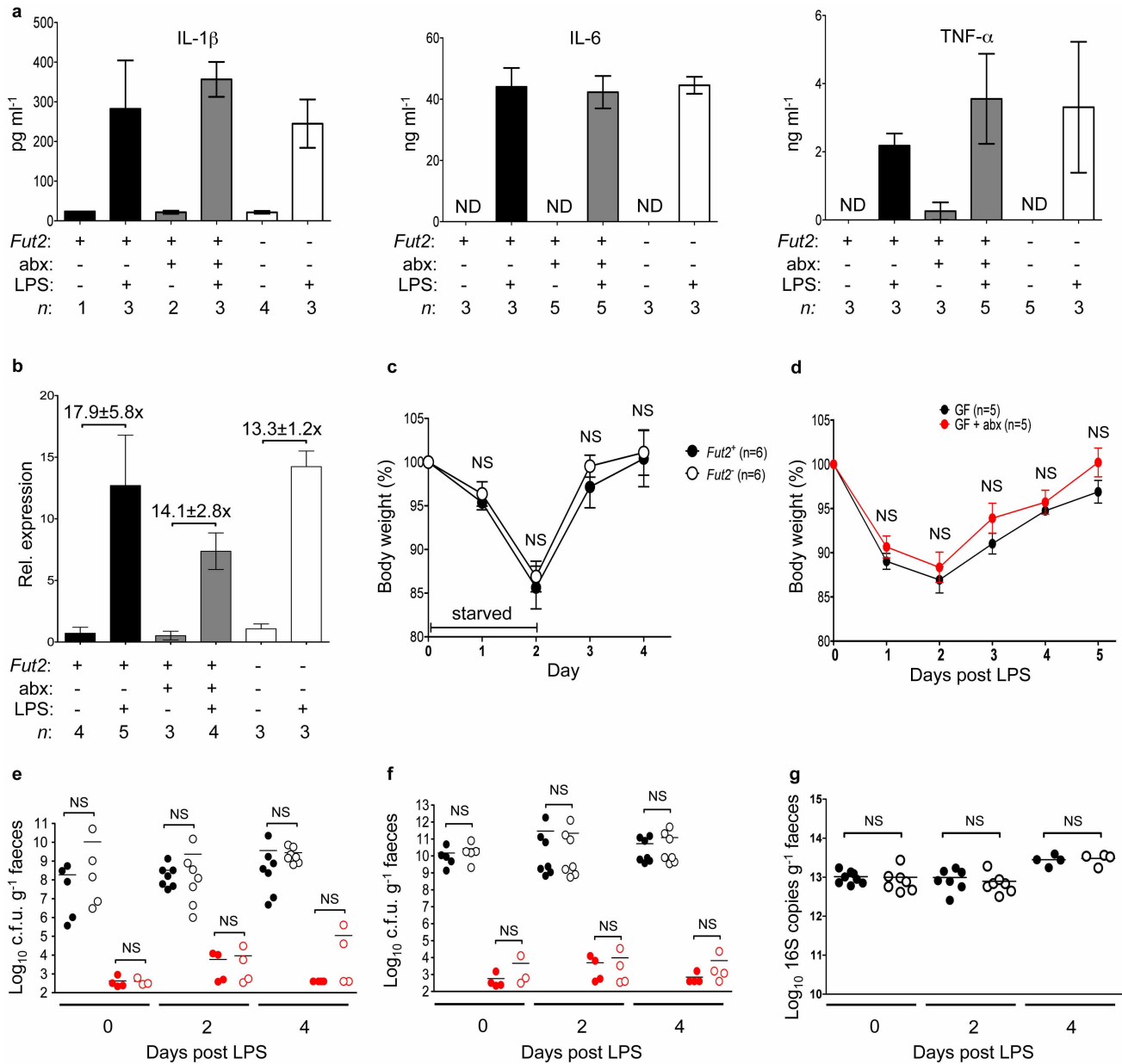
Extended Data Figure 2 | MyD88-dependent pathway for fucosylation of small intestine IECs in response to systemic stimulation of TLRs. **a**, FACS analysis of IECs from three segments of small intestine from the indicated mice. Cells are gated on the FSC/SSC high epithelial cell population. At least two mice per mutant genotype were stained along with two control mice in the

experiments shown. **b**, SPF mice were pre-treated with 20 mg streptomycin and orally infected with *S. enterica* Typhimurium. The small intestine was stained at 24 h post-infection. MyD88 expression was necessary in CD11c⁺ cells but not villin⁺ IECs for *S. enterica* Typhimurium-induced fucosylation. Data are representative of at least two independent experiments.



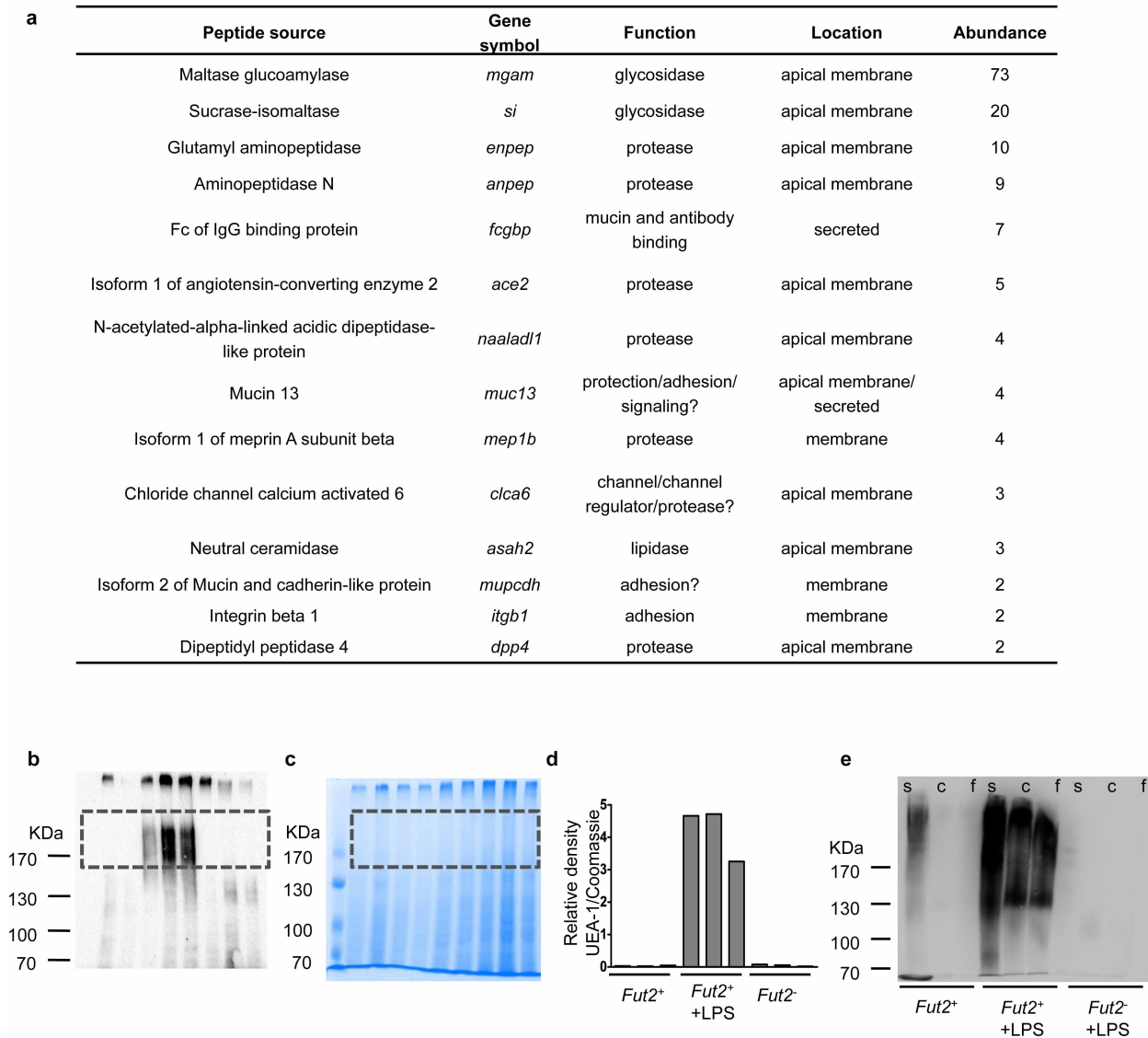
Extended Data Figure 3 | A proposed model for the mechanisms linking inducible fucosylation to the gut microbiota. Systemic microbial agonists activate TLRs on CD11c⁺ dendritic cells (DCs), causing secretion of the cytokine IL-23, which in turn stimulates ROR γ t-dependent ILCs to secrete IL-22. IL-22 causes small intestine epithelial cells to upregulate Fut2. Fucosylated proteins are either secreted into the lumen or expressed on the

cell surface and later shed into the lumen. Fucosidase-expressing bacteria (blue) liberate fucose residues, which they can utilize and share with other bacteria lacking the fucose-cleaving enzyme. Bacterial metabolism of fucose potentially produces metabolites such as short-chain fatty acids (SCFAs). Fucose also directly or indirectly downregulates virulence gene expression by pathobionts (red) or bona fide pathogens²⁷.



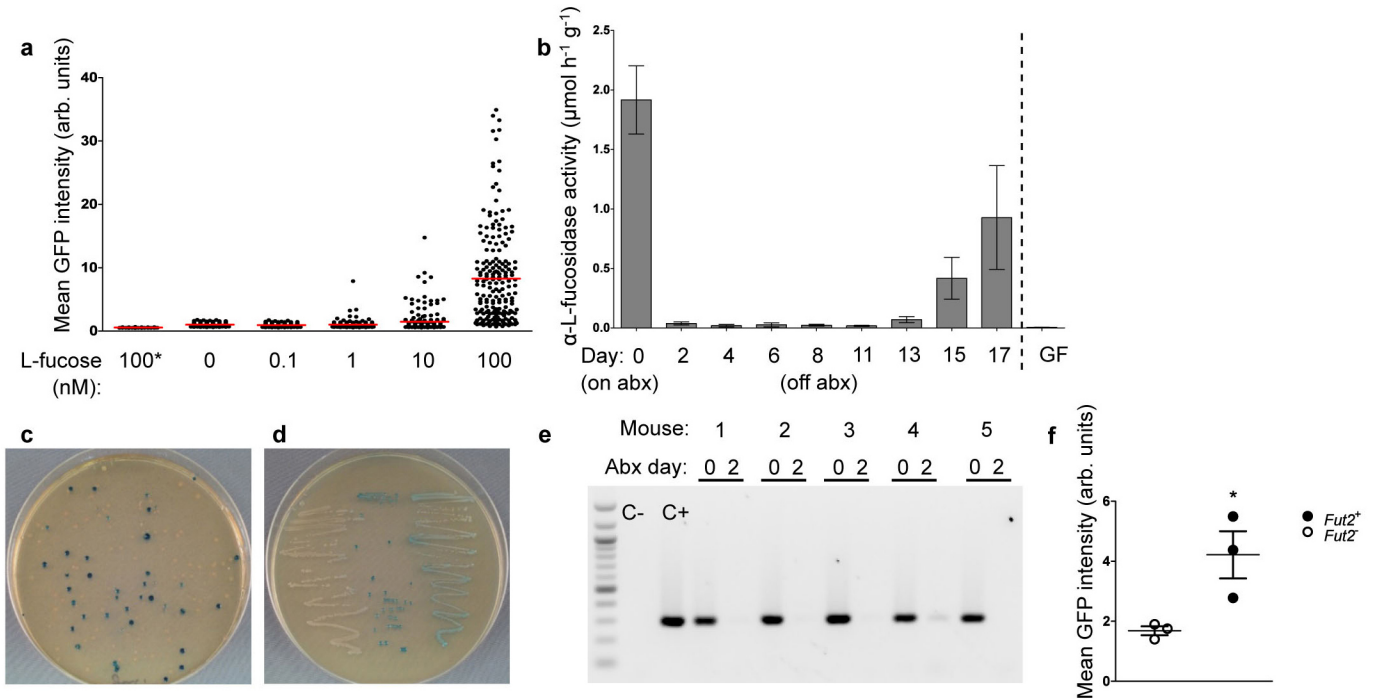
Extended Data Figure 4 | Consequences of LPS injection in Fut2-sufficient and Fut2-deficient BALB/c mice. **a**, Inflammatory cytokines IL-1 β , IL-6 and TNF- α were measured by ELISA in sera of mice before or 2 h after injection with LPS (4 h for IL-1 β). abx, mice on antibiotic water for 2 days before injection. Bars are mean \pm s.e.m.; ND, not detected. Data are combined from three experiments. **b**, Expression of RegIII γ (also regulated by the MyD88–IL23–IL22 pathway). Measurement by qPCR of *reg3g* gene expression in mid-small-intestine tissue, relative to *gapdh* (ddCt method). Numbers indicate mean fold change \pm s.e.m. in LPS-treated versus untreated mice. Differences between LPS-treated Fut2⁺ and abx or Fut2⁻ levels are not significant ($P > 0.05$, two-tailed Student's *t*-test). Data are combined from three experiments. **c**, Weight loss and recovery is not different in Fut2^{+/+} and Fut2^{-/-} mice after simple starvation (mean \pm s.e.m., $P > 0.05$ at all time

points, two-tailed Student's *t*-test; NS, not significant). **d**, Lack of direct toxic effect of antibiotics (abx) measured as the weight loss of BALB/c GF animals treated with LPS i.p. (mean \pm s.e.m., $P > 0.05$ by two-tailed Student's *t*-test at all time points). Data are combined from two experiments. **e–g**, Similar total bacterial loads in Fut2^{+/+} and Fut2^{-/-} mice before and after LPS injection and antibiotic treatment. Total bacterial loads in faeces were estimated by plating on aerobic (**e**) and anaerobic (**f**) non-selective media, and by qPCR for 16S gene copies (**g**). There were no significant differences between Fut2-sufficient (filled circles) and Fut2-deficient (open circles) mice before or after LPS treatment (two-tailed Student's *t*-test). Circles indicate individual mice; horizontal lines indicate means; red circles indicate antibiotic-treated mice. Data are combined from three experiments.



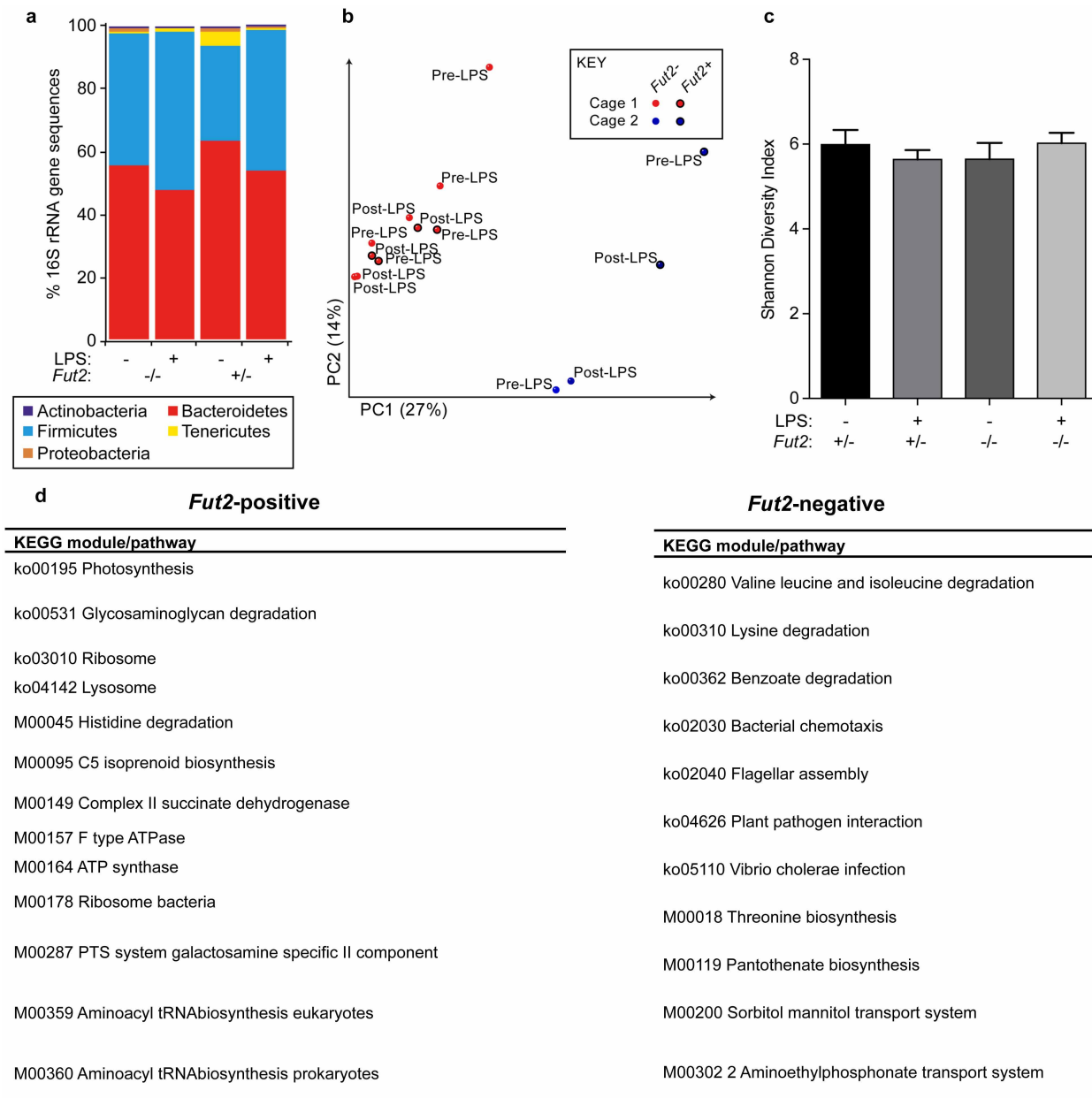
Extended Data Figure 5 | Fucosylated protein in IECs and gut contents.
a, Proteins $\alpha(1,2)$ fucosylated in IECs after LPS injection identified by UEA-1 precipitation and mass spectrometry. Abundance is the number of peptide fragments attributed to each gene. **b**, IECs from $Fut2^+$ untreated, $Fut2^+$ LPS-treated, or $Fut2^-$ untreated mice were isolated, and lysates separated by SDS-PAGE. $\alpha(1,2)$ fucosylated proteins were detected by blotting with UEA-1 lectin

conjugated to horseradish peroxidase (HRP). **c**, Identical gel stained with Coomassie blue for total protein content. **d**, Relative density of the boxed area of each lane from **b** divided by the relative density in **c**. **e**, UEA-1 staining of luminal proteins as in Fig. 3c. Blot is overexposed to show absence of luminal fucosylated proteins in the LPS-treated, $Fut2^-$ mouse. **b–e**, Data are representative of two independent experiments.



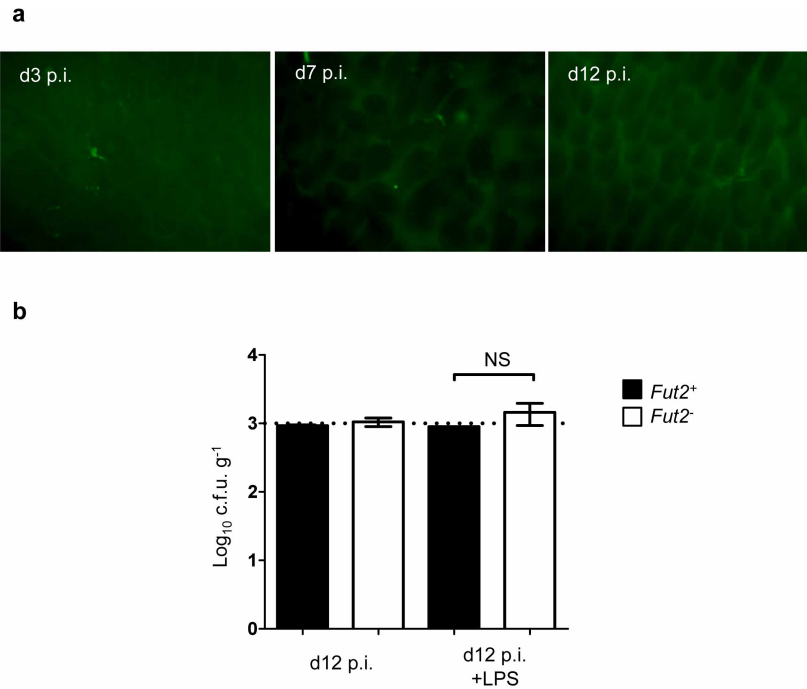
Extended Data Figure 6 | Generation of fucose-sensing reporter bacterial strains. **a**, Reporter *E. coli* were grown to stationary phase in minimal medium containing 10 mM glucose and the indicated concentrations of L-fucose (asterisk indicates promoter-less vector), and GFP fluorescence was measured. **b**, Fucosidase activity is dramatically reduced after 2 days of antibiotics (abx) treatment but recovers after cessation of treatment. Measurement of total α -L-fucosidase activity in faeces. Faecal supernatant was assayed for cleavage of 4-methylumbelliferyl-fucopyranoside substrate by fluorescence. $n = 5$ SPF antibiotics-treated, 3 GF mice. **c**, Faecal homogenates were plated anaerobically on BHIS agar containing 5-bromo-4-chloro-3-indolyl α -L-fucopyranoside, which forms a blue precipitate upon cleavage of the fucosyl residue. Both blue

and white colonies are present. **d**, Pure cultures of *Bacteroides* species were streaked on the same medium as in **c**. *B. uniformis* (left) is not predicted to carry an α -L-fucosidase gene, and remains white; *B. acidifaciens* (middle) and *B. thetaiotaomicron* (right) both express fucosidase activity and develop blue colonies. **e**, Loss of *B. acidifaciens* from the faeces of mice treated with antibiotics (Abx) in water (PCR for the *gyrB* gene). C-, water control; C+, *B. acidifaciens* genomic DNA. **f**, Summary of reporter *E. coli* experiments in SPF mice (representative experiment is shown in Fig. 3e). Points are mean GFP fluorescence from all reporter bacteria measured in each of three independent experiments ($n = 65$ bacteria per mouse; $*P < 0.05$, Student's t test).



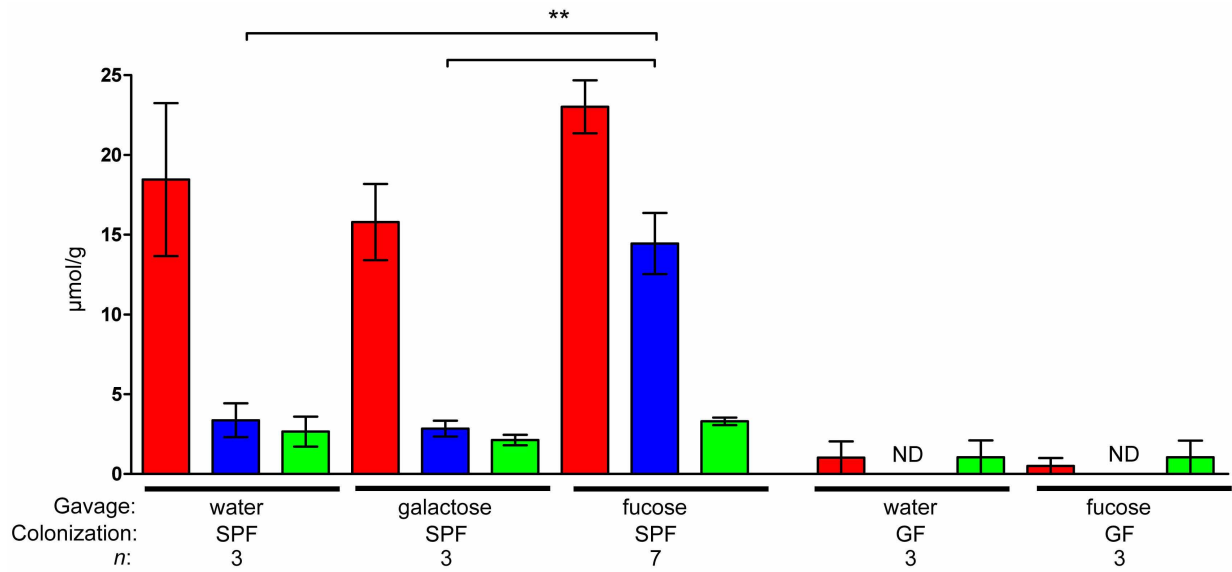
Extended Data Figure 7 | Microbial community structure is impacted by cohousing yet robust to host fucosylation and LPS exposure, whereas microbial gene expression depends on *Fut2*. **a**, Stable relative abundance of bacterial phyla across treatment groups and genotypes, as indicated by 16S rRNA gene sequencing. Values represent the mean abundance of phyla found at >1% relative abundance in at least one sample. **b**, Unweighted UniFrac analysis of the gut microbiota of *Fut2*-deficient (no outline) and *Fut2*-sufficient (black outline) mice. Points are coloured based on kinship and labelled by time

point (before or after LPS exposure). Results are based on 180,000 randomly selected 16S rRNA gene sequences per sample. **c**, Microbial diversity as measured by the Shannon diversity index ($n = 178,100$ sequences per sample). Values are mean \pm s.e.m. ($n = 3$ *Fut2*⁺, 4 *Fut2*⁻ mice per time point). **d**, KEGG modules and pathways expressed in microbiota at higher levels after LPS exposure in *Fut2*-positive (left) and *Fut2*-negative mice (right) ($n = 3$ per group; Humann/LefSe analysis; LDA >2).



Extended Data Figure 8 | Lack of inducible fucosylation and small intestine colonization in *C. rodentium*-infected mice. **a**, *C. rodentium* causes no small intestine fucosylation in SPF mice at day (d)3, day 7, or day 12 post-infection (p.i.). **b**, Small intestine colonization by *C. rodentium* is low regardless of

Fut2 expression and LPS treatment. Small intestine contents were removed by gentle squeezing, homogenized in PBS, and plated. Data are shown as mean \pm s.e.m.; $n = 4$. NS, not significant. Dotted line shows the limit of detection. Data are representative of two experiments.



Extended Data Figure 9 | Effect of exogenous fucose on caecal short-chain fatty acid levels. Caecal short-chain fatty acids were measured after gavaging starved mice with the indicated sugars (100 mM concentration). Fucose

gavage leads to increased propionate production in SPF but not GF mice. Data are shown as mean ± s.e.m. ** $P < 0.01$, Student's two-tailed t -test. ND, not detected. Data are combined from three experiments.

Enhanced neonatal Fc receptor function improves protection against primate SHIV infection

Sung-Youl Ko¹, Amarendra Pegu¹, Rebecca S. Rudicell¹†, Zhi-yong Yang¹†, M. Gordon Joyce¹, Xuejun Chen¹, Keyun Wang¹, Saran Bao¹, Thomas D. Kraemer², Timo Rath², Ming Zeng³†, Stephen D. Schmidt¹, John-Paul Todd¹, Scott R. Penzak⁴†, Kevin O. Saunders¹, Martha C. Nason⁵, Ashley T. Haase³, Srinivas S. Rao¹, Richard S. Blumberg², John R. Mascola¹ & Gary J. Nabel¹†

To protect against human immunodeficiency virus (HIV-1) infection, broadly neutralizing antibodies (bnAbs) must be active at the portals of viral entry in the gastrointestinal or cervicovaginal tracts. The localization and persistence of antibodies at these sites is influenced by the neonatal Fc receptor (FcRn)^{1,2}, whose role in protecting against infection *in vivo* has not been defined. Here, we show that a bnAb with enhanced FcRn binding has increased gut mucosal tissue localization, which improves protection against lentiviral infection in non-human primates. A bnAb directed to the CD4-binding site of the HIV-1 envelope (Env) protein (denoted VRC01)³ was modified by site-directed mutagenesis to increase its binding affinity for FcRn. This enhanced FcRn-binding mutant bnAb, denoted VRC01-LS, displayed increased transcytosis across human FcRn-expressing cellular monolayers *in vitro* while retaining FcγRIIIa binding and function, including antibody-dependent cell-mediated cytotoxicity (ADCC) activity, at levels similar to VRC01 (the wild type). VRC01-LS had a threefold longer serum half-life than VRC01 in non-human primates and persisted in the rectal mucosa even when it was no longer detectable in the serum. Notably, VRC01-LS mediated protection superior to that afforded by VRC01 against intrarectal infection with simian-human immunodeficiency virus (SHIV). These findings suggest that modification of FcRn binding provides a mechanism not only to increase serum half-life but also to enhance mucosal localization that confers immune protection. Mutations that enhance FcRn function could therefore increase the potency and durability of passive immunization strategies to prevent HIV-1 infection.

Antibody protection against viral infection is influenced by effector functions mediated by the Fc domain of the antibody interacting with activating Fc receptors (including FcγRIIIa-mediated ADCC⁴) or inhibitory Fc receptors on the surface of immune cells⁵. FcRn is a multifunctional receptor that plays a role in IgG transport and homeostasis^{1,2}. FcRn binds to the Fc portion of IgG with high affinity at an acidic pH (<6.5). Following endocytosis, free IgG is degraded in the lysosome, while FcRn-bound IgG avoids degradation at low pH and is recycled into the extracellular space, thus prolonging the half-life of IgG^{1,2}. Recently, several monoclonal antibodies with broad and potent neutralizing activity against diverse HIV-1 Env proteins have been shown to confer passive protection against SHIV infection^{3,6–11,29}. To determine whether the protective efficacy of such antibodies can be increased by modulation of FcRn effector function, we evaluated the role of mutations that increase binding to FcRn. To potentiate the effector function of the bnAb VRC01, we introduced mutations in the CH2 and/or CH3 domains of VRC01 and analysed the binding affinity of these mutated antibodies for human FcRn and FcγRIIIa. Five mutants known to enhance binding to human FcRn (Fig. 1a), along with a non-FcRn-binding

mutant (IHH)^{12–16}, were characterized *in vitro*. Similarly to VRC01, all mutants bound to a previously described HIV-1 Env protein, RSC3 (which consists of a resurfaced, stabilized core³ of HIV-1 glycoprotein (gp) 120) (Fig. 1b); this finding is consistent with previously published reports of the effects of these FcRn-binding mutations on other IgGs¹⁴. All mutants also had similar neutralization potency and breadth to VRC01 (Extended Data Table 1). These data confirm that the FcRn-binding mutations did not interfere with the recognition of cognate antigen or with neutralization.

We also compared the ability of the FcRn-binding mutants to bind to human FcRn at physiological or endosomal pH (7.4 and 6.0, respectively). As expected, all mutants except VRC01-IHH bound more strongly to FcRn at either pH than did VRC01 (Fig. 1c). While the binding of the Fc mutants to human FcRn at pH 7.4 by enzyme-linked immunosorbent assay (ELISA) was greater than that of VRC01 at saturating antibody concentrations (>10 μg ml⁻¹), differences in binding potency are best discriminated by using half-maximum binding concentrations (EC₅₀). Analysis of EC₅₀ values revealed greater binding by each of the FcRn-enabled mutants at pH 6.0 than at pH 7.4 (Extended Data Table 2), as expected and consistent with previous reports on the pH dependence of FcRn binding. It has been proposed that higher affinity binding to FcRn at pH 7.4 might inhibit the release of FcRn-bound IgG¹⁷. The FcRn-binding mutants of VRC01 were dissociated at pH 7.4 similar to VRC01 (Extended Data Fig. 1), indicating that VRC01 FcRn-binding mutations enhanced the pH-dependent binding but did not affect release at physiological pH.

FcRn affects the transport of IgG from the basolateral to the apical surface of mucosal epithelial cells¹. We therefore compared the transport of VRC01 and its FcRn-binding mutants across MDCK (Madin-Darby canine kidney) cells that express human FcRn and β₂-microglobulin in a transwell system *in vitro*. When the concentration of monoclonal antibodies in the upper (apical) chamber was measured 2 h after antibody addition to the lower (basolateral) chamber, the transcytosis of all mutants with enhanced FcRn binding was at least 2.5-fold higher than that of VRC01. The highest level of transcytosis was exhibited by the VRC01-LS and VRC01-YTE mutants (Fig. 1d). These results show that mutations which enhance FcRn binding increase the transport of VRC01 in a functional cell culture system. To confirm this effect *in vivo*, we evaluated the pharmacokinetics of these antibody mutants in human FcRn-transgenic mice¹⁸ (*n* = 4 per group). Similar to the cell culture results, all mutants with enhanced binding to human FcRn had a longer half-life than VRC01 (Extended Data Fig. 2).

In addition to virus neutralization, another Fc effector function has been implicated in immune protection. ADCC, which is mediated by IgG binding to FcγRIIIa, can lyse infected cells^{4,19}. Because enhanced

¹Vaccine Research Center, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Building 40, Room 4502, MSC-3005, 40 Convent Drive, Bethesda, Maryland 20892-3005, USA.

²Division of Gastroenterology, Department of Medicine, Brigham & Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, Massachusetts 02115, USA. ³Department of Microbiology, Medical School, University of Minnesota, 420 Delaware Street South East, Minneapolis, Minnesota 55455, USA. ⁴Clinical Pharmacokinetics Laboratory, Pharmacy Department, Clinical Center, National Institutes of Health, Building 10, 10 Center Drive, Bethesda, Maryland 20814, USA. ⁵Biostatistics Research Branch, Division of Clinical Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, 6700A Rockledge Drive, Room 5235, Bethesda, Maryland 20892, USA. †Present addresses: Sanofi, 640 Memorial Drive, Cambridge, Massachusetts 02139, USA (R.S.R., Z.-Y.Y. and G.J.N.); Center for Genetics of Host Defense, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75235-8505, USA (M.Z.); University of North Texas System College of Pharmacy, 3500 Camp Bowie Boulevard, RES-340J, Fort Worth, Texas 76107, USA (S.R.P.).

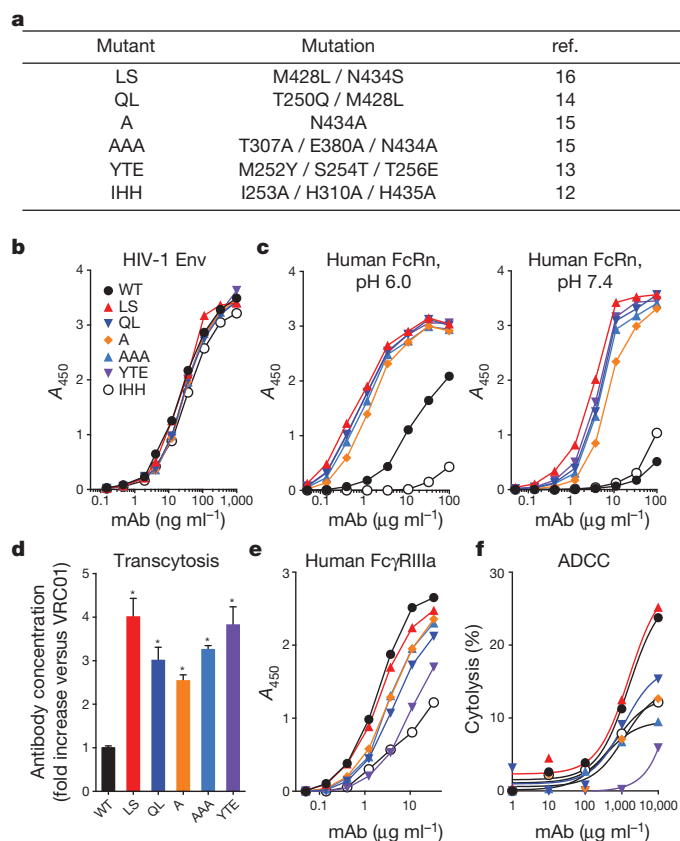


Figure 1 | *In vitro* evaluation of VRC01 and its FcRn-binding mutants. **a**, The FcRn-binding mutations used in this study are shown. The IHH mutant does not bind to human FcRn, and the other mutations enhance binding to human FcRn. **b**, Binding to HIV-1 Env (RSC3, a resurfaced, stabilized core of HIV-1 gp120) was examined by ELISA. **c**, Binding to human FcRn was assessed at pH 6.0 and pH 7.4 by ELISA. **d**, Transcytosis was assessed by adding $1.0 \mu\text{g ml}^{-1}$ of each monoclonal antibody (mAb) to the basolateral side of a monolayer of MDCK cells that express human FcRn, which were grown on a transwell filter plate, and incubating for 2 h. The concentration of each antibody on the apical side was measured by quantitative RSC3 ELISA, and the fold increase in antibody concentration over VRC01 (WT) is shown. *, $P < 0.001$ versus VRC01, two-tailed Student's *t*-test. **e**, Binding to human FcγRIIIa was examined by ELISA. **f**, The ability of each monoclonal antibody to induce ADCC was analysed using human peripheral blood mononuclear cells as effector cells and HIV-1-infected CEM-NK^R cells as targets and staining with the fluorescent dyes PKH26 and carboxyfluorescein succinimidyl ester (CFSE). The percentage killing was calculated from the percentage of CFSE⁻ cells within the PKH26^{hi} population. Assays were performed in duplicate (**b**, **c**, **e**, **f**) or quadruplicate (**d**), and the data points represent the mean values (**b**, **c**, **e**, **f**) or the mean \pm s.e.m. (**d**). The data are representative of two independent experiments (**b**–**f**). A_{450} , absorbance at 450 nm.

FcRn-binding mutations might affect the interactions of IgG with FcγRIIIa and thereby alter ADCC activity¹³, we assessed the FcγRIIIa binding and ADCC effector function of each mutant. We performed ELISA binding assays with human FcγRIIIa and ADCC assays using human peripheral blood mononuclear cells as effector cells and HIV-1-infected CEM-NK^R cells (a natural killer (NK)-cell-resistant human T leukaemia cell line) as targets²⁰. Most of the enhanced FcRn-binding mutants showed lower FcγRIIIa binding and ADCC activity than VRC01. Interestingly, only the VRC01-LS mutant exhibited both FcγRIIIa binding and ADCC activity comparable to VRC01 (Fig. 1e, f). We further examined the binding affinities of VRC01-LS for human FcRn and FcγRIIIa using surface plasmon resonance analysis. Consistent with the ELISA results, VRC01-LS exhibited a 12-fold higher human FcRn binding affinity than VRC01, whereas both displayed similar binding to human FcγRIIIa (Table 1). These data suggested that the increased FcRn binding of VRC01-LS

Table 1 | Comparison of binding affinities for human FcRn and FcγRIIIa

Binding to bnAb		K_{on} (1/nM × s)	K_{off} (1/s)	K_{d} (nM)	Fold K_{d} *
Human FcRn (pH 6.0)	WT	0.000094	0.2790	3,000	1.00
	LS	0.000223	0.0558	250	12.00
	IHH	NB	NB	NB	–
Human FcγRIIIa	WT	0.000080	0.105	1,300	1.00
	LS	0.000099	0.100	1,010	1.29
	IHH	0.000041	0.187	4,500	0.29

* Fold change in K_{d} versus wild type (WT). WT, VRC01; LS, VRC01-LS; IHH, VRC01-IHH (non-FcRn-binding mutant); K_{on} , association rate constant; K_{off} , dissociation rate constant; K_{d} , equilibrium dissociation constant; NB, not bound.

did not affect its FcγRIIIa-mediated effector function. We further examined the binding of these VRC01 variants to two other functional IgG receptors, FcγRIIa and FcγRIIb. We found no difference in the binding of VRC01 and VRC01-LS to either receptor, as determined by ELISA (Extended Data Fig. 3), suggesting that VRC01-LS does not differ from VRC01 with respect to ADCC effector function or immune suppression through FcγRIIa or FcγRIIb, respectively.

Because its ADCC activity was not affected by the FcRn-binding mutation, we selected the VRC01-LS mutant for further analysis in non-human primates. Its half-life *in vivo* was evaluated by injection (10 mg per kg body weight) and compared with VRC01, and VRC01-IHH, in Indian rhesus macaques (*Macaca mulatta*). The serum concentrations of each monoclonal antibody were measured by ELISA, and the half-lives were calculated using a two-compartment model (Fig. 2a). The LS mutant had a 2.5-fold longer half-life (VRC01, 4.65 days; VRC01-LS, 11.80 days) and a 2.5-fold slower clearance rate than the wild type (VRC01, $16.52 \text{ ml day}^{-1} \text{ kg}^{-1}$; VRC01-LS, $6.47 \text{ ml day}^{-1} \text{ kg}^{-1}$). By contrast, VRC01-IHH displayed markedly less persistence than the wild type, with a half-life of <1 day. This result confirmed that the increase in FcRn binding of VRC01-LS prolonged the antibody's half-life in non-human primates.

Since the gut mucosa is a primary site of HIV infection, we examined the accumulation of each monoclonal antibody in the rectal mucosa. Because direct tissue immunohistochemical staining is subject to sampling variation, we used a more quantitative approach: we quantified the total amount of VRC01 or VRC01-LS in rectal tissue homogenates, by using ELISA. Protein was extracted from rectal biopsy samples, and the amount of each monoclonal antibody was normalized to the total amount of tissue protein. After a single injection (10 mg kg^{-1}), VRC01-LS remained detectable in rectal tissue for more than 70 days and persisted at significantly higher levels than VRC01 ($P < 0.001$, two-tailed *t*-test), which was not detected after 28 days (Fig. 2b). VRC01-IHH was not detected in rectal tissues after 14 days. At lower concentrations of monoclonal antibody (0.3 mg kg^{-1}), a similar trend was seen in both the serum and the rectal mucosa. The levels of VRC01 in rectal tissue peaked at day 2, whereas VRC01-LS continued to increase up to day 5, after which it persisted at higher levels than VRC01 (Fig. 3a). In parallel studies, we also compared the relative monoclonal antibody levels in rectal secretions and tissue in cynomolgus macaques (*Macaca fascicularis*). Similarly to rhesus macaques, VRC01-LS circulated for threefold longer in the serum than did VRC01 (Extended Data Fig. 4a). Compared with VRC01, higher levels of VRC01-LS were measured in both rectal secretions and tissue at 28 days post antibody infusion, although the rectal tissue sampling was more consistent than the secretions (Extended Data Fig. 4b, c). We therefore suggest that tissue IgG provides an important indicator of mucosal IgG localization, consistent with a previous report²¹.

To determine whether VRC01 co-localized with FcRn *in vivo*, rectal biopsy samples taken after a single injection were stained either with a polyclonal antibody specific for human FcRn or with an HIV-1 Env probe, RSC3 (ref. 3). FcRn co-localized with VRC01 mainly in mucosal epithelial cells (Fig. 2c). Together, these data demonstrated that VRC01-LS persisted longer in the serum than VRC01 and accumulated in the mucosa as a result of its increased binding affinity for FcRn.

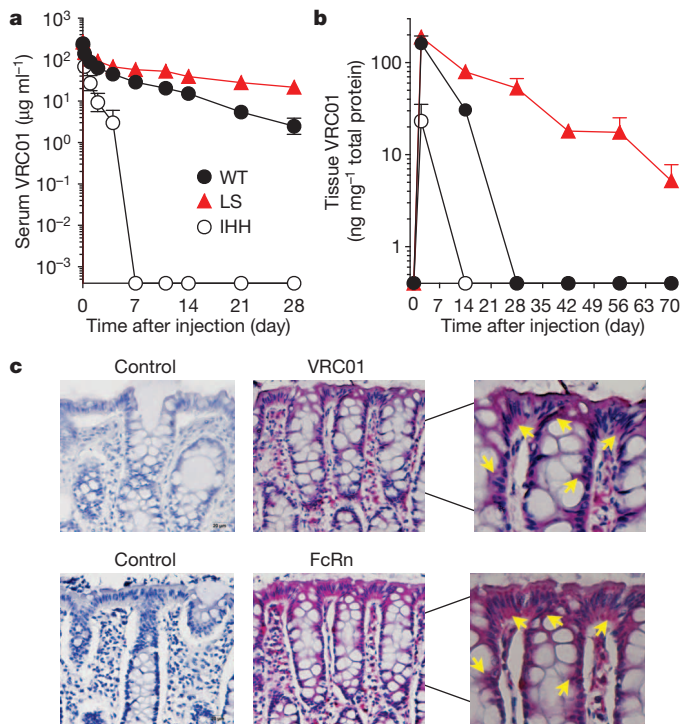


Figure 2 | Pharmacokinetic study in Indian rhesus macaques. **a**, The indicated VRC01 variant monoclonal antibodies were injected intravenously at 10 mg per kg body weight into rhesus macaques on day 0 (VRC01 and VRC01-LS, $n = 4$; VRC01-IHH, $n = 2$), and the serum levels over time are shown. The pharmacokinetic parameters were calculated with a two-compartment model (VRC01 versus VRC01-LS, half-life, 4.7 versus 11.8 days; clearance, 16.5 versus 6.5 $\text{ml day}^{-1} \text{kg}^{-1}$; area under the curve (AUC), 632 versus 1,645 $\text{day} \times \mu\text{g ml}^{-1}$). **b**, The amount of monoclonal antibody per mg of total tissue protein in rectal biopsy samples from each monoclonal-antibody-injected rhesus macaque ($n = 4$ per group) was quantitated and is shown over time. VRC01-LS persisted at higher levels than did VRC01 ($P < 0.001$, two-tailed t -test). The data points are the mean \pm s.e.m. **c**, Immunohistochemical staining of the rectal biopsy samples ($n = 6$ at 10 or 20 mg kg^{-1} dose) after VRC01 administration shows co-localization of VRC01, as measured by indirect binding to an Env probe, RSC3 (VRC01, red; top) and FcRn, stained with an anti-human FcRn polyclonal antibody (FcRn, red; bottom). The sections shown are representative of those observed in the six macaques. The adjacent sections were stained with RSC3 and the anti-human FcRn polyclonal antibody in test samples. Buffer without RSC3 and control rabbit IgG were used as negative controls for VRC01 and FcRn staining, respectively. The arrows indicate VRC01 or FcRn staining in the cytoplasm of epithelial cells. The images were captured at $\times 20$ original magnification, except for the right column, which shows magnified views of the centre column ($\times 40$ original magnification).

To determine whether VRC01-LS can protect against infection more effectively than VRC01, we performed an intrarectal challenge with the CC-chemokine receptor 5 (CCR5)-tropic SHIV BaLP4. Rhesus macaques ($n = 12$ per group) were injected intravenously with a low, sub-optimally protective, dose (0.3 mg kg^{-1}) of VRC01, VRC01-LS or control IgG and challenged intrarectally with SHIV 5 days after passive antibody transfer. All control-IgG-injected macaques became infected within 3 weeks of SHIV challenge, as measured by plasma viraemia. Ten of 12 macaques that received VRC01 were infected, in contrast to 5 of 12 VRC01-LS-injected macaques ($P = 0.0447$, one-tailed Fisher's exact test) (Fig. 3b and Extended Data Fig. 5). Among the infected animals, no significant differences in peak or set-point viraemia were observed between these experimental groups. VRC01-LS therefore conferred superior protection to VRC01 against acquisition of lentiviral infection.

In this study, we show that modification of a monoclonal antibody–FcRn interaction through mutations in the Fc domain can increase the protection afforded by the antibody against mucosal lentiviral challenge.

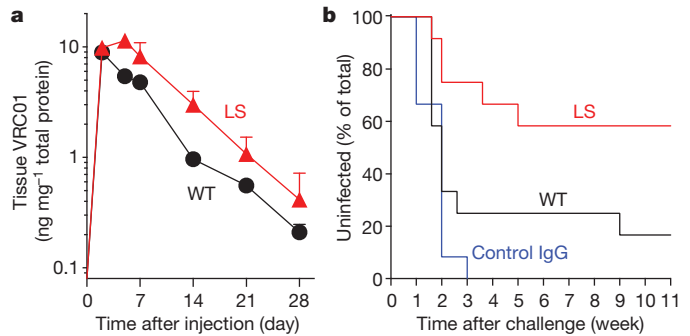


Figure 3 | VRC01-LS affords greater protection against intrarectal SHIV BaLP4 challenge than VRC01. The indicated monoclonal antibodies were administered intravenously to rhesus macaques at a dose of 0.3 mg kg^{-1} on day 0. **a**, Monoclonal antibody levels per mg of total tissue protein in rectal biopsy samples ($n = 4$ per group) are shown over time. The data points are mean \pm s.e.m. **b**, The macaques ($n = 12$ per group) were challenged with SHIV BaLP4 on day 5 after monoclonal antibody administration, and the percentage of uninfected macaques was plotted over time. Normal human IgG was given to the control group. Five of 12 VRC01-LS-treated macaques and 10 of 12 VRC01-treated macaques were infected ($P = 0.0447$, one-tailed Fisher's exact test).

Passive infusion of anti-HIV-1 bnAbs has been shown to prevent mucosal SHIV infection²², and the pharmacokinetics and biodistribution of the bnAbs play important roles in this protection²¹; however, the mechanism that confers mucosal protection has not been defined. Specifically, whether antibody transport and FcRn effector function contribute to mucosal immunity in primates was previously unknown. Our work demonstrates that FcRn plays an important role in IgG transport and provides insight into the mechanism by which FcRn function facilitates protection by IgG against mucosal viral infection. The increased FcRn binding of IgG induced by the LS mutation avoids endosomal degradation, allowing recycling into the serum, thereby increasing the serum antibody levels. The antibody retention in the rectal mucosa is probably due to antibody sequestration in the endosomes of epithelial cells (as indicated by the intracellular localization, Fig. 2c), which is presumably followed either by recycling into the subepithelial mucosal tissue for interactions with professional antigen presenting cells or by transcytosis into the intestinal lumen. Importantly, this mutation also increased the retention of IgG in the intestinal mucosa at the portal of HIV-1 entry, where a higher local concentration can more effectively neutralize local infection.

Our results suggest that active, FcRn-mediated transport, as opposed to simple diffusion, is a major component of IgG compartmentalization in the mucosal tissues of the rectum. FcRn-mediated transport of IgG to mucosal surfaces may play an unrecognized role in the ability of a variety of vaccines to successfully block mucosal viral infections. For example, vaccination with inactive poliovirus (also known as the Salk vaccine) generates an IgG response that can block poliovirus infection in the gut mucosa²³. There are similarities, as well as potential differences, between poliovirus and HIV-1 pathogenesis. Both poliovirus and HIV-1 invade the human body through mucosal tissues, especially the intestinal mucosa, but the receptors for cell entry are different. Poliovirus penetrates the body through M cells, in the Peyer's patches²⁴, whereas HIV-1 targets intraepithelial CD4^+ T cells in the mucosal epithelium. Although FcRn mediates the transcytosis of IgG across the intestinal epithelium *in vivo* as has been shown in a humanized mouse model²⁵, our results demonstrate that FcRn is also a means to retain IgG in tissues. This latter effect may restrain the quantities of IgG that are transported into the lumen and potentially explains why previous studies have failed to find an association between monoclonal antibody levels in mucosal secretions and protection²¹. Macaques with higher levels of bnAbs (VRC01-LS) in rectal tissues had a lower rate of infection than macaques injected with VRC01 (Fig. 3). Thus, it is important to measure IgG in tissues, in addition to in mucosal secretions.

The sexual transmission of HIV-1 is initiated by the replication of one or a small number of virions that penetrate the vaginal or rectal mucosa²⁶. The FcRn-mediated anatomical localization of bnAbs to the rectal mucosa may more effectively block penetration at this site of entry. This localization, combined with the quantitative effects of FcRn on serum IgG levels, explains how increasing the affinity for FcRn improves bnAb efficacy. Notably, FcRn expression has also been reported in the female genital tract²⁷: FcRn knockout in this previous paper was associated with decreased IgG-mediated protection against intravaginal herpes simplex virus (HSV-2) challenge. Because the FcRn knockout effects are global, it is unclear whether this effect was mediated through a systemic or mucosal effect; nonetheless, together, these studies are complementary and support the idea that FcRn mediates mucosal protection by IgG. The potential role of FcRn at other mucosal sites is further supported by the observation that VRC01 accumulates in vaginal tissues and localizes to the basal and parabasal layers of the vaginal epithelium, similarly to FcRn, as determined by immunohistochemistry (Extended Data Fig. 6). Presumably, FcRn-mediated transport increases the monoclonal antibody levels in this tissue, where it might similarly confer protection against infection.

Altering the pharmacokinetic profile of antibodies by increasing half-life and tissue localization has practical implications for the use of passive monoclonal antibody administration for immunoprophylaxis and therapy. One concern about the use of such monoclonal antibody mutations is whether they might elicit immune responses to the FcRn-binding mutation. We found no evidence of increased antibody responses to VRC01-LS or specifically to the LS mutation relative to VRC01 (Extended Data Fig. 7). Because VRC01 is human and xenogeneic to non-human primates, such questions about immunogenicity must eventually be evaluated in humans. In this regard, at least one human monoclonal antibody with an FcRn-binding-site mutation (motavizumab-YTE; M252Y / S254T / T256E) has been tested in healthy human adults and did not induce an appreciable increase in an anti-drug antibody response compared with the wild-type antibody²⁸.

Potent bnAbs with increased FcRn binding can be retained in relevant tissues for extended periods, decreasing the necessary amount of antibody, as well as the frequency of injections, thereby making the treatment or prevention of viral infections by passive antibody transfer more feasible than at present and broadly applicable. In addition to its potential importance for the optimization of passive immune prevention of HIV-1 infection of humans, this mechanism is also likely to be relevant to antibody-mediated protection against other mucosal pathogens such as rotavirus, poliovirus, norovirus and influenza virus.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 March; accepted 25 June 2014.

Published online 13 August 2014.

- Roopenian, D. C. & Akilesh, S. FcRn: the neonatal Fc receptor comes of age. *Nature Rev. Immunol.* **7**, 715–725 (2007).
- Ward, E. S. & Ober, R. J. Multitasking by exploitation of intracellular transport functions: the many faces of FcRn. *Adv. Immunol.* **103**, 77–115 (2009).
- Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329**, 856–861 (2010).
- Hessell, A. J. *et al.* Fc receptor but not complement binding is important in antibody protection against HIV. *Nature* **449**, 101–104 (2007).
- Nimmerjahn, F. & Ravetch, J. V. Fc γ receptors as regulators of immune responses. *Nature Rev. Immunol.* **8**, 34–47 (2008).
- Hessell, A. J. *et al.* Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. *PLoS Pathogens* **5**, e1000433 (2009).
- Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326**, 285–289 (2009).
- Hessell, A. J. *et al.* Broadly neutralizing monoclonal antibodies 2F5 and 4E10 directed against the human immunodeficiency virus type 1 gp41 membrane-proximal external region protect against mucosal challenge by simian-human immunodeficiency virus SHIVBa-L. *J. Virol.* **84**, 1302–1313 (2010).

- Walker, L. M. *et al.* Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466–470 (2011).
- Huang, J. *et al.* Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. *Nature* **491**, 406–412 (2012).
- Moldt, B. *et al.* Highly potent HIV-specific antibody neutralization *in vitro* translates into effective protection against mucosal SHIV challenge *in vivo*. *Proc. Natl Acad. Sci. USA* **109**, 18921–18925 (2012).
- Bitonti, A. J. *et al.* Pulmonary delivery of an erythropoietin Fc fusion protein in non-human primates through an immunoglobulin transport pathway. *Proc. Natl Acad. Sci. USA* **101**, 9763–9768 (2004).
- Dall'Acqua, W. F., Kiener, P. A. & Wu, H. Properties of human IgG1s engineered for enhanced binding to the neonatal Fc receptor (FcRn). *J. Biol. Chem.* **281**, 23514–23524 (2006).
- Hinton, P. R. *et al.* An engineered human IgG1 antibody with longer serum half-life. *J. Immunol.* **176**, 346–356 (2006).
- Petkova, S. B. *et al.* Enhanced half-life of genetically engineered human IgG1 antibodies in a humanized FcRn mouse model: potential application in humorally mediated autoimmune disease. *Int. Immunol.* **18**, 1759–1769 (2006).
- Zalevsky, J. *et al.* Enhanced antibody half-life improves *in vivo* activity. *Nature Biotechnol.* **28**, 157–159 (2010).
- Dall'Acqua, W. F. *et al.* Increasing the affinity of a human IgG1 for the neonatal Fc receptor: biological consequences. *J. Immunol.* **169**, 5171–5180 (2002).
- Roopenian, D. C. *et al.* The MHC class I-like IgG receptor controls perinatal IgG transport, IgG homeostasis, and fate of IgG-Fc-coupled drugs. *J. Immunol.* **170**, 3528–3533 (2003).
- Smalls-Mantey, A. *et al.* Antibody-dependent cellular cytotoxicity against primary HIV-infected CD4⁺ T cells is directly associated with the magnitude of surface IgG binding. *J. Virol.* **86**, 8672–8680 (2012).
- Gómez-Román, V. R. *et al.* A simplified method for the rapid fluorometric assessment of antibody-dependent cell-mediated cytotoxicity. *J. Immunol. Methods* **308**, 53–67 (2006).
- Klein, K. *et al.* Neutralizing IgG at the portal of infection mediates protection against vaginal simian/human immunodeficiency virus challenge. *J. Virol.* **87**, 11604–11616 (2013).
- Kwong, P. D., Mascola, J. R. & Nabel, G. J. Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nature Rev. Immunol.* **13**, 693–701 (2013).
- Onorato, I. M. *et al.* Mucosal immunity induced by enhance-potency inactivated and oral polio vaccines. *J. Infect. Dis.* **163**, 1–6 (1991).
- Sicinski, P. *et al.* Poliovirus type 1 enters the human host through intestinal M cells. *Gastroenterology* **98**, 56–58 (1990).
- Yoshida, M. *et al.* Human neonatal Fc receptor mediates transport of IgG into luminal secretions for delivery of antigens to mucosal dendritic cells. *Immunity* **20**, 769–783 (2004).
- Keele, B. F. *et al.* Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc. Natl Acad. Sci. USA* **105**, 7552–7557 (2008).
- Li, Z. *et al.* Transfer of IgG in the female genital tract by MHC class I-related neonatal Fc receptor (FcRn) confers protective immunity to vaginal infection. *Proc. Natl Acad. Sci. USA* **108**, 4388–4393 (2011).
- Robbie, G. J. *et al.* A novel investigational Fc-modified humanized monoclonal antibody, motavizumab-YTE, has an extended half-life in healthy adults. *Antimicrob. Agents Chemother.* **57**, 6147–6153 (2013).
- Pegu, A. *et al.* Neutralizing antibodies to HIV-1 envelope protect more effectively *in vivo* than those to the CD4 receptor. *Sci. Transl. Med.* **6**, 243ra88 (2014).

Acknowledgements We thank M. Roederer for advice on the design and statistical analysis of animal studies. This research was supported by the Intramural Research Program of the Vaccine Research Center, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), and in part by a grant from the Foundation for the National Institutes of Health with support from the Collaboration for AIDS Vaccine Discovery (CAVD), award OPP1039775, from the Bill & Melinda Gates Foundation. R.S.B. is supported by the NIH (DK044319, DK051362, DK053056 and DK088199) and the Harvard Digestive Diseases Center (DK0034854). T.R. is supported by the German research foundation (DFG; RA 2040/1-1). The findings and conclusions in this report are those of the authors and do not necessarily reflect the views of the funding agencies.

Author Contributions S.-Y.K., Z.-Y.Y., J.R.M. and G.J.N. designed the study, analysed the data and prepared the manuscript. A.P. analysed the data, set up the ADCC assay and provided the material for the ADCC assays. R.S.R. and K.O.S. helped to prepare the manuscript. M.G.J. performed the surface plasmon resonance analysis. X.C. performed DNA cloning and protein purification. T.D.K., T.R. and R.S.B. performed the transcytosis assay. S.B., M.Z. and A.T.H. performed immunohistochemical staining. S.D.S. and J.R.M. performed the neutralization assays. K.W., J.-P.T. and S.S.R. performed the pharmacokinetics and challenge study. S.R.P. analysed the pharmacokinetic data. M.C.N. conducted statistical analyses.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to G.J.N. (Gary.Nabel@sanofi.com) or J.R.M. (jmascola@nih.gov).

METHODS

Generation and characterization of VRC01 Fc mutants. The Fc mutations used in this study are indicated in Fig 1a. The amino acid positions are numbered according to the EU index of human IgG1 (ref. 30). Mutations were introduced to the heavy chain gene of VRC01 by site-directed mutagenesis using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). Mammalian expression plasmids encoding the heavy and light chain genes were transiently co-transfected into the human embryonic kidney cell line 293F, and the supernatant was harvested 5 days later. The monoclonal antibodies were purified from the supernatant by Protein G column chromatography (GE Healthcare).

The glycosylation patterns of VRC01 and VRC01-LS were analysed by using nanoLC-ESI-MS/MS (nano liquid chromatography electrospray ionization mass spectrometry) peptide sequencing technology. In brief, each solution sample was cleaned and digested in-gel with sequencing grade modified trypsin (Promega). The resultant peptide mixture was analysed by an LC-ESI-MS/MS system, in which high performance liquid chromatography (HPLC) with a 75- μm (inner diameter) reverse phase C18 column was coupled in-line with an ion trap mass spectrometer (Thermo Scientific). Comparing the glycosylation patterns of VRC01 and VRC01-LS by mass spectrometry, we focused on the glycan at amino acid position N297, which is the single *N*-glycosylation site in the Fc region of IgG1 isotype antibodies (such as VRC01). A very similar pattern of complex glycan configuration was noted for both monoclonal antibodies (Extended Data Table 3). G1F (GalGlcNAc₂Man₃GlcNAc₂Fuc) and G0F (GlcNAc₂Man₃GlcNAcFuc) are the major glycosylation species found in both VRC01 and VRC01-LS. While quantitative differences were noted in some glycan species, these differences represent small changes and fall within the limits of the expected experimental variability between preparations. More importantly, the LS mutation neither generated nor removed additional potential *N*-linked glycosylation patterns observed in VRC01.

ELISA. RSC3 ELISAs were used to analyse IgG binding. ELISA plates (96 well) were coated with 2 $\mu\text{g ml}^{-1}$ RSC3 in PBS, incubated at 4 °C overnight and blocked with PBS containing 5% BSA at room temperature for 1 h. Serial dilutions of each monoclonal antibody in 2.5% BSA in PBS were added to the plates and incubated at room temperature for 1 h. Horseradish peroxidase (HRP)-conjugated anti-human IgG, Fc γ -specific (Jackson ImmunoResearch Laboratories), was added and incubated at room temperature for 1 h. Tetramethylbenzidine (TMB, Sigma) HRP substrate was added to each well, and the absorbance of the yellow colour that developed after adding 0.5 M H₂SO₄ was measured at 450 nm.

The binding of IgG to human FcRn or human Fc γ receptors was examined as follows. His-tagged human FcRn was purified as previously described³¹, and His-tagged human Fc γ RIIIa, Fc γ RIIa and Fc γ RIIb were purchased from R&D Systems. Nickel-coated plates (96 well, Thermo Scientific) were coated with 2 $\mu\text{g ml}^{-1}$ His-tagged human FcRn or human Fc γ receptor and incubated at room temperature for 1 h. The ELISA was performed as described above. To examine the binding of monoclonal antibodies to human FcRn at pH 6.0, PBS at pH 6.0 was used instead of PBS at pH 7.4. EC₅₀ determination for human FcRn binding was calculated by computer-assisted nonlinear fitting using GraphPad Prism (Graphpad Software).

VRC01 and its mutants were quantified as follows. ELISA plates (96 well) were coated with 2 $\mu\text{g ml}^{-1}$ RSC3 in PBS, incubated at 4 °C overnight and blocked with PBS containing 5% BSA at room temperature for 1 h. Sera were diluted in 2.5% BSA in PBS and added to the plate and incubated at room temperature for 1 h, and the ELISA was carried out as described above. Purified monoclonal antibodies were used as a standard, and the limit of detection was 0.4 ng ml⁻¹.

Anti-VRC01 antibody responses were evaluated as follows. The sera from macaques that had been administered VRC01 or VRC01-LS were added to 96-well ELISA plates coated with 2 $\mu\text{g ml}^{-1}$ VRC01 or VRC01-LS. To test VRC01-LS-specific antibody responses, the sera from the two groups were absorbed with 20 $\mu\text{g ml}^{-1}$ VRC01, and the ELISA was carried out against VRC01-LS. The sera were diluted with TBS containing 5% skim milk, 2% BSA and 0.05% Tween 20, and the ELISA was performed as described above.

Neutralization assays. Antibody-mediated neutralization of pseudotyped HIV-1 isolates was measured using the TZM-bl luciferase reporter gene assay as described previously^{32–34}. The target baseline infection level was ~200,000 relative light units. Neutralization curves were fit using a five-parameter hill-slope equation as described previously³⁴. The concentrations of VRC01 and its mutants that were required to inhibit infection by 50% or 80% are reported as 50% or 80% inhibitory concentrations (IC₅₀ or IC₈₀), respectively.

Biacore binding assays. For surface plasmon resonance analysis, human FcRn and human Fc γ RIIIa (5 $\mu\text{g ml}^{-1}$ in 10 mM acetate, pH 5.5, for 10 min at a 10 $\mu\text{l min}^{-1}$ flow rate) were immobilized on a Biacore CM5 Sensor Chip to an immobilization level of ~500 response units (RU). VRC01 and its Fc mutants were injected for 2 min at various concentrations (0.01–10 μM) and allowed to dissociate for 3 min at a 50 $\mu\text{l min}^{-1}$ flow rate. The human FcRn surface was regenerated by two sequential injections of 1 \times PBS (pH 7.4) for 1 min at an 80 $\mu\text{l min}^{-1}$ flow rate, while the human

Fc γ RIIIa surface was regenerated with two sequential injections of 10 mM glycine (pH 2.1). All binding affinities were calculated using a 1:1 Langmuir fitting equation and Scrubber 2 (BioLogic Software), with an average value determined over the range of concentrations measured.

Animals. mFcRn^{-/-} hFcRn(276)^{Tg/Tg} mice were obtained from Jackson Laboratories (stock number 004919). They were housed and bred in the animal facility of the Vaccine Research Center (VRC) at the National Institute of Allergy and Infectious Diseases. *Macaca mulatta* animals of Indian origin and *Macaca fascicularis* were used in the non-human primate study. All animal experiments were reviewed and approved by the Animal Care and Use Committee of the VRC, and all animals were housed and cared for in accordance with local, state, federal and institutional policies in a facility accredited by the American Association for Accreditation of Laboratory Animal Care at the National Institutes of Health. Animal numbers per group in this study were chosen in accordance with previously published papers^{13,15,16}. Animals were chosen and randomized based on age and weight.

ADCC assay. ADCC assays were performed using a fluorescence activated cell sorting (FACS)-based method, as previously described²⁰. Briefly, HIV-IIIB-infected CEM NK^R-CCR5 cells were used as target cells and human peripheral blood mononuclear cells were used as effector cells. Target cells were double-stained with the fluorescent dyes PKH26 (Sigma) and carboxyfluorescein succinimidyl ester (CFSE; Invitrogen). Target cells (10,000) and serially diluted monoclonal antibody were added to each well of a 96-well plate, mixed thoroughly and incubated at room temperature for 15 min. Effector cells (500,000) were added to each well and mixed thoroughly. The plate was centrifuged for 3 min at 400 r.p.m. and incubated at 37 °C in a 5% CO₂ incubator for 4 h. The cells were washed with PBS, fixed with 1% paraformaldehyde in PBS, and stored at 4 °C overnight. The percentage of CFSE⁺ cells within the PKH26^{hi} population was obtained by using an LSR-II flow cytometer (BD Biosciences) and FlowJo software (TreeStar). Non-stained and single-stained targets were included in every experiment to compensate for single-stained CFSE and PKH26 emissions.

Transcytosis assay. Assessments of transcytosis of VRC01 and its Fc mutants were performed as previously described for IgG³⁵. Briefly, MDCK cells were transfected to express human β_2 -microglobulin and human FcRn or vector controls and were grown to confluence in DMEM containing 10% FBS and 1% penicillin and streptomycin on 12-mm transwells with a pore size of 0.4 μm (Corning) and allowed to polarize for 96 h. Twelve hours before the commencement of transcytosis experiments, the medium was changed to serum-free medium without antibiotics. On the day of the experiment, the measured transepithelial resistance ranged between 150 and 200 Ωcm^2 . Transwells were incubated in Hank's balanced salt solution (HBSS), pH 7.4, for 20 min followed by equilibration at 36 °C and 5% CO₂ in a 12-well plate containing HBSS, pH 6.0, in the input chamber and HBSS, pH 7.4, in the output chamber. pH 6.0-adjusted VRC01 variants at varying concentrations in HBSS were then directly added to the input chamber. After incubating for 2 h at 36 °C and 5% CO₂, the medium in the output chamber was collected, and the monoclonal antibody concentrations were quantified by ELISA. Mycoplasma-free stocks of transfected MDCK cells using the Venor GeM Mycoplasma Detection Kit (Sigma, MP0025) were maintained as aliquots and regularly thawed for use every 3–4 weeks.

Pharmacokinetic experiments in rhesus macaques. Female Indian rhesus macaques weighing between 2.9 and 6.7 kg were randomly assigned to groups according to body weight (four macaques per group for VRC01 and VRC01-LS, and two macaques per group for VRC01-IHH) and were intravenously injected with 10 mg kg⁻¹ monoclonal antibody. Blood was collected before injection on day 0, at 30 min and 6 h after injection, and at 1, 2, 4, 7, 14, 21 and 28 days after injection. The concentration of each monoclonal antibody was quantified by ELISA. Pharmacokinetic parameters were calculated with a two-compartment model using the WinNonlin software (Pharsight).

Immunohistochemical staining. All staining procedures were performed as previously described³⁶, using 5- μm tissue sections mounted on glass slides. Tissues were deparaffinized and rehydrated in deionized water. Heat-induced epitope retrieval was performed using a water bath (98 °C) in EDTA Decloaker (Biocare Medical), followed by cooling to room temperature. Tissue sections were then blocked with Sniper blocking reagent (Biocare Medical) for 30 min at room temperature. RSC3-biotin was diluted at 1:100 in TNB (0.1 M Tris-HCl, pH 7.5, 0.1 M NaCl and 0.05% Tween 20 with DuPont Blocking Reagent) and incubated overnight at 4 °C. After the RSC3-biotin incubation, the sections were washed with PBS containing 0.1% Tween 20 and then incubated with anti-biotin antibody conjugated to alkaline phosphatase in TNB for 2 h at room temperature. After the incubation with anti-biotin antibody, the sections were washed with PBS containing 0.1% Tween 20. Signal was detected with the Warp Red Chromogen Kit (Biocare Medical). The sections were counterstained with Harris Hematoxylin (Surgipath), dehydrated rapidly in a gradient of ethanols and mounted with Permount (Fisher Scientific). Stained sections were examined by light microscopy at ambient temperatures. Light micrographs were

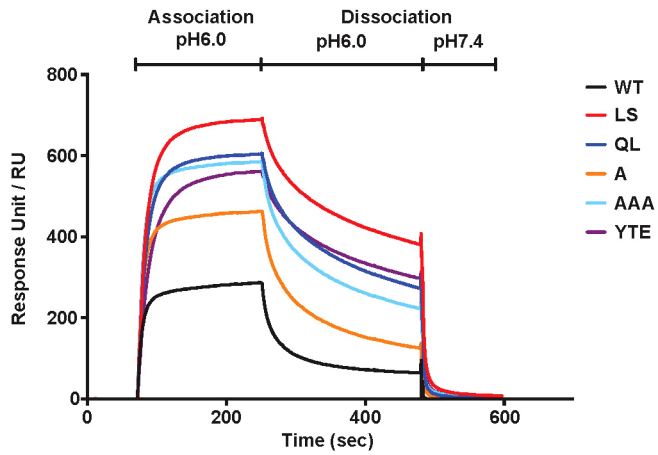
taken using an Olympus BX60 upright microscope with the following objectives: $\times 10$ (0.3 NA), $\times 20$ (0.5 NA) and $\times 40$ (0.75 NA); images were acquired using a SPOT colour mosaic camera (model 11.2; Diagnostic Instruments) and SPOT acquisition software (version 4.5.9; Diagnostic Instruments). Anti-biotin antibody conjugated to alkaline phosphatase was used as a negative control antibody in all instances and yielded negative staining results. For staining of human FcRn, heat-induced epitope retrieval was performed in Diva Decloaker (Biocare Medical) at 96 °C for 20 min, and polyclonal anti-human-FcRn antibody (1:50 dilution; Sigma, HPA012122) and anti-rabbit IgG conjugated to alkaline phosphatase were used as primary and secondary antibodies, respectively.

Quantitation of monoclonal antibodies in rectal biopsies. As previously described³⁷, rectal biopsy samples were added to 1.5-ml tubes containing 100 μ l PBS with EDTA-free protease inhibitor (Roche). The samples were ground on ice for 30 s and centrifuged at 4 °C for 20 min at 12,000 r.p.m., and the supernatants were collected. The total protein amounts were quantified by a BCA Protein Assay Kit (Thermo Scientific), and VRC01 and its Fc mutants were quantified by ELISA. The monoclonal antibody amount per mg of total protein is shown. The trends over days 2–28 were compared between groups by assuming a linear model for each animal and then testing the parameters from the linear models for differences between groups using a two-tailed *t*-test.

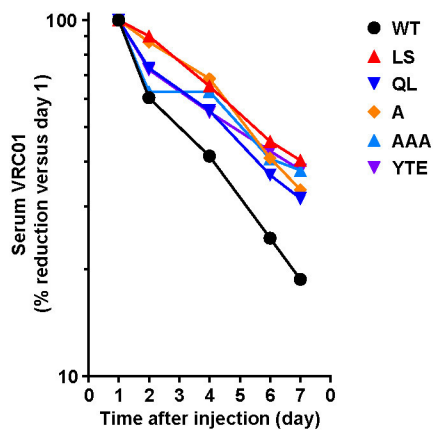
Challenge experiment in rhesus macaques. Male Indian rhesus macaques weighing 3.2–4.8 kg were randomly assigned to groups according to body weight, intravenously injected with 0.3 mg kg⁻¹ each monoclonal antibody and intrarectally challenged with undiluted SHIV BaLP4 5 days after the monoclonal antibody injection. Blood was collected twice a week for up to 4 weeks and then once a week for up to 12 weeks. Plasma was isolated from the blood samples. The plasma viral RNA levels were determined using a modified two-step quantitative reverse transcription

PCR process. Experimental samples were run in parallel with a simian immunodeficiency virus (SIV) *gag* RNA standard on an Applied Biosystems StepOne Real-Time PCR System. The lower limit of detection using this assay was 250 copies of SIV RNA ml⁻¹. The forward primer used was 5'-GTC TGC GTC ATC TGG ATT C-3', and the reverse primer was 5'-CAC TAG GTG TCT CTG CAC TAT CTG TTT TG-3'.

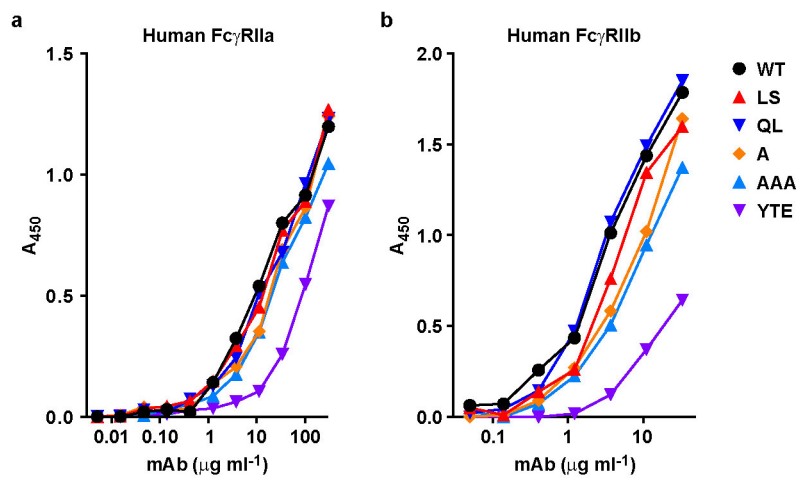
30. Edelman, G. M. *et al.* The covalent structure of an entire γ G immunoglobulin molecule. *Proc. Natl Acad. Sci. USA* **63**, 78–85 (1969).
31. Gastinel, L. N., Simister, N. E. & Bjorkman, P. J. Expression and crystallization of a soluble and functional form of an Fc receptor related to class I histocompatibility molecules. *Proc. Natl Acad. Sci. USA* **89**, 638–642 (1992).
32. Li, M. *et al.* Human immunodeficiency virus type 1 *env* clones from acute and early subtype B infections for standardized assessments of vaccine-elicited neutralizing antibodies. *J. Virol.* **79**, 10108–10125 (2005).
33. Shu, Y. *et al.* Efficient protein boosting after plasmid DNA or recombinant adenovirus immunization with HIV-1 vaccine constructs. *Vaccine* **25**, 1398–1408 (2007).
34. Seaman, M. S. *et al.* Tiered categorization of a diverse panel of HIV-1 Env pseudoviruses for assessment of neutralizing antibodies. *J. Virol.* **84**, 1439–1452 (2010).
35. Claypool, S. M., Dickinson, B. L., Yoshida, M., Lencer, W. I. & Blumberg, R. S. Functional reconstitution of human FcRn in Madin–Darby canine kidney cells requires co-expressed human β_2 -microglobulin. *J. Biol. Chem.* **277**, 28038–28050 (2002).
36. Zeng, M. *et al.* Cumulative mechanisms of lymphoid tissue fibrosis and T cell depletion in HIV-1 and SIV infections. *J. Clin. Invest.* **121**, 998–1008 (2011).
37. Baker, K. *et al.* Neonatal Fc receptor expression in dendritic cells mediates protective immunity against colorectal cancer. *Immunity* **39**, 1095–1107 (2013).



Extended Data Figure 1 | Enhanced FcRn-binding mutants of VRC01 are released from human FcRn at pH 7.4. VRC01 or the indicated FcRn-binding mutants were injected in PBS (pH 6.0) at a concentration of 100 nM over human-FcRn-immobilized Biacore CM5 sensor chips (~500 response units (RU)) and were dissociated using PBS at pH 6.0 followed by PBS at pH 7.4.

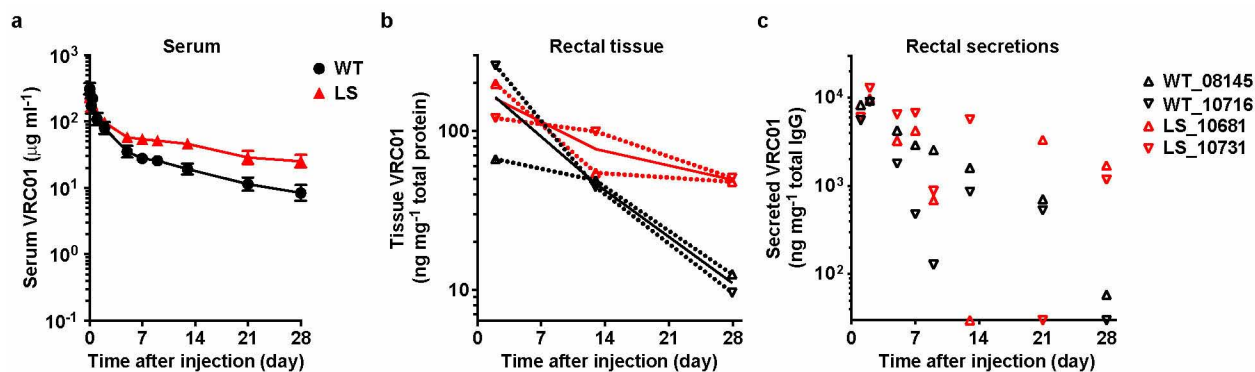


Extended Data Figure 2 | Mutants with enhanced human FcRn binding persist at higher serum concentrations than VRC01 in the human-FcRn(276)-transgenic mouse. The indicated VRC01-derived monoclonal antibodies (2 mg kg^{-1}) were injected intravenously into 6–8-week-old human-FcRn(276)-transgenic mice ($n = 4$, male and female mice evenly distributed). Serum concentrations were assessed by indirect ELISA against RSC3 over time. The percentage of the monoclonal antibodies remaining in the serum is shown compared with the percentage on day 1 (set at 100%).



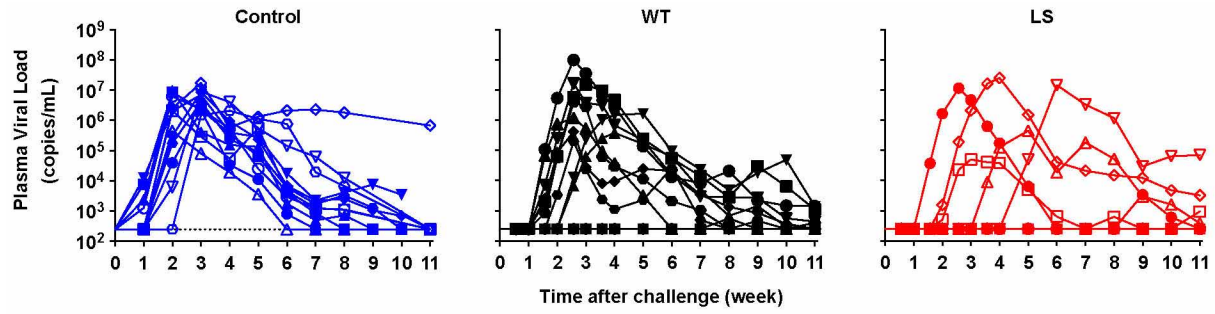
Extended Data Figure 3 | Binding of VRC01, VRC01-LS and other FcRn-binding mutants to Fc γ RIIa and Fc γ RIIb. Binding to the indicated

human Fc receptors was evaluated by ELISA as described in Methods.



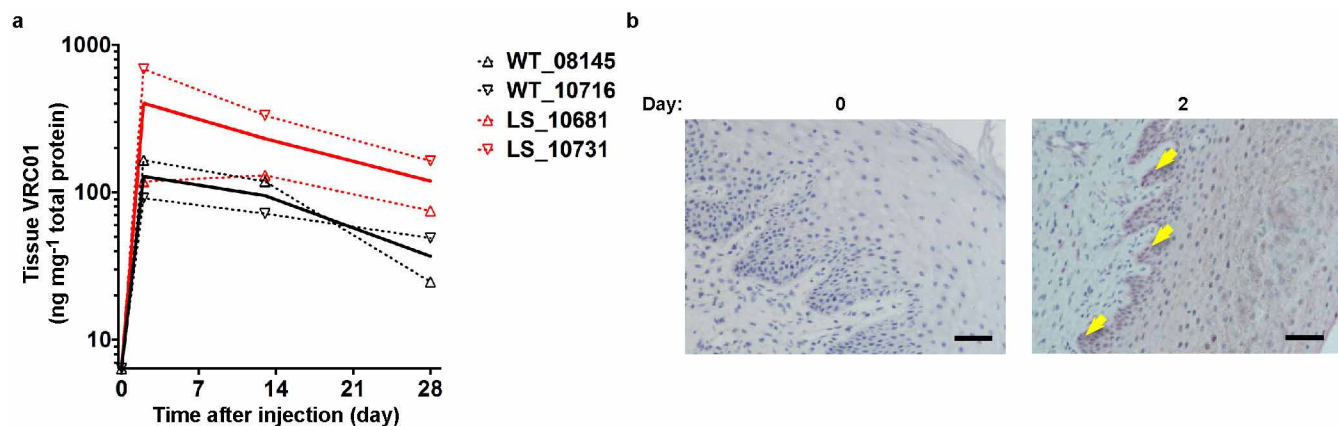
Extended Data Figure 4 | Serum pharmacokinetics, rectal tissue accumulation and mucosal secretion of VRC01-derived monoclonal antibodies in cynomolgus macaques. VRC01 or VRC01-LS (10 mg kg^{-1}) were injected intravenously into cynomolgus macaques. **a**, The serum levels ($n = 4$) were analysed over time. **b, c**, The amounts of monoclonal antibody per mg of total tissue protein in rectal biopsy samples (**b**) and per mg of total IgG in rectal secretions (**c**) from each monoclonal-antibody-injected cynomolgus

macaque ($n = 2$ per group) were quantitated and are shown over time. The values for each macaque and the mean values for the groups are shown as dotted lines and heavier solid lines, respectively (**b**). Pharmacokinetic parameters were calculated with a two-compartment model (VRC01 versus VRC01-LS, half-life, 9.0 versus 30.3 days; clearance, 15.7 versus $3.7 \text{ ml day}^{-1} \text{ kg}^{-1}$; area under the curve (AUC), 896 versus $2,812 \text{ day} \times \mu\text{g ml}^{-1}$).



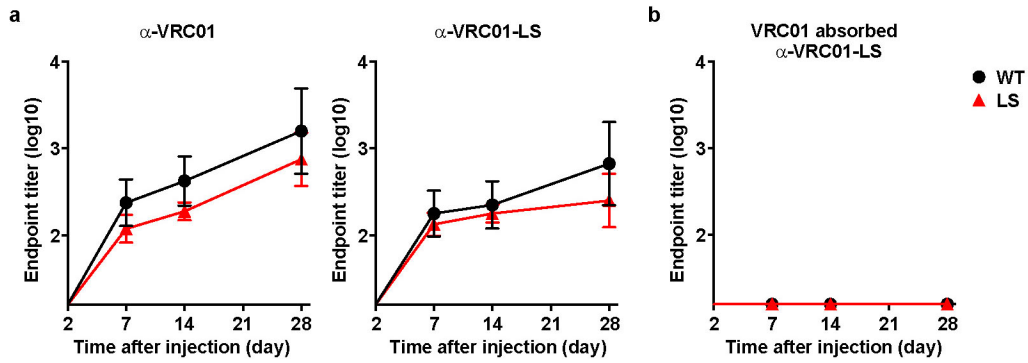
Extended Data Figure 5 | Viral load measurements over time after passive antibody transfer and SHIV challenge of non-human primates. The indicated monoclonal antibodies (0.3 mg kg^{-1}) were administered intravenously to rhesus macaques ($n = 12$). Macaques were challenged with

SHIV BaLP4 5 days later, and the plasma viral loads were measured over time. Normal human IgG was given to the control group. Five of 12 VRC01-LS-injected macaques and 10 of 12 VRC01-injected macaques were infected ($P = 0.0447$, one-tailed Fisher's exact test).



Extended Data Figure 6 | Accumulation and localization of VRC01 in vaginal tissue. **a**, VRC01 or VRC01-LS (10 mg kg^{-1}) were injected intravenously into female cynomolgus macaques ($n = 2$). The monoclonal antibody concentration per mg of total tissue protein in vaginal biopsy samples was quantitated at the indicated times. The values for each macaque and the mean values for the groups are shown as dotted and solid lines, respectively.

b, VRC01 (20 mg kg^{-1}) was injected intravenously into rhesus macaques ($n = 2$), and vaginal biopsy samples were taken and processed. Immunohistochemical staining was performed before and after the antibody dosing. The arrows indicate VRC01 (RSC3 staining, red) in basal and parabasal epithelial cells. The sections are representative of the sections assessed for two macaques. Scale bar, $50 \mu\text{m}$.



Extended Data Figure 7 | Similar antibody responses to VRC01 and VRC01-LS in non-human primates. VRC01 or VRC01-LS (10 mg kg^{-1}) was injected intravenously into Indian rhesus macaques ($n = 4$). **a**, Anti-VRC01 antibody responses (left) or anti-VRC01-LS antibody responses (right) were

evaluated by ELISA. **b**, Sera from animals that were administered VRC01 or VRC01-LS were absorbed with $20 \mu\text{g ml}^{-1}$ VRC01, and binding to VRC01-LS was then evaluated by ELISA.

Extended Data Table 1 | Comparative potency of HIV-1 neutralization by VRC01 and enhanced FcRn-binding mutants

	IC ₅₀ (µg/mL)*					
	WT	LS	QL	A	AAA	YTE
Q23.17	0.063	0.082	0.067	0.074	0.080	0.079
Q842.d12	0.020	0.028	0.020	0.019	0.019	0.021
YU2	0.139	0.166	0.144	0.168	0.154	0.176
JR-FL	0.026	0.031	0.027	0.027	0.027	0.032
7165.18	>50	>50	>50	>50	>50	>50
Du156.12	0.090	0.097	0.085	0.098	0.110	0.093
ZM109.4	0.094	0.154	0.109	0.136	0.084	0.188
SVA.MLV	>50	>50	>50	>50	>50	>50
SIVmac251.30	>50	>50	>50	>50	>50	>50

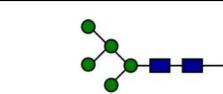
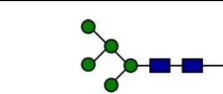
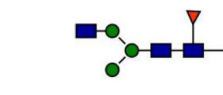
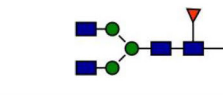
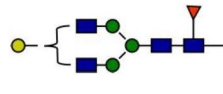
	IC ₈₀ (µg/mL)*					
	WT	LS	QL	A	AAA	YTE
Q23.17	0.162	0.228	0.195	0.204	0.227	0.239
Q842.d12	0.078	0.099	0.082	0.079	0.076	0.089
YU2	0.341	0.439	0.387	0.482	0.409	0.478
JR-FL	0.081	0.099	0.092	0.093	0.093	0.105
7165.18	>50	>50	>50	>50	>50	>50
Du156.12	0.255	0.279	0.263	0.290	0.295	0.284
ZM109.4	0.414	0.548	0.466	0.538	0.429	0.612
SVA.MLV	>50	>50	>50	>50	>50	>50
SIVmac251.30	>50	>50	>50	>50	>50	>50

* Values are colour coded according to their potency: white (>50 µg ml⁻¹), orange (0.1–1.0 µg ml⁻¹) and red (<0.1 µg ml⁻¹).

Extended Data Table 2 | EC₅₀ values for the binding of VRC01 and VRC01-LS to human FcRn, as determined by ELISA

	EC50 (µg/mL)		Fold change
	pH6.0	pH7.4	pH7.4 / pH6.0
WT	~ 20.00	> 100.00	> 5.00
LS	0.69	3.15	4.53
QL	0.87	4.60	5.27
A	1.35	6.69	4.97
AAA	0.95	4.61	4.82
YTE	0.83	4.01	4.86

Extended Data Table 3 | Relative abundance of N297 glycosylation in VRC01 and VRC01-LS*

Glycan		Relative abundance (%)	
		WT	LS
Man ₅ GlcNAc ₂ (Man-5)		0.0	5.4
GlcNAc ₂ Man ₃ GlcNAcFuc		0.0	2.3
GlcNAc ₂ Man ₃ GlcNAc ₂ Fuc (G0F)		87.6	80.6
GalGlcNAc ₂ Man ₃ GlcNAc ₂ Fuc (G1F)		12.4	8.9

*Glycosylation was analysed by nanoLC-ESI-MS/MS (nano liquid chromatography electrospray ionization mass spectrometry) peptide sequencing technology.

Synergistic blockade of mitotic exit by two chemical inhibitors of the APC/C

Katharine L. Sackton^{1*}, Nevena Dimova^{1*}, Xing Zeng^{1*}, Wei Tian^{2†*}, Mengmeng Zhang¹, Timothy B. Sackton³, Johnathan Meaders¹, Kathleen L. Pfaff^{1†}, Frederic Sigoillot^{1†}, Hongtao Yu^{2,4}, Xuelian Luo² & Randall W. King¹

Protein machines are multi-subunit protein complexes that orchestrate highly regulated biochemical tasks. An example is the anaphase-promoting complex/cyclosome (APC/C), a 13-subunit ubiquitin ligase that initiates the metaphase–anaphase transition and mitotic exit by targeting proteins such as securin and cyclin B1 for ubiquitin-dependent destruction by the proteasome^{1,2}. Because blocking mitotic exit is an effective approach for inducing tumour cell death^{3,4}, the APC/C represents a potential novel target for cancer therapy. APC/C activation in mitosis requires binding of Cdc20 (ref. 5), which forms a co-receptor with the APC/C to recognize substrates containing a destruction box (D-box)^{6–14}. Here we demonstrate that we can synergistically inhibit APC/C-dependent proteolysis and mitotic exit by simultaneously disrupting two protein–protein interactions within the APC/C–Cdc20–substrate ternary complex. We identify a small molecule, called apcin (APC inhibitor), which binds to Cdc20 and competitively inhibits the ubiquitylation of D-box-containing substrates. Analysis of the crystal structure of the apcin–Cdc20 complex suggests that apcin occupies the D-box-binding pocket on the side face of the WD40-domain. The ability of apcin to block mitotic exit is synergistically amplified by co-addition of tosyl-L-arginine methyl ester, a small molecule that blocks the APC/C–Cdc20 interaction^{15,16}. This work suggests that simultaneous disruption of multiple, weak protein–protein interactions is an effective approach for inactivating a protein machine.

We identified apcin (Fig. 1a) in an earlier study as an inhibitor of cyclin proteolysis in mitotic *Xenopus* egg extract¹⁷, but its mechanism of action remained unknown. Analysis of the structure–activity relationship (Fig. 1b and Extended Data Fig. 1a) revealed that elimination of one nitrogen in the pyrimidine ring of apcin (apcin-P) reduced activity slightly, but replacement with a morpholino group (apcin-M) eliminated activity. In contrast, elimination of the nitro-imidazole moiety (apcin-A) had little effect. To identify the target of apcin, we coupled apcin-A to beads via its amino group, incubated the beads with mitotic *Xenopus* extract, then removed the beads. We found that Cdc20 was depleted from the extract (Fig. 1c), resulting in stabilization of a cyclin-B-luciferase reporter protein (Fig. 1d). Cyclin degradation was rescued by adding *in vitro*-translated Cdc20 (Fig. 1d), implicating Cdc20 as the target of apcin. Cdc20 binding to apcin-A beads could be competed by free apcin (Fig. 1e), but not by the inactive analogue apcin-M or the APC/C inhibitor tosyl-L-arginine methyl ester (TAME)^{15,16} (Extended Data Fig. 1b). A cyclin B1 amino (N)-terminal fragment (cycB1-NT) also competed for Cdc20 binding to the apcin-A resin, but the same fragment with a mutated D-box did not (Fig. 1f). Among a panel of WD40-containing proteins, Cdc20 binding to apcin-A beads was most robust, followed by Cdh1, with much less binding of other WD40-containing proteins observed (Extended Data Fig. 1c). Although we have not tested the ability of apcin to inhibit Cdh1-dependent ubiquitylation, apcin inhibited

Cdh1-dependent proteolysis in interphase *Xenopus* extract less efficiently than Cdc20-dependent proteolysis in mitotic *Xenopus* extract (Extended Data Fig. 1d). Apcin bound to endogenous Cdc20 in *Xenopus* extract in a dose-dependent manner (Extended Data Fig. 1e) that correlated with its ability to inhibit formation of high-molecular-mass ubiquitin conjugates of cycB1-NT (Fig. 1g) or full-length cyclin B1 (Extended Data Fig. 2a). Kinetic analysis of a reconstituted APC/C-dependent ubiquitylation reaction¹⁶ showed that apcin caused a significant increase ($P = 0.0039$) in the Michaelis constant, K_m (inhibition constant $K_i = 23 \mu\text{M}$), but no reduction in the catalytic rate constant k_{cat} (Extended Data Fig. 2b). Together these results suggest that apcin competitively inhibits APC/C-dependent ubiquitylation by binding to Cdc20 and preventing substrate recognition.

To identify the site on Cdc20 that binds apcin, we soaked apcin into Cdc20 protein crystals and determined the structure of the Cdc20–apcin complex to 2.1 Å resolution (Extended Data Table 1). We found that apcin bound a small pocket on the side of the WD40 domain that has been implicated in binding the D-box (Fig. 2a, b and Extended Data Fig. 3a, b)^{14,18,19}. The binding mode of apcin is consistent with the structure–activity relationship, as the pyrimidine ring and amination nitrogens make hydrogen bonds with backbone atoms from D177. The hydrophobic trichloromethyl group is buried in the pocket occupied by leucine of the D-box (Extended Data Fig. 3b). The nitro-imidazole moiety is positioned facing solvent, explaining why apcin-A retains activity and can be used to isolate Cdc20 when coupled to beads.

We tested whether mutations in the presumptive binding site affect the binding of Cdc20 to apcin-A beads (Fig. 2c and Extended Data Fig. 3c, d). V200 is located at the base of the hydrophobic pocket and mutation to methionine blocked binding of Cdc20 to apcin-A beads and inactivated the ability of Cdc20 to rescue cyclin degradation in a Cdc20-depleted extract. Mutation of D177, P179 or I216, which line the binding pocket, also reduced binding to apcin and function of Cdc20. Mutation of R174, which lies near the pyrimidine ring of apcin, also reduced apcin binding and blocked Cdc20 rescue activity, consistent with a role of this residue in interacting with negatively charged amino acids in the D-box¹⁴. In contrast, mutation of E465, which interacts with the conserved arginine of the D-box¹⁴ but lies distant from the apcin-binding pocket, decreased the ability of Cdc20 to rescue degradation, but had little effect on apcin binding. E180 lies further away and mutation to alanine had no effect on apcin binding or rescue activity. Overall, we observed a strong correlation between effects of mutations on apcin binding and their effects on Cdc20 function for residues that line the apcin-binding pocket.

Cdc20 is recruited to the APC/C through multiple weak interactions^{6,8,20–23}. For example, substrates can promote cooperative Cdc20 binding to the APC/C through a co-receptor interaction in which the substrate is simultaneously recognized by Cdc20 and the APC/C^{6–14}. Consistent with this idea, we found that addition of substrate increased Cdc20 loading onto the APC/C in a concentration- and D-box-dependent manner

¹Department of Cell Biology, Harvard Medical School, 240 Longwood Avenue, Boston, Massachusetts 02115, USA. ²Department of Pharmacology, University of Texas Southwestern Medical Center, 6001 Forest Park Road, Dallas, Texas 75390, USA. ³Department of Organismic and Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts 02138, USA. ⁴Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA. †Present addresses: Key Laboratory of Molecular Biophysics of the Ministry of Education, College of Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China (W.T.); Department of Stem Cell and Regenerative Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA (K.L.P.); Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA (F.S.).

*These authors contributed equally to this work.

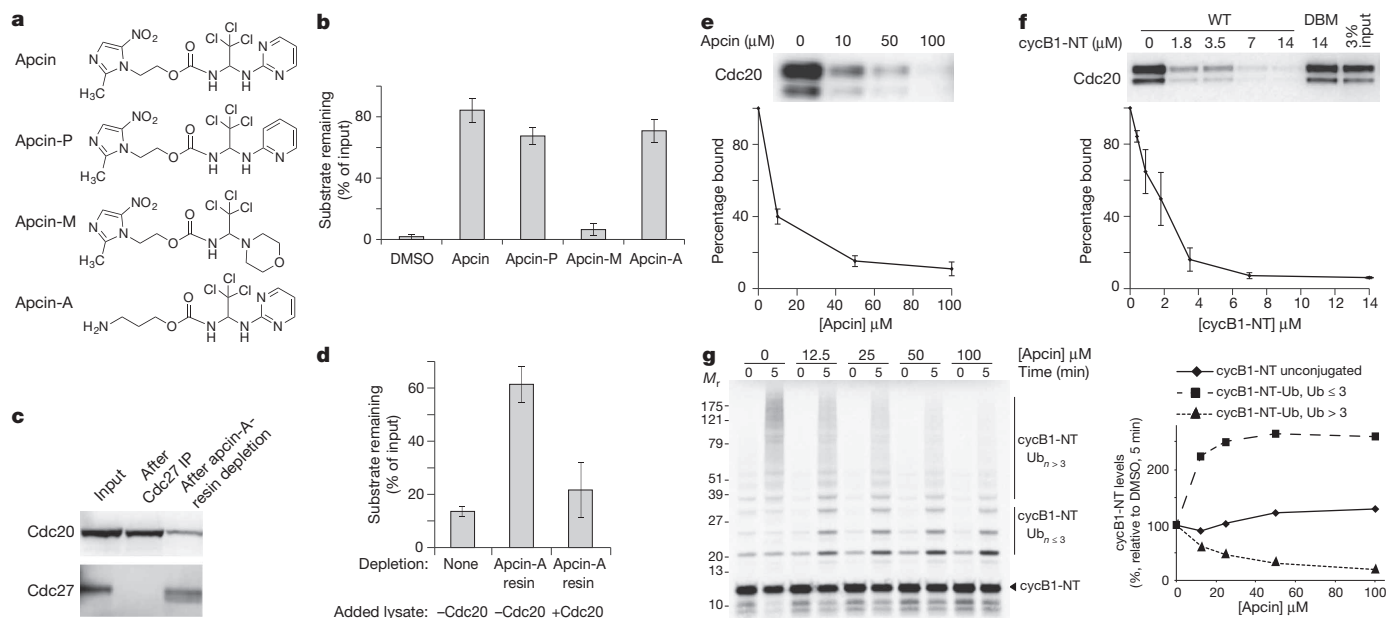


Figure 1 | Apcin binds to Cdc20 and inhibits APC/C-dependent ubiquitylation. **a**, Structures of apcin and derivatives. **b**, Effects of apcin and derivatives (200 μ M) on proteolysis of an N-terminal fragment of cyclin B1 (cycB1-NT) in mitotic *Xenopus* egg extract. Substrate levels were measured at 40 min. $N = 3$ independent experiments. **c**, Apcin-A resin depletes Cdc20 from mitotic *Xenopus* egg extract. Cdc20 and Cdc27 levels were measured by western blotting. **d**, Depletion with Apcin-A resin stabilizes a cyclin-luciferase reporter protein and degradation can be rescued by addition of *in vitro*-translated Cdc20. Substrate levels were measured at 60 min. $N = 3$

independent experiments. **e**, Cdc20 expressed in reticulocyte lysate binds to apcin-A resin and can be competed by free apcin. Cdc20 was detected by western blotting. $N = 3$ independent experiments. **f**, Wild-type (WT) cycB1-NT, but not a D-box mutant (DBM, mutation of R42A, L45A), competes with Cdc20 binding to apcin-A resin. Cdc20 was detected by western blotting. $N = 4$ independent experiments. **g**, Apcin inhibits formation of high-molecular-mass ubiquitin conjugates of [35 S]cycB1-NT in mitotic *Xenopus* extract. Bars, mean \pm s.e.m.

in *Xenopus* extract (Fig. 3a and Extended Data Fig. 4a, b). Substrate-induced loading of Cdc20 could be blocked by addition of apcin (Fig. 3a and Extended Data Fig. 4a), indicating that the leucine-binding pocket of Cdc20 is critical for co-receptor function. Because the binding of Cdc20 to the APC/C was variable in the absence of added substrate, we were not able to assess the effects of apcin under this condition. Thus we cannot exclude the possibility that apcin might also decrease Cdc20 binding to the APC/C in the absence of substrate. The small molecule TAME, which antagonizes the Ile-Arg (IR)-tail interaction between Cdc20 and the APC/C^{5,16}, also antagonized Cdc20 loading. At high concentrations of substrate, the combined use of apcin and TAME was more

effective at blocking Cdc20 binding to the APC/C than either compound used alone, suggesting that simultaneous disruption of multiple interactions between substrate, Cdc20 and APC/C may be an effective strategy for inhibiting APC/C.

Because substrates can be recruited to the APC/C through both D-box-dependent and -independent mechanisms, we compared the ability of apcin to stabilize different APC/C substrates in mitotic *Xenopus* extract. Apcin stabilized cycB1-NT and securin most effectively (Fig. 3b and Extended Data Figs 4c and 5b), with somewhat weaker effects against full-length cyclin B1 (Fig. 3b and Extended Data Fig. 4c). Interestingly, even high concentrations of apcin failed to stabilize cyclin A2 or Nek2A

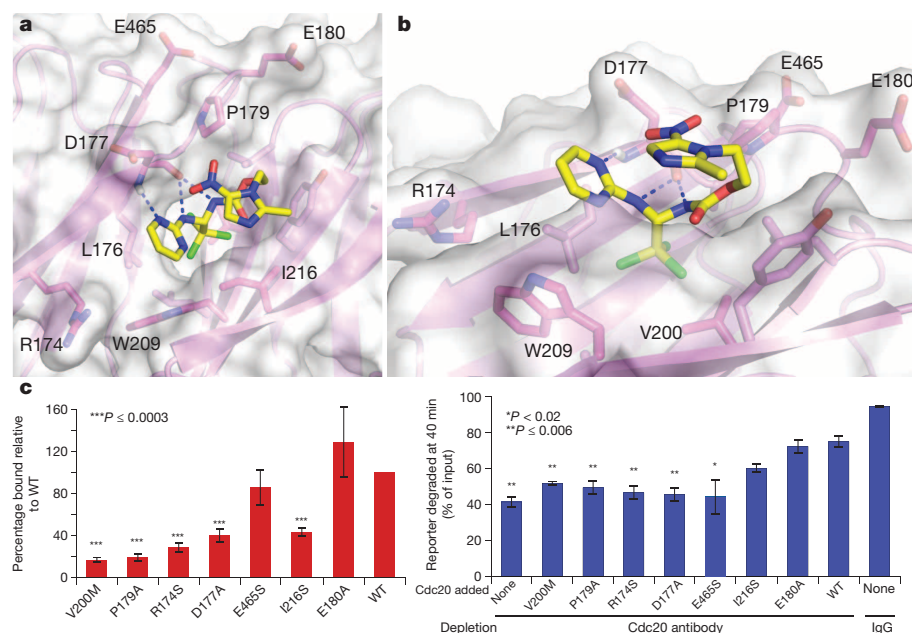


Figure 2 | Apcin binds to the D-box binding site of Cdc20. **a**, Crystal structure of the apcin-Cdc20 complex. Apcin atoms are labelled in yellow (carbon), blue (nitrogen), red (oxygen) and green (chlorine). Cdc20 is shown in magenta. Dotted blue lines indicate hydrogen bonds. **b**, View is rotated to show the position of V200 at the base of the hydrophobic binding pocket. **c**, Mutation of residues in the binding pocket reduces Cdc20 binding to apcin-A resin (red bars) and the capacity of *in vitro*-translated Cdc20 protein to rescue cyclin-luciferase degradation in mitotic *Xenopus* egg extract immunodepleted of Cdc20 (blue bars). Bars, mean \pm s.e.m. of three independent experiments. P values were calculated by an unpaired *t*-test compared with WT.

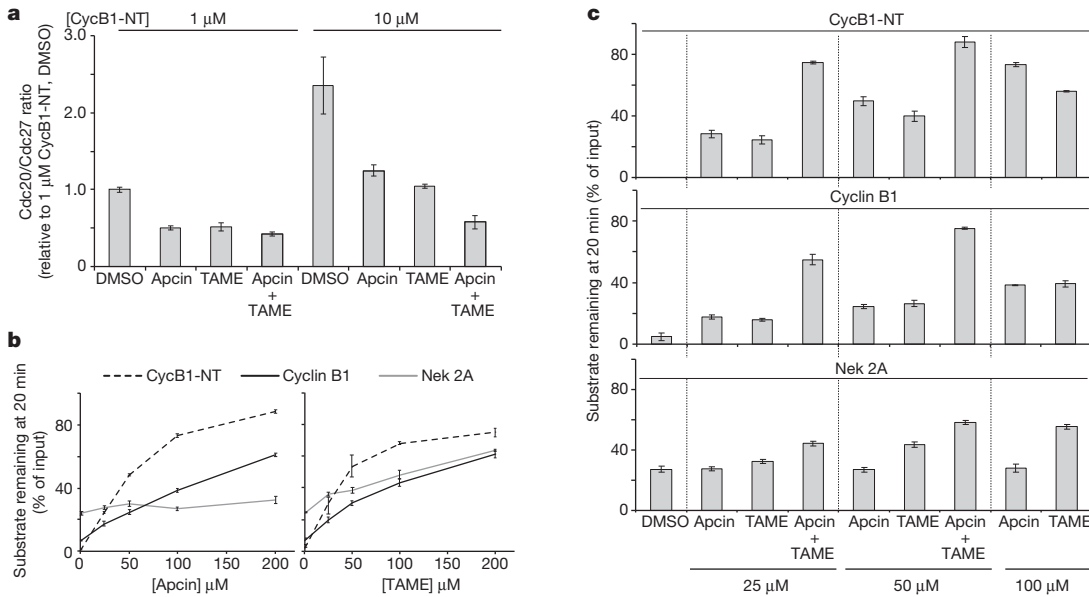


Figure 3 | Effects of apcin on Cdc20 binding to APC/C and stability of APC/C substrates in mitotic *Xenopus* extract. **a**, Apcin blocks co-receptor-dependent binding of Cdc20 to APC/C. Substrate (cycB1-NT, 1 μM or 10 μM), apcin and/or TAME (50 μM each) were added, and APC/C was isolated with anti-Cdc27 antibody. Levels of Cdc20 were assessed by western blotting and normalized to the levels of Cdc27. **b**, Effect of apcin and TAME on APC/C substrate stability. Levels of ³⁵S-labelled substrates were assessed by gel electrophoresis and phosphorimaging. **c**, As in **b** except that combinations of apcin and TAME were examined. Values and error bars in **a–c** represent means ± s.e.m. of three independent experiments.

(Fig. 3b and Extended Data Figs 4c and 5a), suggesting that apcin is highly specific for the D-box binding site. In contrast, TAME inhibited the degradation of all APC/C substrates tested (Fig. 3b and Extended Data Figs 4c and 5a, b), consistent with its ability to block recruitment of Cdc20 to the APC/C directly^{15,16}. The ability of apcin to inhibit formation of high-molecular-mass substrate–ubiquitin conjugates closely correlated with its ability to stabilize each substrate (Extended Data Fig. 6). These results suggest that the leucine-binding pocket of Cdc20 plays an essential role in recruiting the D-boxes of securin and cycB1-NT, and that apcin can effectively compete with these interactions. The ability of full-length cyclin B1 and cyclin A2 to bind Cks1 via Cdk1 may facilitate their D-box-independent recruitment to the APC/C^{24,25}, helping these substrates partially overcome the effects of apcin. Furthermore, the N-terminal region of cyclin A2 appears to bind Cdc20 with higher affinity than cyclin B1 (ref. 26), which may further reduce the effectiveness of apcin. In contrast to these substrates whose degradation is D-box-dependent, Nek2A is recruited directly to the APC/C via a Cdc20-independent mechanism that requires its Met-Arg (MR)-tail rather than a D-box^{27–29}, explaining why apcin fails to inhibit ubiquitylation or degradation of this protein.

Because apcin and TAME inhibit APC/C-dependent proteolysis by distinct mechanisms, we tested the effect of combining the inhibitors on the proteolysis of APC/C substrates in *Xenopus* extract. The combination of apcin and TAME led to synergistic stabilization of cyclin B1, cycB1-NT, securin and cyclin A2, with a much weaker effect for Nek2A (Fig. 3c and Extended Data Figs 4d and 5a, b). For example, combining TAME and apcin at 25 μM each was more effective at stabilizing cyclin B1 than using either compound alone at 100 μM. Apcin slightly enhanced the ability of TAME to stabilize Nek2A (Fig. 3c), suggesting that the leucine-binding pocket of Cdc20 may bind Nek2A, even though this interaction is not essential for proteolysis if the APC/C is not otherwise perturbed.

We next examined the effect of apcin, proTAME¹⁵ (a cell-permeable TAME prodrug), and the combination, on mitotic exit in four different human cell lines. Apcin and proTAME synergized to increase the mitotic fraction in all cell lines examined (Fig. 4a and Extended Data Fig. 7a, b). Apcin-M was inactive whereas apcin-P retained activity (Extended Data Fig. 7c), consistent with effects on Cdc20 binding and APC/C-dependent proteolysis in *Xenopus* extract. In live-cell imaging experiments in RPE1 cells (Fig. 4b, Extended Data Fig. 8a–d and Supplementary Videos 1 and 2),

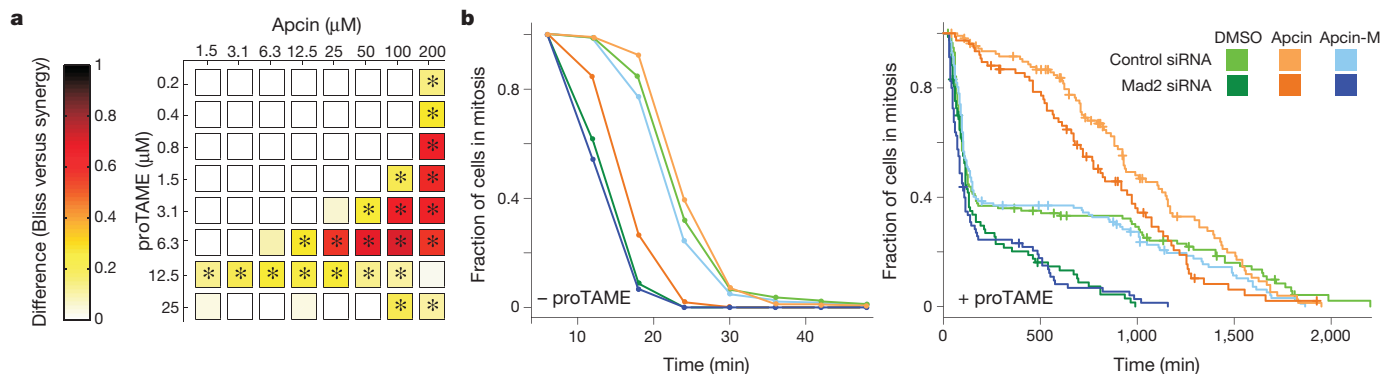


Figure 4 | Apcin synergizes with proTAME to prolong mitotic duration. **a**, RPE1 cells were treated with indicated concentrations of apcin and proTAME for 18 h, fixed and the mitotic index determined by automated high-throughput imaging. The panel shows the difference between the mitotic indices calculated by a Bliss-independence model compared with a synergy model; any positive value indicates synergy. **P* < 0.05 on the basis of analysis of four technical replicates. **b**, Asynchronous RPE1 cells expressing H2B–green fluorescent protein (GFP) were transfected with short interfering RNA (siRNA)

and 24 h later treated with apcin or apcin-M (25 μM) and/or proTAME (6 μM). Cells were then imaged every 6 min for 45 h. Mitotic duration and cell fate (Extended Data Fig. 8b) were determined by manual inspection of the videos and plotted as inverse cumulative frequency (–proTAME) or Kaplan–Meier curves (+proTAME). The hatch marks on the Kaplan–Meier curves indicate censored cells that did not exit mitosis or die in mitosis before the end of the movie or before they migrated out of the field of view. Graphs include the combined results of two independent experiments.

25 μM apcin had no detectable effect on mitotic duration ($P = 0.279$) unless mitosis was artificially shortened by depletion of the spindle assembly checkpoint (SAC) protein Mad2 ($P = 0.0001$). In contrast, in the presence of proTAME, the addition of apcin dramatically slowed the rate of mitotic exit in a synergistic manner (Fig. 4b and Extended Data Fig. 8a–d): the rate of mitotic exit was 63% of that predicted by a multiplicative combination of the single compound effects ($P = 0.016$). Significant synergy was also observed in U2OS cells ($P = 2.0 \times 10^{-8}$; Extended Data Fig. 8e). Addition of the inactive derivative apcin-M had no effect on the rate of mitotic exit in the presence of proTAME ($P = 0.68$; Fig. 4b and Extended Data Fig. 8a–e). The response of cells to proTAME alone was biphasic, because prolongation of metaphase can cause cohesion fatigue³⁰, which reactivates the SAC to block mitotic exit in a subpopulation of cells. Notably, the addition of apcin eliminated the biphasic response. Furthermore, the combined effect of apcin and proTAME was largely preserved when the SAC was inactivated by Mad2 depletion. When modelled quantitatively, the degree of synergy between apcin and proTAME was in fact enhanced in the absence of Mad2, as the rate of mitotic exit was reduced to 38% of the rate predicted by a multiplicative combination of the single compound effects ($P = 8.92 \times 10^{-5}$; Extended Data Fig. 8d). Together these findings suggest that synergistic inhibition of mitotic exit does not rely on the SAC, but instead probably arises from direct pharmacological APC/C inhibition.

The discovery of apcin and the elucidation of its mechanism validate the importance of the leucine-binding pocket in Cdc20 as a D-box co-receptor. Apcin preferentially stabilizes APC/C substrates whose degradation is D-box dependent. However, on its own, low concentrations of apcin do not effectively block mitotic exit, either because substrates can outcompete apcin binding to the leucine-binding pocket, or because substrates can be recruited to the APC/C through other mechanisms. The effectiveness of apcin can be dramatically enhanced by the addition of TAME, which blocks Cdc20 loading through a distinct mechanism, highlighting the importance of multiple weak protein–protein interactions in promoting activator binding and efficient substrate ubiquitylation (Extended Data Fig. 9). Our work highlights the possibility of disrupting the function of a protein machine by simultaneously inhibiting multiple protein–protein interactions. Because dynamic protein complexes regulate virtually all aspects of cell biology, simultaneous targeting of multiple weak interactions may represent a new opportunity for therapeutic targeting of protein complexes that may otherwise be difficult to inhibit with a single compound.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 7 August 2013; accepted 4 July 2014.

Published online 24 August 2014.

- Pines, J. Cubism and the cell cycle: the many faces of the APC/C. *Nature Rev. Mol. Cell Biol.* **12**, 427–438 (2011).
- Primorac, I. & Musacchio, A. Panta rhei: the APC/C at steady state. *J. Cell Biol.* **201**, 177–189 (2013).
- Huang, H. C., Shi, J., Orth, J. D. & Mitchison, T. J. Evidence that mitotic exit is a better cancer therapeutic target than spindle assembly. *Cancer Cell* **16**, 347–358 (2009).
- Manchado, E. *et al.* Targeting mitotic exit leads to tumor regression *in vivo*: modulation by Cdk1, Mast1, and the PP2A/B55 α,δ phosphatase. *Cancer Cell* **18**, 641–654 (2010).
- Yu, H. Cdc20: a WD40 activator for a cell cycle degradation machine. *Mol. Cell* **27**, 3–16 (2007).
- Passmore, L. *et al.* Doc1 mediates the activity of the anaphase-promoting complex by contributing to substrate recognition. *EMBO J.* **22**, 786–796 (2003).
- Passmore, L. A. & Barford, D. Coactivator functions in a stoichiometric complex with anaphase-promoting complex/cyclosome to mediate substrate recognition. *EMBO Rep.* **6**, 873–878 (2005).
- Burton, J. L., Tsakraklides, V. & Solomon, M. J. Assembly of an APC-Cdh1-substrate complex is stimulated by engagement of a destruction box. *Mol. Cell* **18**, 533–542 (2005).
- Carroll, C. W., Enquist-Newman, M. & Morgan, D. O. The APC subunit Doc1 promotes recognition of the substrate destruction box. *Curr. Biol.* **15**, 11–18 (2005).

- Eytan, E., Moshe, Y., Braunstein, I. & Hershko, A. Roles of the anaphase-promoting complex/cyclosome and of its activator Cdc20 in functional substrate binding. *Proc. Natl Acad. Sci. USA* **103**, 2081–2086 (2006).
- Matyskiela, M. E. & Morgan, D. O. Analysis of activator-binding sites on the APC/C supports a cooperative substrate-binding mechanism. *Mol. Cell* **34**, 68–80 (2009).
- Buschhorn, B. *et al.* Substrate binding on the APC/C occurs between the coactivator Cdh1 and the processivity factor Doc1. *Nature Struct. Mol. Biol.* **18**, 6–13 (2011).
- da Fonseca, P. C. *et al.* Structures of APC/C(Cdh1) with substrates identify Cdh1 and Apc10 as the D-box co-receptor. *Nature* **470**, 274–278 (2011).
- He, J. *et al.* Insights into degron recognition by APC/C coactivators from the structure of an Acm1-Cdh1 complex. *Mol. Cell* **50**, 649–660 (2013).
- Zeng, X. *et al.* Pharmacologic inhibition of the anaphase-promoting complex induces a spindle checkpoint-dependent mitotic arrest in the absence of spindle damage. *Cancer Cell* **18**, 382–395 (2010).
- Zeng, X. & King, R. W. An APC/C inhibitor stabilizes cyclin B1 by prematurely terminating ubiquitylation. *Nature Chem. Biol.* **8**, 383–392 (2012).
- Verma, R. *et al.* Ubistatins inhibit proteasome-dependent degradation by binding to the ubiquitin chain. *Science* **306**, 117–120 (2004).
- Chao, W. C., Kulkarni, K., Zhang, Z., Kong, E. H. & Barford, D. Structure of the mitotic checkpoint complex. *Nature* **484**, 208–213 (2012).
- Tian, W. *et al.* Structural analysis of human Cdc20 supports multisite degron recognition by APC/C. *Proc. Natl Acad. Sci. USA* **109**, 18419–18424 (2012).
- Schwab, M., Neutzner, M., Möcker, D. & Seufert, W. Yeast Hct1 recognizes the mitotic cyclin Clb2 and other substrates of the ubiquitin ligase APC. *EMBO J.* **20**, 5165–5175 (2001).
- Vodermaier, H. C., Gieffers, C., Maurer-Stroh, S., Eisenhaber, F. & Peters, J. M. TPR subunits of the anaphase-promoting complex mediate binding to the activator protein Cdh1. *Curr. Biol.* **13**, 1459–1468 (2003).
- Kraft, C., Vodermaier, H. C., Maurer-Stroh, S., Eisenhaber, F. & Peters, J. M. The WD40 propeller domain of Cdh1 functions as a destruction box receptor for APC/C substrates. *Mol. Cell* **18**, 543–553 (2005).
- Thornton, B. R. *et al.* An architectural map of the anaphase-promoting complex. *Genes Dev.* **20**, 449–460 (2006).
- van Zon, W. *et al.* The APC/C recruits cyclin B1-Cdk1-Cks in prometaphase before D box recognition to control mitotic exit. *J. Cell Biol.* **190**, 587–602 (2010).
- Wolthuis, R. *et al.* Cdc20 and Cks direct the spindle checkpoint-independent destruction of cyclin A. *Mol. Cell* **30**, 290–302 (2008).
- Di Fiore, B. & Pines, J. How cyclin A destruction escapes the spindle assembly checkpoint. *J. Cell Biol.* **190**, 501–509 (2010).
- Hayes, M. J. *et al.* Early mitotic degradation of Nek2A depends on Cdc20-independent interaction with the APC/C. *Nature Cell Biol.* **8**, 607–614 (2006).
- Kimata, Y., Baxter, J. E., Fry, A. M. & Yamano, H. A role for the Fizzy/Cdc20 family of proteins in activation of the APC/C distinct from substrate recruitment. *Mol. Cell* **32**, 576–583 (2008).
- Sedgwick, G. G. *et al.* Mechanisms controlling the temporal degradation of Nek2A and Kif18A by the APC/C-Cdc20 complex. *EMBO J.* **32**, 303–314 (2013).
- Lara-Gonzalez, P. & Taylor, S. Cohesion fatigue explains why pharmacological inhibition of the APC/C induces a spindle checkpoint-dependent mitotic arrest. *PLoS ONE* **10**.1371/journal.pone.0049041 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank W. Harper for providing constructs for WD40-containing proteins, T. Gahman for assistance with apcin synthesis and D. Tomchick for assistance with structure refinement. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the US Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357. This work was supported by grants from the National Institutes of Health (GM085004 to X.L. and GM066492 to R.W.K.) and by a grant from the Lynch Foundation to R.W.K.

Author Contributions K.L.S. and X.Z. performed experiments to identify Cdc20 as the target of apcin. K.L.S. characterized binding of apcin to Cdc20, and evaluated effects of Cdc20 mutations on apcin binding and cyclin proteolysis. W.T. purified Cdc20 and performed crystallization and structure determination of the Cdc20–apcin complex. H.Y. and X.L. contributed to structure determination and data analysis. N.D. and X.Z. characterized effects of apcin and TAME on substrate degradation, ubiquitylation and Cdc20 binding to APC/C in *Xenopus* extract. N.D. evaluated binding of apcin to Cdc20 using the thermal shift assay. M.Z. characterized effects of apcin and proTAME on mitotic index in fixed cell assays. K.L.S. and X.Z. characterized effects of apcin and proTAME by live cell imaging. F.S. developed the high-throughput mitotic index assay. K.L.S., X.Z., J.M., K.L.P. and F.S. analysed time-lapse videos. T.B.S. developed statistical models and performed statistical analysis. R.W.K. conceived the project, assisted with experimental design and data analysis, and wrote the manuscript, with assistance from all authors.

Author Information Structure coordinates have been deposited in Protein Data Bank under accession number 4N14. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.K. (randy_king@hms.harvard.edu).

METHODS

Reagents. Commercial antibodies used for western analysis were as follows: anti-Cdc27 (610455, BD Transduction Laboratories), anti-Cdc20 (BA8; sc-93399, Santa Cruz Biotechnology and NB 100-2646, Novus Biologicals) to recognize *Xenopus* Cdc20, anti-Cdc20 (H-175; sc-8358, Santa Cruz Biotechnology) to recognize human Cdc20, and anti-HA-peroxidase (3F10; 12013819001, Roche). Secondary antibodies used included anti-rabbit IgG-HRP (NA934; GE Healthcare) and anti-mouse IgG-HRP (NA931; GE Healthcare). For APC/C immunopurification from *Xenopus* extract, anti-Cdc27 (AF3.1; sc-9972, Santa Cruz Biotechnology) was used. For immunodepletion of *Xenopus* Cdc20, a rabbit polyclonal antibody was generated by Yenzym by immunization with an N-terminal fragment of Cdc20 (residues 1–170; tagged at the carboxy (C) terminus). Chemicals used were cycloheximide (239764, Calbiochem), calcium ionophore A23187, free acid form (100105, Calbiochem), tosyl-L-arginine methyl ester (T4626, Sigma), proTAME (I-440, Boston Biochem), MG262 (I-120, Boston Biochem), apcin (T0506-3874, Enamine), apcin-P (Amb2237944, Ambinter) and apcin-M (Amb1395012, Ambinter). Apcin and apcin-A were also synthesized by Sundia Meditech according to the methods described in the Supplementary Information.

Assessment of substrate degradation in *Xenopus* egg extract. Use of female *Xenopus laevis* to produce cytoplasmic egg extracts was approved by the Harvard Medical School Standing Committee on Animals (protocol number 03231). Interphase *Xenopus* egg extract was prepared from eggs laid overnight according to the protocol of ref. 31 with the exception that eggs were activated with $2 \mu\text{g ml}^{-1}$ calcium ionophore (A23187) for 30 min before the crushing spin. Extract was frozen in liquid nitrogen and stored at -80°C . Interphase extract was induced to enter mitosis by addition of non-degradable cyclin B (MBP- Δ 90) at $20 \mu\text{g ml}^{-1}$ and incubated at $22\text{--}24^\circ\text{C}$ for 30–60 min. MBP- Δ 90 consists of a fusion of the maltose-binding protein (MBP) to *Xenopus* cyclin B1 lacking its N-terminal 90 amino acids³² and was expressed in *Escherichia coli* by inducing cultures at an attenuation ($D_{600 \text{ nm}}$) of 0.6 with $300 \mu\text{M}$ isopropylthiogalactoside (IPTG) for 5 h at $22\text{--}24^\circ\text{C}$. Purification followed New England BioLabs protocol. To promote degradation in interphase extract, Cdh1 protein, expressed in baculovirus, was added to extract at a final concentration of 50 nM in the presence of $75 \mu\text{M}$ roscovitine. Roscovitine addition is necessary to suppress inhibitory phosphorylation of Cdh1 by Cdk1. Extract was then pre-treated with drug (dimethylsulphoxide (DMSO), apcin and/or TAME) for 15 min at 22°C before addition of substrates.

Substrates consisted of human full-length cyclin B1, cyclin A2, securin, Nek2A or an N-terminal fragment (residues 1–88) of human cyclin B1 (cyc B1-NT). Each substrate was amplified with primers by PCR to allow T7-dependent transcription of the PCR product. Substrates were expressed and labelled with [³⁵S]methionine (Perkin Elmer NEG709A500UC) using the TNT system (Promega). To measure degradation of substrates, extract was pre-treated with DMSO or test compounds for 10 min in the presence of $100 \mu\text{g ml}^{-1}$ cycloheximide to prevent re-incorporation of free labelled amino acid. The *in vitro* translation reaction was then added to the *Xenopus* extract at 10% final volume. Extract was then incubated at 24°C , with shaking at 1,250 r.p.m., with samples taken at indicated times. Reactions were quenched with sodium dodecyl sulphate (SDS) sample buffer and processed for SDS-polyacrylamide gel electrophoresis (PAGE) and phosphor imaging (Bio-Rad PMI); quantification was performed using Quantity One software (Bio-Rad).

Assessment of substrate ubiquitylation in mitotic *Xenopus* egg extract. Interphase *Xenopus* egg extract was supplemented with MBP- Δ 90 to promote entry into mitosis. Mitotic extract was then treated for 30 min at 24°C with $20 \mu\text{M}$ ubiquitin-vinyl sulfone (UbVS; U-202, Boston Biochem) to suppress deubiquitylation and with $100 \mu\text{g ml}^{-1}$ cycloheximide to prevent re-incorporation of free labelled amino acid. Subsequently, mixtures containing apcin or DMSO, as indicated, proteasome inhibitor MG262 ($150 \mu\text{M}$) and wild-type ubiquitin ($44 \mu\text{M}$) were added to UbVS-treated extract and incubated for additional 10 min, at 24°C with agitation. *In vitro* translation reactions expressing human full length cyclin B1, cyclin A2, securin, Nek2A, or cycB1-NT with [³⁵S]methionine labelling were also pre-treated with UbVS at $20 \mu\text{M}$ and added to pre-treated extract at 10% final volume. Extract was then incubated at 24°C , with shaking at 1250 r.p.m., and substrate ubiquitylation monitored by taking samples at indicated times. Reactions were quenched with SDS sample buffer and processed for SDS-PAGE and phosphor imaging (Bio-Rad PMI), and quantification was performed using Quantity One software (Bio-Rad).

Measurement of the apcin-Cdc20 interaction in *Xenopus* egg extract using cellular thermal shift assay. The cellular thermal shift assay (CETSA) method³³ was adapted to examine apcin engagement of endogenous Cdc20 in *Xenopus* egg extract. Interphase extract was diluted tenfold and incubated with various concentrations of apcin dissolved in DMSO in a total volume of $200 \mu\text{l}$, with a final DMSO concentration of 1%. After 10-min incubation at $22\text{--}24^\circ\text{C}$, $50 \mu\text{l}$ of the lysate was transferred into PCR tubes (20170-012, VWR) and heated in a PCR machine (Mastercycler gradient, Eppendorf) for 3 min at 46°C , followed by cooling for 3 min at $22\text{--}24^\circ\text{C}$. These conditions were established in preliminary experiments as the

temperature that yielded the greatest degree of Cdc20 precipitation that could be rescued by apcin treatment (data not shown). The heated lysates were centrifuged at 14,000 r.p.m. (20,000g) for 20 min at 4°C to separate the soluble fractions from precipitates. Twenty microlitres of the supernatants were mixed with SDS sample buffer and the fraction of soluble Cdc20 was analysed by SDS-PAGE and anti-Cdc20 western blot. Quantification of soluble Cdc20 used Fuji Imager LAS3000 and ImageJ software. Soluble Cdc20 levels were normalized to soluble Cdc20 in samples treated with the highest concentration of apcin.

Measurement of Cdc20 binding to APC/C in *Xenopus* egg extract. To examine levels of Cdc20 associated with APC/C, the APC/C was immunopurified from mitotic *Xenopus* egg extract. For $100 \mu\text{l}$ extract, $2 \mu\text{g}$ of anti-Cdc27 antibody (AF3.1, Santa Cruz Biotechnology) was cross-linked to $5 \mu\text{l}$ of Affiprep Protein A beads (156-0006, Bio-Rad) and incubated for 1 h at 4°C . Apcin, TAME or DMSO was mixed with extract upon addition to anti-Cdc27-Affiprep Protein A beads in the presence or absence of cycB1-NT containing a HA-tag at the N terminus and His tag at the C terminus, as previously described³⁴. After incubation with extract, beads were washed quickly three times with 20-fold volume of XB (10 mM potassium HEPES, pH 7.7, 100 mM KCl, 0.1 mM CaCl_2 , 1 mM MgCl_2) and combined with SDS sample buffer. For analysis of Cdc20 binding to APC/C, samples were processed for SDS-PAGE and immunoblotting against Cdc20 and APC/C subunit Cdc27. Chemiluminescence was imaged on a Fuji LAS 3000 with Image Reader LAS-3000 software. Levels of Cdc20 were quantified using ImageJ and data normalized to respective Cdc27 levels.

Coupling of apcin-A to affigel-10 resin. Affigel-10 resin (153-6099, Bio-Rad) was washed twice with DMSO and dried. The resin was then mixed with 5 mM or 15 mM apcin-A dissolved in DMSO ($2\times$ volume of dry resin). *N,N*-diisopropylethylamine was diluted 50-fold into the solution. The resin was rotated at $22\text{--}24^\circ\text{C}$ for 2 h and the reaction was quenched with 1/5 resin volume of ethanolamine. The resin was then washed sequentially with isopropanol, water and XB + 0.05% Tween. The resin was stored at 4°C as a 50% slurry in XB + 0.05% Tween.

Cdc20 depletion by apcin-A resin. For a round of depletion of Cdc20 from mitotic *Xenopus* extract, apcin-A resin (15 mM coupling) was incubated with extract at 4°C rotating for 30 min. The volume of resin used was 40% of the extract volume. Three rounds of depletion were performed. To rescue degradation in the depleted extract, reticulocyte lysate containing *in vitro*-translated human Cdc20 or control reticulocyte lysate, treated as for a translation reaction but with no DNA template, was added to the extract at 1/10th extract volume.

Assay for Cdc20 binding to apcin-A resin. Human Cdc20 in pCS2 vector was mutated at the various residues described with the QuikChange Site-Directed Mutagenesis Kit (Agilent Technologies) and custom primers for each sequence. All mutations were confirmed by sequencing. For a single pull-down assay, $5 \mu\text{l}$ of apcin-A resin (5 mM coupling) was incubated with $30 \mu\text{l}$ diluted *in vitro*-translated human Cdc20 or other WD40 proteins ($5 \mu\text{l}$ reticulocyte lysate diluted to $30 \mu\text{l}$ with XB + 0.05% Tween) at 24°C with shaking (1500 r.p.m.) for 30 min. Competitors were pre-incubated with Cdc20 in reticulocyte lysate for 2 min at $22\text{--}24^\circ\text{C}$ before adding to apcin-A resin. Bound and input Cdc20 were detected by western blotting (Fig. 1e, f). Cdc20, other WD40 proteins, ODC and Cdc20 mutants in Fig. 2c and Extended Data Figs 1b, c and 3d were labelled with [³⁵S]methionine and detected by phosphorimaging.

Assessment of kinetics of APC/C-dependent ubiquitylation in a reconstituted system. Measurements of kinetics of ubiquitylation used cycB1-NT and APC/C isolated from *Xenopus* extract, exactly as previously described¹⁶, in the presence or absence of $50 \mu\text{M}$ apcin. The K_i was calculated on the basis of the assumption of a competitive inhibition model according to the equation $K_{m, \text{apcin}} = K_{m, \text{untreated}} (1 + [I]/K_i)$, where $[I] = 50 \mu\text{M}$; $K_{m, \text{apcin}}$ (773 nM) and $K_{m, \text{untreated}}$ (245 nM) were the average values from three independent experiments. The *P* value was calculated by an unpaired *t*-test.

Antibody-based depletion of Cdc20 from *Xenopus* extract. Cdc20 antibody, covalently coupled to Affiprep protein-A beads as described¹⁵, was incubated with mitotic extract at 4°C with rotation for 30 min. The volume of antibody beads used was 20% of the reaction volume. Three rounds of depletion were performed, with separation of extract from beads after each round by centrifugation in spin columns (89868, Thermo-Pierce).

Luciferase assay. A fusion of the N-terminal domain of cyclin B1 to luciferase¹⁷ was added to mitotic extract at $4 \mu\text{g ml}^{-1}$ (Fig. 1d) or to interphase extract at $250 \mu\text{g ml}^{-1}$ for 10 min at $22\text{--}24^\circ\text{C}$ then diluted to a final concentration of $4 \mu\text{g ml}^{-1}$ in mitotic extract (Fig. 2c). The extract was incubated at $22\text{--}24^\circ\text{C}$ and 3 μl samples were taken at 0, 20, 40 and 60 min. The samples were mixed quickly with $30 \mu\text{l}$ luciferin assay buffer (270 μM coenzyme A, 20 mM tricine, 3.67 mM MgSO_4 , 0.1 mM ethylenediaminetetraacetic acid (EDTA), 33.3 mM dithiothreitol (DTT), 530 μM ATP and 470 μM luciferin, pH 7.8), and the level of luminescence was measured on a Wallac 1420 multilabel counter.

Protein purification and crystallization. The coding region of human Cdc20 containing residues 161–477 (Cdc20–WD40) was amplified by PCR and cloned into the modified pFastBac vector. Recombinant baculovirus encoding the N-terminal His6-tagged Cdc20 protein was constructed using the Bac-to-Bac system (Invitrogen) according to the manufacturer's protocols. A tobacco etch virus protease cleavage site was introduced into the N terminus of Cdc20. Sf9 insect cells were infected with the Cdc20 baculovirus and harvested at about 60 h post-infection. His6-tagged Cdc20–WD40 was purified with Ni²⁺-NTA agarose resin (Qiagen) and cleaved with tobacco etch virus protease to remove the His₆-tag. The Cdc20–WD40 protein was further purified by anion exchange chromatography with a Mono-Q column followed by size exclusion chromatography with a Superdex 200 column (GE Healthcare). Purified Cdc20–WD40 was concentrated to 4–5 mg ml⁻¹ in the Superdex 200 column buffer containing 25 mM Tris (pH 8.5), 150 mM NaCl, 1 mM MgCl₂, 5% glycerol and 5 mM *tris*(2-carboxyethyl)phosphine.

The Cdc20–WD40 protein was crystallized at 20 °C using the sitting-drop vapour-diffusion method with a reservoir solution containing 0.1 M 2-(*N*-morpholino)ethanesulfonic acid (MES, pH 6.5), 15% (w/v) polyethylene glycol (PEG) 6000 and 5% MPD. The Cdc20–WD40 crystals were transferred to a new 2 µl sitting drop with a reservoir solution containing 0.1 M MES (pH 6.5) and 20% (w/v) PEG 6000 and soaked for 3–5 h to remove bound MPD. Apicin compound was dissolved into DMSO and added to the same drop at the final concentration of 5 mM. After overnight soaking, the crystals were cryo-protected in a solution containing 0.1 M MES (pH 6.5), 20% (w/v) PEG 6000, 10% glycerol and 5 mM apicin, then flash-cooled in liquid nitrogen. Crystals diffracted to a minimum Bragg spacing (d_{\min}) of about 2.1 Å and exhibited the symmetry of space group $P2_1$ with cell dimensions of $a = 41$ Å, $b = 87$ Å, $c = 48$ Å and $\beta = 113^\circ$, and contained one Cdc20 molecule per asymmetric unit.

Data collection and structure determination. Diffraction data were collected at beamline 19-ID (SBC-CAT) at the Advanced Photon Source (Argonne National Laboratory) and processed with HKL3000 (ref. 35). Phases were obtained by molecular replacement with Phaser using the crystal structure of human Cdc20–WD40 (Protein Data Bank accession number 4GGC) as search model³⁶. Iterative model building and refinements used COOT and Phenix, respectively^{37,38}. The final model for Cdc20–WD40–apicin ($R_{\text{work}} = 16.5\%$, $R_{\text{free}} = 21.3\%$) contained 313 residues, 82 water molecules and one apicin molecule. MolProbity was used for structure validation to show that all models had good geometry, except for one surface residue that was an outlier in a Ramachandran plot³⁹. Data collection and structure refinement statistics are summarized in Extended Data Table 1.

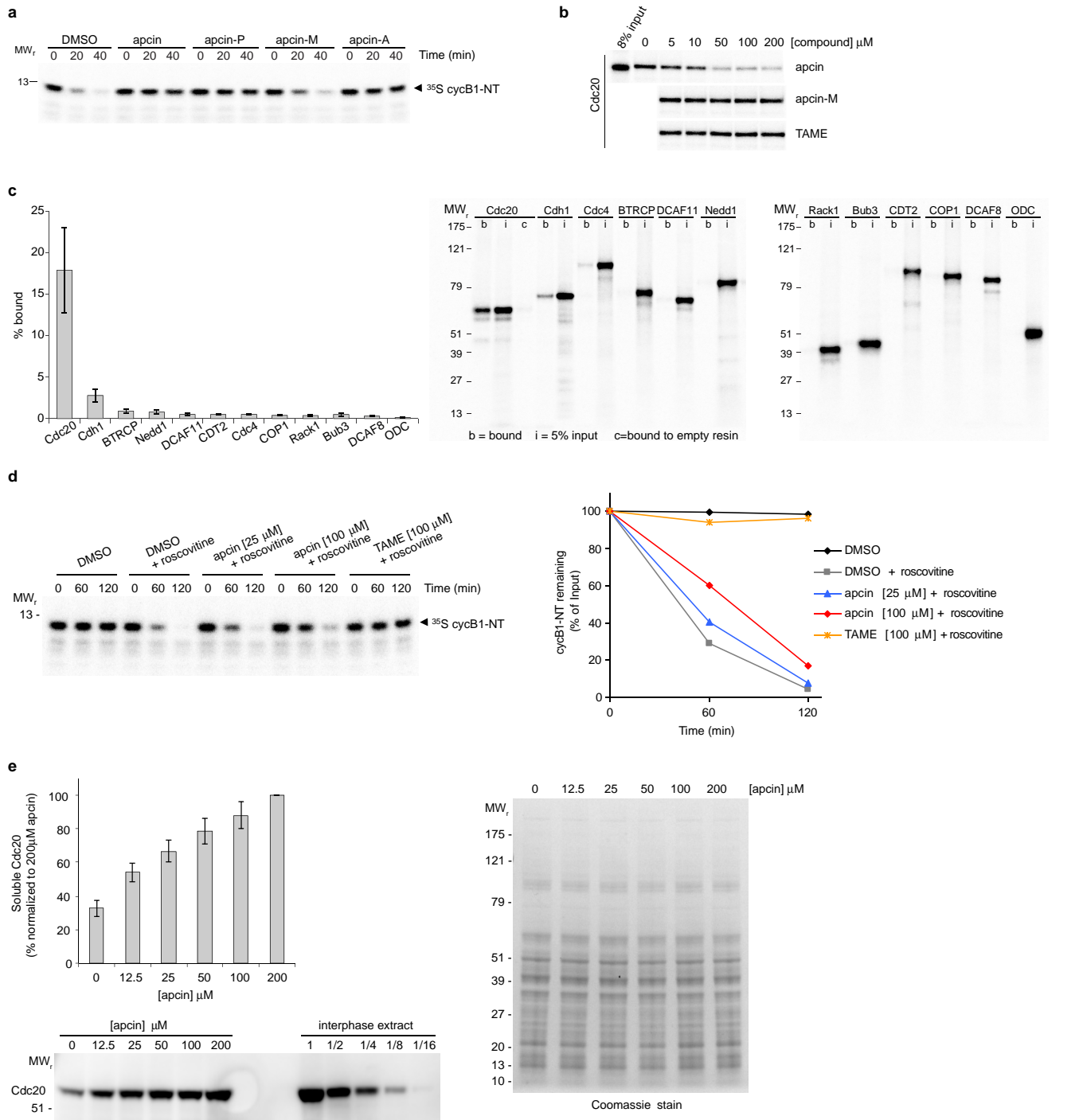
High-throughput image-based assay to measure mitotic fraction. Parental A549, U2OS and hTERT-RPE1 cells were purchased from ATCC. DLD-1 cells were purchased from Sigma. For hTERT-RPE1, A549 and U2OS cells, stable cell lines expressing H2B–GFP were derived using described methods⁴⁰ and used in the experiments. DLD-1 cells were used without further modification. Cell lines were tested for mycoplasma contamination (Lonza kit LT07-218) after they were derived and were found negative. For each cell line, asynchronous cells were re-suspended to a density of 3.75×10^4 cells per millilitre. A WellMate dispenser (Thermo Scientific) was used to distribute 40 µl of suspension to each well of a black, clear-bottom 384-well plate (3712, Corning). Plates were sealed with breathable white rayon sealing tape (241205, Nunc) during plating and subsequent incubation. After 24 h incubation, the cells were treated with indicated concentrations of apicin and proTAME dissolved in DMSO, in four technical replicates. After 18 h, cells were fixed and stained directly without wash steps to avoid loss of mitotic cells, by adding 10 µl of 6× concentrated fixing/staining reagent (60% formalin, 0.6% Triton X-100 and 1.5 µg ml⁻¹ Hoechst 33342 in DPBS). The plates were sealed with aluminium sealing tape (276014, Nunc) and incubated at 22–24 °C for 40 min before imaging. Plates were then imaged at four positions per well using an ImageXpress Micro (Molecular Devices) high-throughput microscope, with a ×10 objective, yielding a total of 16 images per condition (four images × four replicates). Cell images were processed automatically in ImageJ to identify the nuclei, count the number of nuclei and determine

the maximum fluorescence intensity of each nucleus in each image. The output files from ImageJ for each treatment were pooled and the cumulative frequency curves of maximum intensity for the cell population in each treatment were computed using Matlab. An intensity threshold was set on the basis of the mitotic fraction in the wells treated with DMSO to separate mitotic cells from interphase cells. The interphase fraction for each treatment was indicated by the fraction below the threshold on the cumulative frequency plot. Methods for statistically analysing the data are presented in Supplementary Information.

Fluorescence live-cell imaging. For experiments with RPE1-H2B–GFP cells, asynchronous RPE1-H2B–GFP cells were plated in 24-well glass-bottom plates (Greiner BioOne, 662892) 18–24 h before siRNA transfection using RNAiMax (Invitrogen). Cells were transfected with Mad2 siRNA (GGAACAACUGAAAGAUGGdTdT, synthesized by Dharmacon) or non-targeting control siRNA (D-001210-01-20, Dharmacon) at a final concentration of 20 nM. Twenty-four hours after transfection, cells were treated with compounds and imaging was initiated following compound treatment. For U2OS-H2B–GFP cells, the cells were first synchronized by double thymidine block (18 h first block, 8 h release, 17 h second block; thymidine concentration 2 mM). Compounds were added at 7 h after release from the second block, and imaging was initiated after compound treatment. To measure efficiency of Mad2 knockdown by siRNA, western blot samples were each prepared from a single well of the 24-well glass-bottom plates. Twenty-four hours after transfection, cells were collected by trypsinization, pelleting and re-suspension in 2× NuPAGE sample buffer (Invitrogen) + 50 mM DTT.

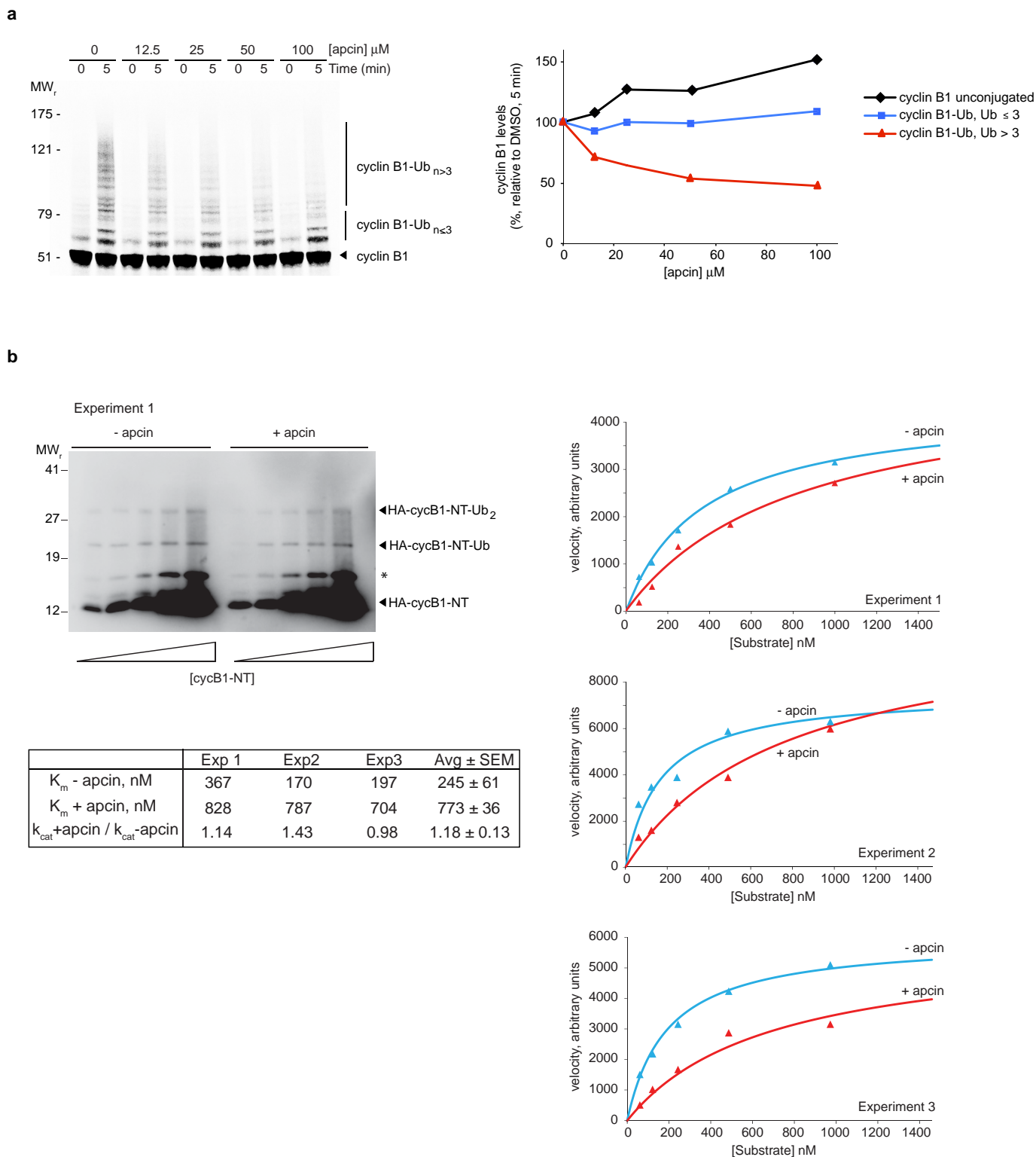
For imaging, plates were inserted into a covered chamber supplied with humidified 5% CO₂ and mounted onto a motorized microscope stage (Prior Scientific). Differential interference contrast and fluorescence images were captured at 6 min intervals for 45 h using a Nikon Ti inverted fluorescence microscope fitted with a 37 °C enclosed incubation chamber and using a ×20 Plan Apo 0.75 numerical aperture objective lens. A Hamamatsu ORCA cooled CCD (charge-coupled device) camera collected the images with 2 × 2 binning using Nikon Elements software (version 3.0). Videos were manually analysed using Nikon Elements software or ImageJ. Mitotic duration was defined as the time from nuclear envelope breakdown until anaphase, in the case of normal mitosis, or until exit from prolonged mitosis as indicated by cytoplasmic blebbing accompanied by changes in chromatin as detected by H2B–GFP. Cell fate was scored as 'division' if two daughter cells were produced by mitosis of any duration, and 'abnormal exit' if a single cell of interphase appearance resulted after that cell was in a mitotic state of any duration. The 'death' fate describes cells that entered mitosis and died while in mitosis. Methods for statistically analysing live cell imaging data are presented in Supplementary Information.

- Murray, A. Cell cycle extracts. *Methods Cell Biol.* **36**, 581–605 (1991).
- Salic, A. & King, R. W. Identifying small molecule inhibitors of the ubiquitin-proteasome pathway in *Xenopus* egg extracts. *Methods Enzymol.* **399**, 567–585 (2005).
- Martinez Molina, D. *et al.* Monitoring drug target engagement in cells and tissues using the cellular thermal shift assay. *Science* **341**, 84–87 (2013).
- Dimova, N. V. *et al.* APC/C-mediated multiple monoubiquitylation provides an alternative degradation signal for cyclin B1. *Nature Cell Biol.* **14**, 168–176 (2012).
- Otwinowski, Z. & Minor, W. Processing X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
- McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
- Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
- Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
- Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
- Sigoillot, F. *et al.* A time-series method for automated measurement of changes in mitotic and interphase duration from time-lapse movies. *PLoS ONE* 0025511 (2011).



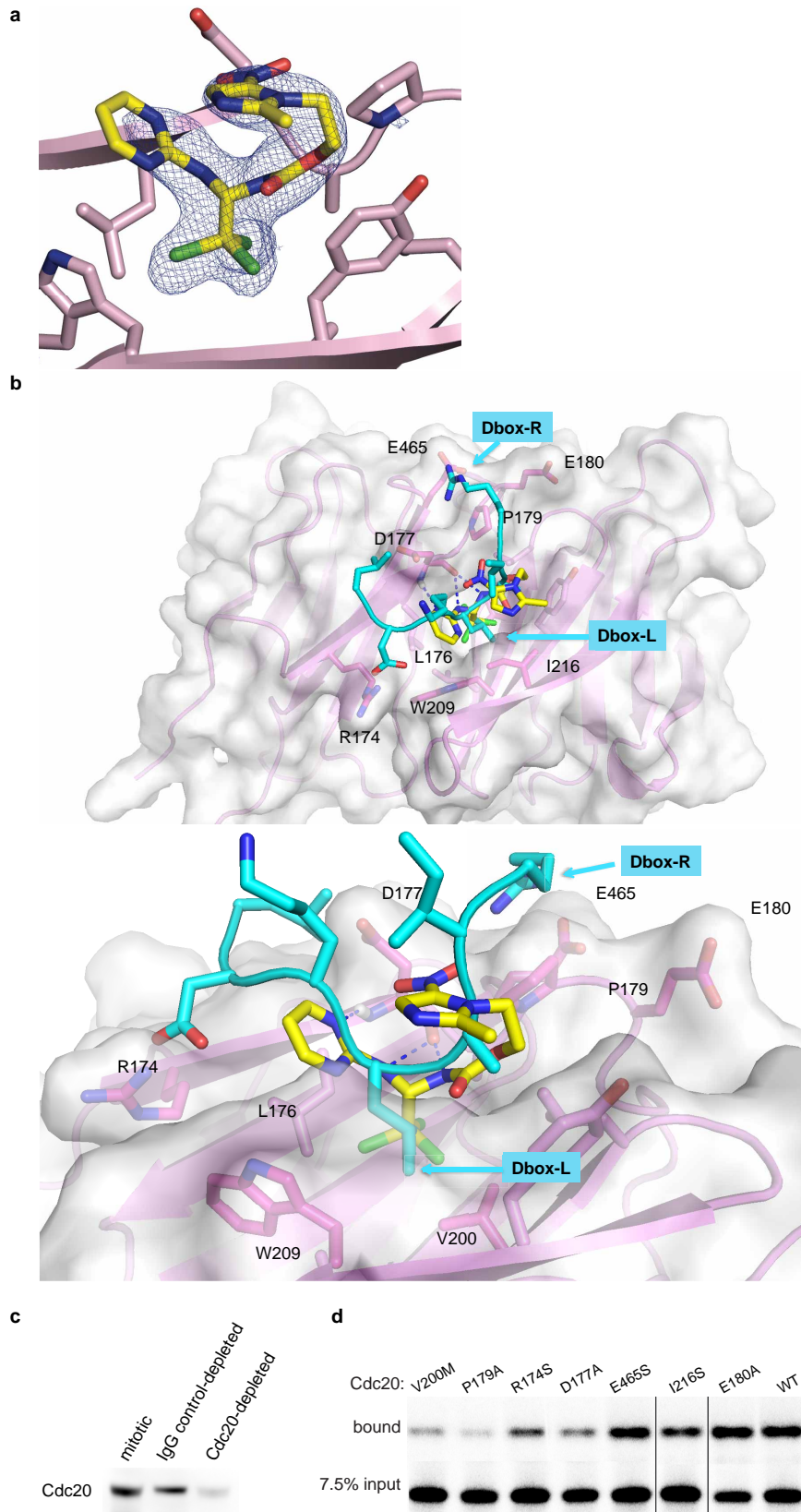
Extended Data Figure 1 | Apcin binds Cdc20 and inhibits APC/C-dependent proteolysis in *Xenopus* extract. **a**, Effects of apcin and derivatives (200 μM) on degradation of a cyclin B1 N-terminal fragment (cycB1-NT) in mitotic *Xenopus* egg extracts. Substrate was expressed in reticulocyte lysate and labelled with [³⁵S]methionine. Samples were analysed by SDS-gel electrophoresis and phosphorimaging. Quantitation of the 40-min time point from three independent experiments is shown in Fig. 1b. **b**, Apcin-M and TAME do not inhibit binding of Cdc20 to apcin-A resin. The experiment was performed as shown in Fig. 1e except that the inactive apcin derivative apcin-M or the Cdc20-IR tail antagonist TAME were also tested and [³⁵S]Cdc20 was detected by autoradiography. **c**, Apcin-A interacts strongly with Cdc20, weakly with Cdh1, and shows little interaction with other WD40-domain proteins. Proteins were expressed in reticulocyte lysate and labelled with [³⁵S]methionine. Left panel, mean value for the percentage bound on the basis of three experiments (error bars, s.e.m.). Right panel, representative autoradiograph of one of three experiments. b, bound; i, 5% input; c, bound to

empty resin (control). **d**, Apcin inhibits degradation of cycB1-NT in Cdh1-treated interphase extract. The addition of roscovitine (75 μM) is required to inhibit Cdk1 activity that suppresses Cdh1-dependent proteolysis. Note that 100 μM apcin is less effective at stabilizing the substrate in Cdh1-activated interphase extract than in mitotic *Xenopus* extract (Fig. 3b). **e**, Apcin binds to endogenous Cdc20 in *Xenopus* extract as measured by a thermal shift assay³³. Extract was incubated with varying concentrations of apcin, heated to 46 °C for 3 minutes, and precipitated proteins removed by centrifugation. The soluble fraction was analysed by SDS-PAGE and western blotting for Cdc20. The left panel shows the percentage of soluble Cdc20 (mean ± s.e.m. from three independent experiments). Western blot from one of three experiments is shown below. For comparison, total Cdc20 from interphase extract (and various dilutions) is shown. Right panel, Coomassie-stained gel of soluble proteins, indicating that there is not an observable non-specific stabilization of proteins induced by apcin addition.



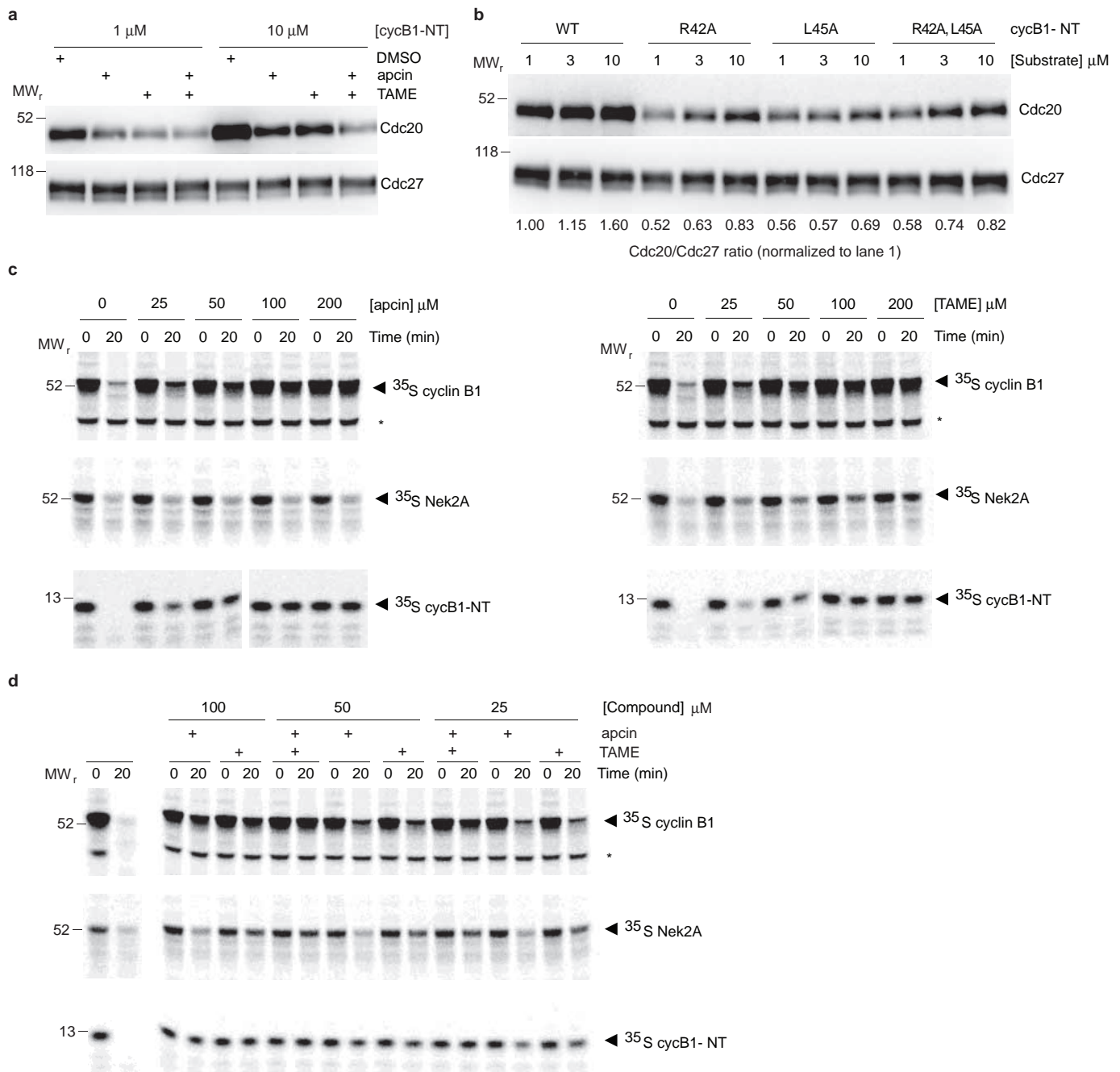
Extended Data Figure 2 | Apcin acts as a competitive inhibitor of APC/C-dependent ubiquitylation. **a**, Apcin inhibits formation of high-molecular-mass ubiquitin conjugates of full-length cyclin B1. Mitotic *Xenopus* extract was pre-treated with the deubiquitinating enzyme inhibitor ubiquitin vinyl sulfone (UbVS, 20 μ M) and proteasome inhibitor MG262 (150 μ M) to stabilize ubiquitin conjugates. [35 S]cyclin B1 was added together with ubiquitin (44 μ M) and samples analysed by SDS-PAGE and phosphorimaging. Right panel shows quantitation of the experiment. **b**, Apcin acts as a competitive inhibitor of APC/C-dependent ubiquitylation. APC/C was purified from mitotic *Xenopus* extracts and the initial rates of ubiquitylation of HA-tagged cycB1-NT were

measured in the presence of methylated ubiquitin to prevent ubiquitin chain elongation. The reaction was stopped at 45 s and the products were detected by anti-HA blot. The left panel shows the anti-HA blot from one experiment; substrate concentrations were 62.5, 125, 250, 500 and 1,000 nM (left to right). Asterisk indicates an SDS-resistant aggregated form of substrate. Quantitation of three independent experiments, and a summary of kinetic parameters, are shown. Note that the effects of apcin in this reconstituted assay performed under initial rate conditions appear distinct from those obtained in crude *Xenopus* extract. See Supplementary Discussion for a more detailed discussion of these differences.



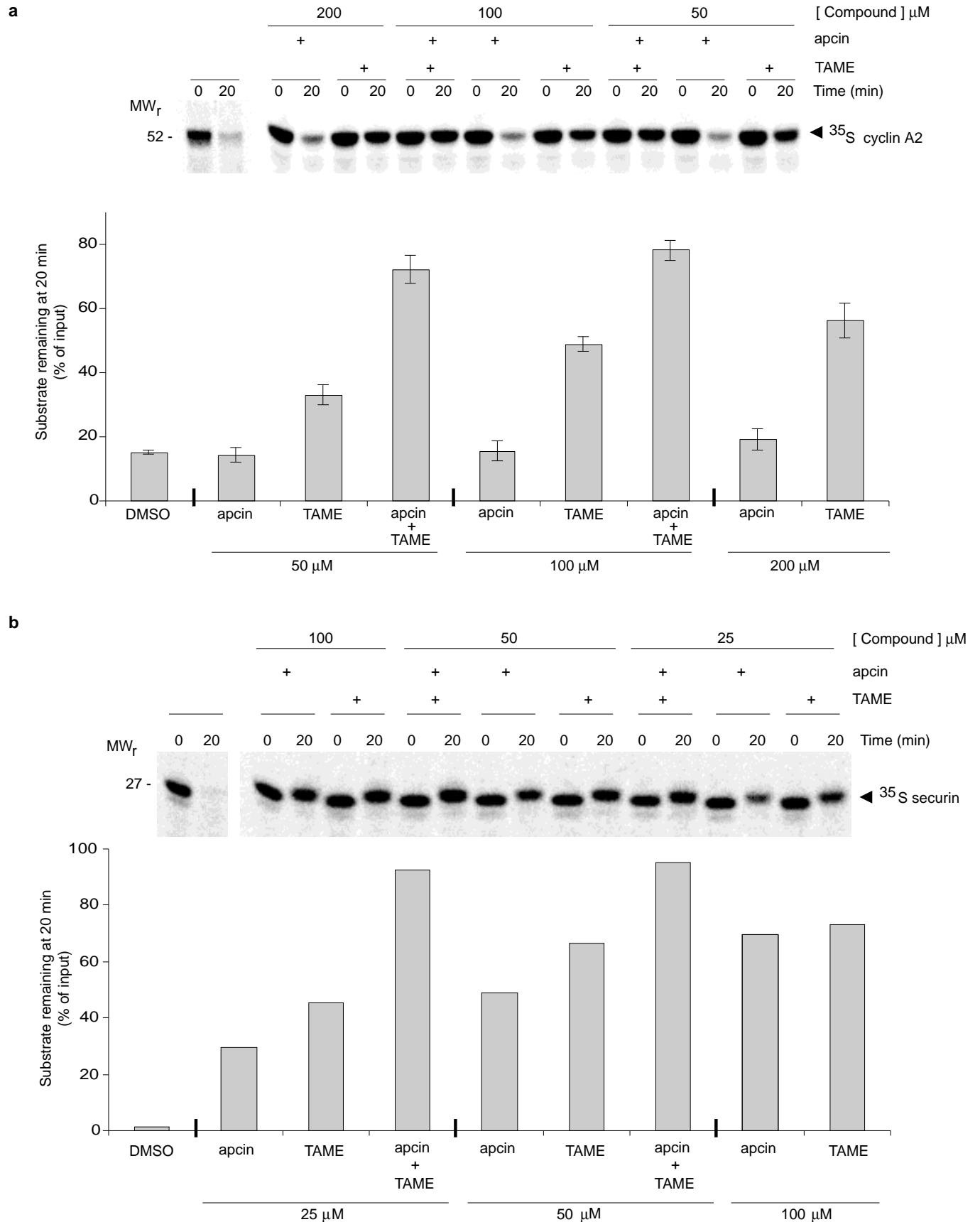
Extended Data Figure 3 | Structure of the apcin-Cdc20 complex. **a**, The $F_0 - F_c$ omit electron density map at the contour level of 3σ is shown. The density for the bound apcin conformation is unambiguous and is consistent with the structure-activity analysis and mutagenesis experiments. **b**, Overlay of the structure of the Cdc20-apsin structure with the structure of a D-box-containing protein bound to Cdh1 (ref. 14). The trichloromethyl group of apcin projects into a hydrophobic pocket that is occupied by the leucine of the RXXL

motif of the D-box. The position of the arginine from the RXXL motif suggests a role for E465 of Cdc20 in a charge-based interaction with the D-box, consistent with our data that the E465S mutation disrupts the ability of Cdc20 to promote substrate degradation more than it perturbs apcin binding. **c**, Depletion with anti-Cdc20 antibody covalently coupled to protein A beads depletes endogenous *Xenopus* Cdc20 from mitotic extract. **d**, Example autoradiogram of data shown in Fig. 2b.



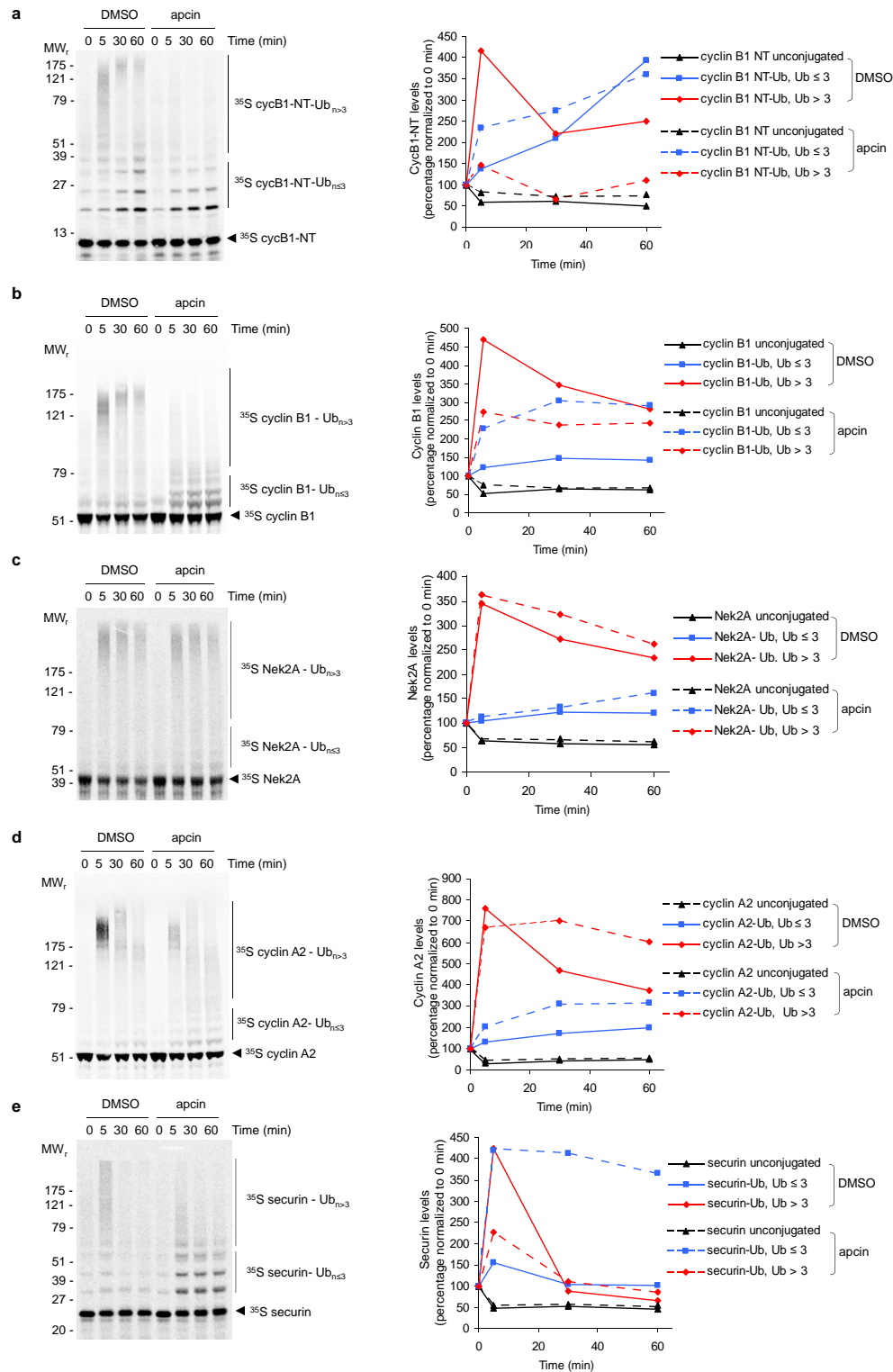
Extended Data Figure 4 | Effects of apcin on Cdc20 binding to APC/C and stability of APC/C substrates in mitotic *Xenopus* extract. **a**, Example western blot for data shown in Fig. 3a. **b**, Substrate-mediated recruitment of Cdc20 to APC/C in *Xenopus* extract is dependent on the D-box motif. Increasing concentrations of wild-type or different D-box mutants of cycB1-NT were introduced into mitotically arrested *Xenopus* extract and the APC/C was isolated with anti-Cdc27 antibodies by immunoprecipitation for 1 h at 4 °C. The immunoprecipitate was separated by SDS-PAGE and analysed by

western blotting against Cdc20 and Cdc27. Levels of Cdc20 were quantitated using ImageJ and normalized to APC/C subunit Cdc27. **c**, Analysis of effects of apcin and TAME on APC/C substrate degradation in mitotic *Xenopus* extract. Levels of ³⁵S-labelled substrates were assessed by SDS-PAGE and phosphorimaging. Asterisk represents a non-specific band. Images show one of three experiments quantitated in Fig. 3b. **d**, Experiment performed as in **c**, but examining the combined effects of apcin and TAME. Image shows one of three experiments quantitated in Fig. 3c.



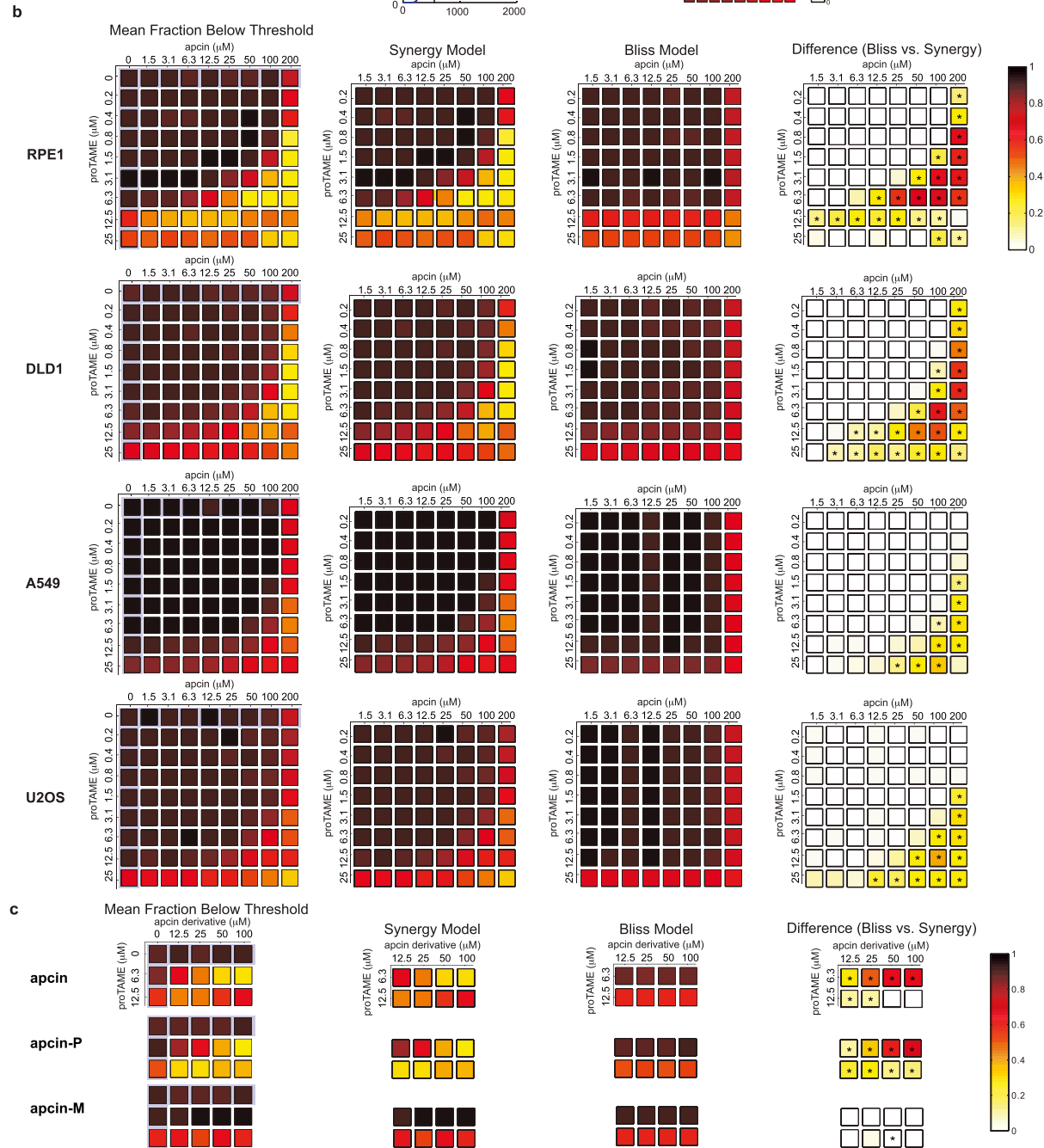
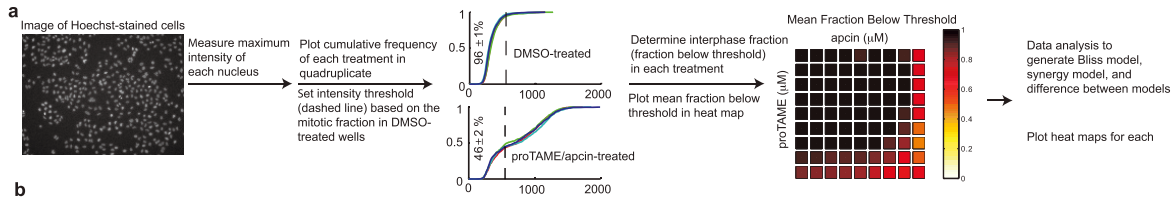
Extended Data Figure 5 | Effects of apcin and TAME on stability of APC/C substrates in mitotic *Xenopus* extract. Apcin and TAME synergize in stabilizing cyclin A2 (a) and securin (b) in mitotic *Xenopus* extract. Levels of ³⁵S-labelled substrates were assessed by SDS-PAGE and phosphorimaging.

The change in mobility of securin between 0 and 20 min is probably a result of mitotic phosphorylation. Error bars in a represent mean and s.e.m. of three independent experiments. Data in b are representative of two independent experiments.



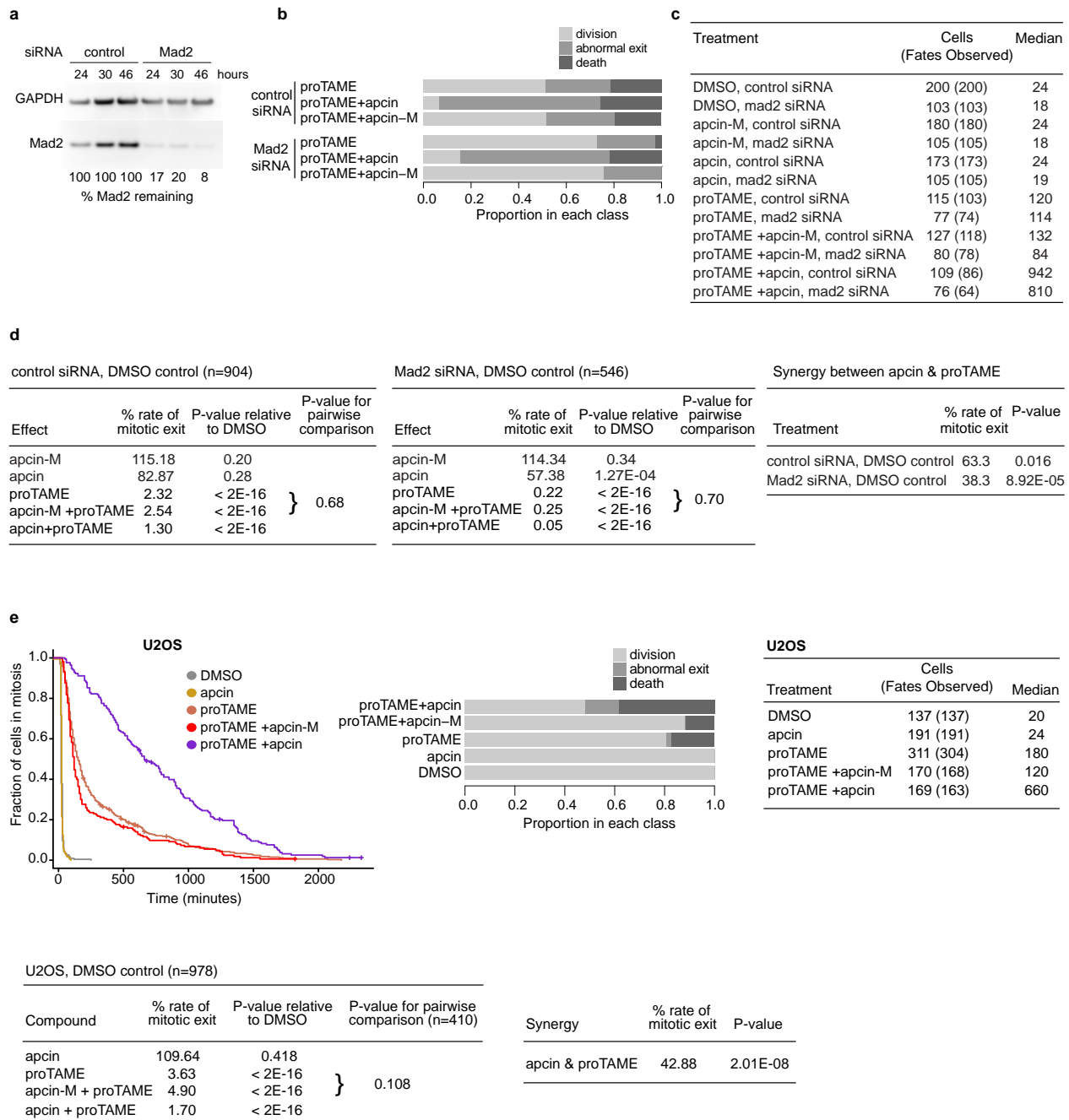
Extended Data Figure 6 | ApcIn has differential effects on substrate ubiquitylation in mitotic *Xenopus* extract that correlate with effects on proteolysis. To examine the profile of ubiquitylated species generated in *Xenopus* extract, deubiquitylation and proteasome-mediated degradation were inhibited by pre-treatment with the general deubiquitylating-enzyme inhibitor ubiquitin-vinyl sulfone (UbVS; 20 μ M), as established previously³⁴, and the proteasome-inhibitor MG262 (150 μ M). Next, wild-type ubiquitin (Ub; 44 μ M) and apcIn (100 μ M) or DMSO were added. ³⁵S-labelled APC/C substrates expressed in reticulocyte lysate were introduced into treated extract and their ubiquitylation at indicated times assessed by SDS-PAGE and phosphor imaging. Levels of substrates that were not modified with ubiquitin (unconjugated), modified with one to three ubiquitins (substrate-Ub, Ub \leq 3)

or modified with more than three ubiquitin moieties (substrate-Ub, Ub > 3) were quantitated by phosphorimaging and plotted relative to radiolabelled protein in the respective region of the gel at 0 min in the apcIn or DMSO sample. **a**, cycB1-NT; **b**, cyclin B1; **c** Nek2A; **d**, cyclin A2; **e**, securin. Note that for cyclin A2, apcIn reduces the amount of Ub conjugates with very high molecular mass (as indicated by inspection of the gel image) but does not reduce the fraction of conjugates modified with more than three ubiquitins (as indicated by the quantitation). These results are consistent with the inability of apcIn to stabilize cyclin A2 in proteolysis assays, given that the proteasome typically requires at least four ubiquitin molecules to be attached to a substrate for efficient proteolysis.



Extended Data Figure 7 | Apcin and proTAME synergize to block mitotic exit in human cells, as measured in a fixed cell assay. **a**, Summary of the fixed-cell imaging assay and data processing methods to determine synergy between apcin and proTAME. See methods for detailed description of the assay. **b**, First column: primary data plotted as a heat map displaying mean fraction below threshold for each drug treatment concentration in each of four cell lines. This threshold is established on the basis of the mitotic index of DMSO-treated cells. Note that a high value in this column indicates a low mitotic index. Effects of single drugs alone are highlighted in pale purple. Second column (labelled ‘Synergy model’): calculated effect of the combination of drugs on the basis of a model that permits synergistic interaction between proTAME and apcin. Because this panel shows calculated values of combination effects, the effects of

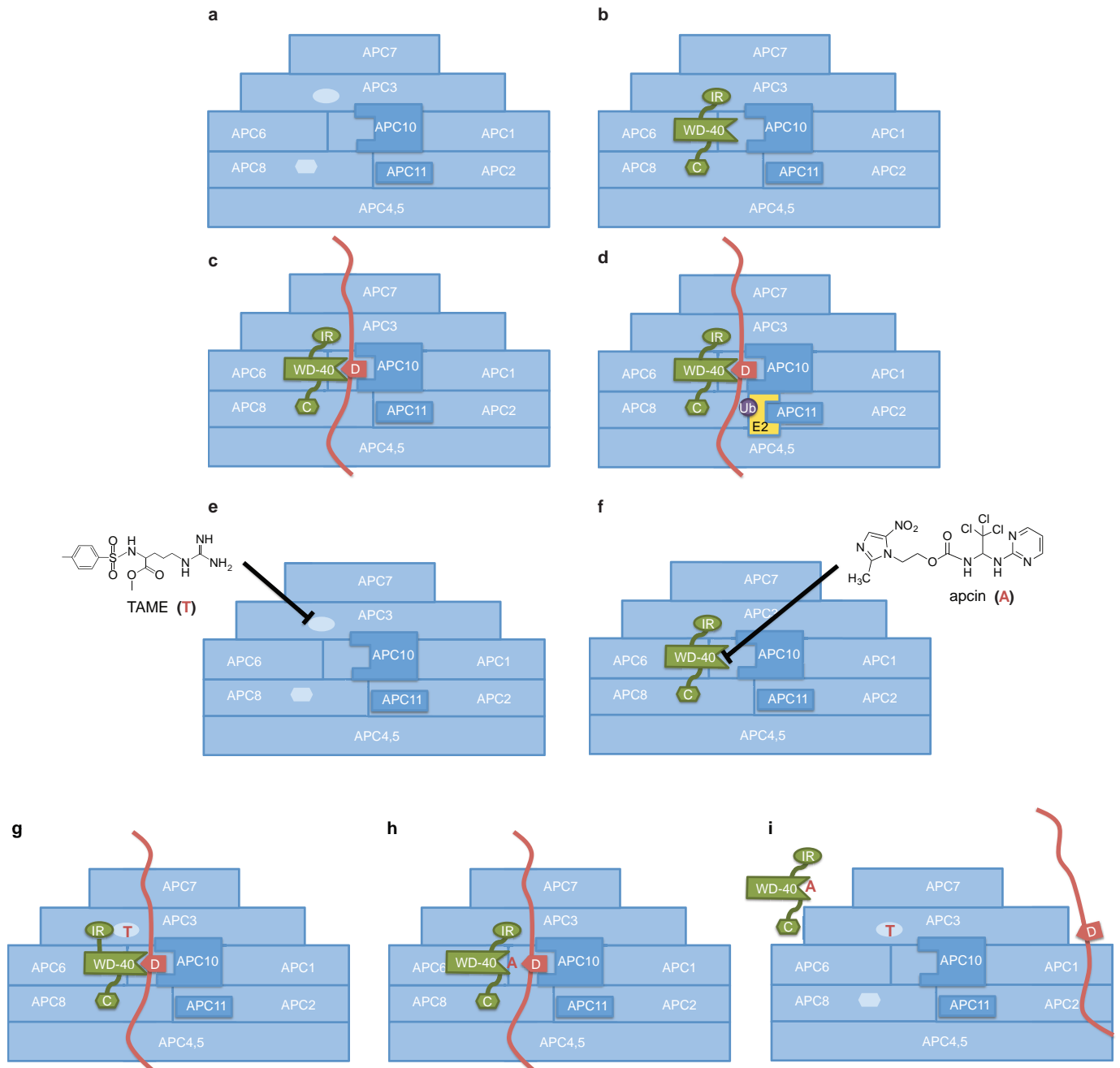
individual drugs are not shown. Note that the synergy model closely parallels the actual data shown in the first column. Third column: calculated effect of the combination of drugs on the basis of a model that permits only multiplicative interaction between proTAME and apcin (Bliss model). Because this panel shows calculated values of combination effects, the effects of individual drugs are not shown. Note that the Bliss model does not closely parallel the actual data shown in the first column. Fourth column: heat map of the difference between the synergy and Bliss model predictions shows the degree of synergy at each drug dose combination (same as Fig. 4a for RPE1 cells). **P* < 0.05 for analysis of four technical replicates. See Supplementary Information for details of the statistical analysis. **c**, Activity of apcin derivatives in the fixed cell assay described in **a**, using RPE1 cells.



Extended Data Figure 8 | Apcin and proTAME synergize to block mitotic exit in human cells, as measured in a live cell assay.

a, Western blot of Mad2 knockdown in RPE1 cells by siRNA from one of the experiments shown in Fig. 4b. **b**, Analysis of cell fate in RPE1 cells for the experiment shown in Fig. 4b. **c**, Combined data from two independent experiments in RPE1 cells shown in Fig. 4b. 'Cells' is the total number of cells analysed, while 'Fates observed' is the subset of cells whose fate was observed, excluding cells that migrated out of view during the movie or were still arrested at the end of imaging. The median is the time on the x axis of the Kaplan–Meier curve corresponding to 0.5 on the y axis 'Fraction of cells in mitosis'. **d**, Statistical modelling of data from experiment in Fig. 4b (RPE1 cells). The two tables on the left show the rate of mitotic exit relative to DMSO control for each of two siRNA treatment subgroups (control siRNA, left; Mad2 siRNA, middle), the *P* value for the comparison to DMSO (Cox proportional hazards model, see Supplementary Methods) and the *P* value for the comparison of proTAME with or without apcin-M. To determine *P* values for the pairwise comparisons of proTAME versus apcin-M with proTAME, we fitted a Cox proportional hazards model similar to that described in the Methods but with just an apcin-M effect and analysed just the subset of cells that were treated with either proTAME only or

apcin-M with proTAME. The table labelled 'Synergy between apcin and proTAME' shows the rate of mitotic exit with both compounds (apcin with proTAME), relative to what would be predicted by a multiplicative combination of the effects of each compound alone. The magnitude of the synergy is roughly doubled when checkpoint activity is reduced by Mad2 siRNA. **e**, Synchronized U2OS H2B–GFP cells were treated with apcin or apcin-M (25 μ M) and/or proTAME (12 μ M). Cells were then imaged every 6 or 10 min for 45 h. Mitotic duration and cell fate were determined by manual inspection of the videos and plotted as Kaplan–Meier curves. The hatch marks on the Kaplan–Meier curves indicate mitotic duration endpoints of censored cells. Graphs include the combined results of five independent experiments. The U2OS model differs from the model used to test RPE1 data in that the U2OS analysis is not stratified by either date or person, and the U2OS data do not include an effect of apcin-M alone (in the absence of proTAME). Pairwise comparison between proTAME and proTAME with apcin-M was tested using a Cox proportional hazards model stratified by date, using data from only the experimental blocks in which both proTAME and proTAME with apcin were tested.



Extended Data Figure 9 | Model of the effects of apcincin and TAME on formation of the APC/C-Cdc20-substrate ternary complex. **a**, Schematic drawing of core APC/C subunits. Not all subunits are indicated, and not all known interactions between subunits are illustrated for the sake of simplicity. The light blue oval and polygon indicate binding sites for Cdc20 on core APC/C subunits. **b**, In the absence of substrate, Cdc20 (green) can bind to the APC/C via the C-box (labelled 'C'), which interacts with APC8, and the IR-tail (labelled 'IR'), which interacts with APC3 (Cdc27). **c**, Binding of substrates (red) that contain a D-box (labelled 'D') can promote formation of a co-receptor interaction between the WD-40 domain of Cdc20 and Apc10.

d, The RING-containing subunit APC11 can recruit the E2 enzyme to conjugate ubiquitin to the substrate. **e**, TAME binds APC3 to interfere with the IR-tail binding site. **f**, Apcincin binds to the leucine pocket of the WD-40 domain of Cdc20. **g**, In the presence of TAME (labelled 'T'), the IR-binding site is disrupted, but Cdc20 can still be recruited to the APC/C through the C-box interaction and co-receptor interaction. **h**, Apcincin (labelled 'A') can disrupt the D-box interaction between the substrate and Cdc20, but Cdc20 can still interact through the C-box and IR-tail interactions. **i**, Combined use of apcincin and TAME disrupts both interactions, cooperatively disrupting the interaction between APC/C, Cdc20 and substrate.

Extended Data Table 1 | Data collection and refinement statistics for Cdc20–apcin structure

Data collection	
Space group	P2 ₁
Cell dimensions	
<i>a, b, c</i> (Å)	41.04, 87.18, 48.04
β (°)	112.72
Resolution range (Å) ^a	40.00 – 2.10 (2.13 – 2.10)
<i>R</i> _{merge} (%)	8.20 (38.9)
<i>I</i> / σ (<i>I</i>)	13.7 (1.8)
Data completeness (%)	92.8 (58.5)
Redundancy	3.5 (1.6)
Energy (eV)	12,684.1
Unique reflections	16,859 (522)
Wilson B-value (Å ²)	27.8
Refinement statistics	
Resolution range (Å)	28.58 – 2.10 (2.23 – 2.10)
No. of reflections <i>R</i> _{work} / <i>R</i> _{free}	16,859/857 (1,915/100)
Data completeness (%)	92.4 (67.0)
Atoms (non-H protein/solvent/inhibitor)	2,430/81/27
<i>R</i> _{work} (%)	16.5 (21.4)
<i>R</i> _{free} (%)	21.3 (29.1)
R.m.s.d. bond length (Å)	0.002
R.m.s.d. bond angle (°)	0.73
Mean B-value (Å ²) (protein/solvent/inhibitor)	29.6/30.0/51.1
Ramachandran plot (%) (favored/additional/disallowed) ^b	96.5/2.9/0.6
Maximum likelihood coordinate error	0.20
Missing residues, by chain	A: 161 – 164

^aData for the outermost shell are given in parentheses. ^bAs defined by the validation suite MolProbity.

Transcriptional interference by antisense RNA is required for circadian clock function

Zhihong Xue¹, Qiaohong Ye¹, Simon R. Anson², Jichen Yang³, Guanghua Xiao³, David Kowbel⁴, N. Louise Glass⁴, Susan K. Crosthwaite² & Yi Liu¹

Eukaryotic circadian oscillators consist of negative feedback loops that generate endogenous rhythmicities¹. Natural antisense RNAs are found in a wide range of eukaryotic organisms^{2–5}. Nevertheless, the physiological importance and mode of action of most antisense RNAs are not clear^{6–9}. *frequency* (*frq*) encodes a component of the *Neurospora* core circadian negative feedback loop, which was thought to generate sustained rhythmicity¹⁰. Transcription of *qrf*, the long non-coding *frq* antisense RNA, is induced by light, and its level oscillates in antiphase to *frq* sense RNA³. Here we show that *qrf* transcription is regulated by both light-dependent and light-independent mechanisms. Light-dependent *qrf* transcription represses *frq* expression and regulates clock resetting. Light-independent *qrf* expression, on the other hand, is required for circadian rhythmicity. *frq* transcription also inhibits *qrf* expression and drives the antiphase rhythm of *qrf* transcripts. The mutual inhibition of *frq* and *qrf* transcription thus forms a double negative feedback loop that is interlocked with the core feedback loop. Genetic and mathematical modelling analyses indicate that such an arrangement is required for robust and sustained circadian rhythmicity. Moreover, our results suggest that antisense transcription inhibits sense expression by mediating chromatin modifications and premature termination of transcription. Taken together, our results establish antisense transcription as an essential feature in a circadian system and shed light on the importance and mechanism of antisense action.

The transcription factors WHITE COLLAR (WC)-1 and -2 form a complex that activates *frq* transcription in the dark (DD) and mediates light-induced *frq* transcription for light-resetting of the clock by binding to light-responsive elements (LREs) on the *frq* promoter^{11–13}. 3' rapid amplification of cDNA ends (RACE) and RNA sequencing showed that *frq* and *qrf* transcripts overlap almost completely (Fig. 1a and Extended Fig. 1a). In the *wc* mutants, *frq* expression was nearly abolished and *qrf* transcript was observed at ~25% of wild-type levels (Fig. 1b), indicating that both WC-dependent and WC-independent mechanisms mediate *qrf* transcription. WC complex binds to the *qrf* promoter region¹⁴. *frq* constructs with point mutations in each of the five putative binding sites in the *qrf* promoter were individually introduced into a *frq*¹⁰ (*frq* and *qrf* null) strain¹⁵. Mutation of only one site, qLRE, markedly decreased the *qrf* level (Fig. 1c and Extended Data Fig. 1b–d). In the *frq*^{qLREmut} strain, the *qrf* level was comparable to that of the *wc* mutant. In a qLRE knock-in strain (*frq*^{KI(qLREmut)}; Extended Data Fig. 1e), *qrf* levels were also much lower than in the control knock-in strain, and WC binding at the *qrf* promoter was completely abolished (Fig. 1c, d and Extended Data Fig. 1f), indicating that qLRE is the only WC binding site in the *qrf* promoter.

Coinciding with low levels of *qrf* in the qLRE mutant, higher than wild-type levels of *frq* messenger RNA and FRQ protein were observed (Fig. 1c and Extended Data Fig. 1d, f, g). *frq* and *qrf* are rapidly light-induced in the wild type (Fig. 1e) but in the *frq*¹⁰; *frq*^{qLREmut} strain, whereas *qrf* induction was completely abolished (Fig. 1f), induction of *frq* in

the light was significantly elevated when compared with the wild type, indicating that light induction of *qrf* represses light-induced *frq*. Circadian condensation rhythms of qLRE mutant strains in DD were near

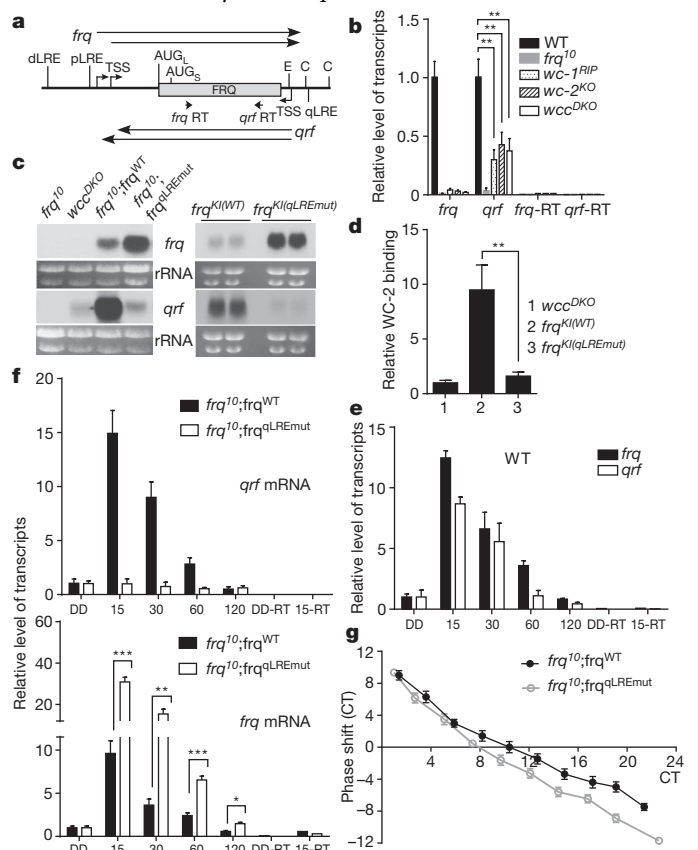


Figure 1 | WC-mediated light-induction of *qrf* represses *frq* to regulate light resetting of the clock. **a**, Diagram showing the *frq* locus. E, EcoRV site; C, ClaI site. **b**, Strand-specific analyses by quantitative PCR with reverse transcription (RT-qPCR), showing the levels of *frq* and *qrf* in constant light conditions (LL). *frq*-RT and *qrf*-RT are control reactions without reverse transcriptase. Error bars show s.d. ($n = 3$). **c**, Northern blot analyses showing the expression of *frq* and *qrf* in LL. **d**, WC-2 ChIP assays showing the WC binding levels at the qLRE of *qrf* promoter in LL. **e**, Strand-specific RT-qPCR analyses showing the levels of *frq* and *qrf* after 2 min of light induction at DD24 in the wild type (WT). The numbers on the x axis represent the time points (minutes) in the dark after light induction. **f**, Strand-specific RT-qPCR analyses showing the levels of *frq* and *qrf* after 2 min of light induction at DD24 in the indicated strains. The numbers on the x axis represent the time points (minutes) in the dark after light induction. **g**, Phase response curves of condensation rhythms after 2 min of light pulse at different circadian time (CT) points (hours). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (unpaired two-tailed t -test with Gaussian distribution assumption).

¹Department of Physiology, The University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, USA. ²Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, UK. ³Department of Clinical Sciences, The University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390, USA. ⁴Department of Plant and Microbial Biology, University of California, Berkeley, California 94720, USA.

normal; however, a light pulse resulted in significantly greater phase shifts in the qLRE mutants than in the controls (Fig. 1g and Extended Data Fig. 1h). These results are consistent with a previous report³ and indicate that *qrf* regulates light resetting of the clock by repressing light-induced *frq*.

Similar levels of *qrf* transcripts seen in the *frq*^{KI(WT)} and *frq*^{KI(qLREmut)} strains at DD24 (24 h in constant darkness) (Extended Data Fig. 2a) indicate that qLRE does not regulate *qrf* expression in DD. In a strain (*frq*¹⁰;frq.aq) in which the promoter of *qrf* was replaced with the quinic acid (QA)-inducible *qa-2* promoter, *qrf* expression was completely abolished in the absence of QA (Fig. 2a). On addition of QA, *qrf* was induced but *frq* levels were significantly decreased, further indicating repression of *frq* by *qrf*.

Several observations indicate that *qrf* expression must be within a certain range to permit a functional clock. Without QA, the *frq*¹⁰;frq.aq strain showed arrhythmic conidiation after the first day (Fig. 2b and Extended Data Fig. 2b). Moreover, a luciferase reporter (*Pfrq-luc*)¹⁶ in the *frq*¹⁰;frq.aq strain showed that the circadian luciferase activity seen in the control strain was abolished, indicating that *qrf* expression is required for clock function (Fig. 2c and Extended Data Fig. 2c, d). As QA concentration increased, circadian conidiation rhythms were gradually restored in the *frq*¹⁰;frq.aq strain, and at 10^{-5} – 10^{-4} M QA, the rhythms were similar to those in the control strain (Fig. 2b and Extended Data Fig. 2b). At higher QA concentrations, however, the amplitudes of the rhythms decreased (Fig. 2b) or became arrhythmic (Extended Data Fig. 2b). In addition, the phase of the rhythms was significantly delayed without QA (Extended Data Fig. 2e), a defect that was also rescued by *qrf* induction. Restoration of circadian rhythms of FRQ expression, FRQ phosphorylation profiles and *frq* mRNA oscillation by QA was also seen in the *frq*¹⁰;frq.aq strain (Extended Data Fig. 2f, g).

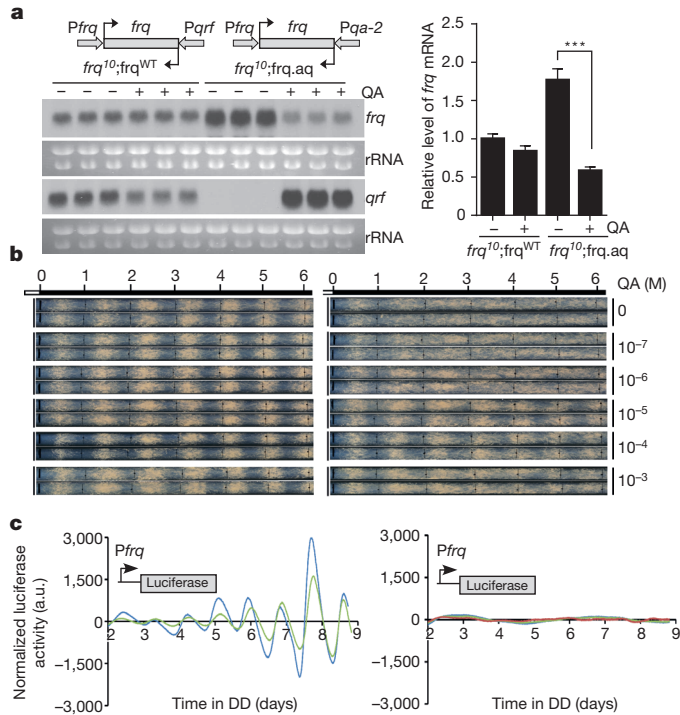


Figure 2 | *qrf* expression is required for circadian rhythmicities. **a**, Northern blot analysis showing the levels of *frq* and *qrf* with/without 1 mM QA (0.1% glucose) in LL. Right panel: densitometric analyses of the results. Error bars show s.d. ($n = 3$). **b**, Race tube analyses of *frq*¹⁰;frq^{WT} (left; τ (10^{-5} M) = 21.6 ± 0.3 hours) and *frq*¹⁰;frq.aq (right; τ (10^{-5} M) = 21.5 ± 0.4 hours) in 0.1% glucose with different concentrations of QA, in DD. **c**, Luciferase reporter assay showing the normalized *frq* promoter activity after two days in DD, for the wild type (left) and for strain *frq*¹⁰;frq.aq (right), in the absence of QA.

*qrf*RNA oscillates in DD in the wild type in antiphase to *frq* (Fig. 3a)³, but WC-2 does not bind to the *qrf* promoter in DD (Extended Data Fig. 3a). Moreover, the qLRE mutation did not affect either *frq* or *qrf* levels in DD (Extended Data Fig. 3b), indicating that the WC complex does not regulate *qrf* transcription in DD. Crucially, a luciferase reporter (*Pqrf-luc*) driven by the *qrf* promoter showed that the *qrf* promoter activity is not rhythmic in a wild-type strain (Fig. 3b and Extended Data Fig. 3c, d).

Several results indicate that *frq* and *qrf* are mutually inhibitory. First, in a *frq*⁹ mutant, only truncated FRQ protein is made, resulting in high *frq* levels¹⁵ but decreased *qrf* levels (Extended Data Fig. 3e). Second, when the *frq*¹⁰;frq^{qLREmut} strain was exposed to light, induction of *frq* resulted in decreased *qrf* (Fig. 3c). Third, low *frq* mRNA levels in the *wcc*^{DKO} mutant led to elevated *qrf* levels in DD (Fig. 3d). These results indicate that *frq* inhibits *qrf* expression.

To determine whether the rhythmic *frq* transcription can drive an anti-phasic rhythm of *qrf* without rhythmic *qrf* promoter activity, we created a luciferase reporter construct (*Pmin-luc-Pfrq*) in which the luciferase

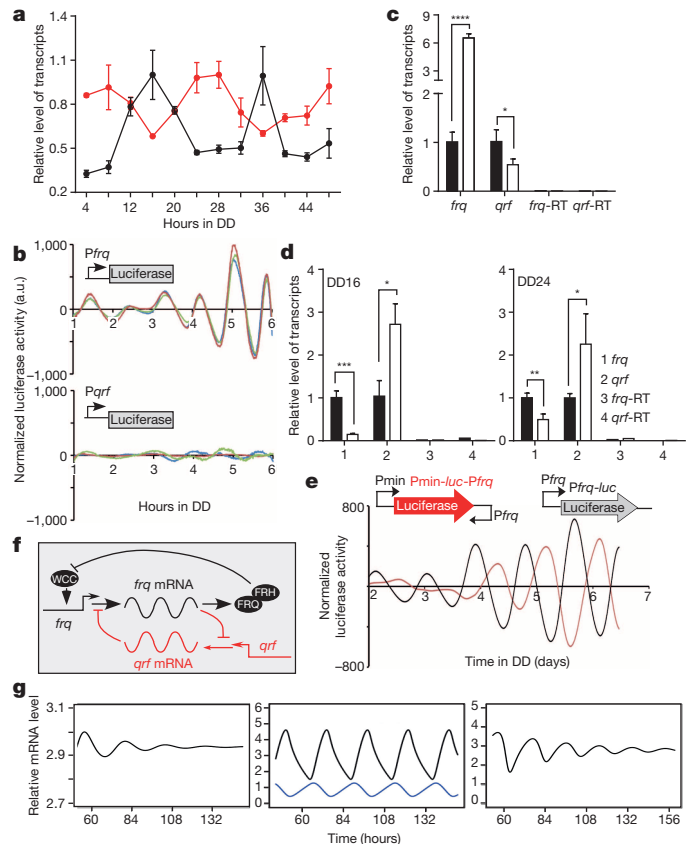


Figure 3 | Mutual inhibition of *frq* and *qrf* transcription forms a double negative feedback loop that is required for clock function. **a**, Strand-specific RT-qPCR results showing the oscillations of *frq* (black) and *qrf* (red) in DD. Error bars show s.d. ($n = 3$). **b**, Luciferase reporter assay showing the normalized *frq* or *qrf* promoter activity after one day in DD in wild-type strains with the *Pfrq-luc* (top) or *Pqrf-luc* (bottom) construct. **c**, Strand-specific RT-qPCR results showing the levels of *frq* and *qrf* in the *frq*¹⁰;frq^{qLREmut} strain at DD24 (filled bars) or 60 min in the dark after 2 min of light-pulse treatment (open bars). **d**, Levels of *frq* and *qrf* in the wild-type (filled bars) and *wcc*^{DKO} (open bars) strains measured by strand-specific RT-qPCR in DD. **e**, Luciferase reporter assays showing the normalized luciferase levels in a wild-type strain that carries the *Pfrq-luc* (black) or *Pmin-luc-Pfrq* (red) construct. **f**, A model of the *Neurospora* oscillator. **g**, Mathematical simulation of relative *frq* levels (black traces) in DD without *qrf* ($k_{19} = 0$) (left), with *qrf* ($k_{16} = 0.5$, $k_{17} = 0.35$, $k_{19} = 0.1$) (middle; *qrf* shown as blue trace), and with *qrf* overexpression ($k_{16} = 0.91$) (right). For a description of the variables, see Methods.

sense mRNA was driven by a constitutive promoter¹⁷ and antisense luciferase mRNA was driven by the *frq* promoter (Fig. 3e). Wild-type strains containing the *Pfrq-luc* or *Pmin-luc* (lacking antisense luciferase RNA) construct were used as controls. Luminescence in the *Pmin-luc* strain was arrhythmic, but the *Pmin-luc-Pfrq* strain exhibited a robust circadian luminescence rhythm in antiphase to that of the *Pfrq-luc* rhythm (Fig. 3e and Extended Data Fig. 4a, b). These results indicate that the antiphase rhythm of *qrf* expression is driven by rhythmic *frq* transcription independently of RNA sequence. Therefore, *frq* and *qrf* transcription forms a double negative feedback loop that results in antiphase rhythms of *frq* and *qrf* (Fig. 3f).

Mathematical modelling (Extended Data Fig. 4c)¹⁸ demonstrated that without *qrf* the *Neurospora* circadian oscillator can only generate a low-amplitude *frq* oscillation that damps out quickly (Fig. 3g). When the double negative feedback loops were introduced into the model, both *frq* and *qrf* levels oscillated robustly with antiphase rhythms. When *qrf* was overexpressed, the *frq* oscillation was not sustained. These results suggest that the previously known circadian feedback loops, although not sufficient to sustain a persisting rhythm, are the source of the rhythmicity that is amplified and sustained by the resonance of the mutually inhibitory and antiphase expression of *frq* and *qrf*.

How do *qrf* and *frq* inhibit each other? Light-induced *frq* transcription, *frq* and *qrf* levels and circadian rhythms were normal in RNA-mediated interference (RNAi) mutants (Extended Data Fig. 5a–d)¹⁹. Convergent transcription is also known to induce DNA and histone H3K9 methylation in the *frq* region^{20,21}, but the deletion of genes required for DNA and histone H3K9 methylation also did not affect *frq* expression (Extended Data Fig. 5e, f).

We introduced a *frq* construct (*qrf*) with the *frq* promoter deleted into the *frq*¹⁰ and *frq*¹⁰; *frq*^{qLREmut} strains at the *csr-1* locus (Extended Data Fig. 6a). This transgene can express normal levels of *qrf* without detectable *frq* expression (Extended Data Fig. 6b). In the *frq*¹⁰; *frq*^{qLREmut}; *qrf* strains, even though *qrf* expression was restored to normal levels, *frq* levels were not rescued (Extended Data Fig. 6c, d), indicating that *qrf* regulates *frq* *in cis*.

WC binding to the *frq* promoter initiates WC-dependent *frq* transcription, but the qLRE mutation did not affect WC binding at the *frq* promoter (Extended Data Fig. 6e, f). However, levels of *frq* pre-mRNA were significantly elevated in the qLRE mutants (Fig. 4a and Extended Data Fig. 6g). Moreover, *qrf* expression in the *frq*¹⁰; *frq*.aq strain did not affect *frq* RNA stability (Extended Data Fig. 6h). These results suggest that *qrf* regulates *frq* after transcriptional initiation.

After transcriptional initiation, the RNA polymerase II (pol II) C-terminal domain (CTD) is phosphorylated at Ser 2 and Ser 5, with Ser 5 and Ser 2 phosphorylation enriched near the 5' and 3' ends of transcribed regions, respectively²². Similar Ser 5 and Ser 2 phosphorylation profiles were seen at a *Neurospora* locus without antisense transcripts (Extended Data Fig. 7a, b). In contrast, both modifications of the CTD peaked at the same position in the middle of the transcribed *frq* region (Fig. 4b and Extended Data Fig. 8). Mutation of the qLRE, which decreases *qrf* expression, resulted in decreases in both Ser 2 and Ser 5 phosphorylation. Phosphorylation of pol II can trigger the methylation of histone H3 on Lys 36 (H3K36me3)²³. H3K36me3 enrichment at the *frq* locus peaked at the same position as CTD phosphorylation, and the qLRE mutation decreased H3K36me3 (Fig. 4b, bottom panel). In the *frq*¹⁰; *frq*.aq strain minus QA, in which *qrf* expression was completely abolished, the distributions of CTD phosphorylation and H3K36me3 on *frq* resembled the control locus lacking antisense transcription (Extended Data Fig. 9a–c). These results suggest that pol II stalls in the middle of the *frq* locus as a result of convergent transcription⁹.

SET-2 methylates H3K36 in *Neurospora*²⁴ and is required for clock function²⁵. Even though *frq* and *qrf* levels were increased only modestly in the *set-2*^{KO} single knockout mutants, their levels were markedly elevated in the *frq*⁹; *set-2*^{KO} double mutant (Fig. 4c and Extended Data Fig. 10a), suggesting that H3K36me3 contributes to the suppression of *frq* and *qrf* transcription. However, the induction of *qrf* still resulted in

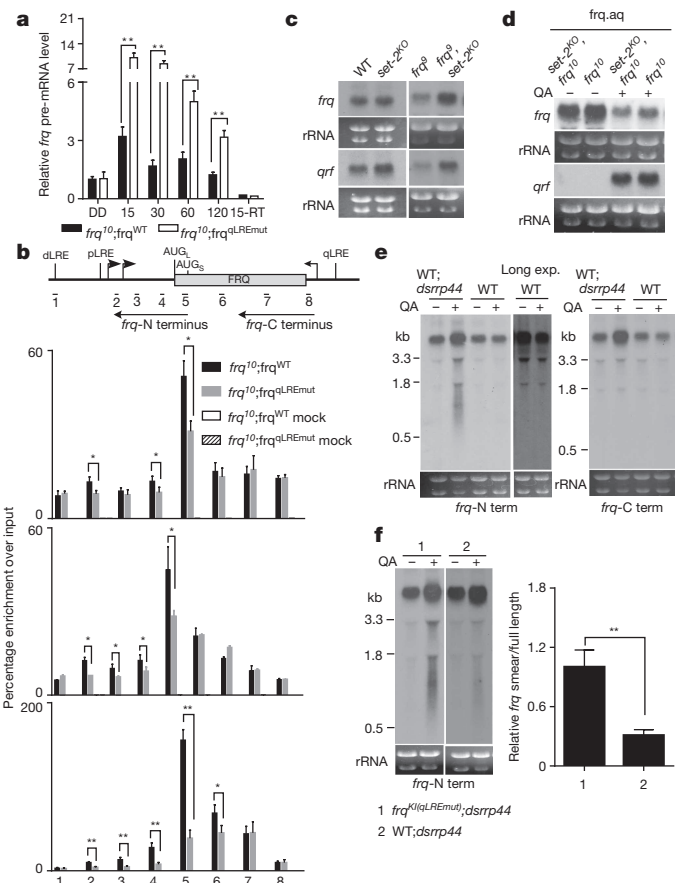


Figure 4 | *qrf* transcription results in pol II stalling, premature termination of transcription, and chromatin modifications.

a, Strand-specific RT–qPCR results showing the levels of intron-containing *frq* after a 2-min light pulse at DD24 in the indicated strains. The numbers on the *x* axis represent the time points (minutes) in the dark after light induction. **b**, ChIP assays showing the relative enrichment of pol II Ser 5 (top histogram), Ser 2 (middle) and H3K36me3 (bottom) in the *frq* locus in LL. IgG was used as the mock control for IP. dLRE, distal LRE; pLRE, proximal LRE; qLRE, *qrf* LRE. **c**, Northern blot analyses showing the levels of *frq* and *qrf* in LL in indicated strains. **d**, Northern blot analyses showing the levels of *frq* and *qrf* with/without 1 mM QA (0.1% glucose) in LL in indicated strains. **e**, Northern blot analysis in LL. The addition of QA results in *rrp44* silencing. *frq*-N term and *frq*-C term are specific for the 5' or 3' half of the *frq* transcripts, respectively (shown in **b**). Long exp., long exposure; kb, kilobases. **f**, Northern blot analysis in LL. The ratios between truncated and full-length *frq* transcripts from three independent experiments are shown. Error bars show s.d. ($n = 3$).

** $P < 0.01$ (unpaired two-tailed *t* test with Gaussian distribution assumption).

a decrease in *frq* in the *set-2*^{KO}, *frq*¹⁰ double mutant strain (Fig. 4d and Extended Data Fig. 10b), indicating the existence of another mechanism that mediates the action of antisense transcription.

We hypothesized that the stalling of pol II as a result of convergent transcription may abort transcription prematurely, which should result in truncated transcripts from their respective 5' ends but not from the 3' ends. Because these truncated transcripts are normally unstable, we first screened a panel of *Neurospora* mutants for the nuclease that degrades such transcripts. The RNA exosome is involved in degrading full-length *frq* transcripts²⁶. Northern blot analysis using an RNA probe (*frq*-N term or *qrf*-N-term) specific for the 5' end of *frq* or *qrf*, respectively, revealed that the silencing of *rrp44*, which encodes the exosome catalytic subunit, resulted in the appearance of a low-molecular-mass RNA smear in the wild-type strain (Fig. 4e and Extended Data Fig. 10c). In contrast, a probe specific for the 3' half of *frq* failed to detect the RNA smear. The amount of the RNA smear was decreased markedly in the qLRE mutant, which has reduced levels of *qrf* expression, (Fig. 4f) and

was not observed for a control gene without antisense transcripts (Extended Data Fig. 10d). Taken together, these results suggest that both premature transcription termination and chromatin modifications contribute to the transcriptional inhibition of *frq* by *qrf*.

We discovered here that the sense and antisense transcription of *frq* forms a double negative feedback loop that is interlocked with the core circadian feedback loops in *Neurospora*. The mutual transcription interference of *frq* and *qrf* results in antiphase oscillations of *frq* and *qrf* that resonate to achieve robust and sustained circadian gene expression. In the silkworm and mouse liver, antisense *per* RNAs exist and were also found to oscillate in antiphase to sense RNAs^{27–30}, suggesting that a similar mechanism may function in animal circadian systems.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 24 September 2013; accepted 10 July 2014.

Published online 17 August 2014.

- Dunlap, J. C. Molecular bases for circadian clocks. *Cell* **96**, 271–290 (1999).
- Berretta, J. & Morillon, A. Pervasive transcription constitutes a new level of eukaryotic genome regulation. *EMBO Rep.* **10**, 973–982 (2009).
- Kramer, C., Loros, J. J., Dunlap, J. C. & Crosthwaite, S. K. Role for antisense RNA in regulating circadian clock function in *Neurospora crassa*. *Nature* **421**, 948–952 (2003).
- Nagano, T. *et al.* The *Air* noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**, 1717–1720 (2008).
- Deuve, J. L. & Avner, P. The coupling of X-chromosome inactivation to pluripotency. *Annu. Rev. Cell Dev. Biol.* **27**, 611–629 (2011).
- Camblong, J., Iglesias, N., Fickentscher, C., Dieppois, G. & Stutz, F. Antisense RNA stabilization induces transcriptional gene silencing via histone deacetylation in *S. cerevisiae*. *Cell* **131**, 706–717 (2007).
- Ogawa, Y., Sun, B. K. & Lee, J. T. Intersection of the RNA interference and X-inactivation pathways. *Science* **320**, 1336–1341 (2008).
- Prescott, E. M. & Proudfoot, N. J. Transcriptional collision between convergent genes in budding yeast. *Proc. Natl Acad. Sci. USA* **99**, 8796–8801 (2002).
- Hobson, D. J., Wei, W., Steinmetz, L. M. & Svejstrup, J. Q. RNA polymerase II collision interrupts convergent transcription. *Mol. Cell* **48**, 365–374 (2012).
- Heintzen, C. & Liu, Y. The *Neurospora crassa* circadian clock. *Adv. Genet.* **58**, 25–66 (2007).
- Crosthwaite, S. K., Dunlap, J. C. & Loros, J. J. *Neurospora wc-1* and *wc-2*: Transcription, photoresponses, and the origins of circadian rhythmicity. *Science* **276**, 763–769 (1997).
- He, Q. *et al.* White collar-1, a DNA binding transcription factor and a light sensor. *Science* **297**, 840–843 (2002).
- Froehlich, A. C., Liu, Y., Loros, J. J. & Dunlap, J. C. White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter. *Science* **297**, 815–819 (2002).
- Smith, K. M. *et al.* Transcription factors in light and circadian clock signaling networks revealed by genomewide mapping of direct targets for *Neurospora* white collar complex. *Eukaryot. Cell* **9**, 1549–1556 (2010).
- Aronson, B. D., Johnson, K. A. & Dunlap, J. C. The circadian clock locus *frequency*: a single ORF defines period length and temperature compensation. *Proc. Natl Acad. Sci. USA* **91**, 7683–7687 (1994).
- Gooch, V. D. *et al.* Fully codon-optimized luciferase uncovers novel temperature characteristics of the *Neurospora* clock. *Eukaryot. Cell* **7**, 28–37 (2008).
- Bell-Pedersen, D., Dunlap, J. C. & Loros, J. J. Distinct cis-acting elements mediate clock, light, and developmental regulation of the *Neurospora crassa eas* (*ccg-2*) gene. *Mol. Cell. Biol.* **16**, 513–521 (1996).
- Hong, C. I., Jolma, I. W., Loros, J. J., Dunlap, J. C. & Ruoff, P. Simulating dark expressions and interactions of *frq* and *wc-1* in the *Neurospora* circadian clock. *Biophys. J.* **94**, 1221–1232 (2008).
- Chang, S. S., Zhang, Z. & Liu, Y. RNA interference pathways in fungi: mechanisms and functions. *Annu. Rev. Microbiol.* **66**, 305–323 (2012).
- Dang, Y., Li, L., Guo, W., Xue, Z. & Liu, Y. Convergent transcription induces dynamic DNA Methylation at disiRNA loci. *PLoS Genet.* **9**, e1003761 (2013).
- Belden, W. J., Lewis, Z. A., Selker, E. U., Loros, J. J. & Dunlap, J. C. CHD1 remodels chromatin and influences transient DNA methylation at the clock gene frequency. *PLoS Genet.* **7**, e1002166 (2011).
- Hsin, J. P. & Manley, J. L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137 (2012).
- Li, B., Howe, L., Anderson, S., Yates, J. R., III & Workman, J. L. The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **278**, 8897–8903 (2003).
- Adhvaryu, K. K., Morris, S. A., Strahl, B. D. & Selker, E. U. Methylation of histone H3 lysine 36 is required for normal development in *Neurospora crassa*. *Eukaryot. Cell* **4**, 1455–1464 (2005).
- Zhou, Z. *et al.* Suppression of WC-independent frequency transcription by RCO-1 is essential for *Neurospora* circadian clock. *Proc. Natl Acad. Sci. USA* **110**, E4867–E4874 (2013).
- Guo, J., Cheng, P., Yuan, H. & Liu, Y. The exosome regulates circadian gene expression in a posttranscriptional negative feedback loop. *Cell* **138**, 1236–1246 (2009).
- Sauman, I. & Reppert, S. M. Circadian clock neurons in the silkworm *Antheraea pernyi*: novel mechanisms of Period protein regulation. *Neuron* **17**, 889–900 (1996).
- Koike, N. *et al.* Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* **338**, 349–354 (2012).
- Menet, J. S., Rodriguez, J., Abruzzi, K. C. & Rosbash, M. Nascent-Seq reveals novel features of mouse circadian transcriptional regulation. *eLife* **1**, e00011 (2012).
- Vollmers, C. *et al.* Circadian oscillations of protein-coding and regulatory RNAs in a highly dynamic mammalian liver epigenome. *Cell Metab.* **16**, 833–845 (2012).

Acknowledgements We thank J. Cha, Y. Dang and H. Yuan for technical assistance, and B. Li for critical comments. Supported by grants from the National Institutes of Health to Y.L. (GM068496, GM062591) and N.L.G. (GM081597), the Welch Foundation (I-1560) to Y.L., the Cancer Prevention Research Institute of Texas (RP101496) to Z. X., and the Biotechnology and Biological Sciences Research Council (BBS/S/C2005/13012) to S.R.A. and S.K.C.

Author Contributions Z.X., S.R.A., S.K.C. and Y.L. designed experiments. Z.X., S.R.A., J.Y. and D.K. performed experiments. Q.Y. provided technical support. Y.L., Z.X., S.R.A., N.L.G. and S.K.C. analysed data. Y.L., Z.X. and S.K.C. wrote the paper.

Author Information RNA-Seq data are available at the Sequence Read Archive under accession number SRP030415. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to Y.L. (E-mail: Yi.Liu@UTsouthwestern.edu).

METHODS

Strains, plasmid constructs, and growth conditions. The wild-type strain used in this study was 87-3 (*ras-1bd*, a). The *frq*¹⁰ strain is a *frq*-null and *qrf*-null mutant¹⁵. The *frq*⁹ strain bears a frame-shift mutation in the *frq* open reading frame (ORF)¹⁵. The mutants *wc-1*^{RIP}, *wc-2*^{KO}, *wcc*^{DKO} (*wc-1*^{RIP}, *wc-2*^{KO}), *dcl*^{DKO} (*dcl-1*^{RIP}, *dcl-2*^{KO}), *qde-1*^{KO}, *qde-2*^{RIP} and *qde-3*^{KO} were generated in previous studies^{12,31}. The *dim-2*^{KO} and *dim-5*^{KO} strains were provided by Q. He³². The *set-2*^{KO} (FGSC no. 15505) strain was from Fungal Genetic Stock Center (FGSC). The *set-2*^{KO}*frq*⁹ double mutant was created in this study by crossing *set-2*^{KO} and *frq*⁹. The *dsrrp44* strain was generated by introduction of a plasmid expressing QA-driven *rrp44*-specific RNA hairpin into a wild-type strain²⁶. The *frq*^{KI(WT)} and *frq*^{KI(qLREmut)} strains were created in this study (Extended Data Fig. 1b, e). The *frq*^{KI(qLREmut)}*dsrrp44* strain was created by crossing *frq*^{KI(qLREmut)} and *dsrrp44*. The *frq*¹⁰; *frq*^{WT}, *frq*¹⁰; *frq*^{qLREmut}, *frq*¹⁰; *frq*.aq strains were obtained by targeting plasmids pKAJ120, pKAJ120^{qLREmut} and pKAJ120.aq, respectively, into the *his-3* locus of the *frq*¹⁰ mutant as described previously¹⁵ (Extended Data Fig. 1c). The *frq*¹⁰; *frq*^{qLREmut}; *qrf* strain was generated by targeting the plasmid pCSR1.qrf into the *csr-1* locus of the *frq*¹⁰; *frq*^{qLREmut} mutant as described previously³³ (Extended Data Fig. 6a).

The pKAJ120 containing the entire wild-type *frq* gene, including its promoter and a *his-3* targeting sequence, was used as the parental plasmid¹⁵. The qLRE in pKAJ120 was mutated as described (Extended Data Fig. 1b) by site-directed mutagenesis to create pKAJ120^{qLREmut}. The fragment between BssHII and SapI sites of pKAJ120 was replaced with an inverted *Neurospora qa-2* promoter to create pKAJ120.aq. The PCR fragment containing entire wild-type *qrf* and its promoter (primer, 5'-TTCATTAAGGTGGGGCAGG-3' and 5'-TTTCCACGCCGCCAGTC-3') was inserted into vector pCSR1 between NotI and PstI sites to generate pCSR1.qrf³³.

Growth conditions were as described previously¹⁵. Liquid cultures were grown in minimal medium (1 × Vogel's, 2% glucose). When QA was used, liquid cultures were grown in low-glucose medium (1 × Vogel's, 0.1% glucose, 0.17% arginine) with indicated concentrations of QA. For *rrp44* knockdown assay, *Neurospora* was cultured into mats in low-glucose medium with 10⁻² M QA for two days. Afterwards, *Neurospora* mats were cut into discs and cultured in flasks in the same medium with shaking. After two days the tissues were harvested. For the mRNA decay assay, the cultured conditions were as described previously²⁶. For rhythmic experiments, the *Neurospora* cultures were transferred from LL to DD at time 0 and were collected in DD at the indicated time (hours). For light induction, *Neurospora* cultures were grown in DD for 24 h with shaking, and then treated with a 1,750 lx light pulse. Afterwards, *Neurospora* cultures were transferred back to DD and collected at the indicated time (minutes). For race tube assay, the medium contained 1 × Vogel's, 0.1% glucose, 0.17% arginine, 50 ng ml⁻¹ biotin and 1.5% agar with indicated concentrations of QA (for results shown in Extended Data Figs 2b, glucose was not present in the medium). Strains were inoculated and grown in constant light at 25 °C for 24 h before being transferred to DD at 25 °C. Calculations of period length and phase were performed as described³⁴.

Analysis of phase response to light. The assay was performed in race tubes containing acetate/casamino acid medium (1 × Vogel's, 1.2% sodium acetate, 0.05% casamino acid hydrolysate, 1.5% agar)³⁵. Race tubes were first grown in LL at 25 °C for 48 h before being transferred to DD. Cultures were then grown in DD at 25 °C for 25 h, and different individual cultures were given a 2-min light pulse (1,750 lx) at different times (2-h intervals) to cover an entire circadian cycle. The amount of phase shift was determined by comparing light-treated cultures (six replicas for each time point) with those of the control cultures (kept in DD). The initial LL to DD transition was defined as circadian time (CT) 12. The phases of the cultures were calculated as the average phase for two consecutive days after the light treatment.

Protein and RNA analysis. Tissue harvest, protein extraction and western blots were performed as described previously¹⁵. For protein separation, 40 µg of total protein was loaded in each lane of SDS-PAGE (7.5%). Total RNA was extracted with Trizol in accordance with the manufacturer's protocol, and then further purified with 2.5 M LiCl as described previously³⁶. Northern blot analyses were performed as described previously using [³²P]UTP-labelled riboprobes¹⁵. Riboprobes were transcribed *in vitro* from PCR products by T3 or T7 RNA polymerase (Ambion) with the manufacturer's protocol. The primer sequences used for the template amplification were *frq*-N term (5'-TAATACGACTCATATAGGG (T7 promoter) GGCAGGGTACGATTGGATT-3' and 5'-GGGTAGTCGTGACTTTGTGACG-3'), *frq*-C term (5'-TAATACGACTCACTATAGGG (T7 promoter) CCTTCGT TGGATATCCATCATG-3' and 5'-GAATTCCTGCAGGGAAGCCGG-3'), *qrf*-N term (5'-AATTAACCCCTACTAAAGGG (T3 promoter) GAATTCCTGCA GGAAGCCGG-3' and 5'-CCTTCGTGGATATCCATCATG-3') and NCU01953 N-term (5'-TAATACGACTCACTATAGGG (T7 promoter) GTGCCAAAGAG TTGGCCATTC-3' and 5'-CTTGACCACAACTGTTGAAC-3'). Except for results shown in Fig. 4e, f, riboprobes of *frq*-C term and *qrf*-N term were used for detecting *frq* and *qrf*, respectively. For Fig. 4e, f and Extended Data Fig. 10, the riboprobes were hydrolysed for 25 min in 40 mM Na₂CO₃, 60 mM NaHCO₃,

10 mM 2-mercaptoethanol; the reaction was then stopped with an equal volume of 0.2 M sodium acetate pH 6.0, 1% acetic acid, 10 mM dithiothreitol (DTT). For western blotting and northern blotting, densitometry analyses were performed with Image J.

Quantitative PCR and RT-qPCR were performed as described previously²⁶. For strand-specific RT-qPCR, several modifications of the protocol were made. The primer sequences for strand-specific RT reactions were *frq* (5'-GCTAGCTTCAG CTAGGCATC (adaptor) CGTTGCCTCCAACCTCACGTTTCTT-3'), *qrf* (5'-CC TCAGCTCTGACGAGTCGTAC (adaptor) GTCATGGAGCCCTCTGGTCTTG GT-3'), *frq* pre-mRNA (5'-GCTAGCTTCAGCTAGGCATC (adaptor) TTGAA CCGTAGGGAGGAGGAGAG-3') and *β-tubulin* (5'-CTCGTTGTCAATGCAG AAGGTC-3'). The RT reaction was performed by mixing the primers of specific strand and *β-tubulin*. *β-tubulin* was used for internal control. The primer sequences for the qPCR step of the RT-qPCR assay were *frq* (5'-AGCTTCAGCTAGGCA TCCGTT-3' and 5'-GCAGTTTGGTTCCGACGTGATG-3'), *qrf* (5'-CAGCTC GTACGAGTCGTACGTC-3' and 5'-ATCTTCCGATGTTGTGCGAGCGT-3'), *frq* pre-mRNA (5'-AGCTTCAGCTAGGCATCTTGAACG-3' and 5'-ACGGCATC TCATCCATCTCACCA-3') and *β-tubulin* (5'-ATAACTTCGTCTTCGGCCA G-3' and 5'-ACATCGAGAACCTGGTCAAC-3').

Strand-specific RNA sequencing. Conidia (~10⁶ ml⁻¹) were inoculated into 30 ml of Vogel's minimal medium with 2% sucrose and incubated at 25 °C for 20 h with shaking. The mycelia were filtered through Whatman paper and washed once with ice-cold phosphate-buffered saline. About 200 mg was put into a 2.0-ml cryo vial with silicon beads and frozen in liquid nitrogen. Total RNA was isolated from frozen mycelia in 1.0 ml of TRIzol using a BioSpec beater. Total RNA (5 µg) was run out on a 1.0% agarose gel to check for degradation. The 3' paired-end RNaseq library was prepared from 10 µg of total RNA as described³⁷. In brief, the poly(A)⁺ fraction was isolated on oligo(dT)₂₅ magnetic beads (Invitrogen) and then treated to chemical fractionation with ZnCl (Ambion) followed by another round of poly(A)⁺ enrichment. First-strand cDNA was generated from an anchored oligo(dT)₂₀ primer followed by second-strand synthesis with DNA polymerase I. The addition of paired-end adaptors and size isolation with ampureXP beads (Ambion) was followed by PCR enrichment as described³⁷. The library has an average insert size of 200 ± 36 bp. The library was sequenced to a depth of 22 × 10⁶ 36-bp paired-end reads on an Illumina GAIIx sequencer.

Nascent nuclear transcripts isolation. Liquid cultures were grown in 1 l of minimal medium (1 × Vogel's, 2% glucose). The dried frozen tissue (~4 g) was ground in liquid N₂, 8 ml of cold buffer A (1 M Sorbitol, 7% Ficoll, 20% glycerol, 5 mM magnesium acetate, 3 mM CaCl₂, 50 mM Tris-Cl pH 7.5, 3 mM DTT, 10 µg ml⁻¹ leupeptin, 10 µg ml⁻¹ pepstatin, 10 µg ml⁻¹ phenylmethylsulphonyl fluoride (PMSF)) was mixed with tissue powder on ice for 5–10 min with stirring. The mixture was filtered through four layers of pre-wetted miracloth (buffer A) into small flask on ice. Buffer A was added to the filtered mixture to a volume of 8 ml. Pre-chilled buffer B (10% glycerol, 5 mM magnesium acetate, 25 mM Tris-Cl pH 7.5, 10 µg ml⁻¹ leupeptin, 10 µg ml⁻¹ pepstatin, 10 µg ml⁻¹ phenylmethylsulphonyl fluoride) (16 ml) was slowly added with stirring. The mixture obtained was gently layered on 10 ml of pre-chilled buffer A/B (2.5:4) in a 50 ml screw-cap tube. The supernatant was collected after centrifuging at 3,000g for 7 min at 4 °C. Supernatant (33 ml) was gently layered on 5 ml of pre-chilled buffer D (1 M sucrose, 10% glycerol, 5 mM magnesium acetate, 25 mM Tris-Cl pH 7.5, 1 mM DTT, 10 µg ml⁻¹ leupeptin, 10 µg ml⁻¹ pepstatin, 10 µg ml⁻¹ PMSF). The pellet (purified nuclei) was collected after centrifugation at 9,400g for 15 min at 4 °C. The pellet was suspended in 1 ml of Trizol. Total RNA was extracted with Trizol in accordance with the manufacturer's protocol. Contaminated DNA was removed with a TURBO DNA-free kit, following the manufacturer's protocol. Treated RNA (50 ng) was used for strand-specific RT-qPCR.

Chromatin immunoprecipitation (ChIP) assay. The ChIP assay was performed as described previously, with several modifications³⁸. First, 500 µg of total cell lysate was used for each immunoprecipitation assay. The antibodies used in this study were WC-2 (ref. 12) (2 µl per assay), Pol II CTD (Abcam no. ab26721, 2 µg per assay), Pol II S2P-CTD (Abcam no. ab5095, 2 µg per assay), Pol II S5P-CTD (Abcam no. ab5131, 2 µg per assay), histone H3 (Abcam no. ab1791, 2 µg per assay) and H3K36me3 (Abcam no. ab9050, 2 µg per assay). Each experiment was performed independently three times, and immunoprecipitation with IgG or *wcc*^{DKO} extract was used as the negative control. qPCR was used for measuring the immunoprecipitated DNA. Primer sequences for qPCR were *frq* dLRE (5'-AGAGTTTGGC CGGACAACCAGTA-3' and 5'-GCTTCACCCGAAAGTATCTTGAGCCT-3'), *frq* pLRE (5'-GTCGCAGAGACCCCTGAACCTT-3' and 5'-TCCACAGATGC ACAGAAATCG-3') and *qrf* qLRE (5'-ATCGACTACTAGTCCCGGTTCTGTT-3' and 5'-TTGCTGATAATGCGCTGAGGGTCT-3'). For the experiments in Fig. 4b and Extended Data Figs 8 and 9 the primer sequences were 1 (5'-AGAGTTTGG CCGGACAACCAGTA-3' and 5'-GCTTCACCCGAAAGTATCTTGAGCCT-3'), 2 (5'-GGGTAGTCGTGACTTTGTGACG-3' and 5'-ACCGACTTTAGGTTG

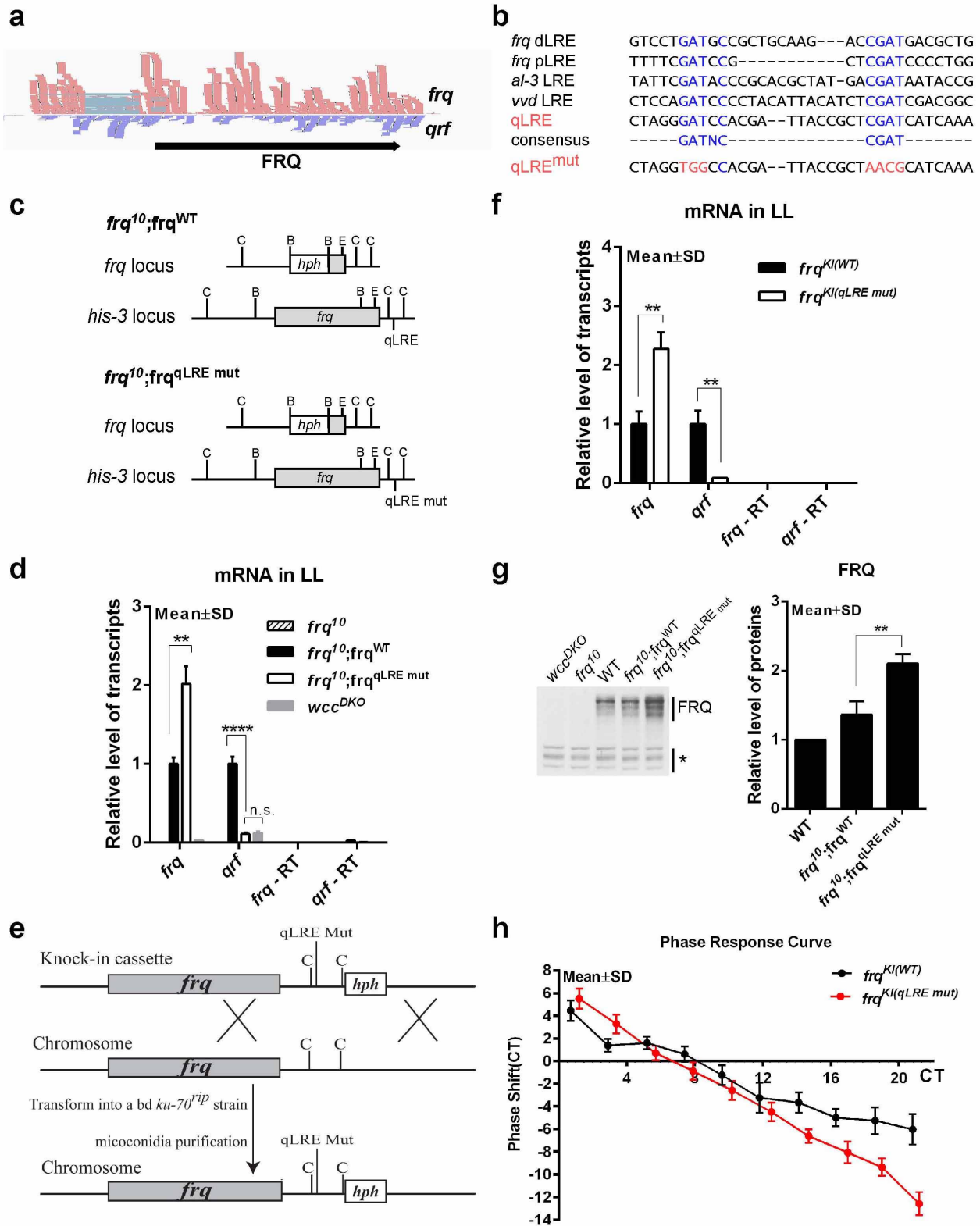
TGTG-3'), 3 (5'-ACGGCCTTTCTCTGTTTACC-3' and 5'-GCAGAGTTGGGTCGGATTTA-3'), 4 (5'-CCGTCTAGCGTGCTTCTT-3' and 5'-GGCACTAATGAGGTTTCGAGATT-3'), 5 (5'-TTACTTCATCTTCCGCACTGG-3' and 5'-GGCAGGGTTACGATTGGATT-3'), 6 (5'-ACTGTTACCGGAAAGATCAG-3' and 5'-GGCCATGTTGGTTCCTT-3') 7 (5'-TTGCACCGATCTTTCAGGAG-3' and 5'-CTGCAGCACATGTTCACTTC-3') and 8 (5'-GGCAGAGAGTGGCTATAGTA-3' and 5'-CCGCTCTTCTCTCACATAG-3'). For the experiments in Extended Data Fig. 7 the primer sequences were 1 (5'-CAAGTGGGCCCTACAGTTATT-3' and 5'-CCAATCCACTTCCCTTTCCA-3'), 2 (5'-CACCGTTCGTCGCTTACT-3' and 5'-GGTATCCGCTCTGTCCCTTTG-3'), 3 (5'-TCAAGATACGGAGCGCAAGAATG-3' and 5'-GATTGTAGCTGTTCCACTCTC-3'), 4 (5'-TGACCGACCCAGTTGATTTC-3' and 5'-GTGGCACTTATCACCTTCTT-3'), 5 (5'-CGACCGTTAGGAGACCAAATAG-3' and 5'-AATGGCTTCTTGTGGTTAGA-3'), 6 (5'-CTTAGTCGGATCAGTGGCTATT-3' and 5'-TTCGTAGATGCCGTGAGATG-3'), 7 (5'-CGTGGACTTCTGCTACTCTT-3' and 5'-CCTTTCAGTCCTCCTCTTT-3') and 8 (5'-GTCTGGGAGCTTCTGTTGAATA-3' and 5'-TAAAGTGGTGAACGACCTCATC-3').

Luciferase reporter assay. The luciferase reporter assay was performed as described previously¹⁶. The reporter construct in Fig. 2c (*Pfrq-luc*) was generated by inserting the BamHI–NotI fragment of *Pfrq-luc-I* (a gift from J. Dunlap) into corresponding sites of pBARKS1 (ref. 16). The reporter construct in Fig. 3b (*Pqrf-luc*) was generated by inserting the PCR fragment containing the *qrf* promoter and the luciferase gene into NotI–EcoRI sites of pBARKS1. For the PCR fragment, the *qrf* promoter was directly fused with the luciferase gene ORF. The primer sequences were *qrf* promoter (5'-ATCGATTTCATTAAGGTGGGG-3' and 5'-GGTGGTGGGTGGTGGGTGGAG) and luciferase gene (5'-ATGGAGGACGCCAAGAATCAAG-3' and 5'-TCAGAGCTTGGACTTGCCGCC-3'). The plasmids *Pfrq-luc* and *Pqrf-luc* were transformed into wild-type or *frq*¹⁰;frq.aq strains with Ignite selection as described previously³⁹. The plasmid *Pmin-luc* was generated by inserting the luciferase gene ORF into the XmaI site of pDE3dBH. The plasmid *Pmin-luc-Pfrq* was generated by inserting the inverted *frq* promoter (the same as *Pfrq-luc-I*) downstream of the luciferase gene ORF of *Pmin-luc*. The non-regulated basal promoter in pDE3dBH (upstream of the EcoRI site) was used as the minimal promoter to control the expression of the luciferase gene in *Pmin-luc* and *Pmin-luc-Pfrq* as described previously¹⁷. For Fig. 3e, the plasmids *Pfrq-luc-I*, *Pmin-luc-Pfrq* and *Pmin-luc* were transformed into the wild-type strain by targeting the *his-3* locus. Under our experimental conditions, luciferase signals are highly variable during the first 1–2 days in the LumiCycle and only become stabilized afterwards, which is probably due to an artefact caused by the transfer of the cultures from light to dark. Thus, the results presented were recorded after 1–2 days in DD.

Mathematical modelling. The mathematical simulation of relative *frq* and *qrf* RNA levels in DD was performed on the basis of a previously described model^{18,40–42}, with several modifications. The equations used here are listed in Extended Data Fig. 4c.

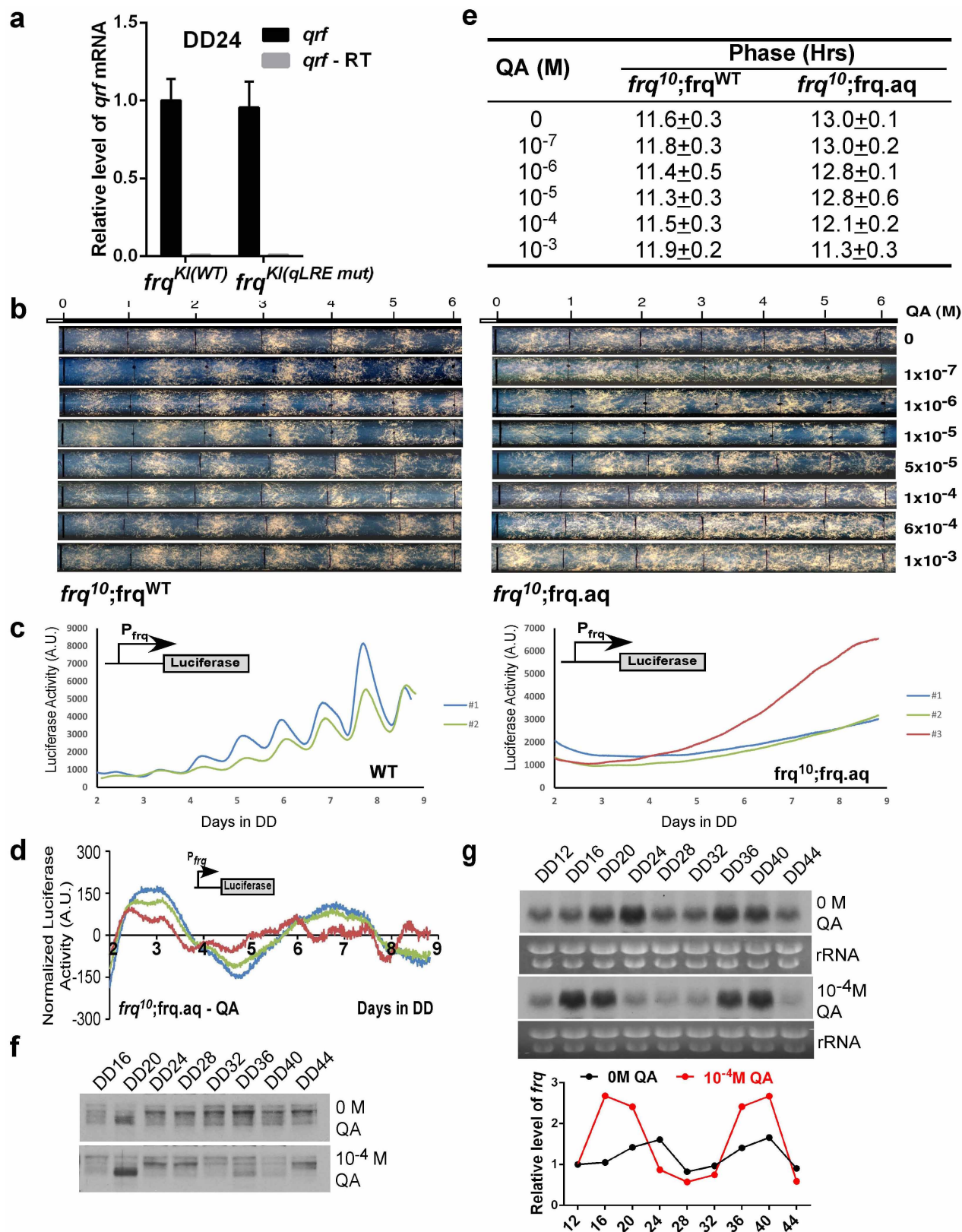
The differential equations (1–8) were solved numerically by the Gear method and XPPAUT^{43,44}. The parameters were as follows. The values of $k_2, k_3, k_6, k_7, k_8, k_9, k_{11}, k_{12}, k_{13}, k_{14}, k_{15}, K, K_2$ and k_{01} were from ref. 18. The values of k_4 and k_5 were from ref. 40, 41. The value of k_{10} was from ref. 42. The values of other parameters were determined in this study. For Fig. 3g, $k_{19} = 0$ (left panel); $k_{17} = 0.35, k_{19} = 0.1$ (middle panel); and $k_{19} = 0.1, k_{17} = 0.35, k_{16} = 0.91$ (right panel).

- Choudhary, S. *et al.* A double-stranded-RNA response program important for RNA interference efficiency. *Mol. Cell. Biol.* **27**, 3995–4005 (2007).
- Zhao, Y. *et al.* Ubiquitin ligase components Cullin4 and DDB1 are essential for DNA methylation in *Neurospora crassa*. *J. Biol. Chem.* **285**, 4355–4365 (2010).
- Bardiya, N. & Shiu, P. K. Cyclosporin A-resistance based gene placement system for *Neurospora crassa*. *Fungal Genet. Biol.* **44**, 307–314 (2007).
- Liu, Y., Garceau, N., Loros, J. J. & Dunlap, J. C. Thermally regulated translational control mediates an aspect of temperature compensation in the *Neurospora* circadian clock. *Cell* **89**, 477–486 (1997).
- Dharmananda, S. *Studies of the Circadian Clock of Neurospora crassa: Light-induced Phase Shifting*. PhD thesis, Univ. California, Santa Cruz (1980).
- Barlow, J. J., Mathias, A. P., Williamson, R. & Gammack, D. B. A simple method for the quantitative isolation of undegraded high molecular weight ribonucleic acid. *Biochem. Biophys. Res. Commun.* **13**, 61–66 (1963).
- Yoon, O. K. & Brem, R. B. Noncanonical transcript forms in yeast and their regulation during environmental stress. *RNA* **16**, 1256–1267 (2010).
- He, Q. & Liu, Y. Molecular mechanism of light responses in *Neurospora*: from light-induced transcription to photoadaptation. *Genes Dev.* **19**, 2888–2899 (2005).
- Pall, M. L. The use of Ignite (Basta; glufosinate; phosphinothricin) to select transformants of bar-containing plasmids in *Neurospora crassa*. *Fungal Genet. Newsl.* **40**, 57 (1993).
- Ruoff, P., Loros, J. J. & Dunlap, J. C. The relationship between FRQ-protein stability and temperature compensation in the *Neurospora* circadian clock. *Proc. Natl Acad. Sci. USA* **102**, 17681–17686 (2005).
- Ruoff, P., Vinsjevik, M., Monnerjahn, C. & Rensing, L. The Goodwin oscillator: on the importance of degradation reactions in the circadian clock. *J. Biol. Rhythms* **14**, 469–479 (1999).
- Yu, Y. *et al.* A genetic network for the clock of *Neurospora crassa*. *Proc. Natl Acad. Sci. USA* **104**, 2809–2814 (2007).
- Gear, C. Simultaneous numerical solution of differential-algebraic equations. *IEEE Trans. Circ. Syst.* **18**, 89–95 (1971).
- Ermentrout, B. XPPAUT. *Scholarpedia* **2**, 1399, <http://www.scholarpedia.org/article/XPPAUT> (2007).



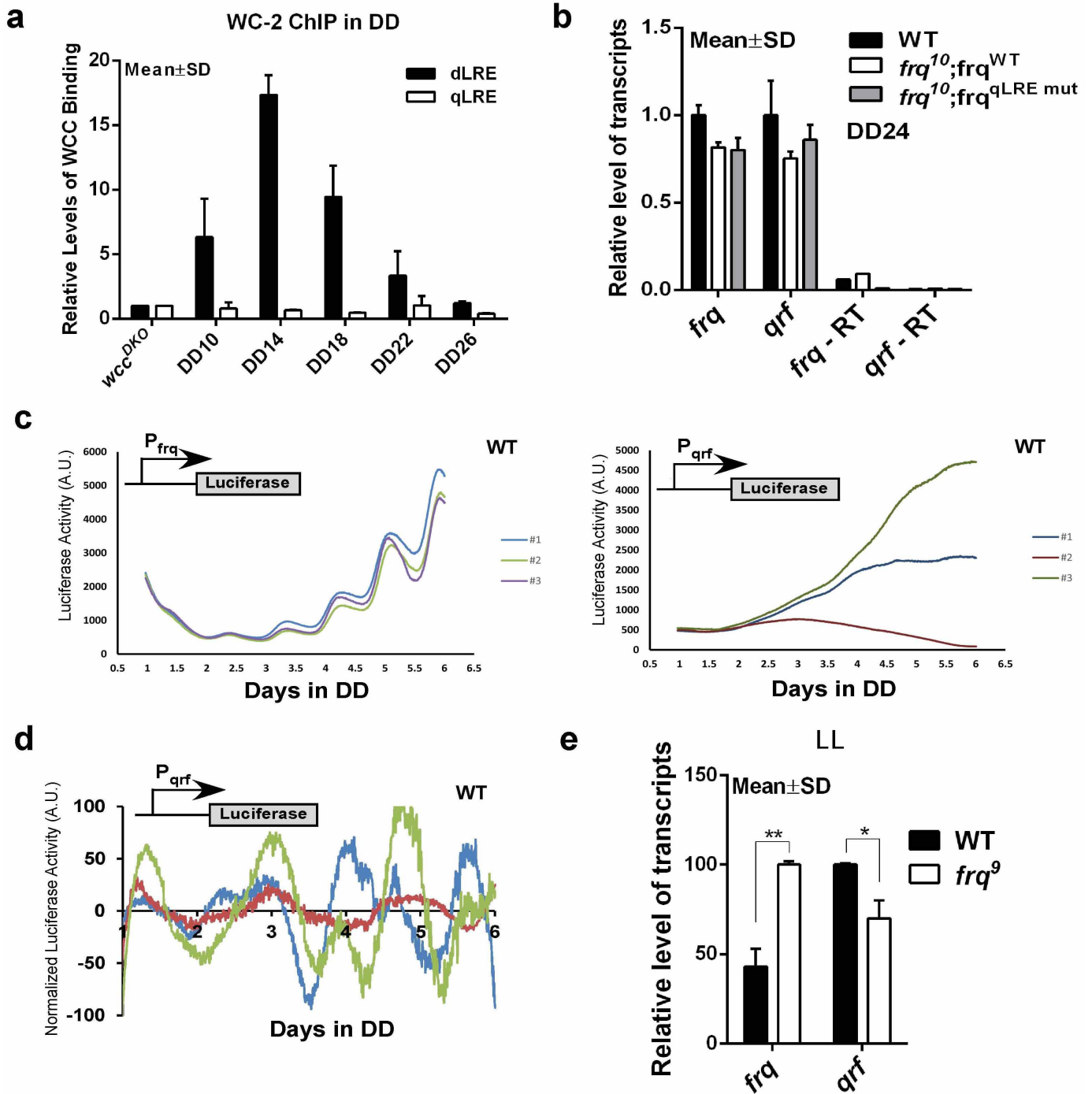
Extended Data Figure 1 | Light-induced *qrf* expression represses *frq* transcription and regulates light resetting of the clock. **a**, Strand-specific RNA-Seq result of the *frq* locus, showing the overlapping *frq* and *qrf* transcripts. **b**, Sequence alignment of known LRE elements. The qLRE and the mutated regions in the *qrf* promoter are shown. **c**, Diagrams showing the chromosomal modifications in the indicated loci in the *frq*¹⁰;*frq*^{WT} and *frq*¹⁰;*frq*^{qLRE mut} strains. C, ClaI; B, BglII; E, EcoRV. **d**, Strand-specific RT-qPCR results showing the expression levels of *frq* and *qrf* in indicated strains in LL. Error bars show standard deviations ($n = 3$). ** $P < 0.01$; n.s. indicates that the difference is not statistically significant. *frq*-RT and *qrf*-RT represent the non-RT reaction

control. **e**, Diagram describing the strategy used to obtain the knock-in strains by homologous recombination. **f**, Strand-specific RT-qPCR results showing the expression levels of *frq* and *qrf* in the indicated knock-in strains in LL. **g**, Western blot results showing the FRQ expression levels in the indicated strains in LL. The densitometric analysis of western blot results from three independent experiments is shown at the right. Error bars show standard deviations. ** $P < 0.01$. **h**, Phase response curves of circadian conidiation rhythms of the indicated knock-in strains after 2 min of light pulse at different circadian time (CT) points.



Extended Data Figure 2 | *qrf* expression is required for circadian rhythmicities. **a**, Strand-specific RT-qPCR results showing the expression levels of *qrf* in the indicated knock-in strains at DD24. **b**, Race tube analyses of the *frq*¹⁰;*frq*^{WT} and *frq*¹⁰;*frq*^{.aq} strains in medium containing 0% glucose and 0.17% arginine with the indicated concentrations of QA in DD. The lack of glucose in medium is known to allow more efficient expression from the *qa-2* promoter. The black lines on race tubes indicate the daily growth fronts. **c**, The unnormalized luciferase activity of the experiments in Fig. 2c. **d**, The

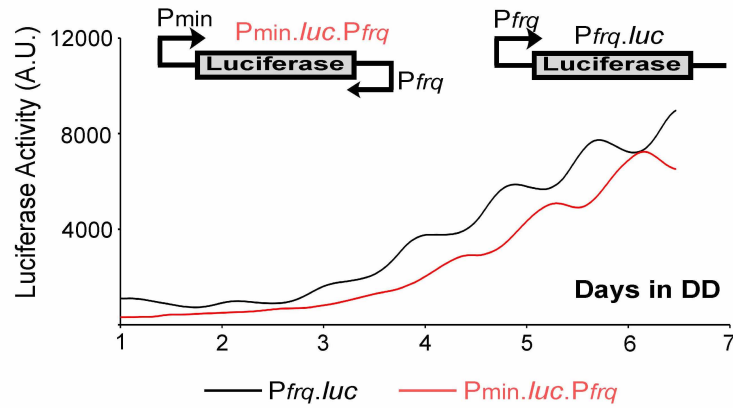
normalized result of Fig. 2c on a $\times 10$ scale, showing that fluctuation of the luciferase activity in the *frq*¹⁰;*frq*^{.aq} strain is random and not rhythmic. **e**, The phases of the first conidiation band in DD of the race tube results shown in Fig. 2b. **f**, Western blot analysis showing the FRQ expression profiles in the *frq*¹⁰;*frq*^{.aq} strain in the presence and absence of QA in DD at the indicated time points. **g**, Northern blot analysis showing *frq* expression profiles in the *frq*¹⁰;*frq*^{.aq} strain in DD at the indicated time points. The densitometric analysis is shown below.



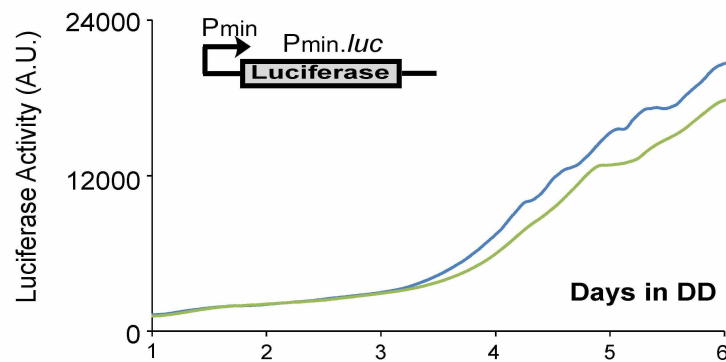
Extended Data Figure 3 | Regulation of *qrf* expression. **a**, WC complex does not bind to the *qrf* promoter in DD. WC-2 ChIP assays showing the relative levels of WC binding at the *frq* (dLRE) and *qrf* (qLRE) promoters at the indicated time points in DD. **b**, Strand-specific RT-qPCR results showing the levels of *frq* and *qrf* at DD24 in the indicated strains. **c**, The unnormalized

luciferase activity of the experiments in Fig. 3b. **d**, The normalized luciferase activity of the *Pqrf-luc* construct in Fig. 3b shown on a $\times 10$ scale. **e**, Densitometric analyses of three independent northern blot results, indicating that levels of *qrf* transcripts are decreased in the *frq*⁹ strains as a result of increased *frq* expression.

a



b



c

$$\frac{d[frqmRNA]}{dt} = \frac{k_1[WC1_n]^2}{K + [WC1_n]^2} - k_4[frqmRNA] - k_{19}[qrfmRNA] + k_{01} \quad (1)$$

$$\frac{d[FRQ_c]}{dt} = k_2[frqmRNA] - (k_3 + k_5)[FRQ_c] \quad (2)$$

$$\frac{d[FRQ_n]}{dt} = k_3[FRQ_c] + k_{14}[FRQ_n \cdot WC1_n] - [FRQ_n](k_6 + k_{13}[WC1_n]) \quad (3)$$

$$\frac{d[wc1mRNA]}{dt} = k_7 - k_{10}[wc1mRNA] \quad (4)$$

$$\frac{d[WC1_c]}{dt} = \frac{k_8[FRQ_c][wc1mRNA]}{K_2 + [FRQ_c]^2} - (k_9 + k_{11})[WC1_c] + k_{02}[wc1mRNA] \quad (5)$$

$$\frac{d[WC1_n]}{dt} = k_9[WC1_c] - [WC1_n](k_{12} + k_{13}[FRQ_n]) + k_{14}[FRQ_n \cdot WC1_n] \quad (6)$$

$$\frac{d[FRQ_n \cdot WC1_n]}{dt} = k_{13}[FRQ_n][WC1_n] - (k_{14} + k_{15})[FRQ_n \cdot WC1_n] \quad (7)$$

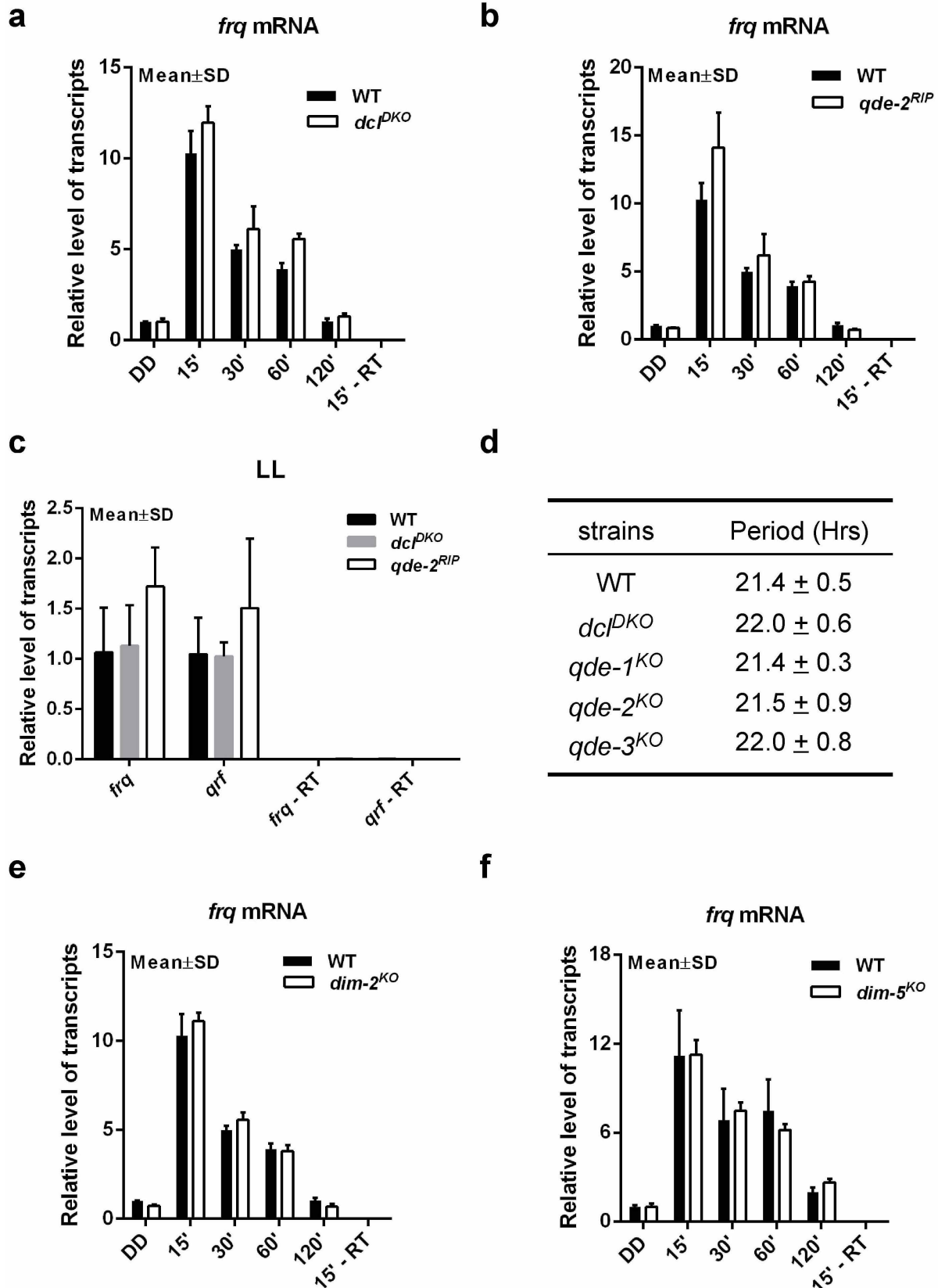
$$\frac{d[qrfmRNA]}{dt} = k_{16} - k_{17}[frqmRNA] - k_{18}[qrfmRNA] \quad (8)$$

Dimensionless value (h^{-1})

k_1	k_2	k_3	k_4	k_5	k_6	k_7	k_8	k_9	k_{10}	k_{11}	k_{12}
1.2	1.8	0.15	0.23	0.27	0.07	0.16	0.8	40	0.1	0.05	0.02
k_{13}	k_{14}	k_{15}	k_{16}	k_{17}	k_{18}	k_{19}	K	K_2	k_{01}	k_{02}	
50	1.0	5	0.5	0.35	0.23	0.1	1.25	1.0	0.64	0.1	

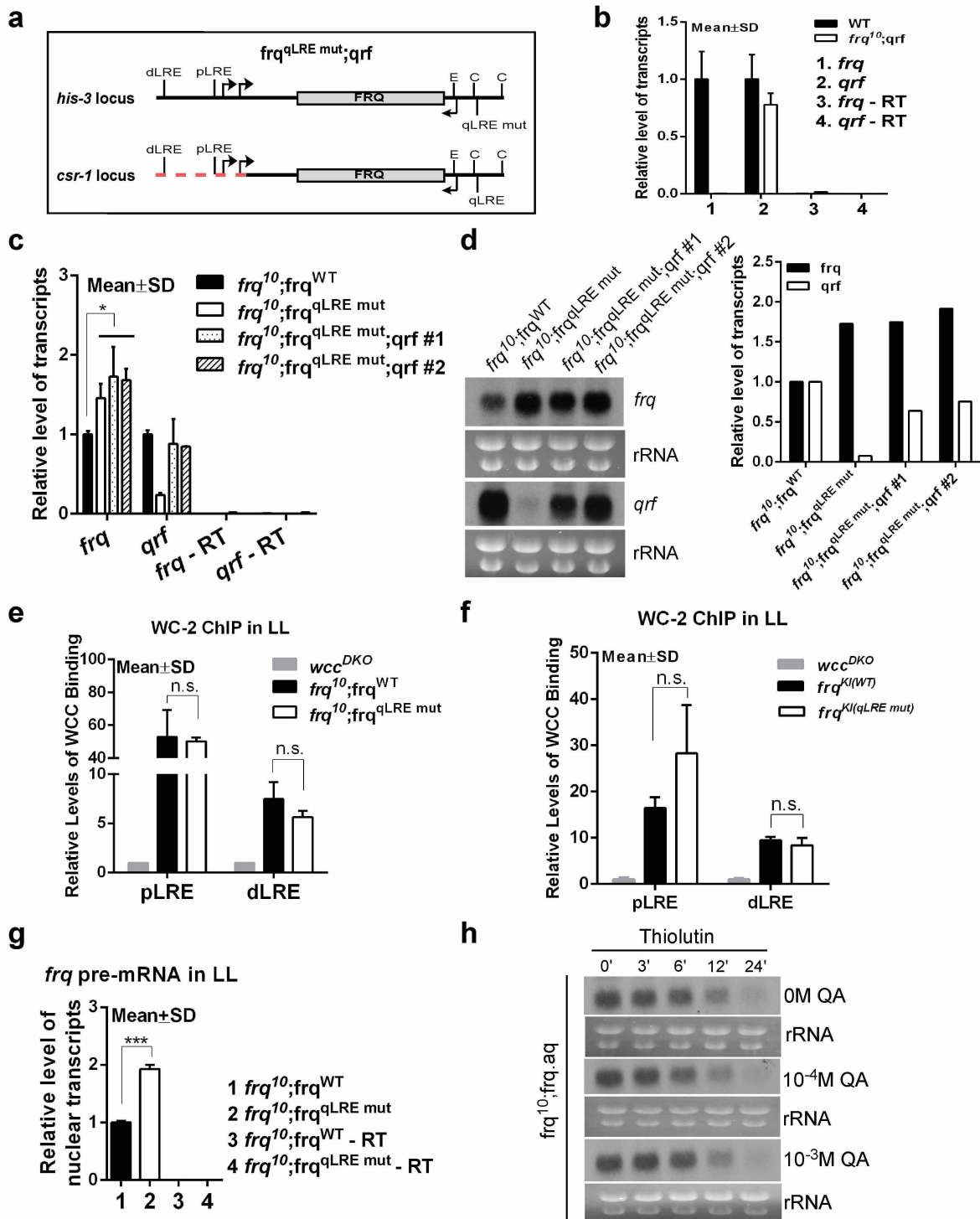
Extended Data Figure 4 | Rhythmic transcription of *frq* drives rhythmic *qrf* expression and mathematical modelling of the *Neurospora* circadian oscillator. **a**, The unnormalized luciferase activity of the experiments in Fig. 3e. **b**, The unnormalized luciferase activity of the wild-type strain carrying the *Pmin-luc* construct. Results of two independent transformants are shown. **c**, Mathematical modelling of the *Neurospora* circadian oscillator with the

double negative feedback loop. The differential equations used in the model are shown. The model is identical to a previously developed model²³ with the exception of equation (1), which in this case includes the inhibition of *frq* transcription by *qrf*, and equation (8), which includes the inhibition of *qrf* transcription by *frq*. The rate constants used in the simulations are listed below.



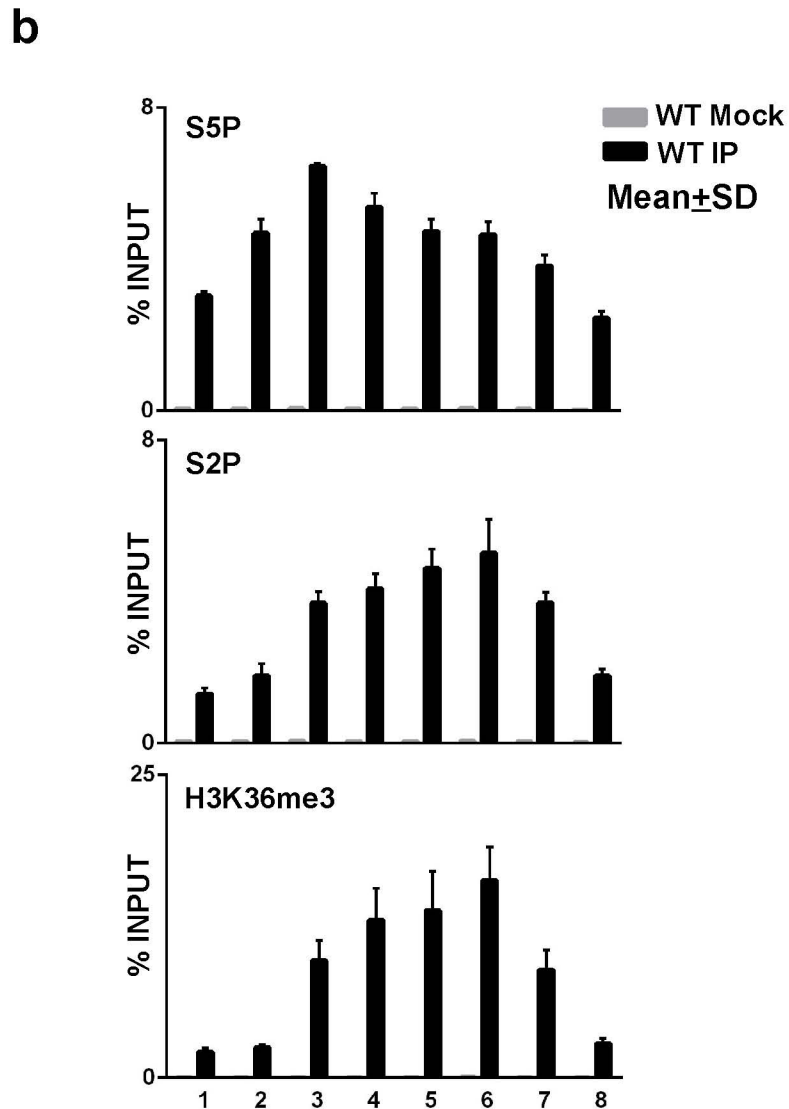
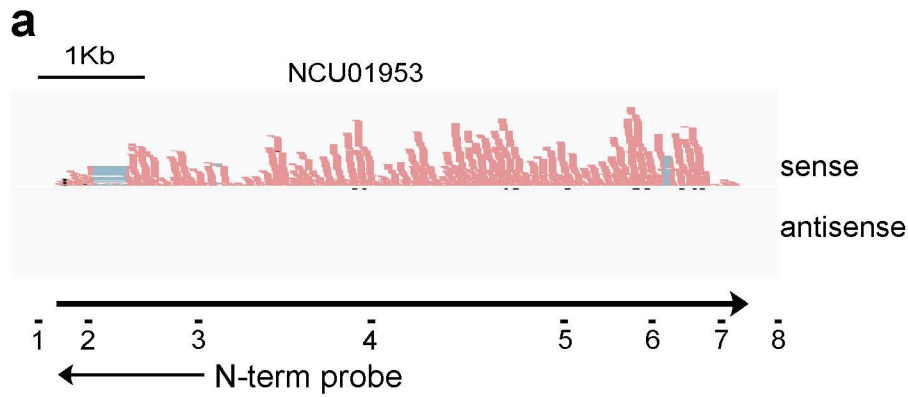
Extended Data Figure 5 | Neither RNAi nor DNA methylation pathways have a significant role in the clock. a, b, Strand-specific RT-qPCR results showing the induction of *frq* after a light pulse are similar in the indicated strains at DD24. b, Strand-specific RT-qPCR results showing the induction of *frq* after a light pulse is similar in the indicated strains at DD24. c, Strand-specific RT-qPCR results showing similar expression levels of *frq* and *qrf* in the

indicated strains in LL. d, A table showing the period lengths of the conidiation rhythms in the wild type and in different RNAi mutants. e, Strand-specific RT-qPCR results showing that the induction of *frq* after a light pulse is similar in the indicated strains at DD24. f, Strand-specific RT-qPCR results showing that the induction of *frq* after a light pulse is similar in the indicated strains at DD24. Error bars show standard deviations ($n = 3$).



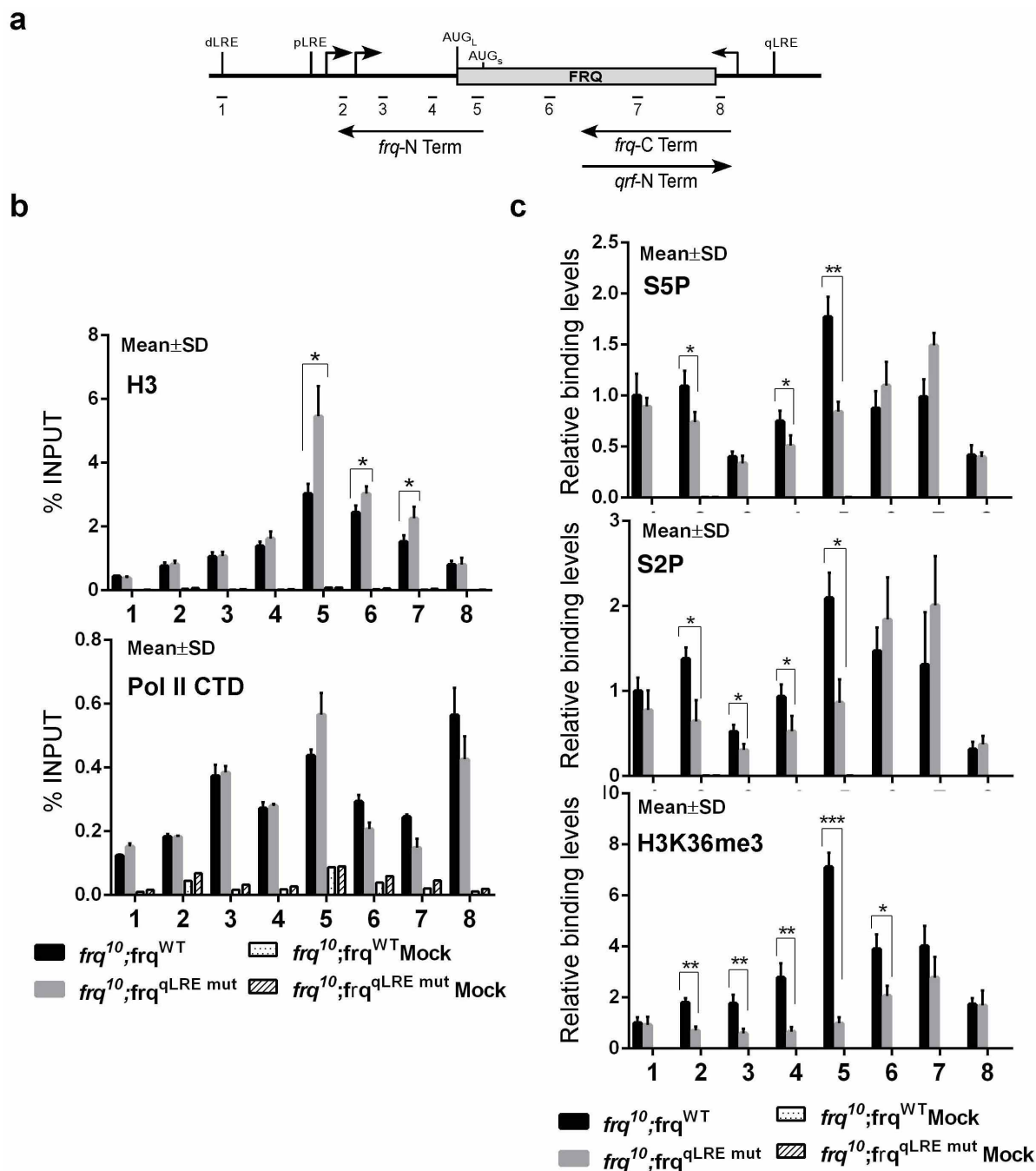
Extended Data Figure 6 | Characterization of the *qrf* action on *frq* expression. **a**, Diagram showing the chromosomal modifications in the *frq*^{qLREmut};*qrf* strains that allow the expression of *qrf* in trans. The *frq*^{qLREmut} construct is at the *his-3* locus, and *qrf* is expressed only from the *csr-1* locus. The red dashed line indicates that the *frq* promoter region is deleted in the *qrf* construct to abolish *frq* expression. **b**, Strand-specific RT-qPCR results showing that only *qrf* is expressed from the *qrf* construct in the *frq*¹⁰;*qrf* strain. **c**, *qrf* expression does not repress *frq* transcription in trans. Strand-specific RT-qPCR results showing the levels of *frq* and *qrf* transcripts in the indicated strains in LL. Error bars show standard deviations. **P* < 0.05 (*n* = 3). **d**, Northern blot results showing that expression of *qrf* in trans in the *frq*¹⁰;*frq*^{qLREmut};*qrf* strains does not repress *frq* expression. Densitometric analysis of the northern blot results is shown at the right. **e**, WC-2 ChIP assays showing the relative WC binding levels at the *frq* promoter in LL in the indicated strains. The *wc*

double mutant (*wcc*^{DKO}) was used as a negative control for ChIP. n.s. indicates a lack of statistical significance (*n* = 3). **f**, WC-2 ChIP assays showing the relative WC-2 binding levels at the *frq* promoter in LL in the indicated strains. The *wc* double mutant (*wcc*^{DKO}) was used as a negative control for the ChIP assays. n.s. indicates a lack of statistical significance (*n* = 3). **g**, Strand-specific RT-qPCR results showing that the mutation of the qLRE element in the *qrf* promoter results in significant increases in light-induced *frq* pre-mRNA expression. Nascent nuclear RNA was used. **h**, Northern blot results showing that the stability of *frq* mRNA is not affected by the transcription of *qrf*. The *frq*¹⁰;*frq*.aq strain that can induce *qrf* expression in the presence of QA was used. Thiolutin, a transcription inhibitor³⁰, was added in the culture to block *frq* transcription so that *frq* mRNA stability could be determined. Cultures were harvested at the indicated time points after the addition of thiolutin.



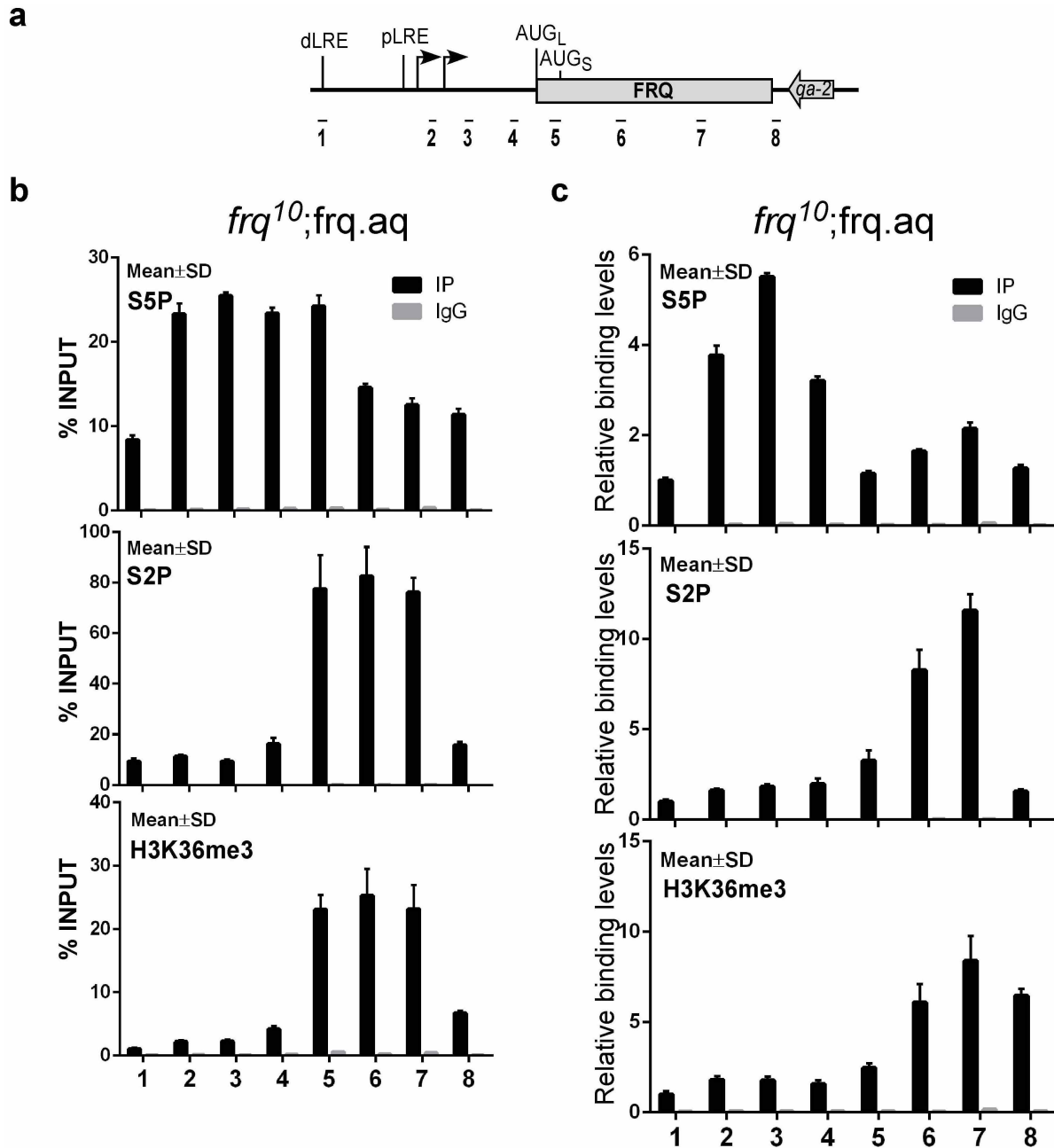
Extended Data Figure 7 | Modified Pol II CTD distribution profiles of a locus lacking antisense transcripts. a, Top panel: strand-specific RNA-Seq results of the NCU01953 locus. Bottom panel: the positions of primers used in ChIP assays and the position of the riboprobe used in northern blot analysis.

b, ChIP assays showing the relative enrichment of Pol II Ser 5 phosphorylation, Pol II Ser 2 phosphorylation, and H3K36me3 at the NCU01953 locus in the wild-type strain. The scale on the y axis is the enrichment percentage of immunoprecipitation (IP) over input. IgG was used as the mock control for IP.



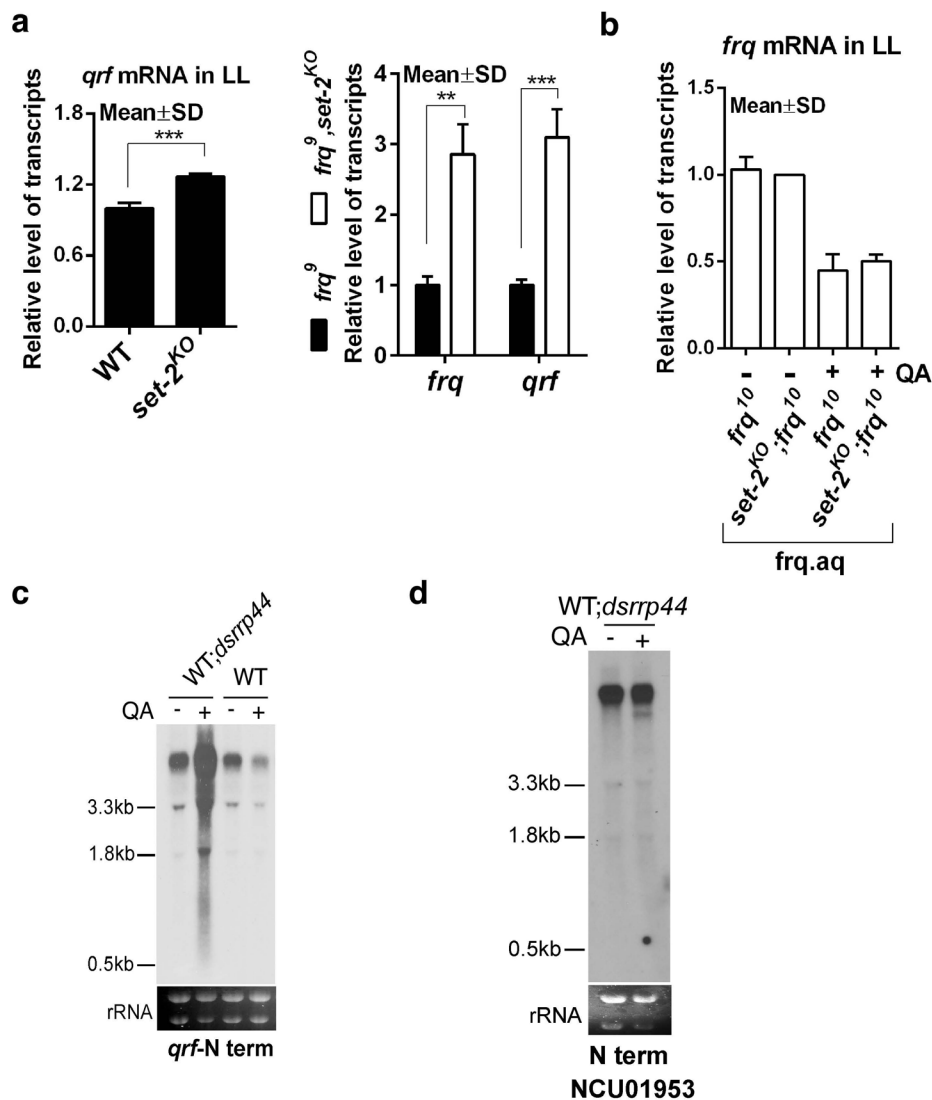
Extended Data Figure 8 | Modified Pol II CTD distribution profiles at the *frq* locus. **a**, Diagram showing the positions of primers at the *frq* locus used for ChIP assays and the positions of riboprobes used for northern blot analyses. **b**, ChIP assays showing the relative enrichment of histone H3 and Pol II CTD in the *frq* locus. The scale on the y axis is the enrichment percentage of immunoprecipitation (IP) over input. IgG was used as the mock control

for IP. **c**, ChIP assays showing the relative enrichment of Pol II Ser 5 phosphorylation and Pol II Ser 2 phosphorylation (ChIP data in Fig. 4b were normalized by Pol II CTD ChIP results in b), and H3K36me3 (ChIP data in Fig. 4b were normalized by histone H3 ChIP results in b). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ ($n = 3$).



Extended Data Figure 9 | Complete abolition of *qrf* expression changes the distributions of modified pol II CTD to a normal profile. **a**, Diagram showing the positions of primers used for ChIP assays. **b**, ChIP assays showing the relative enrichment of Pol II Ser 5 phosphorylation, Pol II Ser 2 phosphorylation, and H3K36me3 in the *frq* locus in the *frq¹⁰;frq.aq* strain in

medium with 2% glucose. Under these conditions, *qrf* expression is completely abolished (shown in Fig. 2a). **c**, ChIP assays showing the relative enrichment of Pol II Ser 5 phosphorylation and Pol II Ser 2 phosphorylation (ChIP data in **b** were normalized by Pol II CTD ChIP results), and H3K36me3 (ChIP data in **b** were normalized by histone H3 ChIP results).



Extended Data Figure 10 | Mechanism of *qrf* action on *frq* expression.

a, Densitometric analyses of three independent experiments shown in Fig. 4c. **b**, Densitometric analyses of independent experiments shown in Fig. 4d. **c**, Northern blot analysis using a N-term probe (marked in Extended Data

Fig. 8a) specific for the 5' half of *qrf* transcripts in the WT and WT;*dsrrp44* strains in LL. **d**, Northern blot analysis using a N-term probe (marked in Extended Data Fig. 7a) specific for the 5' half of NCU01953 transcripts in the WT;*dsrrp44* strain in LL.

CORRIGENDUM

doi:10.1038/nature13840

Corrigendum: Immune clearance of highly pathogenic SIV infection

Scott G. Hansen, Michael Piatak Jr, Abigail B. Ventura, Colette M. Hughes, Roxanne M. Gilbride, Julia C. Ford, Kelli Oswald, Rebecca Shoemaker, Yuan Li, Matthew S. Lewis, Awbrey N. Gilliam, Guangwu Xu, Nathan Whizin, Benjamin J. Burwitz, Shannon L. Planer, John M. Turner, Alfred W. Legasse, Michael K. Axthelm, Jay A. Nelson, Klaus Früh, Jonah B. Sacha, Jacob D. Estes, Brandon F. Keele, Paul T. Edlefsen, Jeffrey D. Lifson & Louis J. Picker

Nature **502**, 100–104 (2013); doi:10.1038/nature12519

The Acknowledgements section of this Letter should have included the following sentence: “We acknowledge the contribution of M. A. Jarvis to the design, construction and initial *in vitro* characterization of all the strain 68.1-derived RhCMV vectors used in this study, including both previously published RhCMV/SIV vectors and a previously unpublished vector expressing an *Mycobacterium tuberculosis* Ag85B–ESAT6 fusion protein, used as a negative control in this study.” S.L.P. and J.M.T. did not reply when asked to approve the wording of this Corrigendum. The HTML version of the paper has been corrected.

CAREERS

@NATUREJOBS Follow us on Twitter for the latest on jobs and careers go.nature.com/e492gf

NATUREJOBS BLOG The latest on careers news and tips blogs.nature.com/naturejobs

NATUREJOBS For the latest career listings and advice www.naturejobs.com



DEVELOPMENT HUBS

Rust no more

Ohio's investments into research are ushering the state from a manufacturing past into a knowledge-economy future.

BY CAMERON WALKER

For generations, Ohio's economy has been centred around industrial manufacturing and agriculture, thanks in large part to a location on the western side of the historically factory-rich region of the United States

and the eastern reaches of the vast fields of the corn belt. But in the past decade, this Midwestern state has begun to transform itself into a manufacturing hub of a different kind. A US\$2-billion infusion of state funding to develop the state's technology economy, along with a cluster of research institutions and a steady

proliferation of start-ups, have resulted in an outpouring of research and tangible goods in an array of sectors from pharmaceuticals and agriculture to medical testing and devices.

Much of the state's research, development and technology transfer take place in its largest cities — especially Cleveland, Columbus (the capital) and Cincinnati. That is largely because they contain the most institutions, including Ohio State University in Columbus and the Cleveland Clinic and Case Western Reserve University in Cleveland. Researchers there and elsewhere in the state attracted \$685 million in biomedical funding from the US National Institutes of Health (NIH) last year, putting Ohio among the ten states that receive the most such funding annually. In 2012, a combination of NIH funding, further federal grants, industry funding and private money gave Ohio's researchers a grand total of \$1.3 billion to spend on bioscience, according to Battelle, a non-profit research organization in Columbus, and the Biotechnology Industry Organization in Washington DC.

UP, UP, UP

All this investment has been producing jobs in both academic and industrial bioscience research, even during the recent recession. Between 2007 and 2012, biosciences employment in the state's private sector rose by 2.2% to a total of 48,294 jobs, while the national rate remained essentially flat, says Battelle's Ryan Helwig. In several sectors, including pharmaceuticals; agricultural feedstock and chemicals; and medical devices, which alone involves 4,523 jobs and just in the Cleveland area, employment grew by double digits during that period, outpacing national hiring in these sectors. Although these jobs include sales and manufacturing positions, companies such as Neuros Medical near Cleveland are also seeking research scientists. In 2015, the company plans to recruit biomedical engineers, clinical managers and field clinical specialists, among others, as it continues pivotal clinical trials on an implantable nerve-blocking device to treat pain for people who have had amputations.

Much of the growth in medical devices and other private-sector ventures stems from research initiated at the state's universities and institutions and from researchers who have followed an entrepreneurial path. One of these is cardiologist Marc Penn, founder and chief medical officer of three biomedical companies, including Cleveland HeartLab, which develops and performs diagnostic tests to assess ►

STATE OUTLINE: FILIP BJORKMAN/SHUTTERSTOCK; HELIX FACTORIES: LUCIANO LOZANO/KONIMAGES/CORBIS

► cardiovascular-disease risk. Although many institutions are going through a steep learning curve when it comes to technology transfer, he says, “the opportunities outweigh the risks” for researchers interested in working in Ohio.

Perhaps the state’s biggest system for converting its bioscience research into goods, services and businesses is a \$2.1-billion taxpayer-supported initiative called Third Frontier, based in Columbus. Launched in 2002, it aims to support the development of technology-based products, companies, industries and jobs; one programme designed to commercialize biomedical research received more than \$151 million in state funds in 2013.

Between 2009 and 2012, Third Frontier invested in 1,080 companies, including 205 in agricultural bioscience and medical technologies. Its investments were directly responsible for creating or retaining 7,780 jobs during this period.

Two of the most recent recipients are in turn funding bioscience research in Ohio and nationally. In June, Third Frontier awarded the Harrington Discovery Institute at University Hospitals in Cleveland — the non-profit arm of The Harrington Project for Discovery & Development — \$25 million to advance its efforts to commercialize drug discoveries. As well as giving physician-scientists in the United States a base amount of \$100,000 over two years, the institute connects them with entrepreneurial and pharmaceutical development teams to help to shepherd their findings to market. Some 31 projects are en route to becoming drugs, with a biomedical start-up around each. The Harrington Project plans to support at least 50 such projects at any given time, says Baiju Shah, chief executive of BioMotiv, The Harrington Project’s for-profit arm. Shah says that BioMotiv may be seeking project managers, particularly those who have experience in managing pharmaceutical development, to lead those fledgling businesses.

Third Frontier has also supported Ohio’s growing neurotechnology field. In June, it awarded \$21 million over four years to the Neurotechnology Innovations Translator, a for-profit partnership between Ohio State and other research institutions, health-care companies and venture-capital firms among others that aims to create and grow neuroscience technology businesses and will begin funding projects in early 2015.

“There are a lot of opportunities with the companies that we’re developing to attract scientists and engineers,” says Ali Rezai, a neurosurgeon at Ohio State and principal investigator for the translator. He estimates that the companies that emerge from the project will

“There are a lot of opportunities with the companies that we’re developing to attract scientists and engineers.”



At the Global Center for Health Innovation, researchers can showcase their innovations and technologies.

generate at least 160 jobs, including positions for researchers and engineers.

Ohio State is also expanding its faculty as part of an initiative that seeks to produce collaborative research, teaching and outreach in theme areas of food production and security, energy and the environment, and health. In the next 10 years, it intends to increase the number of tenured and tenure-track faculty members by 8–10%, or about 300, with a particular focus on bioscience disciplines related to the initiative’s themes.

At present, the university has set aside \$60 million for tenure-track positions in data analytics, emerging and re-emerging diseases, sustainable materials and food to improve health, with the first recruits expected on campus by autumn 2015.

In June, an effort was launched to bring Ohio’s research institutions into a network that will help researchers to develop and run clinical trials in areas such as neuro-oncology, neurosciences, paediatrics and infectious diseases. Organizers of the Cleveland-based Ohio Clinical Trials Collaborative, a multi-institution partnership that has received two years of state funding, intend for it to become self-sustaining by 2016.

It has already made many national and international industry contacts and has close to 30 trial opportunities in development, says John Peterson, the project’s global business-development director. As more industry-sponsored trials come to the state, the collaborative and participating hospitals will need more research and support staff, he says.

ON DISPLAY

Several large multinational and national companies have headquarters or outposts in Ohio, from Procter & Gamble to Battelle. Last autumn, Cleveland opened the 2-hectare Global Center for Health Innovation adjacent

to the new Cleveland Convention Center. The centre has four floors of permanent showrooms that companies and institutions — including Philips and University Hospitals — are occupying. “That’s another symbolic commitment of the state and the region” to bioscience, says Aram Nerpouni, president and chief executive of BioEnterprise, a Cleveland-based business formation, recruitment and acceleration programme.

From 2002 to this year, Cleveland’s biomedical cluster grew from 300 companies to more than 700. Around the state, business-support organizations such as the Youngstown Business Incubator and TechColumbus — as well as bioscience-specific efforts, including the Global Cardiovascular Innovation Center in Cleveland — help entrepreneurs to launch start-ups.

Before Goutham Narla, a medical geneticist at Case Western’s School of Medicine, came to Cleveland from Mount Sinai Hospital in New York City to explore a potential job, the state of Ohio had never been on his radar. But now he and his wife, cancer biologist Analisa DiFeo, love it. Along with its world-class researchers, the region has an ideal combination of ambition and amiability. It is not so laid-back that nothing gets done, but also “not so cut-throat that you’re wondering if you can trust anyone, even your colleagues”, Narla says.

The five graduate students who came with Narla from Mount Sinai have enjoyed the region’s lower cost of living and easier entrée to top-shelf research institutions and researchers. Early-career scientists often find it difficult to tap into the high-level networks based on the east and west coasts, Narla says, but that is not the case in Ohio. “That’s a neat thing here — you have a lot of access, because people are more welcoming.” ■

Cameron Walker is a science writer in Santa Barbara, California.

DUMPSTER DIVING

The silent treatment.

BY ALVARO ZINOS-AMARO

Daniel Weatherby, 45, wealthy and fast dying from complications of a malignant neoplasm in the pharynx — throat cancer — was cut off from any living relatives, and unattached. But the reality of his aloneness and imminent death didn't particularly bother him. He was going to be cryogenically preserved.

"Would you like another pillow?" the nurse asked, and waited for him to gesture with his head.

Daniel had lost the use of his throat weeks ago. A man who prized silence above all else, he now wondered whether in a way the disease — if it weren't for the dying part — might not be a blessing in disguise, as he was no longer expected to talk. But lately he spent a lot of time drifting in and out of consciousness; his hands often turned blue; and his mouth was continually dry, his lips cracked. All this he could do without. He nodded and the nurse propped him up on a second foam cushion. She used glycerin swabs and lip balm to soothe his pain.

"Anything else I can do for you?"

Daniel ignored her. He had realized soon after being admitted to the hospital that one way of dismissing a caregiver was simply to stop acknowledging them.

A few moments later, when he peed himself and died, Daniel was pleased that at least he did so quietly.

He was dreaming that he was back on the sprawling Wisconsin farm of his youth. Then his eyes adjusted to the brightness of a small grey room, without windows or doors.

The first thing he felt was a complete absence of pain. He cleared his throat, which seemed to be working fine. He had been healed. Tears welled in his eyes.

He blinked. Two tall, identically dressed figures appeared by his bedside.

"What is this place?" Daniel asked.

The man and woman, eyes closed, did not reply.

Daniel tried to get out of bed, but was so weak he almost passed out.

"What year is this?"

The woman inclined her head in Daniel's direction, and then the man followed suit. Their eyes remained closed.

"Hey, you two, I'm right here! What's going on?"

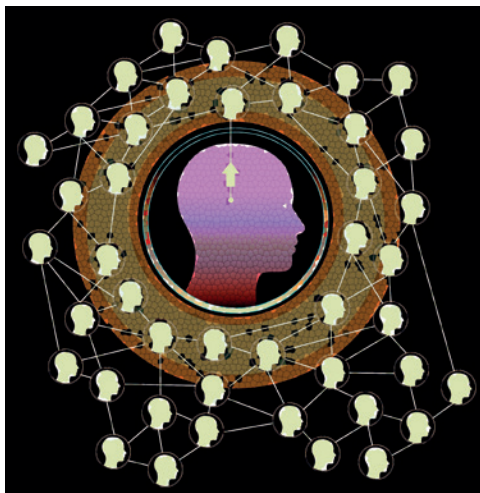
The man smiled first, then the woman.

The year, by your calendar, is 2189, an inhumanly clear voice said in Daniel's head.

"Please talk to me the old-fashioned way." *The old ways are gone*, the voice said. *They were problematic and had to go.*

Daniel's head began to throb. This was more than dizziness.

We are very pleased about the condition of your mind, the voice said. *Spacious, unadulterated.*



Daniel gritted his teeth. "Stop doing whatever you're doing. *It's hurting me.*"

The examination is now over, the voice said.

His head returned to normal. Or did it? Daniel blinked. Now there were eight individuals, four men and four women. "Look, I really appreciate what you've done for me," he mumbled. "Bringing me back to life and all. But I need some fresh air, assuming it still exists in 2189. Can you show me outside?"

This time he managed to stand up, naked and all.

We've done more than that, Daniel. We've made you immortal.

A vision flashed through Daniel's mind. Idyllic parks, bountiful sunshine. Pristine mountains, lakes, oceans, forests, rivers. Extinct species restored. A thriving planetary ecosystem. Everyone smiling and in perfect health. No disease, no crime. No poverty in the world, no poverty of the mind. Everyone happy, all the time. Everyone generous and kind —

And of a kind.

"You're all ... the same?"

One voice, said the voice.

Daniel bolted from the bedside, intending to knock over two of

his captors on his way to the farthest wall, which he planned to bang on with his bare arms and fists until someone let him out. Instead, he passed right through the figures and tumbled to the floor.

We have no wish to assimilate your consciousness into ours, said the voice. *We prize that about you, that you have not heard our voice before. For we are always in need of neuronal space, and a single unconnected mind like yours would be of great value to our network.*

Daniel struggled into an upright position. "What the hell do you want with my mind?"

The eight figures opened their eyes, revealing no pupils, just whites. Daniel preferred them the way they'd been before. "There were no buildings in what you showed me," he said. "We're underground, aren't we?"

A perfect world, he realized, in which all the machinery had been hidden away. And not just the machinery — but everything that was unwanted.

The refuse.

In order to maintain our calm, the voice said, *we must continuously shed that part of ourselves that would lead us back to the old ways. We must shed our jealousy, greed, violence and lust. Our pain and suffering.*

The figures disappeared, until only one remained.

We sift through the biological scraps that survive from the past and occasionally get lucky. As with you. Perfectly preserved. Functional. Your mind will allow us to store the collected anguish of two billion people — a full tenth of the planet's population — for years. You will perform a great service for humanity, Daniel Weatherby. Because of this, and because you will perform it forever, we wished to exercise a minimum of compassion and inform you of your fate. Thank you for your contribution.

"I don't want to —"

Daniel's voice choked off. Everyone was gone.

He was alone in the room.

But not alone. Endless horrors unspooled inside his mind.

Tears fell from Daniel's eyes as he realized he would never again speak. ■

Alvaro Zinos-Amaro is co-author, with Robert Silverberg, of *When the Blue Shift Comes* (March 2014). His short fiction has previously appeared in *Nature*, *Analog* and *Galaxy's Edge*.

JACEY