THIS WEEK

The power of treaties

International weapons conventions may not be perfect, but they are a vital mechanism for making wars less barbaric and less frequent — a cause that should galvanize scientists and others.

The apparent use of sarin to kill more than 1,400 civilians in Damascus on 21 August may highlight the limits of the Chemical Weapons Convention (CWC) — but that does not mean that the world can afford to turn its back on such treaties.

The role of international treaties in restricting the proliferation of nuclear, chemical and biological weapons has not had a good press in recent years. Conventional wisdom tends to scorn the value of such 'pieces of paper' in real politick. Critics from both the left and the right heap derision on their selective reach and implementation.

Yet these treaties are crucial to everyone who is interested in making wars less barbaric and less frequent. Pieces of paper they may be, but large powers adhere to their contents with care, as do the smaller ones who crave international respectability.

As Hans Blix, the former chief United Nations weapons inspector, has pointed out, even regimes that are regarded as political outliers are highly sensitive to treaty adherence. That is why, for example, North Korea withdrew from the Treaty on the Non-Proliferation of Nuclear Weapons (NPT) in 2003, before resuming its nuclear-weapons programme. It is also why Syria is one of just seven countries (including Egypt and Israel) not to join the CWC.

These treaties also matter to scientists, or ought to, because only scientists have the technical expertise and institutional basis to devise their content and implementation. That is not true of diplomats or soldiers, or of other parties, such as the unfortunate victims of chemical weapons in Syria. As the US physicist Herbert York said of his peers' role in the development of the first, limited nuclear test-ban treaty half a century ago: if not us, who?

The CWC has its origins in the 1925 Geneva Protocol, which held the use of chemical weapons to be illegal. Since it came into effect in 1997, the CWC has been a considerable success, with its secretariat supervising the destruction of some 78% of known chemical-weapons stocks in signatory nations — a figure that is expected to reach 99% by 2017.

Three questions will remain after that milestone is reached: what to do about non-signatories; how to deal with non-state actors; and how to extend the CWC's progress to the problematic field of biological weapons.

The position of non-signatories, at least, should become less tenable. As this group diminishes in number, precedent suggests that the terms of the treaty are likely to become accepted as international law, opening violators up to possible criminal prosecution. It is not unreasonable to hope that pressure on non-signatories will eventually bear fruit.

Non-state actors — terrorists among them — have never really fallen under the remit of treaties. However, the CWC does include provisions, thus far never invoked, for 'challenge inspections' to be conducted when weapons violations are suspected. It is conceivable that such inspections could be used if non-state actors were suspected of stockpiling chemical agents.

The last question regards the relationship between the CWC and its older but weaker cousin, the Biological Weapons Convention (BWC). The CWC office has several hundred staff in The Hague in the Netherlands, but the BWC has only a tiny secretariat in Geneva, Switzerland — and no verification regime. That is partly on account of long-standing US resistance to the idea: because biological agents grow or die, they are hard to inventory, and sceptics contend that a verification regime would result in an orgy of commercial larceny.

"International treaties ... are more true to the cause of peace than a fusillade of cruise missiles."

One suggestion is to merge the two (see L. K. Sydnes *Nature* **496,** 25–26; 2013) and adapt the CWC's powerful verification infrastructure to tackle the BWC's mission. But diplomats are sceptical that two international treaties could ever be successfully merged. What is clear is that biologists — whose international representation is more fragmented than that of chemists or physicists — could work harder towards the augmentation of the BWC, and the eventual development of a verification regime.

In the meantime, large political powers need to be less selective in their pursuit of disarmament treaties, and more forthcoming in providing the resources necessary for their implementation. It is hard for Britain and the United States to strengthen the CWC, for example, while they continue to drag their feet in implementing their existing nuclear obligations under the NPT.

International treaties, in the end, will never be entirely fair, or equitable, or implemented consistently. They are nonetheless more impressive than the barrage of platitudes that passes for political discourse on international security — and more true to the cause of peace than a fusillade of cruise missiles. The BWC, CWC and NPT are all imperfect but they are the instruments that we have in our hands. They can each play a part in making war less likely, as well as less ghastly. ■

Nuclear error

Japan should bring in international help to study and mitigate the Fukushima crisis.

The radioactive water leaking from the site of the wrecked Fukushima Daiichi nuclear power plant in Japan is a stern reminder that we have not seen the end of the world's largest nuclear crisis since the Chernobyl meltdown in Ukraine in 1986. After an earthquake and tsunami crippled the Fukushima plant in March 2011, it became clear that efforts to decontaminate the area would be long-lasting, technically challenging and vastly expensive. Now it turns out that the task has been too big for the owner of the plant, the

Tokyo Electric Power Company (TEPCO). The Japanese government on 3 September announced a plan to take over the clean-up, but its intervention is overdue.

In the two and a half years since the accident, TEPCO has repeatedly failed to acknowledge the nature and seriousness of problems with safeguarding nuclear fuels in the three destroyed reactors at Fukushima. Each day, some 400,000 litres of water are being funnelled into the reactor cores to prevent the rods from overheating. Only in recent months has TEPCO admitted that some contaminated water is leaking into the reactor basement and, through cracks in the concrete, into the groundwater and the adjacent sea. Few independent measurements of radiation exposure are available, and it is worryingly unclear how these leaks might affect human health, the environment and food safety. But the problems do not stop there. There are now almost 1,000 storage tanks holding the used cooling water, which, despite treatment at a purification system, contains tritium and other harmful radionuclides. The leaks make clear that this system is a laxly guarded time bomb.

It is no secret that pipes and storage tanks sealed with rubber seams have a habit of leaking. TEPCO's reliance on routine patrols to detect any leaks has been careless, if not irresponsible. That the company, in response to the latest incidents, intends to refit the tanks with sensors and extra safety controls just underlines the makeshift way in which the storage facilities were set up in the first place. Meanwhile, the fate of the constantly amassing polluted water is undecided. Proposals earlier this year to dump it into the sea understandably met with fierce opposition from local fisheries.

Given the government's past actions and information policies, one might doubt whether it would be any more competent than TEPCO at managing the situation and communicating it to the public. Over the weekend, it turned out that radiation doses near the leaking tanks are 18 times larger than first reported: leakage that started as a mere 'anomaly' has turned into a genuine crisis. Japan should start consulting international experts for help. The United States, Russia, France and the United Kingdom — to name but a few — all have know-how in nuclear engineering, clean-up and radiation health that would serve Japan well. An international alliance on research and clean-up would help to restore shattered public trust in the usefulness and effectiveness of monitoring and crisis-mitigation.

"An international alliance on research and clean-up would help to restore shattered public trust."

The most important impacts of the leaks will be those on the sea off Fukushima and the larger Pacific Ocean, which must be closely monitored. After assessments by US and Japanese scientists in 2011 and 2012, two major questions remain unanswered. How much radioactivity is still entering the sea? And, given the high levels of radioactivity that have been measured in some species long after the accident, when will fish and seafood from the region be safe to consume? The leaks make it more urgent to find answers to these questions.

To make reliable assessments of any environmental effects, scientists need to be able to collect data on contamination of marine food webs with all long-lived radionuclides, and particularly with caesium-137, strontium-90 and plutonium-239. They also need to know the sources of contamination, and to study the transport of radionuclides in groundwater, sediments and ocean currents. Current Japanese Prime Minister Shinzo Abe and his government have promised to boost science; they should encourage and support researchers from around the world in collecting and sharing information. Chernobyl was a missed opportunity for post-accident research — in that sense at least, Fukushima could do much better. ■

The nitrogen fix

A simple iron complex offers a chance to update how the global supply of ammonia is made.

When it comes to the natural processes of plants, photosynthesis tends to hog attention. If researchers could efficiently copy the ability to convert sunlight to energy, chemists promise, our energy problems would be over. They have not managed it yet.

They have had more luck with harnessing and mimicking the less-heralded, but just as important, process of nitrogen fixation — the conversion of nitrogen from the air into ammonia, which can be used by plants to make DNA, RNA and proteins, and by industry to make fertilizers and explosives. On page 84, chemists announce the discovery of an important piece of the puzzle. Jonas Peters and his colleagues at the California Institute of Technology (Caltech) in Pasadena have identified a small iron complex that efficiently catalyses the conversion.

The discovery comes a full century after the chemist Carl Bosch opened his nitrogen works in Oppau, Germany, and in doing so sealed the deaths of millions of people, and the birth and survival of billions more. Bosch had worked out how to scale up a laboratory reaction to combine nitrogen from the air and hydrogen from natural gas into synthetic ammonia. Textbooks talk about the Haber process, named after German chemist Fritz Haber, who made the theoretical breakthrough, but it is more properly called the Haber–Bosch process. Both Haber and Bosch won Nobel prizes for their work.

It is hard to overstate the impact of the Haber–Bosch process. A figure published in 2008 (the centenary of Haber's patent) shows how the increase in world population since 1960 has kept step with increases in the use of nitrogen fertilizer (J. W. Erisman *et al. Nature Geosci.* **1,** 636–639; 2008). Population growth through access to fertilizer and therefore food was one of Haber's goals in developing his process. The other was to give Germany mastery of the science of munitions. Both goals demanded that the industrial supply of fixed nitrogen grew from the few hundreds of thousands of tonnes available per year at the start of the twentieth century, when it relied on natural resources such as guano and mineral saltpetre (potassium nitrate and sodium nitrate).

Bosch had to treat nitrogen and hydrogen under massive pressure and heat to make the conversion to ammonia. In industry, the process is still done in the same expensive and energy-intensive way.

Crucially, the synthesis described by the California group unfolds under mild, environmentally friendly conditions, just as it does in nature. (Well, when the conversion is done in the soil — another way to fix nitrogen naturally is through the searing flash of a lightning strike.) Peters and his colleagues examined the enzymes and cofactors that make ammonia among the roots of plants such as legumes. Iron has for decades been known to be important in these cofactors, but exactly how and why have been a mystery. For a while, attention switched to molybdenum, which chemists showed could also help to make ammonia, but biochemical and spectroscopic data have renewed the focus on iron. The finding from the Caltech scientists supports this: the iron complex they identify can do the job with no need for molybdenum.

It took less than five years for Bosch to commercialize Haber's discovery, and to revolutionize the industrial supply of ammonia. It will probably take longer for researchers to build on the latest work, but at least now they have a platform.

The stakes have always been high. In the nineteenth century, Peru and Chile fought a war over guano. When Germany was denied access to Chile's saltpetre during the First World War, the Haber–Bosch process gave it — and the world — an alternative, which it grasped with both hands. All the time, legumes such as alfalfa, peanut and clover have been quietly and efficiently doing their thing. A century after Bosch, they could help to write a new chapter in the ammonia story. And photosynthesis? Watch this space. ■

➲ **NATURE.COM**
To comment online, click on Editorials at:
go.nature.com/xhunqv


ABBY LAPOINTE

Deep-sea trawling must be banned

Industry interests should not be allowed to derail a European Union vote on whether to prohibit a destructive fishing technique, says **Les Watling**.

Trawling the bottom of the ocean, dragging heavy metal equipment along the seabed at high speed, is the most destructive form of deep-sea fishing in the world. The fishing industry loves it because it is very effective. But it is indiscriminate and leaves behind a trail of devastation.

This month, the European Union (EU) is scheduled to vote on a proposed ban on deep-sea bottom trawling. If passed, the ban would be the first of its kind, although it would build on existing prohibitions on trawling in shallower water. It could give the seas some breathing space and fish stocks a chance to recover.

There has been intense lobbying against the proposed ban by a fishing industry that has strong ties with the governments of several countries and support from Ifremer, the predominant French fisheries-research organization. Lobbying groups have threatened legal action against scientists for publishing data deemed to be critical of the industry. The EU Fisheries Committee includes Members of the European Parliament from French, UK and Spanish areas in which deep-sea fishing vessels are docked. These ties have slowed the committee's consideration of the proposed ban to a crawl: it has postponed its vote by several months from the originally scheduled date. By contrast, the EU environment committee, which is not so dominated by members from fishing ports, voted 58 to 1 in favour of the ban in March.

In the run-up to the Fisheries Committee vote, now set for 18 September, science has been dragged into the political fight. The fishing lobby has published a series of influential pamphlets that start with the famous phrase "the inconvenient truth". In essence, the pamphlets suggest that it is possible to use a lighter version of trawling equipment to trawl in the deep sea without doing much damage; that stocks of the target species are not being depleted; and that non-target catch is made up of just a few abundant species that are not in any kind of trouble. Many of the 'truths' listed are quoted by European politicians. Yet they are bunk.

If the European Parliament passes the ban, it will have global force and influence on conservation. As such, responsible scientists cannot let these claims go unchallenged. Here is a rundown of why the assertions are wrong.

Claim: damage can be limited by making a lighter trawl. In fact, trawls must be heavy to reach the sea floor at 800–1,500 metres, the zone in which most deep-sea fishing occurs. They need to be strong in order not to come apart when pulled through the water at 3 knots (5.6 kilometres per hour) or more. Deep-sea organisms are known to be delicate and fragile, often consisting of as much water as tissue. By analogy, it makes no difference if you are run over by a small car that weighs one tonne or a large truck that weighs several tonnes. The flesh of the body is no match for the strength of steel, however light the equipment.

Claim: catches of deep-sea fish species of interest are reaching a sustainable level (maximum sustainable yield; MSY). What this does not take into account is that population assessments are based on industry data for ten index species, so what we know about fish stocks we get from fisheries catches. A vessel is not likely to trawl in an area that does not produce fish, so the abundance data are biased. Yet these are the data used to determine whether the species is at MSY.

Unfortunately, all catches of deep-sea fish in Europe are now at about 20% or less of their peak levels. Because the catches have levelled out at these low values, it is claimed that the populations are at MSY. However, it is well known that the French catch, for example, comes from a smaller area than previously, and it is possible that, rather than 'doing fine', as the range shrinks, the measured populations experience increasing effective mortality rates. Fishing could cause a complete collapse, as happened for cod stocks off eastern Canada in the early 1990s. Furthermore, there is no information on the populations of non-index species, dozens of which are threatened and are routinely caught.

THE BAN COULD GIVE THE SEAS SOME **BREATHING SPACE** AND FISH STOCKS A CHANCE TO **RECOVER.**

Claim: by-catch is limited and no by-catch species are in trouble. Actually, studies suggest that trawls for 3 target species can catch more than 100 non-target species. Reported catch of the non-commercially valuable Baird's slickhead (*Alepocephalus bairdii*), for example, which accounts for more than one-third of by-catch by weight, has declined precipitously, and is now at about 6% of its 2002 high. This decline would be enough for the fish to be classed as endangered by the International Union for Conservation of Nature.

The final claim in the published fisheries propaganda is economic, not scientific. But let us examine it.

Claim: the French deep-sea fishing sector is operating sustainably and is an important employer. In fact, it has benefited from large subsidies, from both the EU and the French government. Since 2004, the three companies involved — Scapêche, Euronor and Dhellemmes — have received about €15 million (US$20 million) but showed more than €11 million in losses. The three firms employ a relatively small number of people on their boats: about 112 in total, or 0.5% of the entire French fisheries-vessels crew.

The vote this month is a chance to reverse much of the damage done to fisheries. Deep-water trawling should be consigned to history. ■

➲ NATURE.COM
Discuss this article online at:
go.nature.com/9lxngp

Les Watling *is professor of biology at the University of Hawaii at Manoa in Honolulu.*
e-mail: watling@hawaii.edu

RESEARCH HIGHLIGHTS

Selections from the scientific literature

PLANETARY SCIENCE

Close companion for Uranus

Planetary scientists have discovered the first object known to share its orbit with Uranus.

The body, measuring some 60 kilometres in width and dubbed 2011 QF_{99}, stays just ahead of Uranus as the two orbit the Sun. Computer simulations indicate that 2011 QF_{99} will remain near the planet for around 3 million years.

It is part of a group of outer Solar System bodies in unstable orbits — they become temporarily trapped by giant planets and later return to wandering the Solar System's fringes, report Mike Alexandersen of the University of British Columbia in Vancouver, Canada, and his colleagues.
Science 341, **994–997 (2013)**

BACTERIOLOGY

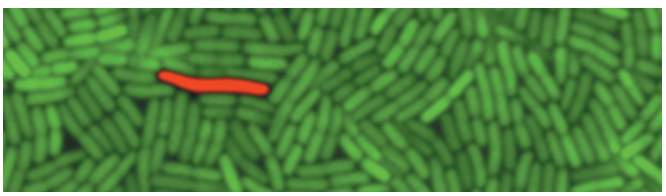
Random signal triggers tolerance

Researchers have determined that randomly produced amounts of a signalling molecule can prompt a multidrug-tolerant state found in one in every million bacterial cells.

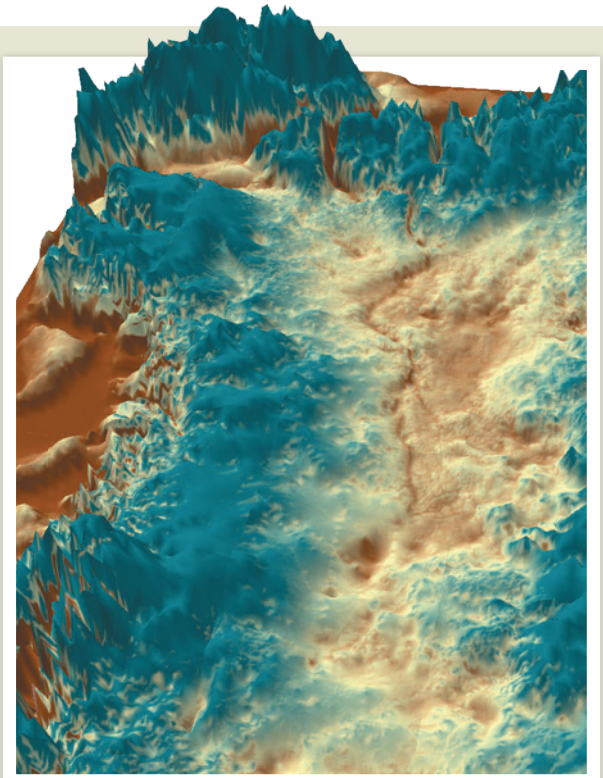
Populations of many disease-causing bacteria contain 'persisters': slow-growing but genetically unchanged cells, which can resuscitate and multiply after antibiotic treatment. Such persisters (**pictured**; in red) are implicated in chronic and recurrent infections such as tuberculosis.

Kenn Gerdes at Newcastle University, UK, and his colleagues showed that the mechanism behind persistence in *Escherichia coli* can be traced to the signalling molecule (p)ppGpp, which is known to initiate a drug-tolerant stress response in bacteria.

Levels of this molecule vary between cells, and the rare cells that produce (p)ppGpp amounts above a certain threshold grow slowly and resist antibiotics. Why some cells produce more (p)ppGpp than others remains unclear, but the authors propose a "microstarvation" model in which the cells with the least nutrients go into persistence. The finding suggests an application for drugs that inhibit (p)ppGpp synthesis.
Cell 154, **1140–1150 (2013)**



ETIENNE MAISONNEUVE



J. BAMBER/UNIV. BRISTOL

GEOLOGY

Greenland's Grand Canyon

Researchers have used radar to penetrate Greenland's ice sheet, revealing a monstrous canyon in the bedrock beneath.

The canyon (**pictured**, looking northwest from central Greenland) runs for at least 750 kilometres to the northern edge of the sheet and is up to 800 metres deep in parts, reports a team led by Jonathan Bamber of the University of Bristol, UK.

Water may have carved the massive canyon at some point before about 3.5 million years ago, when Greenland's ice sheet began growing. This feature could act as a drainage route and might explain why Greenland — unlike Antarctica — has few lakes beneath its ice.
Science 341, **997–999 (2013)**

NANOTECHNOLOGY

Super-bright nanocrystals

A dollop of infrared light can greatly improve nanocrystals that convert infrared photons to visible light.

Such crystals are made of sodium yttrium fluoride doped with ions of elements such as thulium, which emit radiation in the visible spectrum. If too much thulium is added to the mix, the luminosity drops.

But Jiangbo Zhao and Dayong Jin at Macquarie University in Sydney, Australia, and their colleagues have discovered that using high levels of infrared light to irradiate the crystals prevents this 'quenching' of the visible light emission and makes the crystals shine 1,000 times brighter.
Nature Nanotechnol. **http://dx.doi.org/10.1038/nnano.2013.171 (2013)**

MODEL ANIMALS

The social lives of lab mice

Mice raised in labs have more-complex social lives than previous models of group behaviour have suggested.

Alon Chen and Elad Schneidman at the

Weizmann Institute of Science in Rehovot, Israel, and their colleagues used cameras to track the locations of groups of four mice in an arena at night by painting their fur with fluorescent colours.

Models that assumed the mice act as individuals or interact in pairs (the typical model of mouse sociality) did poorly at describing the group's movement as a whole. Only models with interactions between three mice gave good approximations of the observed behaviour.

Using this model, the researchers found that mice raised in standard lab environments were less individualistic than those who had lived in larger groups and in more complex enclosures.

eLife 2, **e00759 (2013)**

CLIMATOLOGY

Geoengineering has its limits

Blocking solar radiation cannot restore a high-carbon world to its preindustrial climate.

Ben Kravitz of the Pacific Northwest National Laboratory in Richland, Washington, and his colleagues analysed results from 12 climate models simulating quadrupled atmospheric carbon dioxide levels, and a matching reduction in solar radiation.

Blocking the sunlight kept average global temperatures in check while preventing 97% of Arctic sea-ice loss, but the effects on temperature and precipitation varied by region, with the Arctic experiencing more warming and the tropics more cooling, as well as less rainfall in places.

The authors suggest that solar geoengineering cannot simultaneously correct for global and regional impacts from rising greenhouse-gas levels.

J. Geophys. Res. **http://doi.org/nmz (2013)**

NEUROBIOLOGY

Protecting neurons from Parkinson's

An enzyme that adds a phosphate group to a protein called α-synuclein — the phosphorylated form of which tangles up inside neurons in Parkinson's disease — also targets the protein for destruction.

Hilal Lashuel and his colleagues from the Swiss Federal Institute of Technology in Lausanne injected rat brains with the genes for both α-synuclein and the enzyme PLK2.

When they increased amounts of α-synuclein, the rats lost specific neurons and developed Parkinson-like symptoms. But both effects were avoided, and levels of neuronal α-synuclein fell, when they overexpressed the gene for PLK2 at the same time.

They showed that the PLK2 enzyme protects neurons from α-synuclein toxicity by shepherding the protein into autophagy, a cellular process for clearing debris.

Proc. Natl Acad. Sci. USA **http://doi.org/nmx (2013)**

ANIMAL BEHAVIOUR

Puppeteer squid haunts the deep

The deep-sea squid *Grimalditeuthis bonplandi* (**pictured**) uses its tentacles to mimic smaller sea creatures, attracting prey.

The animal's tentacle stalks are fragile in comparison to those of most squid, and its tentacle tips lack the suckers or hooks that are usually used to catch and manipulate prey.

Henk-Jan Hoving, then at the Monterey Bay Aquarium Research Institute in Moss Landing, California, and his colleagues analysed videos of seven *G. bonplandi* recorded by remotely operated submarines, the first observations ever made of the squid in its natural habitat. The creatures moved their tentacles in a manner that resembled the movements of smaller animals such as fish or other cephalopods. The researchers suggest that the motions cause vibrations, water movements or bioluminescence in nearby microorganisms. The squid then consume curious creatures that come to investigate.

Proc. R. Soc. B 280, **20131463 (2013)**

For a longer story on this research, see go.nature.com/byyusm

MBARI

GENETICS

A library of human knockouts

The functions of all 20,000 human protein-coding genes are still not known, partly because large gene-knockout libraries are not available for human cells.

One strategy for making such a library is to use a retroviral gene trap that disrupts gene expression when inserted into different genes. But this does not work for human cells, which have two copies of most genes. Now a team led by four researchers — Tilmann Bürckstümmer of biotech firm Haplogen in Vienna, Austria, and Thijn Brummelkamp, Giulio Superti-Furga and Sebastian Nijman of the Research Center for Molecular Medicine of the Austrian Academy of Sciences in Vienna — has overcome this problem.

The team used a cancer-cell line that has only one copy of most genes (near-haploid) to assemble a knockout library of 3,396 human genes. The collection should enable screens of a range of phenotypes, the team reports.

Nature Meth. **http://doi.org/nmw (2013)**

COMMUNITY CHOICE

The most viewed papers in science

MARINE MICROBIOLOGY

Marine plastic fantastic for microbes

HIGHLY READ on pubs.acs.org/journal/esthag in August

A new ecosystem is emerging on plastic marine debris in the world's oceans: the 'plastisphere'.

Tracy Mincer at the Woods Hole Oceanographic Institution and Linda Amaral-Zettler of the Marine Biological Laboratory in Woods Hole, both in Massachusetts, and their colleague sequenced the microbial communities on plastic litter floating in the sea. They revealed a complex food web living in biofilms on two types of plastic, made up of microbes that were genetically distinct from those in the surrounding seawater. Unlike those in seawater samples, the communities on plastic had an even spread of diversity and were not dominated by just one or two types.

Environ. Sci. Technol. **http://doi.org/m4q (2013)**

CORRECTION

The Research Highlight 'Soil life predicts nutrient flow' (*Nature* **500,** 380; 2013) stated an incorrect affiliation. Franciska de Vries is now at the University of Manchester, UK.

NATURE.COM

For the latest research published by *Nature* visit:
www.nature.com/latestresearch

SEVEN DAYS *The news in brief*

POLICY

Badger blitz

A controversial badger cull has begun in the United Kingdom, it was reported on 27 August. The move is aimed at stemming the spread of bovine tuberculosis (TB), which badgers (*Meles meles*) can carry and pass to cattle. Farmers will be allowed to kill around 5,000 badgers over 6 weeks in the English counties of Gloucestershire and Somerset. The cull has angered animal-rights campaigners, and some scientists have argued that it will not significantly reduce TB in cows. The UK government says that previous trials and successful culls in other countries support the campaign (see *Nature* **490,** 317–318; 2012).

Bee rift deepens

The European Commission is facing legal action over its decision to temporarily ban the use of three neonicotinoid insecticides linked to declines in bee populations. On 27 August, agricultural company Syngenta in Basel, Switzerland, said it was challenging the commission's ban on thiamethoxam. Bayer CropScience of Monheim am Rhein, Germany, has also filed legal action on that ban, as well as on restrictions on clothianidin and imidacloprid. The commission moved to ban all three chemicals for two years in May, following a review by the European Food Safety Agency. See go.nature.com/df6hdp for more.

Marijuana truce

The US government will no longer interfere when states legalize marijuana for medical or personal use. The substance is illegal under federal law, but some states permit medical use, and last year, Colorado and Washington became the first to legalize recreational use by those aged 21 and older (see go.nature.com/rtr3ou). The US Department of Justice said on 29 August that it would not challenge states' decisions to regulate marijuana production, sales and distribution, provided that the states impose strict law-enforcement controls.

CLIF TAYLOR

New Mexico's dark skies honoured

One of North America's most famous archaeoastronomy sites has been honoured for its efforts to preserve the night sky for future sky-watchers. On 28 August, the Chaco Culture National Historical Park in New Mexico was formally designated a Dark Sky Park by the International Dark-Sky Association in Tucson, Arizona. Chaco joins 11 other sites that restrict light pollution and protect high-quality astronomical observations. The area was a major centre for the ancestral Pueblo cultures between about AD 850 and 1250, and the hill Fajada Butte (pictured) in the park is known for carvings that may have been used to mark solstices and equinoxes. Other archaeoastronomy features at the park include a pictograph that is thought to depict the AD 1054 supernova that formed the Crab Nebula.

EVENTS

Japan's rocket woes

The Japan Aerospace Exploration Agency is investigating a failure that halted the launch of the Epsilon-1 rocket 19 seconds before take-off on 27 August. It had been hoped that the three-stage, 24-metre rocket would vault Japan into the lucrative satellite-launch industry. But the lift-off from the Uchinoura Space Center, which had already been postponed from 22 August, was cancelled, grounding the rocket and its cargo — a space telescope intended to observe planets including Venus, Mars and Jupiter.

Fukushima fixes

The government of Japan vowed on 3 September to spend billions of yen safeguarding the stricken Fukushima Daiichi nuclear plant following a series of dangerous — and embarrassing — radioactive leaks. Taking control of the clean-up operation, ministers said that they would spend some ¥47 billion (US$470 million) freezing the ground around the plant's perimeter and making repairs to water storage tanks, the source of the leaks. The news comes days after Fukushima's operator, the Tokyo Electric Power Company, admitted that one leak at the plant was 18 times more radioactive than first thought. See go.nature.com/yc4rxa and page 5 for more.

Chemical weapons

The World Health Organization (WHO) on 27 August issued revised guidelines for treating

DOROTHY EDWARDS, THE VIRGINIAN-PILOT/AP

victims of chemical exposure. Released as a draft, the guidance updates existing recommendations for identifying and categorizing patients, and protecting health-care workers from contamination. The agency had been working on the revisions since 2012, but rushed them out in response to the suspected use of chemical weapons in Syria last month.

China Moon rover

China will launch its first Moon lander at the end of the year, the state news agency Xinhua reported on 28 August. The Chang'e-3 mission will carry a rover and other equipment to analyse the lunar surface and to make space observations. China launched lunar orbiter Chang'e-1 in 2007, and carried out high-resolution mapping of the Moon with orbiter Chang'e-2 in 2012. Chang'e-3 is slated to launch at the end of the year from the Xichang Satellite Launch Center in southwest China.

RESEARCH

Dolphins infected

An outbreak of a virus is thought to have caused mass strandings of bottlenose dolphins (*Tursiops truncatus*, **pictured**) between New York and North Carolina. On 27 August, the US National

Oceanic and Atmospheric Administration said that 333 dolphins had beached themselves since 1 July, and that of 33 animals examined, 32 showed signs of antibodies to cetacean morbillivirus, a cousin of the pathogen that causes measles in humans. Agency scientists estimate that the outbreak could last for a further several months.

FACILITIES

Green institute

The Helmholtz Institute for Renewable Energy in Erlangen, Germany officially came into being on 1 September, following an initial cooperative agreement last month between national research centres in Jülich and Berlin, and the Friedrich Alexander University of Erlangen-Nuremberg. The partners have committed a total of €5.5-million (US$7.3-million) per year, and the state of Bavaria will contribute €38 million for initial building and operational costs. With plans to develop advanced solar and hydrogen-based energy-storage technologies, the institute will support Germany's ambitious programme to replace fossil fuels with renewable sources of energy (see *Nature* **496,** 156–158; 2013).

PEOPLE

Italians honoured

Particle physicist and Nobel prizewinner Carlo Rubbia and stem-cell researcher Elena Cattaneo have been made permanent members of Italy's parliament by President Giorgio Napolitano. Each Italian president is empowered to appoint up to five senators for life, for "their outstanding achievements in the social, scientific, artistic and literary fields", but scientists have rarely been awarded this privilege. Cattaneo has taken a strong public position against a controversial, unproven stem-cell therapy known as the Stamina method, for which parliament had agreed in May to fund a trial.

FUNDING

Funding fall-out

Many US scientists are spending more time chasing less grant money than they were three years ago, according to a survey released on 29 August by the American Society for Biochemistry and Molecular Biology in Rockville, Maryland. The study examines the impact of declines in government science funding since 2010 — including this year's across-the-board cuts known as sequestration. Of the 3,731 researchers who responded to the online survey, just over half reported turning away promising potential lab members for financial reasons. Budget pressures have driven 18% to consider moving to another country to pursue research.

Brain-injury fund

The US National Football League (NFL) has agreed to a US$765-million legal settlement with more than 4,500 former players who say that the league ignored or concealed known risks of concussion-related injuries, including depression and early onset of dementia. US judge Anita Brody, who must still approve the settlement, announced the terms in court documents filed on 29 August. In addition to concussion-related compensation and legal expenses, the money will go towards medical exams and a medical-research programme for retired NFL players and their families.

COMING UP

6 SEPTEMBER
NASA's Lunar Atmosphere and Dust Environment Explorer (LADEE) launches from Wallops Flight Facility in Virginia. The mission will orbit the Moon to study the structure and composition of its atmosphere. See go.nature.com/ns43sq for more.

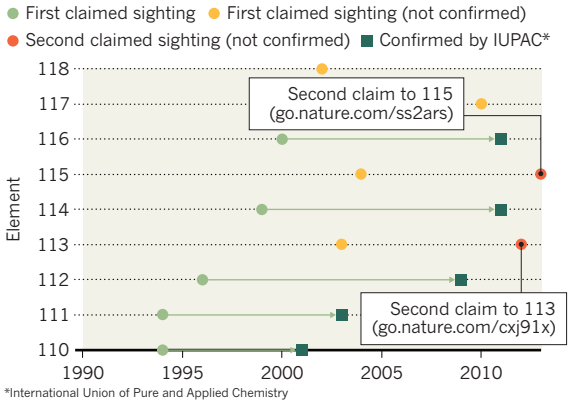
SOURCE: IUPAC

TREND WATCH

Researchers in Sweden have reported strong evidence for the formation of element 115, temporarily called ununpentium. The team detected radioactive-decay products that may have originated after element 115 formed from the collision of calcium ions with americium nuclei (see go.nature.com/ss2ars). The claim follows a possible sighting of the element in 2004 by scientists in Russia; the International Union of Pure and Applied Chemistry, which names elements, has yet to rule on it.

THE SEARCH FOR NEW ELEMENTS

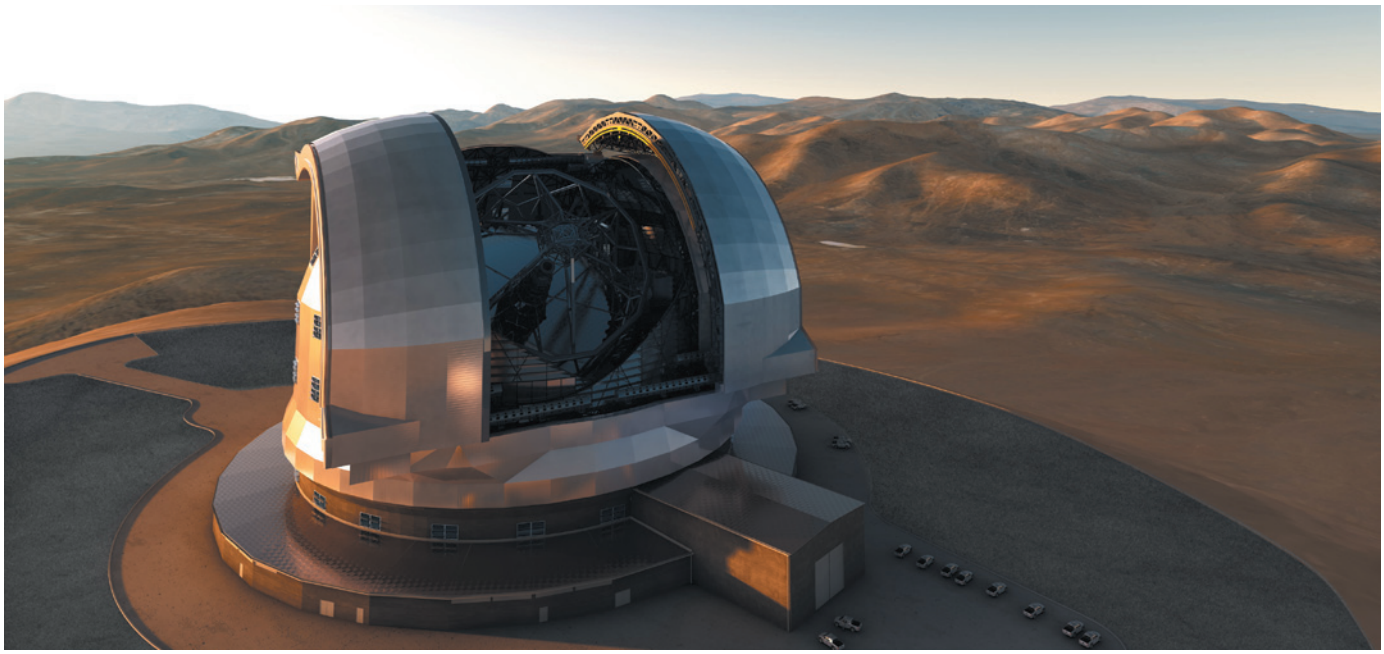
In the past two decades, chemists have created five new elements; elements 113 and 115 are likely to be confirmed soon.



*International Union of Pure and Applied Chemistry

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

NEWS IN FOCUS



ESO

The 39-metre Extremely Large Telescope in Chile (artist's impression) will suffer setbacks without Brazilian money.

ASTRONOMY

Brazil delays stargazing pact

Reluctance to pay entrance fees stalls European Southern Observatory's giant telescope.

BY RAFAEL GARCIA

On 28 August, Jorge Meléndez stepped into a room full of journalists to announce a remarkable discovery: an 8.2-billion-year-old star that was depleted in elements such as iron and aluminium in almost exactly the same way as the Sun — a hint that the older star could host terrestrial planets. In an age when astronomers are obsessed with finding another Earth, Meléndez had found something nearly as exciting: a solar twin.

Even more significant than the discovery were its circumstances. Meléndez, an astronomer at the University of São Paulo, Brazil, had found the star using an elite telescope that belongs to the European Southern Observatory (ESO) — a sign of a functioning, if fragile, transatlantic co-dependency.

In 2010, Brazil signed an agreement with ESO, becoming the organization's first non-European member. Brazil's membership fees would allow ESO to begin construction of its €1.1-billion (US$1.5-billion), 39-metre Extremely Large Telescope (ELT). In return, Brazil's small but growing cadre of astronomers would get access to ESO's existing telescopes in Chile. Meléndez's discovery came after just two nights of observation at the Very Large Telescope, ESO's premier observatory, in the Atacama Desert. He has been granted a further 88 nights at ESO's La Silla Observatory, where he is looking for planets that orbit solar twins. "Before the ESO agreement, it would have been impossible for me to do this," he says.

But the relationship has started to fray. Nearly three years after the agreement was signed, Brazil's Congress still has not ratified it. The country has made nominal membership payments of €4 million a year, maintaining access to ESO telescopes for astronomers such as Meléndez, but has not coughed up its full entrance fees and annual dues, which are expected to total about €270 million over a decade. ESO officials say that further delays will cause the nation to miss out on lucrative construction contracts for the ELT, which will be solicited by the end of this year. "The longer Brazil waits, the more it risks compromising the opportunity for Brazilian companies to be awarded such large contracts," says Tim de Zeeuw, director-general of ESO at its headquarters in Garching, Germany.

No one in Brazil's executive government is pushing the agreement. The science minister who signed it, Sergio Rezende, left the cabinet just days later, and his successor did nothing with the document. It was not until this February, more than a year after the arrival of current minister Marco Antonio Raupp, that the agreement was formally sent to Congress for ratification. But Raupp, whose emphasis has been on innovation and applied science ▶

over basic research since his term began, is not seen as a champion of ESO membership.

That has essentially left Brazilian astronomers to lobby Congress on their own. And not all of them are in favour of joining ESO; some see it as an expensive step that will bankrupt more modest, home-grown efforts to nurture Brazilian astronomy.

João Steiner, an astronomer at the University of São Paulo, is the most vocal critic of ESO membership. He argues that Brazil's fees — which are based on the country's gross domestic product — are comparable to those of Italy and the United Kingdom, even though Brazil has fewer astronomers, and fewer still with research programmes mature enough to compete for observation time. "It doesn't make sense for us to subsidize European science with Brazilian taxpayers' money," says Steiner.

He would prefer Brazilian astronomers to work with more affordable telescopes. These include Gemini, two 8-metre telescopes in Chile and Hawaii, run by six countries including Brazil; the Southern Astrophysical Research Telescope (SOAR), a 4.1-metre instrument in Chile part-owned by Brazil; and a 1.6-metre telescope at Brazil's own Pico dos Dias Observatory near Itajubá.

But Marcio Maia, an astronomer at the National Observatory in Rio de Janeiro, says that his country should be ambitious. "If Brazil wants to remain in the Stone Age of astronomy, we can do it at the city of Itajubá, where there is nothing to be seen, we can keep our meagre share at Gemini or we can use SOAR, a telescope with no good instrument," he says. "We can only learn how to do cutting-edge astronomy by entering the competition and fighting for it."

Steiner acknowledges that joining ESO would bring benefits, but says that the fees are too high at a time when Brazil is unlikely to take on expensive projects. In July, the government announced a 10-billion-real (US$4.3-billion) cut to overall federal expenses. And the Ministry of Science, Technology and Innovation has little room in its own budget; it says that it would have to commit 12% of its general funding to honour the ESO agreement if Congress does not approve new spending for it. Such a commitment is unlikely from Raupp, who has blamed the delay in ratification on the lack of consensus. "The strong divide in the Brazilian astronomical community is certainly an important fact that cannot be ignored," says José Roberto Ferreira, a spokesman for Raupp.

But contrarians such as Steiner seem to be in the minority, according to a 2010 poll by the Brazilian Astronomical Society. In interviews of 180 of the society's 660 astronomers, 75% supported the ESO agreement. Another 17% (mostly theoretical cosmologists) declined to vote; just 8% were against membership.

Even with strong support from astronomers, however, it will be difficult to persuade Congress to approve membership by the end of 2013.

If Brazil fails to ratify the agreement, ESO could boost its funding by adding other member states. Russia, Poland, Canada and Australia have all expressed interest. "We are being regularly approached by other countries," says de Zeeuw. But it would be years before money from any new member would arrive. If ESO wants to build the ELT, it has few choices but to wait for Brazil. ■

CLIMATE CHANGE

Floods spur mountain study

Himalayan nations take action in response to changing climate and its deadly effects.

BY JANE QIU

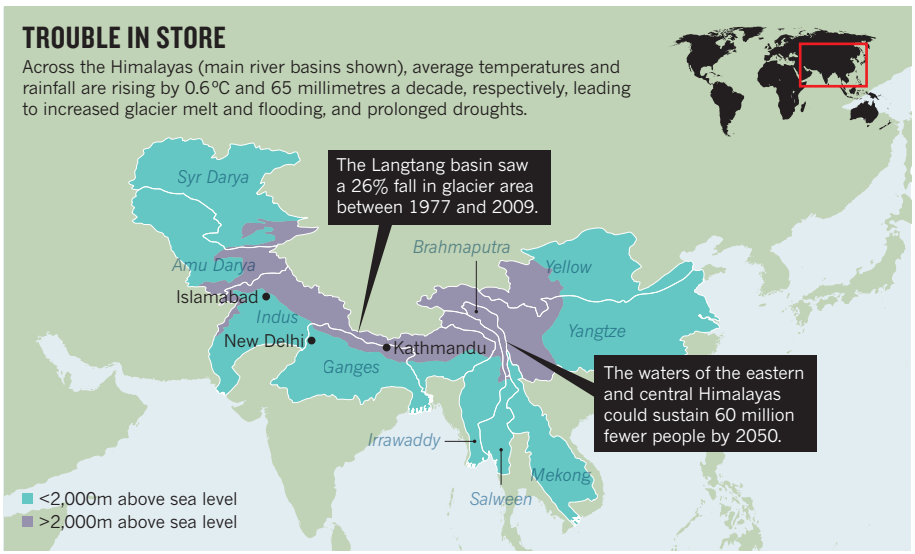
A snapshot of weather-related disasters in the Himalayas suggests that things are amiss on the roof of the world.

This spring, for example, western Nepal was hit by a severe drought, leading to crop failures and exacerbating an already serious food crisis. In June, the same region was devastated by its worst floods in 50 years, caused by unusually intense monsoon rains. The deluge wreaked havoc in the northern Indian states of Uttarakhand and Himachal Pradesh, killing at least 5,700 people and causing an estimated loss of US$2 billion.

After decades of such climate-related incidents, the eight member countries of the International Centre for Integrated Mountain Development (ICIMOD), a regional body, have agreed that enough is enough. On Monday, they launched a three-year study that aims to comprehensively assess the current state of the Hindu Kush Himalayas, the enormous area sustained by the world's highest mountain range, and to make recommendations on how it might be safeguarded and developed.

And none too soon, say many scientists. As climate change tightens its grip, "disasters will become increasingly frequent", says Vinod Tewari, a geologist at the Wadia Institute of Himalayan Geology in Dehradun, India. The Himalayas are getting warmer at a rate of 0.6 °C each decade, three times the global average[1]. Rainfall there is increasing at a rate of 65 millimetres per decade and the monsoon season is getting wetter[1]. However, winters are getting drier (see 'Trouble in store').

As a result of the warming, most Himalayan glaciers are retreating rapidly. Glacier lakes are becoming larger and more numerous, inundating pastures and threatening downstream communities. The changing climate is also "taking a toll on alpine pastures and forests", says Maharaj Pandit, an ecologist at the University of Delhi in New Delhi. Some plants are shifting to higher altitudes[2], others are on the verge of extinction, and the incidence of invasive



species is rising. "If the warming trend persists, we will see drastic changes in the ecosystems, with devastating consequences on biodiversity and the livelihood of mountain communities," he adds.

India and Nepal are not the only nations affected. Hindu Kush is a 3,500-kilometre mountain chain that spans eight countries — from Afghanistan to Myanmar. Known as Asia's water tower, the mountains provide ecosystem services and support the livelihoods of more than one-fifth of the global population.

"The impact of climate change is compounded by many other challenges that face the Himalayas," says David Molden, ICIMOD's director general. In the past few decades, the region has seen increasing population growth, energy shortages, pollution, environmental degradation and disasters — all of which challenge traditional livelihoods. "Once self-sufficient people now face dire economic poverty," he says.

To mitigate climate change, it is crucial to factor in all those challenges, says Molden. This is where ICIMOD's assessment comes in. It will, he says, review the current state of knowledge on glaciers, biodiversity, water resources and pollution, identify trends of changes in climate, demography and land use, and address issues such as poverty, hydropower development, natural disasters and food security. The result will be a set of policy recommendations on, for example, sustainable grazing, increased collaboration between nations, deforestation and flood prevention.

The Himalayas are of global significance, says Volker Mosbrugger, director of the Senckenberg Research Institute and Natural History Museum in Frankfurt, Germany. The report will help to manage the region's wealth of resources sustainably and empower mountain communities to tackle climate change, he says.

Pandit lauds the project as "timely and important", but says that it is the enforcement that matters. "Whatever decisions or recommendations are given, they have to reach local authorities." ■

1. Shrestha, U. B., Gautam, S. & Bawa, K. *PLoS ONE* **7,** e36741 (2012).
2. Telwala, Y., Brook, B. W., Manish, K. & Pandit, M. K. *PLoS ONE* **8,** e57103 (2013).

BECK DIEFENBACH

Non-native eucalyptus trees dominate the wild forest that covers San Francisco's Mount Sutro.

ECOLOGY

Forest management plans in a tangle

Conservation fight flares over invasive California eucalyptus.

BY DANIELLE VENTON

In the middle of metropolitan San Francisco stands an army — and many Bay Area residents want it to stay garrisoned there. It is a forest of non-native, invasive blue gum eucalyptus, along with Monterey cypress and pine, that has covered the city's Mount Sutro since the late nineteenth century.

For more than a decade, the University of California, San Francisco (UCSF), which sits at the base of the hill and owns the 25-hectare reserve, has sought to manage the forest. Its aim is to reduce fire hazards and to encourage native grasses, wild flowers and brush by selectively thinning the trees. It also wants to restore the original natural habitat — and some conservationists and locals are having none of it.

On 29 August, as San Francisco's characteristic fog burned off, about 40 protesters stood in front of the UCSF campus while a crew with chainsaws and electric weed-cutters cleared blackberry bushes, ivy vines and small eucalyptus trees near roads and buildings in what the university is calling "urgent fire measures" recommended by the city's fire department. The protesters were not pleased to see the trees come down, even in the name of fire safety. "Here we have a beautiful, wild forest right in the middle of our city and the university wants to destroy it," says Paul Rotter, a resident of the area since the 1970s who regularly walks the hill.

Selective tree-clearing on Mount Sutro would help the larger remaining trees to fight the beetle and fungal pests currently afflicting the forest, say UCSF consultants. The area's great horned owls could also have more foraging opportunities. Reducing competition for sunlight should encourage the growth of large trees, which, say consultants, would enhance carbon storage and reduce the risk of the catastrophic fires that threaten lives and property, and pour carbon into the atmosphere.

That these reasons hold little water with the protesters highlights an emerging fissure among environmentalists and ecologists. For some, a hardline devotion to preserving native ecosystems is giving way to a more post-modern idea of what constitutes a natural ▶

MORE ONLINE

THE YOMIURI SHIMBUN/AP IMAGES

TOP NEWS

Revelations on severity of Fukushima leaks prompt calls for government action go.nature.com/yc4rxa

MORE NEWS

- *Arabidopsis* database forced to charge users go.nature.com/ekoarw
- Crop pests move with climate change go.nature.com/pbnvqn
- TB sequencing studies map mutations that confer antibiotic resistance go.nature.com/2tgwgg

NATURE PODCAST

Brain training in older adults; quantum key distribution; and silver nanoparticles nature.com/nature/podcast

landscape. "Mount Sutro is part of a larger story," says Richard Hobbs, an ecologist at the University of Western Australia in Crawley. "What some people see as a weed-filled blot on the landscape, others see as something extremely valuable, worthy of managing in its own right. People are increasingly moving away from the belief that a native ecosystem is always best."

That idea grates with many restoration ecologists, says Hobbs. Yet studies increasingly suggest that altered ecosystems need not be bad for biodiversity or ecosystem function. Non-native pine trees provide habitat for threatened cockatoos in Western Australia, for example. And in Scotland, old industrial waste heaps — known as shale bings — are now home to rare and protected plants and animals.

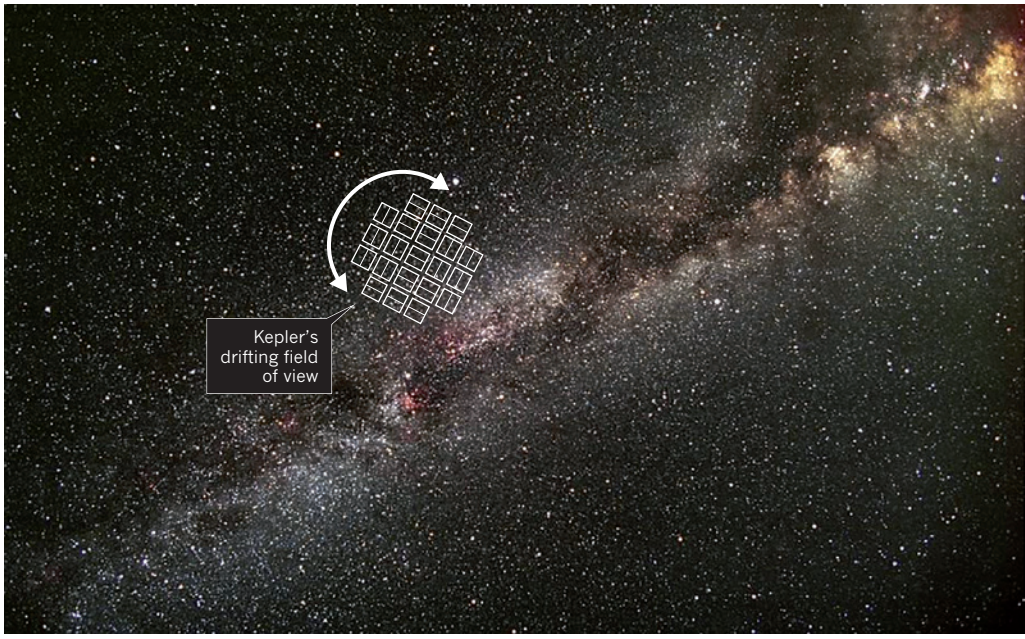
In the early 1990s, Patricia Kennedy of Oregon State University in Corvallis helped to develop management guidelines for northern goshawks. She found that the raptors do not strictly need old-growth forests; land used for timber harvesting can work, too. She says that, at the time, accepting the idea felt like a move to the "dark side". "The whole culture in wildlife biology and conservation circles has been that you can't approximate Mother Nature," she says.

But those ideas are changing today, with altered ecosystems such as Mount Sutro's providing a case in point. In the late 1880s, Adolph Sutro, a mayor of San Francisco, planted the tree-less hill with imported blue gum eucalyptus, as well as Monterey pine and cypress. The eucalyptus quickly took over, and today the forest feels like a primeval jungle — a tangle of almost exclusively introduced species. Joe Mascaro, an ecologist at Stanford University in California who has been publicly critical of UCSF's management plans, says that Mount Sutro has long since given way to a completely new ecosystem. "Restoring it to an original state would be borderline impossible, so why stop the succession that is already in place?"

Resistance to such a heretical idea runs deep among ecologists, but growing numbers are embracing altered ecosystems in the name of pragmatism. "You can reach more win–win situations if you don't insist on purity," says Katharine Suding, an ecologist at the University of California, Berkeley, who specializes in restoring human-affected areas. "It doesn't have to be a natural versus non-natural dichotomy."

For UCSF, finding a middle ground between native and non-native conservation ideals is proving difficult. But the university should get used to it, says Hobbs.

"There is a lot of tension about how to deal with situations like these right now," he says. "With so much non-native habitat, the old views — that everything must be natural — no longer apply." ■



CARTER ROBERTS/NASA

Kepler's field of view is now drifting because of a failure of two of its four reaction wheels.

ASTRONOMY

NASA ponders Kepler's future

Spacecraft could continue to hunt for planets — or take on alternative tasks, such as asteroid spotting.

BY RON COWEN

NASA just can't quit Kepler.

On 15 August, the agency announced that it would stop trying to revive the failed reaction wheels that gave the planet-hunting telescope its precise pointing ability. That essentially brings an end to the main goal of the 4-year-old mission, which has found 3,548 candidate planets by looking for tiny dips in starlight that indicate a planet's passage, or transit, across that star.

But the agency left room for hope: two weeks earlier, it had asked astronomers to submit ideas by 3 September on how the hobbled spacecraft might still perform good science. *Nature* has learned about some of the options in the running, out of the dozens of proposals expected.

Ideas range from a survey of potentially hazardous near-Earth objects to a study of Jupiter-sized exoplanets in large orbits. Kepler scientists will sort through the proposals and decide by 1 November which ones, if any, to recommend to NASA headquarters for further review.

To secure funding from the space agency, the Kepler team will have to show that the studies could not be done by other telescopes. This will be no easy task — especially given that engineers are not sure how well Kepler can perform with just two of its four spinning reaction wheels, which act as stabilizing gyroscopes.

"We're in a real quandary," says Kepler principal investigator Bill Borucki at NASA's Ames Research Center in Moffett Field, California. "We just don't know what Kepler can do."

With three working wheels (a fourth was a spare), Kepler was able to exactly counterbalance the persistent push of sunlight, locking on to targets with such precision that light from a particular star always fell on the same tiny fraction of an individual pixel. But the wheels have a history of poor performance, and in July 2012 one failed — followed by another in May (see go.nature.com/4w1ufr). Although the craft's thrusters can still act as a crude version of a third wheel,

NATURE.COM
For more on Kepler and its discoveries, see:
go.nature.com/8pmjuf

they cannot replicate the pointing accuracy that three wheels provided, and the telescope's focus will drift. Over time, starlight will start to fall on different pixels with slightly different sensitivities. "Every single day, it's as if you're going to use a different detector, a different telescope," says Kepler scientist William Welsh of San Diego State University in California.

Kepler's drift could be minimized by keeping it pointed in the same plane in which the craft orbits the Sun. But that presents a complication. Some of the best science is expected to come from follow-up observations of the field of about 150,000 stars that Kepler has been focused on, and that star field does not lie in the plane.

In one proposal, offered up by Welsh and his colleagues, the craft would continue to stare at this original star field to search for Jupiter-sized planets. Such bodies are sufficiently large that when they pass in front of their parent star they produce a dip in light that can be detected by Kepler even in its compromised state.

Welsh's group would target Jupiters for which Kepler has recorded only a few transits — those that take more than a year to orbit their star. It usually takes a minimum of three transits to confirm the existence of a planet. Catching the third transit could make the difference between a possible and a definitive discovery.

The craft is too shaky to discover an Earth analogue from scratch, but Welsh suggests that it might also be possible for Kepler to add statistical significance to Earth-sized candidates for which transits have already been captured. And David Hogg, an astronomer at New York University, believes that, over the course of many months, Kepler's drift could be used to map out the different light responses of the pixels. That calibration, if detailed enough, could be enough for Kepler to resume its hunt for Earth analogues, says Hogg.

Daniel Fabrycky, an astronomer at the University of Chicago in Illinois, has an alternative follow-up study in mind. He and his colleagues have proposed looking at planetary systems in which densely packed planets are affected by one another's gravitational pulls — creating periodic cycles in which the timing of transits are first advanced and then delayed. The light dip during a transit reveals only the size of the eclipsing planet, but knowledge of transit-time variation yields the planet's mass, which is crucial for working out the density and composition of the bodies. Like Welsh, Fabrycky wants Kepler to zero in on planetary systems with long orbits, for which the full cycle of these transit-timing variations has not yet been seen.

But Andrew Gould, an astronomer at Ohio State University in Columbus, says that he is sceptical about using the craft to simply follow up on its original tasks when its pointing precision has been degraded by a factor of as much as 1,000. "People really have to break out and come up with new ideas," he says.

Perhaps taking such advice to heart, Fabrycky's team has a second proposal: putting Kepler to work not as a planet hunter, but as a sentinel for near-Earth objects, including asteroids several hundred kilometres in diameter that might be on a collision course with Earth. A survey of space rocks would take advantage of Kepler's large field of view. And at least part of the study could be completed with Kepler looking for targets within its orbital plane, so as to optimize its pointing.

"We're in a real quandary. We just don't know what Kepler can do."

Gould has proposed another scheme, in which Kepler would survey stars towards the Milky Way's central bulge for signs of planets, using a technique known as microlensing.

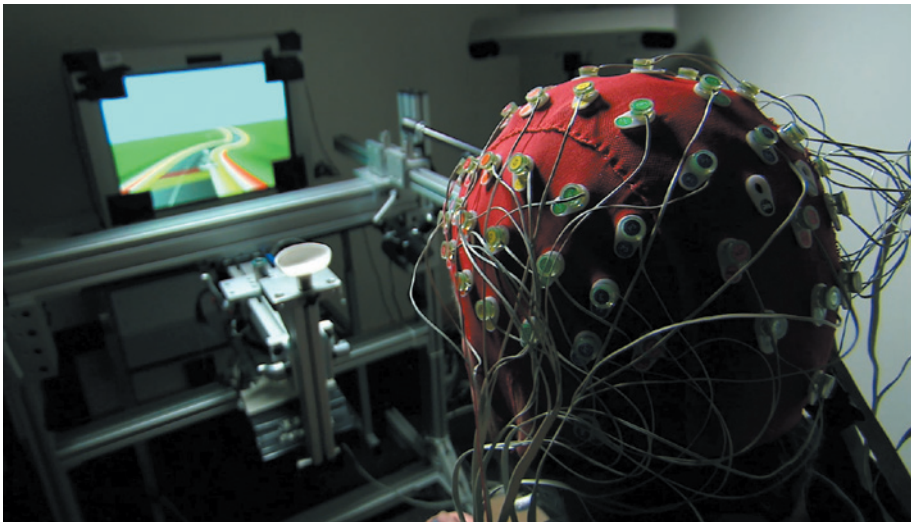
Microlensing relies on a prediction of Einstein's theory of general relativity: the gravity of any massive object bends light. Like a magnifying lens, a foreground star bends and brightens light from stars behind it. A single foreground star, or microlens, produces a characteristic brightening curve, but if that lensing star has a planet, the curve will have an additional wiggle.

Researchers have already used microlensing to reveal some 40 planets towards the centre of the Galaxy, but the observations typically do not reveal masses. By observing microlens planets using Kepler and ground-based telescopes at the same time, differences in transit duration and brightness emerge that can yield the planets' mass. However, the survey could be performed for only about five weeks of the year because of limited chances to view the Galactic Centre without interference from the Sun.

COMPETING CHOICES

If any of the proposals recommended by the Kepler team seems worthwhile to NASA, they will be examined early next year by a review panel of external scientists. At that stage, a repurposed Kepler would face its biggest hurdle — a competition for the limited pot of funds against nine other astrophysics missions, including the Hubble Space Telescope and the Fermi Gamma-ray Space Telescope. On receiving recommendations from the review panel, NASA will make its final funding decisions next June.

Not everyone is rooting for Kepler. Doug Finkbeiner, an astronomer at the Harvard-Smithsonian Center for Astrophysics in Cambridge, Massachusetts, wants NASA to support missions that are still healthy. He has used Fermi to discover two galaxy-sized bubbles of ionized gas blowing from the centre of the Milky Way, and is counting on continued funding for the γ-ray telescope. "My very biased and self-interested perspective is that I hope we let Kepler die," he says. ■



GLOBE TV

Improved mental skills that come from playing a video game are mirrored by increased brain activity.

NEUROSCIENCE

Gaming improves multitasking skills

Study reveals plasticity in age-related cognitive decline.

BY ALISON ABBOTT

Sixty-five-year-old Ann Linsey was starting to worry about how easily she got distracted from whatever she was doing. "As you get older, it seems harder to do more things at once," she says. Then she enrolled in a study to test whether playing a game could improve fading cognitive skills in older people — and was impressed by what it did for her. "I was frustrated because I felt I was losing my faculties. Now I've learnt how to focus my attention."

Commercial companies have claimed for years that computer games can make the user smarter, but have been criticized for failing to show that improved skills in the game translate into better performance in daily life[1]. Now a study published this week in *Nature*[2] — the one in which Linsey participated — convincingly shows that if a game is tailored to a precise cognitive deficit, in this case multitasking in older people, it can indeed be effective.

Led by neuroscientist Adam Gazzaley of the University of California, San Francisco, the study found that a game called NeuroRacer can help older people to improve their capacity to multitask — and the effect seems to carry over to tasks in everyday life and is still there after six months. The study also shows how patterns of brain activity change as those cognitive skills improve.

NeuroRacer is a three-dimensional video game in which players steer a car along a winding, hilly road with their left thumb, while keeping an eye out for signs that randomly pop up. If the sign is a particular shape and colour, players have to shoot it down using a finger on their right hand. This multitasking exercise, says Gazzaley, draws on a mix of cognitive skills just as real life does — such as attention focusing, task switching and working memory (the ability to temporarily hold multiple pieces of information in the mind).

Gazzaley and his colleagues first recruited around 30 participants for each of six decades of life, from the 20s to the 70s, and confirmed that multitasking skills as measured by the game deteriorated linearly with age. They then recruited 46 participants aged 60–85 and put them through a 4-week training period with a version of NeuroRacer that increased in difficulty as the player improved.

After training, subjects had improved so much that they achieved higher scores than untrained 20-year-olds, and the skill remained six months later without practice.

The scientists also conducted a battery of cognitive tests on the participants before and after training. Certain cognitive abilities that were not specifically targeted by the game improved and remained improved — such as working memory and sustained attention. Both skills are important for daily tasks, from reading a newspaper to cooking a meal.

That is significant, says Gazzaley. "NeuroRacer doesn't demand too much of those particular abilities — so it appears that the multitasking challenge may put pressure on the entire cognitive control system, raising the level of all of its components."

The team also recorded brain activity using electroencephalography while participants played NeuroRacer. As their skills increased, so did activity in the prefrontal cortex of the brain, which is associated with cognitive control, in a manner that correlated with improvements in sustained-attention tasks. Activity also increased in a neural network linking the prefrontal cortex with the back of the brain.

The industry that has grown around selling brain-training computer games has polarized opinion about the effectiveness of brain-training packages, says cognitive neuroscientist Torkel Klingberg of the Karolinska Institute in Stockholm. "Some companies are not based on real science, and have made unrealistic claims," he says. "On the other hand, some psychologists have claimed that working memory and attention are fixed and can't be trained up."

But Gazzaley's study confirms that cognitive function can be improved — if you design training methods properly, says Klingberg, who is a consultant for Cogmed, a company he founded in 1999 to market computer-based training methods, particularly for people with attention-deficit disorders[3].

Last year, Gazzaley also co-founded a company, called Akili, for which he is an adviser. It is developing a commercial product similar to NeuroRacer, which remains a research tool, and will seek approval from the US Food and Drug Administration to market it as a therapeutic agent. A 'games' approach might also help people with particular cognitive deficits, such as depression or schizophrenia, adds Daphne Bavelier, a cognitive neuroscientist at the University of Geneva in Switzerland, who develops computer games to improve brain function and who also advises Akili.

Gazzaley cautions against over-hyping: "Video games shouldn't now be seen as a guaranteed panacea." But Linsey, for her part, is happy with what the game did for her and about her own contribution. "It's been exciting to discover the older brain can learn — and I'm glad my own brain helped make the discovery." ■

1. Owen, A. M. *et al. Nature* **465,** 775–778 (2010).
2. Anguera, J. A. *et al. Nature* **501,** 97–101 (2013).
3. Klingberg, T. *Trends Cogn. Sci.* **14,** 317–324 (2010).

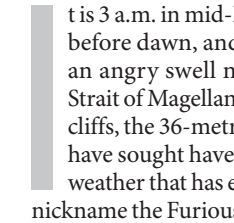
CORRECTION

The News Feature 'Egg engineers' (*Nature* **500,** 392–394; 2013) wrongly said that second-generation primordial germ cells (PGCs) are often not normal — they usually are. It is first-generation PGCs that often display abnormalities.

How Eric Karsenti's quest to understand the cell launched a trip around the world.

BIOLOGY ON THE HIGH SEAS

BY CLAIRE AINSWORTH



Eric Karsenti spent a total of three months aboard *Tara* during her 938-day plankton-sampling expedition.

JULIEN GIRARDOT

It is 3 a.m. in mid-December 2010, an hour before dawn, and *Tara* is anchored amid an angry swell near the entrance to the Strait of Magellan. In the lee of Argentina's cliffs, the 36-metre schooner and her crew have sought haven against the impending weather that has earned these latitudes the nickname the Furious Fifties. But as the winds gather to hurricane force, they snap the safety cord that eases tension on the anchor chain. As the crew struggles to replace it, the storm crescendoes, plucking *Tara*'s anchor from the seabed twice before it can be secured again.

Bracing himself inside the aluminium hull is Eric Karsenti. Compact and bright-eyed beneath a bushy mop of white hair, he has the air of a seasoned seafarer. But he is also an accomplished molecular cell biologist who has spent most of his career studying microtubules — rod-like structures that form part of the cell's internal scaffolding. Approaching retirement, he has switched fields, borrowed a famous fashion designer's yacht and launched a 2.5-year expedition around the world to survey ocean ecosystems in unprecedented breadth and detail. Motivated by a love of adventure as much as by science, Karsenti has found his share of both.

"This was really crazy," Karsenti says later, reflecting on the trip from the safety of the European Molecular Biology Laboratory (EMBL) in Heidelberg, Germany. "I was on the boat for the worst leg of the whole expedition."

The project, called Tara Oceans, set sail from Lorient, France, in September 2009 for a 115,000-kilometre voyage to collect plankton — microscopic marine organisms — at 154 distinct sites around the world. Findings from the expedition are now starting to be published, and the bulk of the data will soon be made publicly available.

Although other surveys, such as the 2004–06 Global Ocean Sampling Expedition, piloted by genomics impresario Craig Venter, have sampled microscopic life in the seas (see *Nature* **446,** 240–241; 2007), the Tara Oceans project is taking a broader approach — to "study it all"[1]. Instead of focusing only on microbes, the scientists collected billions of organisms, from millimetre-scale zooplankton down to viruses 100,000 times smaller.

These marine organisms exert tremendous power over the planet, collectively forming a giant engine that drives the cycling of elements such as carbon, nitrogen and oxygen. Photosynthetic marine microbes produce about as much of the world's oxygen as do land plants. Ocean ecosystems are also hives of evolutionary activity in which countless viruses shuttle genes between organisms.

Understanding how these complex marine ecosystems work requires a holistic approach, says Karsenti. Tara Oceans scientists are identifying the plankton through a range of techniques including genomics, proteomics and automated high-throughput imaging. To link the organisms to their environments, the researchers also measured properties such as the temperature, pH and salinity of the water around each sample, which they plan to cross-reference with the biological data.

"THIS WAS REALLY CRAZY. I WAS ON THE BOAT FOR THE WORST LEG OF THE WHOLE EXPEDITION."

Although the project is limited by its ability to sample areas at only one time, "the data they collect could be used in revolutionary ways", says Jack Gilbert, a microbial ecologist at Argonne National Laboratory in Illinois. By working out how the different plankton species interact with each other and the environment, the Tara Oceans project members hope to understand how ecosystems emerge from the sum of the interactions between their parts. This huge data set, they say, will help researchers to tackle big issues, such as calculating the biodiversity in the oceans, predicting how marine organisms will respond to environmental shifts and, perhaps, gaining insight into how evolution acts on networks of organisms in ecosystems or of molecules in cells[2].

Given his research background, Karsenti would seem an odd fit for this holistic enterprise. For decades, molecular cell biology was a byword for reductionism. But through his work on microtubules, Karsenti, like a growing number of biologists, came to see that reductionism could reveal only half the story, and that cells must be understood in terms of how the interactions between their components form a greater whole.

Although best known as a lab scientist, Karsenti has been interested in the sea since childhood, when he spent summers on France's Brittany coast. Later, he taught sailing to help fund his education. As his studies progressed, however, he left his fascination with the sea mostly behind. At Paris Diderot University, Karsenti studied the typical range of science disciplines, but he also developed an interest in statistical physics, a sideline that would later change the course of his research. After obtaining a PhD in immunology and cell biology from the Pasteur Institute in Paris in 1979, Karsenti joined the lab of cell biologist Marc Kirschner, then at the University of California, San Francisco, to study the cell cycle and cell division. Karsenti was interested in mitosis — the process of cell division — particularly in the mitotic spindle, an intricate structure that forms during cell division to ensure that new cells inherit only one copy of each chromosome.

Under a microscope, the spindle, formed from microtubules, looks like rigging strung from each end of the cell. Hundreds of proteins collaborate to control this elaborate molecular machine, from the motor proteins that crawl along its filaments to the signalling molecules that synchronize its movements with specific points in the cell-division cycle.

The early 1980s was the heyday of classical molecular cell biology and its reductionist approach, and Karsenti found himself purifying proteins and performing biochemical experiments to identify individual molecules and determine their roles in the cell cycle and mitosis. But it became clear to him that this approach, although extremely valuable, could not fully explain the spindle's complex behaviour. "I started to realize that you cannot understand this by simply identifying molecules: you need to understand how they work together," he says.

In 1985 Karsenti set up his own lab at the EMBL. Returning to his interest in statistical physics, he started collaborating with EMBL biologists Thomas Surrey (now at the London Research Institute) and François Nédélec, as well as Stanislas Leibler, a physicist at the Rockefeller University in New York. They were interested in a phenomenon known as self-organization, and in how it could be applied to biology. Physicists, chemists and mathematicians have long studied self-organization to understand how disordered systems can give rise to ordered, dynamic structures such as ocean currents and gyres. These structures are not organized from the top down, but instead emerge spontaneously from the interactions of individual components. By the 1990s, biophysicists had started to apply these concepts to molecular structures in cells. Karsenti says their work was revelatory. "I suddenly saw how

you can with formulas encapsulate the whole biological process," he says.

He also saw links to his own work on the mitotic spindle, which meets three key criteria for a self-organizing system. First, it seems to form a stable structure or pattern, but is actually highly dynamic: its individual tubules are constantly building and demolishing themselves. Second, it requires energy to keep itself going. And third, it involves interactions between fixed, or deterministic, factors — such as the forces exerted by the motor proteins — and random, or stochastic, factors, such as the probability of a motor protein meeting a microtubule. The final product, the spindle, cannot be explained by any simple cause, but instead emerges from this network of interactions.

Karsenti was among a number of researchers who were applying these principles to build computational models of spindle formation. In 2001, he and his colleagues showed how individual interactions between microtubules and motor proteins could result in the emergence of different structures, such as vortices or star-shaped 'asters', at the ends of the spindle[3]. Altering the conditions of these interactions changed the organization: higher concentrations of motor proteins led to more aster formation, for example, whereas slowing down the rate at which the motors move along the microtubules reduced the number of asters.

"I was struck by the similarity to what happens in the formation of fish schools and flocks of birds that can also be modelled using stochastic simulation," says Karsenti. "Yet, both phenomena occur at entirely different scales."

ANCHORS AWEIGH

By 2006, with compulsory retirement looming, Karsenti says, "I had the feeling that I had to do something different." Keen to raise public awareness of biology and evolution, he took some inspiration from Charles Darwin's *The Voyage of the Beagle* (1839). "He really describes in this book how travelling helped him to ask important, fundamental questions about the evolution of Earth and the evolution of life," says Karsenti.

Launching a similar expedition seemed a promising idea — and a timely one, given the impending bicentennial, in 2009, of Darwin's birth (see *Nature*'s Darwin 200 special, www.nature.com/darwin). Karsenti discussed the idea with a friend and colleague of more than 30 years — Christian Sardet, a biologist at France's Oceanological Observatory of Villefranche. They floated a plan to produce a reality-television show centred around such a voyage. When that fell through, Sardet introduced Karsenti to Gaby Gorsky, an oceanographer at the observatory, and the three repaired to a local watering hole, Le Cockpit. "We started to dream," says Gorsky. "If we can get an interesting boat, what kind of project can we push forward?"

Both Gorsky and Sardet work on plankton, and the discussions about their research reawakened Karsenti's interest in ocean life. The three decided that it would be fascinating to study planktonic ecosystems as a whole. Other researchers had already begun to apply the principles of cellular systems biology to ecosystems, using total gene sequences, or 'metagenomes', of microbial communities to pick apart all the metabolic processes they are carrying out. Ecosystem modellers were showing how ocean circulation, together with differences in factors such as nutrient content and temperature, cause microbial species to self-organize into defined geographical distributions[4]. Karsenti could see the basic principle of deterministic and stochastic factors working together to create a dynamic system.

> "HE WAS ALWAYS READY TO DO THE DIFFICULT STUFF. HE WAS THE FIRST TO CLEAN THE DISHES OR THE TOILET."

The models were elegant, but they lacked data. Karsenti, Sardet and Gorsky aimed to rectify that deficit — all they needed was to find a boat and money. Through some of Karsenti's old sailing contacts, they found *Tara*.

A schooner is an unlikely candidate for oceanographic research, but *Tara* had a sound science pedigree. She was built in 1989 as a polar-exploration yacht and later owned by Peter Blake, a legendary explorer from New Zealand. After Blake was murdered by pirates in 2001, the schooner was bought and renamed by Etienne Bourgois, a keen yachtsman and the son of the French fashion designer Agnès B. The family established Tara Expéditions, a non-profit organization that uses the vessel to support and promote environmental research. The foundation's outreach work, as well as the boat herself, made *Tara* the perfect choice for a scientific expedition aimed at popularizing biology. And, Karsenti says, "A sailing boat is much more romantic than a motor boat."

Thanks to his connections in the French competitive yachting community, Karsenti was able to arrange a meeting with Bourgois. The pair co-founded the Tara Oceans project and eventually secured financial backing from a range of funders, including the French government and the EMBL. Karsenti now needed to gather his scientific team.

He cast his net wide to secure scientists with the right expertise, including oceanographers, microbial ecologists, taxonomists, cell and systems biologists, bioinformaticians and ecosystem modellers. The project snowballed, eventually including more than 100 researchers, who took turns working aboard *Tara*. Also taking part were non-scientists such as French journalist Vincent Hilaire, who blogged from the ship about the expedition's adventures (see 'Gathering data, logging adventure'). Some 5,000 children also visited *Tara* over the course of the voyage, and the project's outreach continues online with projects such as *The Plankton Chronicles*, a video series about the strange organisms the team found.

NATURE.COM
For a slideshow from the Tara Oceans expedition, visit:
go.nature.com/jckavb

STORMY SCIENCE

Doing science aboard *Tara* posed a unique set of challenges. A sailing boat was cheaper and more environmentally friendly than a motor vessel, but dangling delicate measuring equipment hundreds of metres down into the water demanded a stable platform, hard to achieve with a boat of *Tara*'s relatively small size. And because Tara Oceans selected discrete, identifiable water masses as sampling stations, the boat had to get to the right place at the right time and under the right weather conditions. Satellite remote sensing allowed the team to monitor oceanographic features such as currents and gyres in near real-time, and detailed forecasts helped to predict whether the elements were going to cooperate.

They often did not. On the white-knuckle run down the coast of Argentina in December 2010, for example, the researchers wanted to sample from three different water masses. For three weeks they endured a succession of low-pressure systems sweeping through the area. *Tara*'s sturdy design meant they could weather even the storms that nearly robbed them of their anchor, but they had to time their stops carefully. "We tried to reach the sampling station when we knew the wind would go down to 20 knots," says Karsenti. "Then we had 20 hours to do the sampling, then bwough! It started again."

Cramming all the work into the brief weather windows then stowing the samples in *Tara*'s bows as she ploughed to the next station was physically and emotionally demanding. Karsenti — on board for four legs of the voyage, a total of three months — led by example. "He was always ready to do the difficult stuff," says Gorsky. "He was the first to clean the dishes or the toilet — no problem. He was never hiding behind his status as leader."

Back on land, the challenge for researchers has been integrating and interpreting the genomic, imaging and environmental data from the project, and making it accessible to the research community. They analysed 27,882 samples, using DNA and RNA sequencing to estimate the diversity of organisms present and study their gene functions and ecological roles. The researchers are also developing high-throughput imaging systems to scan samples and automatically identify and quantify the species present — a task that

SOURCE: TARA EXPÉDITIONS

GATHERING DATA, LOGGING ADVENTURE

In addition to scientists and sailors, *Tara*'s passengers included journalists who regularly wrote about their adventures during the 2.5-year expedition.

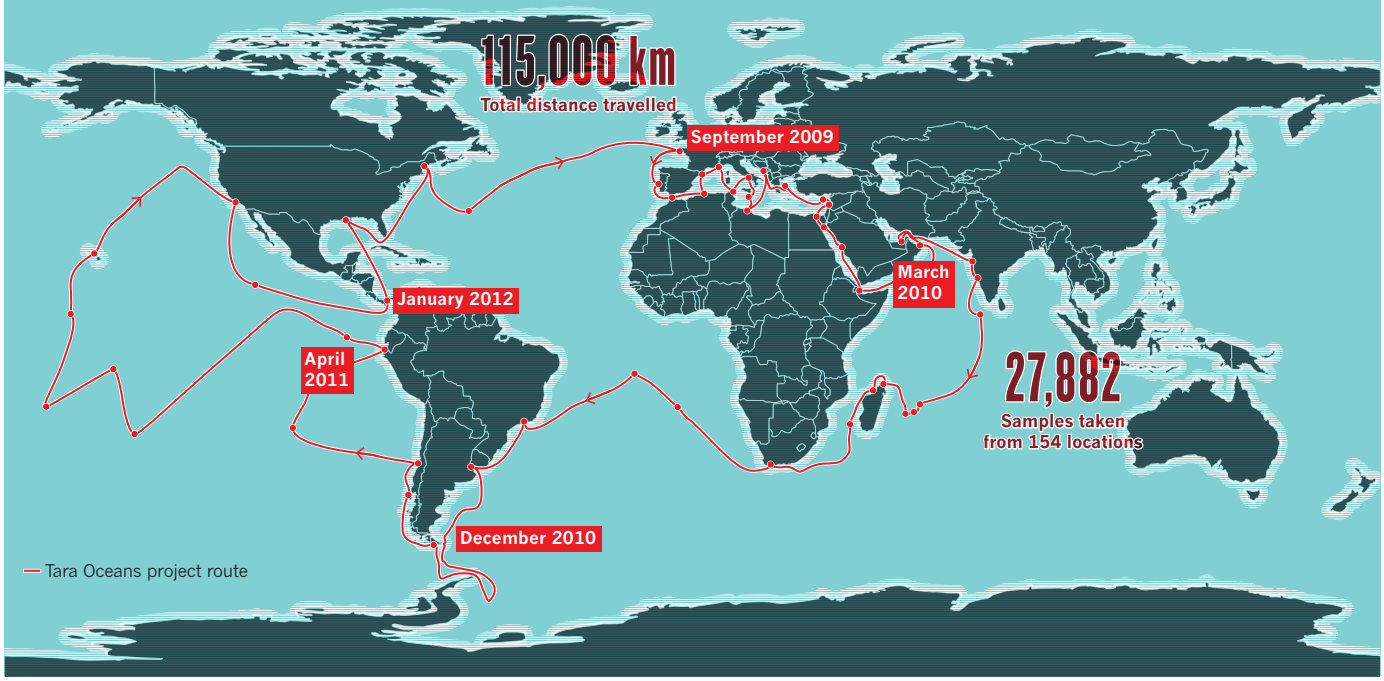
September 2009: *Tara* sets sail from the Brittany coast: *"It is now that the travelling and the scientific work begin."* — Sacha Bollet

March 2010: The crew locates a massive plankton bloom in the Gulf of Oman. *"Tara found herself transformed into a proper beehive, what with everyone lending a hand and busying around enthusiastically."* — Jérôme Bastion

December 2010: *Tara* hunkers down near Argentina. *"The gusts of wind were so strong that we could hardly stand up. Around the boat, the water was fuming."* — Vincent Hilaire

April 2011: The crew reaches its 100th sampling station, 2,200 kilometres off the coast of Ecuador. *"The endurance of the teams and the upkeep of the equipment have proven to be the key to the success of the Tara Oceans expedition."* — Anna Deniaud

January 2012: *Tara* scientists take samples from both ends of the Panama Canal to compare the biodiversity of the oceans it connects. *"The new team will continue the work of their predecessors, smoothly performing their first sampling station."* — Yann Chavance



would take an eternity to do by hand.

Preliminary data are rolling in, and the numbers presented by Karsenti at the 8th Annual Genomics of Energy and Environment meeting in Walnut Creek, California, this March, are impressive. Genome data suggest that the oceans hold hundreds of thousands or even a million different kinds of eukaryotes, nearly all of which are not yet known to science. The researchers are now using single-cell sequencing to study these organisms in more detail. There is also an abundance of diverse viruses, including a number of giant viruses — enigmatic viruses that are larger than many bacterial cells. One litre of water from the ocean's upper, sunlit reaches contains about 45 million of them.

SEA OF POSSIBILITY

Beyond the big numbers, the team is working to deduce potential ecological relationships. Seeing which organisms occur alongside others in samples, for example, can suggest possible interactions. Hiroyuki Ogata, a Tara Project collaborator and microbiologist at the Mediterranean Institute of Microbiology in Marseille, France, and his team dug through Tara Oceans data and found that a family of giant viruses called Megaviridae occurred alongside filamentous organisms known as oomycetes[5]. Evidence of gene transfer between the two organisms gives the first hints, say the researchers, that oomycetes might be hosts for giant viruses. Karsenti says that such discoveries show how Tara Oceans data could help to unpick the parasitic and symbiotic relationships that have shaped evolution in the ocean.

Tara Oceans is now taking part in an arctic sampling mission, and scientists around the world will be able to access data from its first global circuit later this year, when the raw sequences, together with the environmental measurements, will be released in an open-access database hosted by the EMBL European Bioinformatics Institute in Hinxton, UK.

Researchers will then be able to compare the Tara Oceans data with results from other large marine surveys, such as Venter's Global Ocean Sampling Expedition and Spain's 2010–11 Malaspina project, which focused on samples from deep marine ecosystems. They can also compare the data with the plethora of smaller, more local marine-ecological studies that sample the same area over time so that scientists can see how the ecosystems change. Gilbert says that it is crucial that the data from Tara Oceans be viewed in the context of other longitudinal studies, because the static view from a single voyage is limited. "This has always been the criticism of these kinds of biogeographic surveys," he says.

Robert Friedman, chief operating officer at the J. Craig Venter Institute in San Diego, California, agrees: "For us to truly understand the oceans, one was not sufficient; two is probably not sufficient. We're going to need many more."

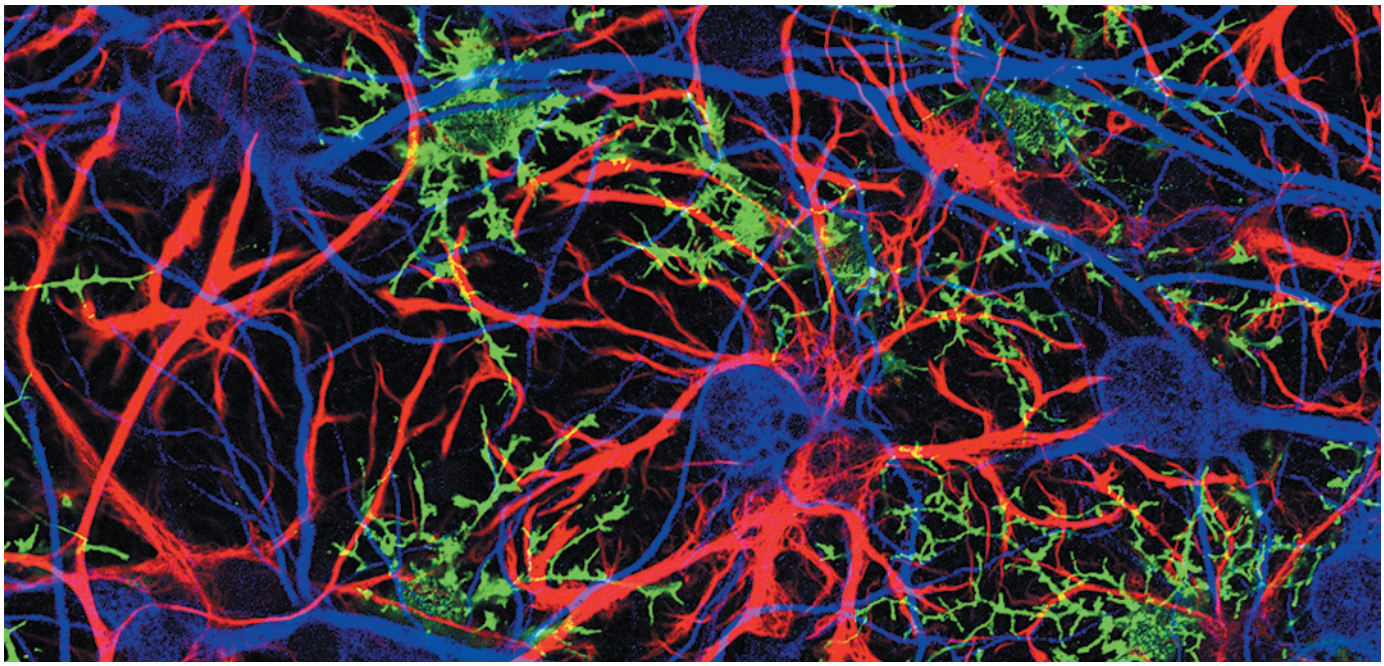
Karsenti hopes other researchers will take up that challenge and that some of them will, as he did, step outside their comfort zones to do so. Today's cell and developmental biologists, he suggests, should look beyond their yeasts, fruitflies and mice, and observe the extraordinary creatures inhabiting the seas. Perhaps future cell biologists will look first to the oceans, inspired by the tale of a sailor who peered into the heart of a single cell and caught a glimpse of the world. ■

Claire Ainsworth *is a freelance science writer in Southampton, UK.*

1. Karsenti, E. *et al. PLoS Biol.* **9,** e1001177 (2011).
2. Karsenti, E. *Mol. Syst. Biol.* **8,** 575 (2012).
3. Surrey, T., Nédélec, F., Leibler, S. & Karsenti, E. *Science* **292,** 1167–1171 (2001).
4. Follows, M. J., Dutkiewicz, S., Grant, S. & Chisholm, S. W. *Science* **315,** 1843–1846 (2007).
5. Hingcamp, P. *et al. ISME J.* **7,** 1678–1695 (2013).

COMMENT

JONATHAN COHEN



Astrocytes (red) and immature oligodendrocytes (green), types of glial cell, intertwine with neurons (blue) from the brain's hippocampus.

Map the other brain

Glia, the non-neuronal cells that make up most of the brain, must not be left out of an ambitious US mapping initiative, says **R. Douglas Fields**.

The Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative announced by US President Barack Obama in April seeks to map and monitor the function of neural connections in the entire brains of experimental animals, and eventually in the human cerebral cortex. Several researchers have raised doubts about the project, cautioning that mapping the brain is a much more complex endeavour than mapping the human genome, and its usefulness more uncertain.

I believe that exploring neural networks and developing techniques with which to do so are important goals that should be vigorously supported. But simply scaling up current efforts to chart neural connections is unlikely to deliver the promised benefits — which include understanding perception, consciousness, how the brain produces memories, and the development of treatments for diseases such as epilepsy, depression and schizophrenia[1].

A major stumbling block is the project's failure to consider that although the human brain contains roughly 100 billion neurons, it contains billions more non-electrical brain cells called glia[2]. These reside outside the neuronal 'connectome' and operate beyond the reach of tools designed to probe electrical signalling in neurons. Dismissed as connective tissue when they were first described in the mid-1800s, glia have long been neglected in the quest to understand neuronal signalling.

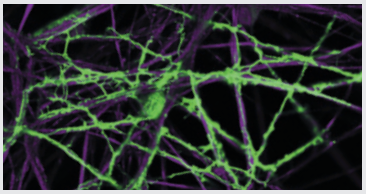
Research is revealing that glia can sense neuronal activity and control it[3]. Various studies also indicate that glia operate in diverse mental processes, for instance, in the formation of memories. They have a central role in brain injury and disease, and they are even at the root of various disorders — such as schizophrenia and Alzheimer's — previously presumed to be exclusively neuronal. That the word 'glia' was not uttered in any of the announcements of the BRAIN Initiative, nor written anywhere in the 'white papers' published in 2012 and 2013 in prominent journals outlining the ambitious plan[1,4], speaks volumes about the need for the community of neuroscientists behind the initiative to expand its thinking.

All major glial cell types in the brain — oligodendrocytes, microglia and astrocytes — communicate with each other and with neurons by using chemical neurotransmitters and gap junctions, channels that permit the direct transfer between cells of ions and small molecules (see 'Roles of glial cells').

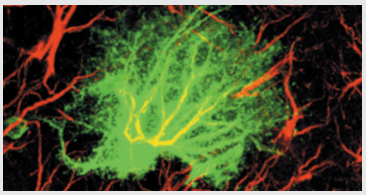
Oligodendrocytes produce the myelin sheath, the insulating material that ▶

THE OTHER HALF

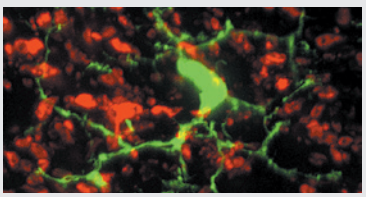
Roles of glial cells



Oligodendrocytes (green)
- Form myelin electrical insulation, increasing conduction velocity by at least 50 times.
- Provide vital metabolic support for axons (purple).
- Involved in multiple sclerosis, amyotrophic lateral sclerosis and the inhibition of repair after spinal-cord injury.



Astrocytes (red and green)
- Ensheath synapses, regulate neuronal excitability and synaptic transmission.
- Respond to injury by secreting extracellular matrix proteins.
- Implicated in neurogenesis, cell migration, and many neurological and psychiatric disorders.



Microglia (green)
- Highly motile and responsive to nervous-system injury and infection.
- Monitor electrical activity in neurons and prune synaptic connections (red).
- Involved in almost all nervous-system diseases and in certain psychiatric conditions.

OLIGODENDROCYTES: R. DOUGLAS FIELDS; ASTROCYTES: ULRIKA WILHELMSSON/ERIC BUSHONG/MARK ELLISMAN; MICROGLIA: BETH STEVENS

surrounds nerve fibres called axons. This sheath greatly increases the speed at which electrical signals are transmitted through axons — a crucial feature because nervous-system function depends on information being transmitted at high speed.

Biologists have known for decades that microglia — the immune cells of the brain — respond to infection and brain trauma by clearing away diseased or damaged tissue and releasing substances that stimulate repair. But last year, a study[5] of how eye–brain connectivity is established in developing mice showed that microglia also prune back synapses and rewire neural connections in a healthy brain, depending on the individual's visual experience shortly after birth.

Various studies[3] have revealed that astrocytes regulate the transmission of electrical signals across synapses by modifying the concentration of extracellular potassium; controlling local blood flow; releasing and taking up neurotransmitters and other neuromodulatory substances; delivering nutrients to neurons; and altering the geometry and volume of space between brain cells. All of these things influence nervous-system communication and plasticity.

When experts on neuronal plasticity and computational neuroscience came together with glial experts at a workshop in February entitled Glial Biology in Learning and Cognition[6], held at the US National Science Foundation in Arlington, Virginia, our unanimous conclusion was that neurons working alone provide only a partial explanation for complex cognitive processes, such as the formation of memories. The complex branching structure of glial cells and their relatively slow chemical (as opposed to electrical) signalling in fact make them better suited than neurons to certain cognitive processes. These include processes requiring the integration of information from spatially distinct parts of the brain, such as learning or the experiencing of emotions, which take place over hours, days and weeks, not in milliseconds or seconds.

Pioneering research on glia is even offering glimpses into the mechanisms that underlie learning, other forms of plasticity and information processing. For instance, my research group has established that, *in vitro*, the impulse traffic in an axon can control its myelination by glia[7]. Because myelination determines the speed at which an electrical signal is transmitted through an axon, it can influence how many inputs fire on to a neuron at the same time — a process that is fundamental to most learning and synaptic plasticity. When people acquire new skills, from juggling to playing computer games, the structure of specific myelinated regions of their brains changes[8].

We are only beginning to learn about the diversity, connectivity and function of astrocyte networks. Various studies suggest that astrocytes have anatomical and physiological properties that may impose higher-order organization on information processing in the brain. In the grey matter of the cerebral cortex and hippocampus, astrocytes are organized in non-overlapping domains. The significance of this tile-like organization is unknown, but one human astrocyte (an intricate, bush-like cell) can encompass, and therefore influence, two million synapses[9]. In fact, human astrocytes are markedly different from those of other animals. Mice that have had a proportion of their astrocytes replaced with human ones show increased synaptic plasticity and faster learning[9] in a range of tests.

Neuroscientists have known for several decades that glia cause certain diseases. Nearly all cancers originating in the brain derive from glia (which, unlike mature neurons, undergo cell division). In multiple sclerosis, the myelin sheaths around axons become damaged, and in HIV-associated neurological conditions, the virus infects astrocytes and microglia, not neurons. Many neurological disorders that were once considered to be exclusively neuronal in nature are now also known to involve glia, including Rett syndrome, a neurodevelopmental condition that includes autism-like symptoms, the motor-neuron disease amyotrophic lateral sclerosis, Alzheimer's disease and chronic pain. The same is true for various other developmental and psychiatric conditions such as schizophrenia, depression and obsessive–compulsive disorder[2,3].

Mapping and monitoring the entire human cerebral cortex is expected to require an investment comparable to that needed for the Human Genome Project: US$3.8 billion[1]. The public will not support such a major expenditure of public funds for a scientific research project if the promised benefits are oversold, if they do not understand the benefits or if the costs are not realistic. This was demonstrated in 1993 by the US Congress's cancellation of a multibillion-dollar project to build a superconducting super collider, the construction of which was already under way in Waxahachie, Texas.

AGENDA FOR SUCCESS

Unlike the quest to capture the Higgs boson or to land a human on the Moon, few people question the need to understand the human brain. The brain is the greatest enduring mystery of the human body, and intellectual disability, brain cancer, spinal-cord injury, senile dementia and mental illnesses such as schizophrenia and depression touch the lives of almost everyone. But on its current trajectory, the BRAIN Initiative risks failure — both scientifically and in terms of public support.

At the World Science Festival in New York in June, one participant argued that information about glia will be a by-product of neuronal mapping and of the development of technologies to trace and record neuronal connections. This is implausible. New methods to monitor electrical signalling in neurons, such as voltage-sensitive dyes or nanoparticles that sense the electrical potential across membranes, will be of little

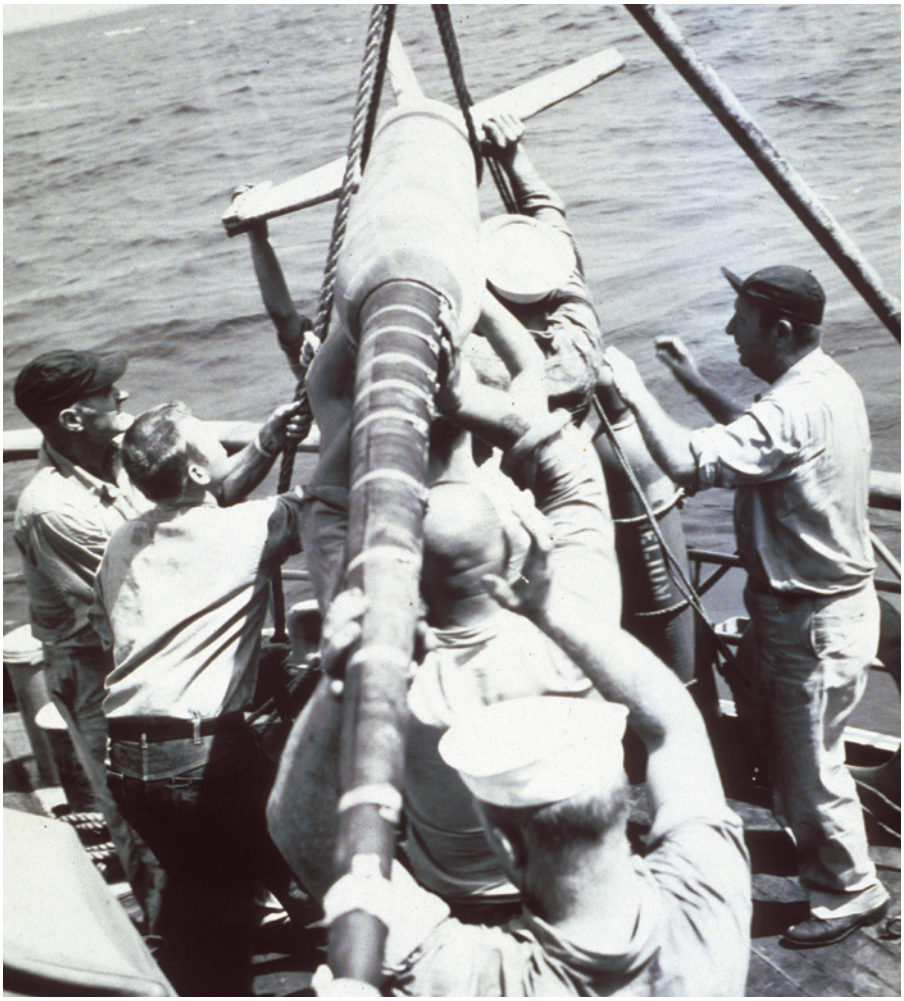
use for understanding cells that do not communicate using electrical impulses. It is this view that has perpetuated our comparative ignorance about glia.

Moreover, the exclusion of glia from the BRAIN Initiative underscores a more general problem with the project: the assumption that enough measuring of enough neurons will in itself uncover 'emergent' properties and, ultimately, cures for diseases[1,4]. Rather than simply materializing from measurements of "every spike from every neuron"[1,4], better understanding and new treatments will require hypothesis-directed research. The 302 neurons and 7,000 connections that make up the nervous system of the roundworm *Caenorhabditis elegans* were mapped in the 1970s and 80s. More than two decades later, little is understood about how the worm's nervous system produces complex behaviours.

In any major mapping expedition, the first priority should be to survey the uncharted regions. Our understanding of one half of the brain (the part comprised of astrocytes, oligodendrocytes and microglia) lags a century behind our knowledge of neurons. I believe that answers to questions about the brain and public support for a large-scale study are more likely to come from expanding the search into this unknown territory. As a first step, tools such as optogenetic methods and mathematical models are needed to assess the number, distribution and properties of different kinds of glial cell in different brain regions, and to identify how glia communicate with each other and with neurons, and what developmental and physiological factors affect this. This exploration into the 'other brain' must be done together with the proposed studies of neurons. It cannot be achieved as a by-product of them. ■

R. Douglas Fields *is chief of the Nervous System Development and Plasticity Section at the US National Institutes of Health in Bethesda, Maryland.*
e-mail: fieldsd@mail.nih.gov

1. Alivisatos, A. P. *et al. Science* **339,** 1284–1285 (2013).
2. Kettenmann, H. & Ransom, B. R. (eds) *Neuroglia* 3rd edn (Oxford Univ. Press, 2013).
3. Fields, R. D. *The Other Brain* (Simon & Schuster, 2009).
4. Alivisatos, A. P. *et al. Neuron* **74,** 970–974 (2012).
5. Schafer, D. P. *et al. Neuron* **74,** 691–705 (2012).
6. Fields, R. D. *et al. The Neuroscientist* (in the press).
7. Wake, H. Lee, P. R. & Fields, R. D. *Science* **333,** 1647–1651 (2011).
8. Zatorre, R., Fields, R. D. & Johansen-Berg, H. *Nature Neurosci.* **15,** 528–536 (2012).
9. Han, X. *et al. Cell Stem Cell* **12,** 342–353 (2013).



The US research vessel *Explorer* towed a magnetometer to map fields over the sea floor in 1960.

How plate tectonics clicked

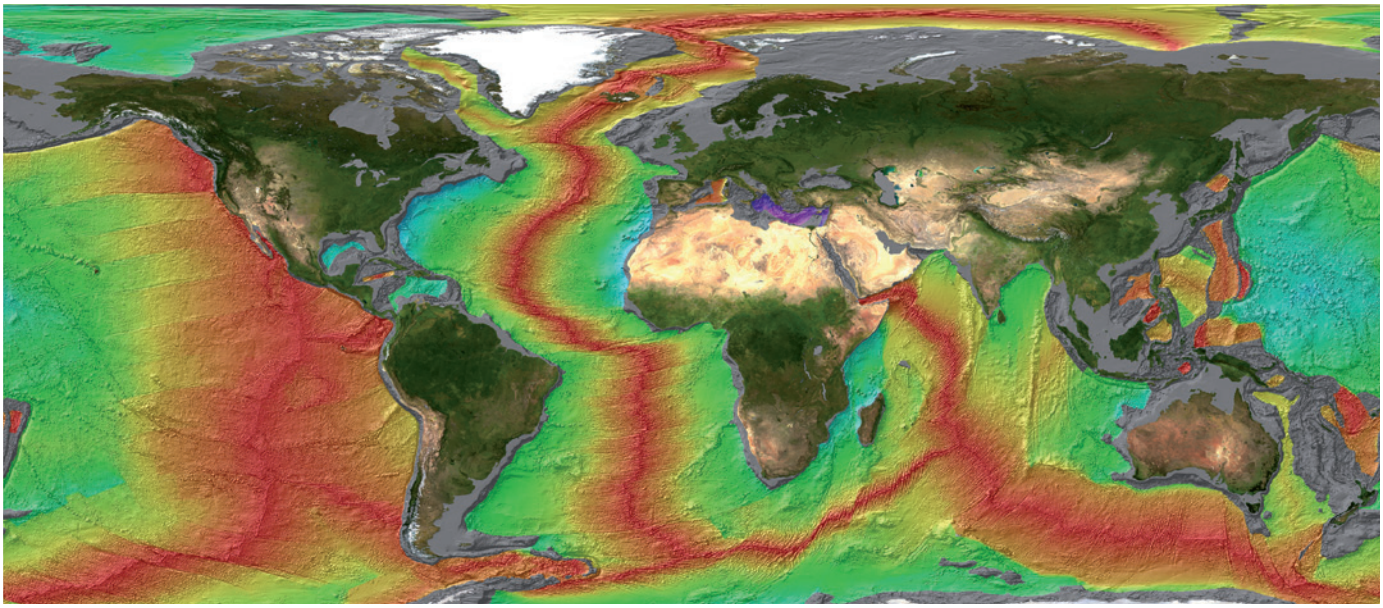
Fifty years after a paper linked sea-floor magnetic stripes with continental drift, **Naomi Oreskes** explains its legacy as a lesson in achieving scientific consensus.

By the time German geophysicist Alfred Wegener proposed continental drift in 1912, palaeontologists had long accepted that past connections between now-separate lands explained the spread of similar fossils and rock layers across them. Geologists, too, knew of slabs of Alpine rock that had been displaced hundreds of kilometres during mountain building.

But the arguments for continental motions did not gel until the 1960s, when a drastic expansion of geophysical research, driven by the cold war, produced evidence that reopened and eventually settled the debate.

One influential study was published[1] in *Nature* 50 years ago this week. British geologists Frederick Vine and Drummond Matthews interpreted stripes of alternating magnetic-field polarity in ocean bedrock as evidence of a spreading sea floor that pushed continents apart. Acceptance that large crustal motions were a reality soon followed, culminating in the theory of plate tectonics.

In its slow convergence of ideas and evidence, the history of plate tectonics holds lessons for today's debates about ▶



ELLIOT LIM/CIRES/NOAA/NATL GEOPHYS. DATA CENT.

The age of ocean rocks increases (red to purple, 0–280 million years) with distance from ridges, where crust is formed, revealing the spread of the sea floor.

human-induced climate change. Although science is always evolving, and our attention is drawn to controversy at the research frontier, it is the stable core of 'consensus' knowledge that provides the best basis for decision-making.

MANTLE CONVECTION

Wegener stands out because his solution was close to the one that we now accept, and because our individualist culture encourages us to look for heroes to credit and discrete events to celebrate. But he was not alone in trying to explain commonalities in fossils and rock strata. In the English-speaking world, two of the most important players in developing theories of continental-scale crustal mobility were South African field geologist Alexander du Toit and British geochronologist Arthur Holmes.

Du Toit articulated the case in his aptly named 1937 book *Our Wandering Continents* (Oliver and Boyd). He acted as a clearing house for geologists around the globe, who sent him maps, rocks and fossils. Holmes, working with the Irish geochemist John Joly, suggested that crustal motion was driven by radioactivity and the heat that it emanates, advocating mantle convection as a means of dissipating radiogenic heat and driving continental drift[2]. Holmes's 1944 textbook *Principles of Physical Geology* (Thomas Nelson & Sons) was an introduction to the subject for many students.

The discussion was joined by Dutch geodesist Felix Vening Meinesz, who worked in the 1930s in the Indonesian archipelago and, with US geologists Harry Hess and Maurice Ewing, in the Caribbean. Meinesz found that Earth's gravitational field was weaker than normal above some of the ocean's deepest regions, which he explained in terms of the buckling of low-density crust into the mantle, dragged down by descending convection currents, and he discussed these ideas with Hess.

During the Second World War, Hess found himself in the US Navy, fighting in the Pacific theatre. He did not return immediately to tectonics after the war, but others did, including several British geophysicists led by P. M. S. Blackett and Keith Runcorn. In an effort to understand the origins of Earth's magnetic field, they discovered that magnetic minerals pointed in different directions at different times in geological history, as if the positions of the poles had changed. Hess was drawn back to the topic after realizing that these 'apparent polar-wandering paths' could be explained by the movements of the continents.

"More than two dozen scientists did the key work that created the theory of plate tectonics."

OCEAN SPREADING

Hess suggested that rising mantle-convection cells would drive apart the ocean floor above them, increasing the separation of continents to either side. The idea, which his colleague Robert Dietz christened 'sea-floor spreading'[3], explained the old geological observations and the new geophysical ones, but it did not gain immediate traction. That would take further geomagnetic information.

Blackett, a socialist who opposed nuclear proliferation, turned to geomagnetism after the war to distance himself from military work[4]. But military concerns — particularly the demands of submarine warfare in the atomic age — drove geophysical exploration of the ocean floor, leading to the discovery in the late 1950s of sea-floor magnetic stripes.

The stripes were a surprise. In the report of the discovery, oceanographers Ronald Mason and Arthur Raff admitted to being at a loss for an explanation. Others were less stymied. Vine and Matthews, as well as Canadian geophysicist Lawrence Morley, independently had the same idea. If the sea floor was spreading, then magnetic stripes would be expected: rock formed at mid-ocean ridges would take on Earth's magnetic field, the polarity alternating as the field periodically reversed.

It was one thing to say that the oceans were widening, another to link it to global crustal motion. More than two dozen scientists, including women such as Tanya Atwater and Marie Tharp, did the key work that created the theory of plate tectonics as we know it — explaining continental drift, volcanism, seismicity and heat flow around the globe[5].

In 1965, Canadian geologist Tuzo Wilson proposed a type of 'transform' fault to accommodate the spreading sea floor around mid-ocean ridges, which was confirmed by US seismologist Lynn Sykes. Other seismologists demonstrated that in deep-ocean trenches, slabs of crust were indeed being driven into the mantle, and geophysicists worked out how these crustal 'plates' move and relate to the features of continental geology.

Vine and Matthews' work is part of a larger story of the growth of Earth science in the twentieth century, made possible by improved technology and greater governmental support after the Second World War. Nearly all seismic and marine geophysical data at the time were collected with military backing, in part because of their cold-war security significance.

This era marked a change in the character of modern science. Research today is expensive and largely government-funded; almost

all major scientific accomplishments are the collective achievement of large teams. This reality — more prosaic than the hagiography of lonely genius — reminds us that although great individuals are worthy of recognition, the strength and power of science lies in the collective effort and judgement of the scientific community.

CONSENSUS MATTERS

In recent months, several of my colleagues in climate science have asked me whether the story of plate tectonics holds lessons for their field in responding to those who disparage the scientific evidence of anthropogenic climate change. I believe that it does.

Many critics of climate science argue that expert agreement is irrelevant. Science, they claim, advances through bold individuals such as Wegener or Galileo Galilei overturning the status quo. But, contrary to the mythology, even Isaac Newton, Charles Darwin and Albert Einstein worked within scientific communities, and saw their work accepted. In glorifying the lone genius, climate-change dissenters tap into a rich cultural vein, but they miss what consensus in science really is and why it matters.

Consensus emerges as scientific knowledge matures and stabilizes. With some notable exceptions, scientists do not consciously try to achieve consensus. They work to develop plausible hypotheses and collect pertinent data, which are debated at conferences, at workshops and in peer-reviewed literature. If experts judge the evidence to be sufficient, and its explanation coherent, they may consider the matter settled. If not, they keep working. History enables us to judge whether scientific claims are still in flux and likely to change, or are stable, and provide a reasonable basis for action.

And maturity takes time. Scientific work, compared with industry, government or business, has no deadline. Perhaps for this reason, when Wegener died in 1930, according to his biographers he was confident that other scientists would one day work out how the continents moved, and that this mechanism would be along the lines of his proposal — as indeed it was. Du Toit and Holmes were similarly convinced.

The equanimity of these men speaks to their confidence in science as a system. They perceived what historian–philosopher Thomas Kuhn articulated in *The Structure of Scientific Revolutions* (University of Chicago Press, 1962): that science is a community affair and that knowledge emerges as the community as a whole accepts it. A debate comes to a close once scientists are persuaded that a phenomenon is real and that they have settled on the right explanation. Further discussion is not productive unless new evidence emerges, as it did for continental drift.

Anthropogenic climate change has the consensus of researchers. Political leaders who deny the human role in climate change should be compared with the hierarchy of the Catholic church, who dismissed Galileo's arguments for heliocentrism for fear of their social implications. But what of scientists who in good faith reject the mainstream view?

Harold Jeffreys is an intriguing example. An eminent professor of astronomy at the University of Cambridge, UK, Jeffreys rejected continental drift in the 1920s and plate tectonics in the 1970s. He believed that the solid Earth was too rigid to permit mantle convection and crustal motion. His view had a strong mathematical basis, but it remained unchanged, even as evidence to the contrary mounted.

If society had faced a major decision in the 1970s that hinged on whether or not continents moved, it would have been foolish to heed Jeffreys and to ignore the larger consensus, backed by half a century of research. As an early advocate of an immature theory, Wegener was different. There were substantial differences of opinion about crustal mobility among scientists in the 1920s. By the 1970s, work such as Vine and Matthews' study had brought consensus.

Fifty years on, history has not vindicated Jeffreys, and it seems unlikely that it will vindicate those who reject the overwhelming evidence of anthropogenic climate change. ■

J. R. CANN

Frederick Vine and Drummond Matthews (1970).

Naomi Oreskes *is professor of the history of science at Harvard University in Cambridge, Massachusetts.*
e-mail: oreskes@fas.harvard.edu

1. Vine, F. J. & Matthews, D. H. *Nature* **199,** 947–949 (1963).
2. Oreskes, N. *The Rejection of Continental Drift: Theory and Method in American Earth Science* (Oxford Univ. Press, 1999).
3. Dietz, R. S. *Nature* **190,** 854–857 (1961).
4. Nye, M. J. *Blackett: Physics, War, and Politics in the Twentieth Century* (Harvard Univ. Press, 2004).
5. Oreskes, N. (ed.) *Plate Tectonics: An Insider's History of the Modern Theory of the Earth* (Westview Press, 2001).

High-rise apartments tower in front of a breeze-block sprawl in Rio de Janeiro, Brazil.

ALEX ROBINSON/ AWL IMAGES/CORBIS

POPULATION

Crowd control

Hania Zlotnik assesses two polarized takes on population growth and planetary capacity.

In June, the United Nations Population Division announced that the world's population could grow from 7.2 billion today to 9.6 billion by 2050, assuming that global fertility continues to decline. Such growth constitutes a fundamental challenge for humanity, and now two thoughtful but very different takes on it explore the implications.

In *Countdown*, Alan Weisman, a journalist probing whether a sustainable balance between nature and the human population can be achieved, offers a key message to guide future action. He avers that no matter what environmental, ecological or social problem we face, it will be easier to solve with fewer people. His book provides an array of examples on how to reduce population growth and, in the process, improve prospects for future generations. He makes a strong case for slowing global population growth — and even for reducing overall population numbers — as a prerequisite for achieving a sustainable future.

10 Billion
STEPHEN EMMOTT
Allen Lane: 2013.

Countdown: Our Last, Best Hope for a Future on Earth?
ALAN WEISMAN
Little, Brown: 2013.

Stephen Emmott's *10 Billion* takes for granted that the population will continue to grow, and is much less sanguine about humanity's chances of avoiding looming crises. Emmott, head of computational science for Microsoft Research in Cambridge, UK, leads an interdisciplinary group of scientists engaged in addressing fundamental problems through complex modelling. His slim, even terse book — based on his 2012 stage show, which presented his view on the "unprecedented planetary emergency we've created" — primarily examines the transformation of the global environment by human activity, a transformation that includes climate change, increasing water shortages and growing urbanization. Emmott's assessment of the capacity of people and technology to prevent the global crises that confront us is grim.

Weisman's book, by contrast, offers hope. *Countdown* emphasizes that it is possible to curb population growth by voluntarily limiting family size. He describes how societies as different as those of Costa Rica, Iran, Japan, Kerala in India, and Thailand, have

managed to reduce the number of children per woman through the voluntary use of contraception, buttressed by strategies to promote the adoption of a small-family norm — for example, by engaging opinion makers such as government leaders, religious authorities and media icons. Weisman explains that providing contraceptives to the quarter of a billion women worldwide who wish to prevent pregnancy but are not using effective contraception would be the first order of business.

It will also be essential to shift attitudes away from prizing high fertility. To illustrate the barriers preventing rapid fertility reductions, Weisman cites the examples of Gaza, Niger, Pakistan and the Philippines, where a combination of religious beliefs, pro-natalist ideologies, the low status of women, weak health systems and a lack of easy access to effective contraception can lead to an average of five or more children per woman. He describes in harrowing detail the consequences of persistent high fertility, with respect to both its detrimental effects on the well-being of individuals and the negative impact of rapidly increasing numbers of people on the survival of other species and the integrity of essential ecosystems.

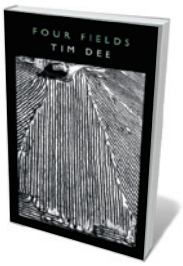
"Weisman describes in harrowing detail the consequences of persistent high fertility."

An important part of the narrative in *Countdown* hinges on how past predictions of an imminent population collapse caused by limitations in natural resources have not come to pass, largely because technology has removed those limitations, at least temporarily. Both *10 Billion* and *Countdown* provide ample evidence that such revolutionary technology is itself causing new forms of environmental stress. The strains of wheat and rice that led to the 'Green Revolution', for instance, made it possible to feed an ever-growing world population; but the crops are thirstier, which is depleting water reserves, and also demand the use of fertilizers, which are a source of pollution.

Both Emmott and Weisman consider that curbing consumption is essential for long-term sustainability, and both despair at the lack of a concerted effort to achieve it. Whereas Emmott does not present ways to address our failures, Weisman's emphasis on expanding access to contraception as the next-best strategy is both pragmatic and workable, as past efforts have shown. It is to be hoped that his message may be heeded sooner rather than later. ■

Hania Zlotnik *is a population specialist. She is the former director of the United Nations Population Division.*
e-mail: hania.zlotnik@hotmail.com

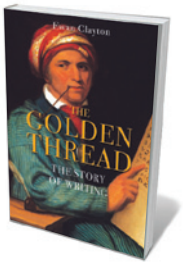
Books in brief



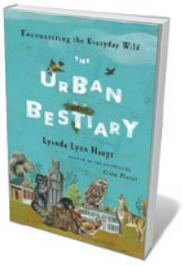
Four Fields
Tim Dee JONATHAN CAPE *(2013)*
In this vivid 'field study', radio producer and nature writer Tim Dee explores four plots of land, in England, Ukraine, the United States and Zambia. To Dee, these are grassy arenas where nature and humanity clash and merge. He is a fiercely focused witness to that drama, whether monitoring grasshoppers in the irradiated zone around Chernobyl, or on a Cambridgeshire fen listening to cuckoos calling, "deepening the place, summoning the curve of the Earth". With a lyricism richly and strangely his own, Dee maps a topography that has as much to do with the mind as with Earth.

The Bitterest Pills: The Troubling Story of Antipsychotic Drugs
Joanna Moncrieff PALGRAVE MACMILLAN *(2013)*
Psychiatrist Joanna Moncrieff delivers a meticulous, balanced history of antipsychotic drugs. In the mid-twentieth century, early forms such as chlorpromazine were created to tackle the biological roots of severe mental illness. But as she shows, new-generation 'atypicals' are marketed to treat conditions such as depression, and US spending on antipsychotics reached almost US$17 billion in 2010. Moncrieff sees some antipsychotics as having a place in treating schizophrenia, but argues that questions about efficacy and side effects such as obesity suggest they should be handled with care.

Eugene Braunwald and the Rise of Modern Medicine
Thomas H. Lee HARVARD UNIVERSITY PRESS *(2013)*
A pioneering 1971 study on myocardial infarction revealed that a heart attack is a process rather than a sudden karate chop, and hence eminently treatable. Eugene Braunwald's finding has helped to cut US heart-attack death rates from more than 30% to less than 10%. The story of the cardiologist's life and work, set against a background of rapid change in medicine, is told engagingly by medic Thomas Lee. Braunwald has lived several lives since his 1938 escape from Nazi-controlled Austria, from glory years at the US National Institutes of Health to his active ninth decade.



The Golden Thread: The Story of Writing
Ewan Clayton ATLANTIC *(2013)*
Calligrapher Ewan Clayton examines one of the earliest forms of communications technology: writing, from its inked and incised beginnings some 5,000 years ago to its digitized present. Clayton, who advised Xerox Corporation's Palo Alto Research Center in California on the use of documents, is erudite on everything from Herculaneum's ancient election 'posters' to contemporary German graffiti. Here too are Isaac Newton's first notebook, the geometry of pen nibs, 'muscular penmanship' in the nineteenth-century United States — and a palpable relish for the dance of words on a page.



The Urban Bestiary: Encountering the Everyday Wild
Lyanda Lynn Haupt LITTLE, BROWN *(2013)*
From coyotes in lifts to woodpecker holes in the *Discovery* space shuttle, odd encounters at the nature–human interface abound in this 'bestiary'. But naturalist Lyanda Lynn Haupt has a serious agenda, arguing that an appreciation of the displaced beasts re-entering our porous cities is crucial to human well-being. As she recounts the precarious existence of an opossum or the "cognitive toolbox" of a crow, Haupt urges a new sensitivity towards urban wildlife. The taxa are US-centric; the message is universal. **Barbara Kiser**

NASA/STSCI

The Hubble Space Telescope cruises above Earth in May 2009.

SPACE EXPLORATION

Probing the Universe

A celebration of uncrewed space missions is less than international in scope, finds **John Zarnecki**.

The space age has revolutionized our knowledge and understanding of the Universe, near and far. The past 60 years have seen a profusion of methods for space exploration — from Earth-orbiting satellites that, freed from the absorbing atmosphere, have opened up almost the entire electromagnetic spectrum to our view, to mostly robotic probes that have allowed us to study most types of object in our Solar System close up.

In *Dreams of Other Worlds*, Chris Impey and Holly Henry submit 12 of these space missions to close scrutiny. They describe how the missions were achieved and place them in a historical and even cultural context — this is, after all, a collaboration between an astronomer and a literary scholar. For those with a scientific background, this provides a generally refreshing perspective compared with many texts. For example, the chapter on the Mars Exploration Rovers invokes the work of H. G. Wells and Orson Welles, and, more surprisingly, that of T. S. Eliot, Virginia Woolf and Vita Sackville-West. And, to help to explain the Spitzer infrared space telescope and some of its discoveries, the authors make use of Jules Verne's 1870 science-fiction classic *Twenty Thousand Leagues Under the Sea* and James Cameron's films *Avatar* (2009) and *Aliens of the Deep* (2005).

Selecting just 12 missions is a bit like picking your all-time best football team or top ten artists. It is impossible. Impey and Henry's list includes examples that are well known beyond the scientific community, such as the Hubble Space Telescope and the Voyager spacecraft, as well as the less widely known Hipparcos star-mapping mission and the Chandra X-ray Observatory. Whatever your view of their choices, they are all well-analysed and presented in a scholarly yet engaging way. The authors remind us — if we need reminding — that in addition to these missions being wonderful technological achievements, science is at their core. Whether that is the high-energy Universe represented by X-ray astronomy, or the cold Universe revealed through infrared observations — from the interior of the Sun to the outer reaches of our Solar System — Impey and Henry are able guides. They explain the scientific imperative of these missions in a way that is accessible and interesting to specialists and generalists.

Dreams of Other Worlds: The Amazing Story of Unmanned Space Exploration
CHRIS IMPEY AND HOLLY HENRY
Princeton University Press: 2013.

I do take issue with the book's subtly US-centric world view. There is no doubt that scientifically, the US space programme has been the dominant force. The Soviet Union achieved many great firsts — such as the first successful lunar soft lander, the 1966 Luna 9, as well as the first image of the far side of the Moon, and Venera 7, the first Venus lander. But, with some notable exceptions, the scientific output of the Soviet, and subsequently Russian, programme has been relatively poor. More recently, the European Space Agency, Japan, India and China have had significant successes, including the first close fly-by of a cometary nucleus — that of Halley's comet — by Europe's Giotto spacecraft in 1986 and the return to Earth of particles from the asteroid 25143 Itokawa by the Japanese Hayabusa spacecraft in 2010.

Not to include a single mission from the Soviet Union, Russia or Japan seems to me myopic. One could easily make a case for the Venus landers Venera 13 and 14, which survived for up to two hours on the surface of the planet; or the Franco–Soviet Vega balloons, which drifted for up to two days in the atmosphere of Venus; or the Hayabusa mission.

Impey and Henry do include two European-led missions in their round dozen. These are Hipparcos, which between 1989 and 1993 measured the positions of more than 100,000 stars and other objects to unprecedented accuracy, and the Solar and Heliospheric Observatory (SOHO), launched in 1995 and still operating, observing both below and far beyond the Sun's visible surface. But, oddly, the authors first refer to SOHO as NASA's, later writing that "SOHO was conceived by the European Space Agency; fourteen countries and more than three hundred engineers were involved in its design and construction," with NASA handling launch and ground operations. Europe has also contributed hardware and finance to the Hubble Space Telescope from the very start — but there is barely a mention of this.

The authors do point out that space science is a truly international activity and has generally been carried out in an open, generous and embracing spirit, despite the occasional blip. However, the subtle but pervasive 'space nationalism' of *Dreams of Other Worlds* tarnishes what is otherwise an informative and engaging read. ■

John Zarnecki *is director of science at the International Space Science Institute in Berne. He has more than 30 years of involvement in space-science missions, including the Hubble Space Telescope, the Giotto mission and the Cassini–Huygens mission.*
e-mail: john.zarnecki@open.ac.uk

Correspondence

Brain projects need stronger foundation

The US BRAIN Initiative and the European Commission's Human Brain Project might be more usefully compared to former US President Richard Nixon's 'War on Cancer' than to the Human Genome Project (*Nature* **499,** 253, 272–274; 2013).

Despite hundreds of millions of dollars being spent after Nixon's 1971 National Cancer Act, the 'war' is still far from over. Back then, cancer scientist Sol Spiegelman remarked: "An all-out effort at this time would be like trying to land a man on the Moon without knowing Newton's laws of gravity."

The same might be said of these huge brain-mapping ventures: we have not yet cracked the neural code, and we have only the most rudimentary understanding of the signature nonlinear dynamics of brain function.

A sound understanding of underlying scientific principles was essential to successful mega-projects such as the Moon landing and the Manhattan Project. Without this, brain-mapping efforts may be premature — the equivalent of mapping the structure of snowflakes, while diverting efforts away from understanding the scientific principles that generate them.

The Human Genome Project was of immense benefit to science, but in and of itself it was a feat of engineering. It is unlikely that the project will provide an appropriate blueprint for perhaps the biggest mystery the human brain has ever pondered.

Dean Buonomano *Brain Research Institute, University of California, Los Angeles, USA.*
dbuono@ucla.edu

Step up funding to halt forensic folly

Forensic science faces challenges in many countries besides the United Kingdom (*Nature* **500,** 5; 2013), particularly in terms of its research funding and culture. If the world's leading economies want science and justice to support each other more effectively and thus allow courts to make informed decisions, forensic science needs to be adequately funded.

Leading funding organizations tend to favour transformational over applied research such as forensic science. Forensic research consequently misses out, even in this age of 'impact' rating.

Genuine breakthroughs in forensic science depend on advances in basic disciplines such as analytical chemistry, molecular biology and biochemistry. But limited funds mean that forensic science has all too often had to depend on the transfer of technology from other fields, combined with statistical, anecdotal or case-study-based evidence.

Making inferences from single-event observations is poor scientific practice, building a 'house of cards' that can lead to the wrong verdict in court. This may be how case law is formed, but it is not a route to robust science.

If it is to contribute effectively to criminal proceedings, forensic research must be conducted by well-resourced and properly trained scientists.

Mark Tibbett *National Soil Resources Institute, Cranfield University, Cranfield, UK.*
mark.tibbett@cranfield.ac.uk

Science luminaries are often religious

Young Earth creationists are easy to lampoon (see G. Branch *Nature* **500,** 149; 2013). However, using reasoned arguments might hold more sway with the US creationist movement.

PZ Myers, author of *The Happy Atheist* (which Branch reviewed), should remember that the majority of those who helped to establish the disciplines that we now practise as modern science were religious believers, including Nicolaus Copernicus, Rene Descartes, Blaise Pascal, Robert Boyle, Isaac Newton, Carl Linnaeus, Edward Jenner, Michael Faraday, Charles Babbage, Joseph Lister, William Thomson and Arthur Stanley Eddington — to name but a few, and excluding a long list of contemporary names. Half of the 10 most influential scientists of the past 350 years chosen for the Royal Society's commemorative stamps in 2010 were religious believers.

Robert White *University of Cambridge, UK.*
George Ellis *University of Cape Town, South Africa.*
Denis Alexander *Faraday Institute for Science and Religion, St Edmund's College, Cambridge, UK.*
dra24@hermes.cam.ac.uk

Temper Italy's strict lab-animal law

The approval in July of a law on animal experimentation in Italy could irreversibly jeopardize the country's medical research, particularly in drug development, and exclude Italy's scientists from international funding (see *Nature* **499,** 258–259; 2013; and *Nature* http://doi.org/nk3; 2013).

Initially intended to replace a European Union directive that regulates the use of laboratory animals with sensible cautionary guidelines, the draft law was repeatedly amended to satisfy antivivisectionists. The law now forbids xenotransplantation (a widely used method in which human tumour cells are implanted into mice to test cancer therapies) and bans the use of animals for studying drug abuse.

A government initiative is urgently needed to modify the regulations while maintaining rigorous and humane animal experimentation.

Livio Trusolino, Andrea Bertotti *University of Turin School of Medicine; and Institute for Cancer Research and Treatment, Candiolo, Turin, Italy.*
livio.trusolino@ircc.it

Ferret H7N9 flu model questioned

We question the relevance to human pandemics of studies designed to investigate transmissibility of the avian influenza A(H7N9) virus between ferrets (M. Richard *et al. Nature* http://doi.org/njc; 2013; see also R. A. M. Fouchier *et al. Nature* **500,** 150–151; 2013).

Earlier studies have demonstrated ferret transmission of wild-type A(H7N9) virus (H. Zhu *et al. Science* **341,** 183–186; 2013; and Q. Zhang *et al. Science* **341,** 410–414; 2013). However, the results are in striking contrast with the evidence that human-to-human transmission is extremely infrequent.

The discrepancy between ferret and human transmission in some strains undermines the purported value of 'gain-of-function' (GOF) experiments, which track genetically modified variants of A(H7N9). Yet these studies and a proposal to do GOF experiments (R. A. M. Fouchier *et al. Nature* **500,** 150–151; 2013) have been published without seriously questioning the relevance of the ferret model.

A naturally occurring virus could differ subtly from any produced experimentally. In that case, human transmissibility, antigenicity and drug resistance would need to be assessed by studies of the actual strain of virus, and data from GOF experiments could in fact be misleading.

The irony of GOF studies is that these results are likely to be useful for public health only if the pandemic arises from a lab accident, as many fear could happen if this work proliferates.

Marc Lipsitch* *Harvard School of Public Health, Boston, Massachusetts, USA.*
mlipsitc@hsph.harvard.edu
**On behalf of 4 co-authors.*
Two authors declare competing financial interests: see go.nature.com/eco2zz for details and for full author list.

AUTISM

A long genetic explanation

Topoisomerase enzymes facilitate gene transcription by resolving DNA tangles. Malfunction of these enzymes seems to compromise the expression of very long genes, potentially mediating neurodevelopmental disorders. SEE ARTICLE P.58

ROBERT N. PLASSCHAERT
& MARISA S. BARTOLOMEI

Topoisomerases are a family of enzymes that catalyse the unwinding and unknotting of DNA sequences. By introducing transient 'nicks', these enzymes can relieve the topological pile-up of DNA that is caused by processes such as replication and transcription. On page 58 of this issue, King *et al.*[1] provide evidence that topoisomerases are required for the proper expression of extremely long genes in neurons. This insight has implications for our understanding of the fundamentals of both transcription and neurodevelopmental disorders*.

Genomic imprinting is an evolutionarily conserved mammalian phenomenon in which expression of a gene occurs preferentially from one parental chromosome. In human neurons, for example, the *UBE3A* gene is expressed only from the maternal chromosome, and deletion or mutation of this allele (gene copy) causes a severe neurodevelopmental disorder called Angelman syndrome[2].

In a previous study[3], King and colleagues screened for small molecules that, when applied to neurons in culture, activated the normally silenced paternal allele of *UBE3A*. Surprisingly, they found that topoisomerase inhibitors activated this allele. This ability of the enzymes to correct gene 'dosage' by derepressing silent alleles seemingly held much promise for the treatment of disorders that involve imprinted genes or that map to the X chromosome in females. But the mechanism underlying this activation remained unclear.

The earlier report also showed that topoisomerase inhibitors reduced the expression of *UBE3A-ATS*, a very long transcript expressed in neurons. *UBE3A-ATS* overlaps with paternal *UBE3A* on the opposite strand and is associated with the silencing of this allele (Fig. 1a). Repression of *UBE3A-ATS* by topoisomerase inhibitors, therefore, suggested that topoisomerases are involved in maintaining the expression of extremely long genes.

King *et al.* now report that, indeed, treatment of mouse and human neurons with topoisomerase inhibitors results in the widespread silencing of very long genes (those longer than 67 kilobases). This repression depends on the dose of the inhibitor and is highly correlated with increased gene length. Sustained repression of topoisomerase expression using short hairpin RNA (shRNA) sequences also resulted in reduced expression of long genes, excluding the possibility of off-target effects of the inhibitors.

To investigate the mechanism of topoisomerase action, King and colleagues mapped genome-wide binding sites of RNA polymerase II (Pol II) — the enzyme that catalyses DNA transcription — before and after treatment with topoisomerase inhibitors. The authors noted a significant enrichment of Pol II in promoter regions after treatment and a corresponding paucity of Pol II in the body of long genes (Fig. 1b). For short genes, however, Pol II density across the gene body was slightly increased. These results suggest that topoisomerases are specifically involved in the elongation step of transcription during the expression of long genes.

King *et al.* also found that topoisomerase inhibitors decreased the expression of an impressive proportion (27%) of long genes that are candidates for an association with autism spectrum disorders (ASDs). The authors show that the inhibitors significantly downregulate the collective expression of such ASD candidate genes, further supporting the link between topoisomerase mutations and reduced expression of long ASD genes. Consistently, recent work[4,5] has uncovered rare *de novo* mutations in topoisomerases from patients with ASD.

Notably, these findings suggest a possible role for topoisomerases in other genetic disorders in which the causal gene is exceptionally long. It is plausible that mutations that reduce the expression of these enzymes could allow the appropriate expression of all but a few very long genes. For instance, *CFTR*, the

*This article and the paper under discussion[1] were published online on 28 August 2013.

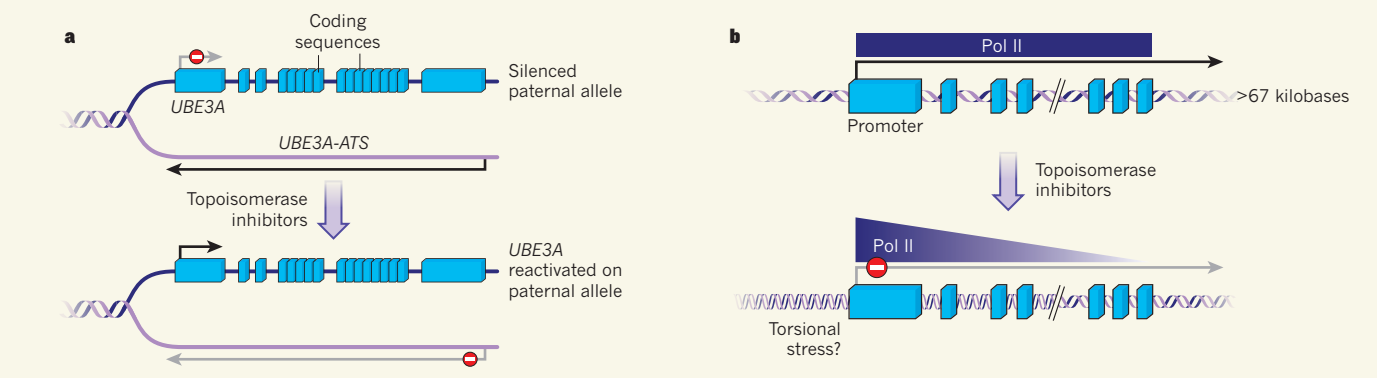


Figure 1 | Topoisomerase inhibitors and transcription. **a**, Treating neurons with topoisomerase inhibitors results in activation of the normally silent paternal allele of the *UBE3A* gene. It also represses transcription of the long transcript *UBE3A-ATS* in the antisense direction. **b**, King *et al.*[1] show that these inhibitors repress the transcription of all genes that have very long sequences (more than 67 kilobases), increasing the binding of the Pol II enzyme at their promoter sequences while decreasing its binding to their coding sequences. The authors hypothesize that topoisomerase inhibition causes a build-up of torsional stress associated with the transcription of long genes, resulting in their repression.

gene mutated in cystic fibrosis, spans more than 200 kilobases. And *DMD*, the causal gene in many forms of muscular dystrophy, spans a staggering 2.2 megabases. It would be surprising if topoisomerases did not contribute to these disorders in certain rare cases at least.

Nevertheless, the work also shows that topoisomerase inhibitors are not a panacea for disorders that would benefit from the activation of a normally silent allele. As these inhibitors are likely to have a widespread effect on the expression of all long genes, even appropriate targets such as *UBE3A* are unlikely to be activated without nonspecific effects. Furthermore, the delicate interplay between gene dosage and traits associated with such disorders makes a broad regulator of transcription such as topoisomerase a less attractive target for drug design.

Despite these concerns, King and colleagues' paper presents an intriguing and fundamentally novel role for topoisomerases in gene regulation. Although these enzymes were known to be required for the expression of longer transcripts in yeast[6], the present work cements their importance in mammals. Moreover, the studies show a requirement for both topoisomerase I and II enzymes in the expression of long genes, whereas only topoisomerase II is required in yeast. Such observations reveal an increased significance for topoisomerases in the regulation of the expanded and much more complex mammalian genome.

By excluding other possibilities, this work also strongly indicates that topoisomerases modify DNA topology during the expression of long transcripts. In prokaryotes (bacteria and archaea), transcription causes dynamic DNA supercoiling, which acts as both a negative and a positive regulator of transcription[7]. Transcription-associated supercoiling also occurs in mammals[8], but its functional relevance is unclear. King and co-authors' findings point towards the possible importance of overcoming supercoiling during the transcription of long genes. Thus, modifying the action or recruitment of topoisomerases should provide a way of exploring how the expression of long mammalian genes is modulated. ■

Robert N. Plasschaert *and* **Marisa S. Bartolomei** *are in the Department of Cell and Developmental Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, Pennsylvania 19104, USA.*
e-mail: bartolom@mail.med.upenn.edu

1. King, I. F. *et al. Nature* **501,** 58–62 (2013).
2. Kishino, T., Lalande, M. & Wagstaff, J. *Nature Genet.* **15,** 70–73 (1997).
3. Huang, H.-S. *et al. Nature* **481,** 185–189 (2011).
4. Neale, B. M. *et al. Nature* **485,** 242–245 (2012).
5. Iossifov, I. *et al. Neuron* **74,** 285–299 (2012).
6. Joshi, R. S., Piña, B. & Roca, J. *Nucleic Acids Res.* **40,** 7907–7915 (2012).
7. Peter, B. J. *et al. Genome Biol.* **5,** R87 (2004).
8. Kouzine, F., Sanford, S., Elisha-Feil, Z. & Levens, D. *Nature Struct. Mol. Biol.* **15,** 146–154 (2008).

QUANTUM INFORMATION

Sharing quantum secrets

A cost-effective architecture for quantum cryptography has been demonstrated in which a single receiver positioned at a network-hub node is shared by many end users to exchange secret encryption keys. **SEE LETTER P.69**

RUPERT URSIN & RICHARD HUGHES

Keeping a secret has never been easy. Throughout history, ideas for encrypting messages have spurred people to come up with ways of breaking the keys that encrypt the messages. Perhaps the most famous example is the race that occurred during the Second World War between the German Enigma cipher machines and the British Colossus — the world's first electronic, digital computer. In the 1980s, the game changed with the invention of a cryptographic technique known as quantum key distribution (QKD)[1] that uses the laws of quantum physics to guarantee secure communication. So far, however, QKD has been demonstrated only for point-to-point communications and relatively simple networks. On page 69 of this issue, Fröhlich *et al.*[2] describe a method that brings the advantages of QKD to as many as 64 end users, who can share a quantum key, and thus a secret. The results are illustrative of the worldwide research progress towards a practical 'quantum Internet'.

Through the seminal research of Auguste Kerckhoffs in the nineteenth century[3] and Claude Shannon in the twentieth century[4], it is possible to reduce the problem of secure communication to the secure transfer of a secret encryption key. Unfortunately, neither Kerckhoffs nor Shannon explained how secure key distribution could be performed. Today, it relies on the unproven difficulty of cracking certain hard mathematical problems, such as factoring large integers. But factoring methods are continually improving, making the security lifetime of this method hard to predict.

Quantum cryptography avoids these issues. In this technique, the key is encoded into quantum states, such as the polarization of a series of single photons that are passed between two parties trying to share secret information. Heisenberg's uncertainty principle dictates that a third party trying to decode the key cannot look at these photons without changing or destroying the information they carry. In this case, it does not matter what technology the third party has: it will never be able to break the laws of quantum physics and decrypt the key.

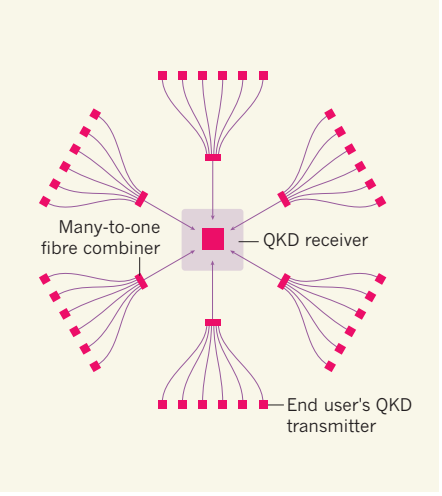


Figure 1 | A quantum access network. Fröhlich and colleagues[2] have developed a cost-effective quantum key distribution (QKD) architecture in which a single QKD receiver is placed at a trusted network-hub node and is shared by many end users' QKD transmitters. The QKD single-photon signals from the users are sent over optical fibres, through a many-to-one combiner, and routed to the network hub.

More-recent research[5] shows that QKD is secure, even if quantum mechanics turns out to be only an approximate theory describing the world. QKD is the first quantum-information application to reach the level of a commercial technology[6,7]. Present-day commercial QKD systems have been developed with a view to incorporating them into existing telecommunication infrastructures at the metro-area scale.

But we live in a networked world and QKD is intrinsically a point-to-point protocol. Several research groups have previously investigated how the advantages of QKD could be brought to multi-party networks (for example, see refs 8 and 9). These 'trusted QKD networks' amount to a mesh of point-to-point QKD links between nodes within which QKD-generated keys must be physically secured against adversaries (hence the need for the nodes to be trusted). However, this approach involves tremendous duplication of resources, with each node requiring QKD receivers to accept incoming photons, and QKD transmitters to send keys on to other nodes. The resulting high

cost and limited scalability have been major obstacles to the adoption of QKD as a cybersecurity technology.

In their study, Fröhlich *et. al.* introduce and demonstrate a cost-effective way to bring the advantages of QKD to multiple end users. This quantum access network, as they dub it, locates a single QKD receiver — with its expensive single-photon detectors — in a network hub, which serves multiple end users, each of whom has a QKD transmitter (Fig. 1). Each user's QKD photons are routed to the one receiver within the network hub over optical fibres. With the expensive single-photon detectors shared among many end users, the approach is cost-effective and reduces the hardware requirements of each user.

Fröhlich *et al.* show that their approach is scalable to as many as 64 end users, provided that the optical fibre between each user and the hub carries only the QKD single-photon signals. But in real-world access networks, these fibres are already 'lit up' in both directions with data-bearing optical signals, which introduce large numbers of photons that could degrade the QKD performance. However, a previous study[10] using a single end user QKD transmitter demonstrated that QKD can successfully co-exist with this bi-directional optical traffic on an access network's optical-fibre links. Taken together with Fröhlich and colleagues' result, this provides strong evidence that QKD could be cost-effectively deployed as an overlay to existing fibre-to-the-home and other access networks.

The present study represents a perfect example of how conceptually new ideas can lead to new quantum technologies, and brings the advantages of a global quantum Internet a step closer to the consumer. ■

Rupert Ursin *is at the Institute for Quantum Optics and Quantum Information (IQOQI), Austrian Academy of Sciences, A-1090 Vienna, Austria, and at the Vienna Center for Quantum Science and Technology, Faculty of Physics, University of Vienna.*
Richard Hughes *is in the Physics Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.*
e-mails: rupert.ursin@univie.ac.at; rxh@lanl.gov

1. Bennett, C. & Brassard, G. *Proc. IEEE Int. Conf. Comput. Syst. Signal Process.* 175–179 (1984).
2. Fröhlich, B. *et al. Nature* **501,** 69–72 (2013).
3. Kerckhoffs, A. *J. Sci. Militaires* **9,** 5–38, 161–191 (1883).
4. Shannon, C. E. *Bell Syst. Tech. J.* **28,** 656–715 (1949).
5. Colbeck, R. & Renner, R. *Nature Commun.* **2,** 411 (2011).
6. www.idquantique.com
7. http://magiqtech.com
8. Peev, M. *et al. New J. Phys.* **11,** 075001 (2009).
9. Sasaki, M. *et al. Opt. Express* **19,** 10387–10409 (2011).
10. Choi, I., Young, R. J. & Townsend, P. D. *New J. Phys.* **13,** 063039 (2011).

EVOLUTIONARY BIOLOGY

Mimicry all the way down

Morphological mimicry among organisms has long been recognized as an adaptive strategy, but mimicry also occurs at the molecular level. One emerging example is microbial pathogens' use of structural mimics that engage host–cell receptors.

MATTHEW F. BARBER & NELS C. ELDE

A century and a half ago, in the rainforests of the Amazon, Henry Walter Bates discovered that certain edible butterflies mimic the appearance of poisonous ones to avoid predation. Such adaptation to approximate existing forms, now referred to as Batesian mimicry, illustrates the astounding potential for evolutionary change driven by natural selection. But hidden from the celebrated naturalist's view were similar instances of mimicry occurring at the submicroscopic level. A growing collection of investigations is now revealing the extraordinary extent to which pathogenic bacteria and viruses mimic host functions. Writing in *Cell Host & Microbe*, Drayman *et al.*[1] provide new evidence for structural mimicry of host ligands for cellular receptors.

Several pathogenic microorganisms are known to produce proteins that mimic the form and functions of host proteins to exploit cellular machinery and counter immune defences. Indeed, molecular mimics have been identified that subvert all major cellular processes — including cell-cycle progression, membrane trafficking, cytoskeleton dynamics and signal transduction — to promote infections. Certain microbes seem to be especially adept at deploying molecular mimicry[2,3]. For instance, among the hundreds of genes carried by some large DNA viruses are several that encode proteins with definitive amino-acid sequence similarity to host proteins. These genes are often acquired from host genomes by a process known as horizontal gene transfer; once captured, selective pressure may lead to the accrual of changes to these genes that increase the fitness of the pathogen, thereby enhancing the advantage conferred by mimicry.

Many mimics, however, cannot be recognized by sequence identity owing to either extensive divergence from transferred host genes or convergent evolution[2,4], whereby a similar structure or function arises in the absence of common ancestry. As such, the use of X-ray crystallography and related approaches to resolve high-resolution protein structures from pathogen mimics has been central to furthering our understanding of molecular mimicry[4]. In this spirit, Drayman *et al.* used existing structural information to screen for novel examples of pathogen mimicry, focusing on potential mimicry of ligands for host receptors.

The authors used established algorithms[5] for scoring similarities between protein structures

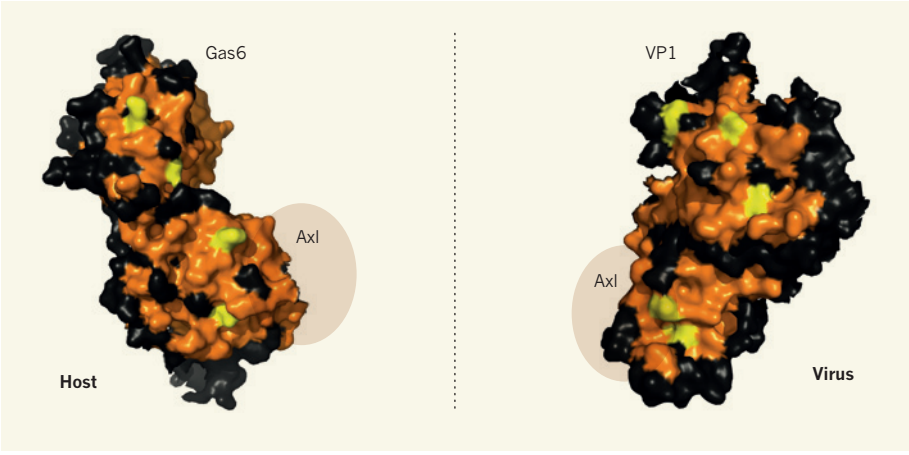


Figure 1 | Molecular mimics. Drayman and colleagues' comparison[1] of the structures of bacterial and viral surface proteins with those of host-cell ligands reveals examples of mimicry at the molecular level. Their study identifies, among other cases, that a surface protein of the simian virus SV40, called VP1 (right), mimics the protein Gas6 (left), which is a ligand of the receptor Axl. Shown are surface representations of VP1 (PDB-3BWQ) and Gas6 (PDB-2C5D), highlighting the docking site of Axl to Gas6 and that proposed for VP1.

to identify surface proteins from viral and bacterial pathogens that share structural features with host-cell ligands. Reasoning that interactions of ligand mimics with host receptors might facilitate pathogen adhesion and access to host cells, the researchers then tested whether putative mimics could directly bind to their cognate receptors. Despite the small data set used for the comparisons, the approach identified several candidates of mimicry, highlighting the potential of melding structural and evolutionary biology not only for delineating mechanisms of pathogen mimicry, but also for identifying unknown mimics and their targets.

One example of mimicry proposed by the authors was a structural feature of the simian polyoma virus SV40 protein VP1, which resembles Gas6, a ligand of the receptor tyrosine kinase Axl (Fig. 1). The identification of Axl as a probable target of ligand mimicry hints at its importance in viral infections. Axl function has also been implicated in infections with Kaposi's sarcoma-associated herpesvirus and Ebola virus[6,7], but its precise role is unclear. This case illustrates how, although informative, the use of structural approaches as a predictive tool for studying mimicry provides only a partial view of the underlying function of the identified mimics. Furthermore, because functional mimicry can evolve in the absence of any structural similarity, many cases of convergent mimicry cannot be uncovered by structural comparisons. To move forward from structure-based screens, we will need to decipher the molecular basis of mimic functions and their contributions to pathogenesis.

Drayman and colleagues' study exemplifies how increasingly powerful tools for exploring the biology of molecules and genomes are helping researchers to uncover additional examples of molecular mimicry, some as exquisite as the classically described cases of morphological mimicry. Molecular mimicry, much like its Batesian counterpart, has major consequences for the evolution of both hosts and pathogens. Mimics pose a daunting challenge to hosts — how does a host differentiate between its native ligands and pathogen mimics, while maintaining core functions?

Research is beginning to illuminate how the rapid accumulation of protein-coding mutations, a conspicuous sign of positive selection, alters host proteins at crucial binding interfaces with pathogen mimics. See-sawing adaptations between hosts and pathogens at these protein surfaces can lead to molecular 'arms races', characterized by intense bouts of positive selection over millions of years[8]. Studying such evolutionary patterns can provide important insight into the potential for hosts to combat the challenge of mimicry[9]. Thus, the thriving alliance between evolutionary and experimental approaches[10] holds great promise for enhancing our knowledge of the evolution of pathogen mimicry. In this light, it is exciting to consider how Bates's observations of rainforest butterflies long ago might help to inform our understanding of infectious disease today. ■

Matthew F. Barber *and* **Nels C. Elde** *are in the Department of Human Genetics, University of Utah, Salt Lake City, Utah 84112, USA.*
e-mails: mbarber@genetics.utah.edu; nelde@genetics.utah.edu

1. Drayman, N. *et al. Cell Host Microbe* **14,** 63–73 (2013).
2. Elde, N. C. & Malik, H. S. *Nature Rev. Microbiol.* **7,** 787–797 (2009).
3. Doxey, A. C. & McConkey, B. J. *Virulence* **4,** 453–466 (2013).
4. Stebbins, C. E. & Galán, J. E. *Nature* **412,** 701–705 (2001).
5. Holm, L. & Sander, C. *Trends Biochem. Sci.* **20,** 478–480 (1995).
6. Liu, R. *et al. Blood* **116,** 297–305 (2010).
7. Brindley, M. A. *et al. Virology* **415,** 83–94 (2011).
8. Daugherty, M. D. & Malik, H. S. *Annu. Rev. Genet.* **46,** 677–700 (2012).
9. Elde, N. C., Child, S. J., Geballe, A. P. & Malik, H. S. *Nature* **457,** 485–489 (2009).
10. Dean, A. M. & Thornton, J. W. *Nature Rev. Genet.* **8,** 675–688 (2007).

BIOGEOCHEMISTRY

As different as night and day

An analysis of northern ecosystems shows that the effects on plant growth of rising night-time temperatures are opposite to those of increasing daytime temperatures — a finding that has implications for carbon-cycle models. SEE LETTER P.88

CHRISTOPHER STILL

An under-appreciated aspect of climate change is the fact that Earth is warming at a higher rate at night than during the day: over the past 50 years, daily minimum temperatures have increased about 40% faster than daily maximum temperatures[1]. This asymmetric warming may have important biological consequences, particularly for fundamental ecosystem metabolic processes that are strongly sensitive to temperature variations, such as photosynthesis and respiration. On page 88 of this issue, Peng *et al.*[2] document regionally significant, and in many cases opposing, effects of year-to-year (interannual) variations in daytime and night-time temperatures on plant growth and carbon cycling in land regions of the Northern Hemisphere.

Photosynthesis is driven by light and thus happens only during the day, whereas plant and microbial respiration occurs continuously. Therefore, faster night-time warming presumably affects respiration more than it affects photosynthesis, and this could have far-reaching implications for how ecosystems react to expected increases in warming in coming decades. But remarkably little research has been

PETER ESSICK/AURORA PHOTOS/ROBERT HARDING PICTURE LIBRARY

Figure 1 | Boreal forest by night. Climate change is causing nights to warm faster than days. Peng *et al.*[2] report that the metabolic effects of warmer nights oppose the effects of warmer days in certain Northern Hemisphere ecosystems, such as boreal forests.

done on how asymmetric warming influences ecological function, especially at large scales. To address this issue, Peng and colleagues have analysed satellite-derived data sets of plant greenness, which is a proxy for plant growth.

The authors found that ecosystems in cool, wet temperate and boreal regions such as northwestern North America and Japan, and those in cold regions such as Siberia and the Tibetan plateau, seem to have benefited most from daytime temperature increases over the period considered (1982–2009). By contrast, ecosystems in dry temperate regions, such as central Eurasia and western China, showed the opposite effect: increasing daytime temperatures correlated with decreasing plant greenness. These contrasting responses broadly agree with expectations for ecosystems in which plant growth is limited primarily by temperature (cool, wet climates) or moisture (warm, dry climates).

More intriguingly, Peng and colleagues found that ecosystems in many of the boreal and wet temperate regions grew less well in response to increases in night-time minimum temperatures — the opposite effect to their response to increasing daytime maximum temperatures (Fig. 1). Conversely, in many arid and semi-arid regions, such as the grasslands of China and North America, increasing night-time minimum temperatures correlated positively with plant greenness.

Peng *et al.* used a statistical approach to control for other contributing environmental variables, such as solar radiation and precipitation. This allowed them to isolate the interannual greenness responses to daytime maximum and night-time minimum temperature variations. The authors confirmed the statistical validity of their findings using other techniques, and also analysed the sensitivity of the greenness response to alternative interpolated climate data sets and at individual weather-station locations. Importantly, the different analyses all confirmed the same broad conclusions.

A strength of this study is that the researchers explored ecosystem responses to asymmetric warming using a variety of other large-scale data sets, and found similar patterns. One data set was for the net exchange of carbon between land and the atmosphere — a quantity that integrates photosynthesis and respiration, and which was inferred from a multi-year analysis[3]. Peng and co-workers found that this quantity correlated positively with daytime temperature variations for cool and wet boreal ecosystems, but negatively with night-time temperatures for these ecosystems. They also observed that the amplitudes of the seasonal cycles of carbon dioxide levels measured at Point Barrow, Alaska, and Mauna Loa, Hawaii, vary in the same way with daytime and night-time temperature variations in boreal regions, but not in temperate areas.

Peng *et al.* focused only on boreal and temperate ecosystems. The response to asymmetric warming of tropical and subtropical ecosystems, which account for most CO_2 exchange between the land and the atmosphere, is not clear and merits further investigation. Previous work[4] at a well-studied tropical forest revealed a negative correlation between tree growth and annual mean daily minimum temperatures, a response broadly similar to Peng and colleagues' findings for boreal forests. Tropical forests are thought to be vulnerable to warming[5], with some evidence[6] suggesting that they are already near high-temperature thresholds above which growth could be restricted. Future research could help to fill major gaps in our understanding of thermal tolerance and acclimation in tropical and subtropical plant species, and thus their response to warming[5,7].

So what are the physiological mechanisms that drive large-scale correlations between temperature variations and ecosystem metabolism? The commonly discussed mechanisms involve biochemical responses to temperature, but with some interesting twists. For example, the positive correlation found between night-time minimum temperatures and greenness in semi-arid grasslands is puzzling, but might be related to greater night-time plant respiration that stimulates increased daytime photosynthesis[8]. Increases in night-time respiration have also been invoked in a pioneering study[9] of nocturnal warming that documented different plant responses in grassland: the dominant grass species declined in response to increases in night-time temperature during spring, whereas other plant species that use a different photosynthetic pathway increased in number.

A research agenda to investigate these mechanisms further should include manipulative field and mesocosm experiments (in which small parts of a natural ecosystem are enclosed and warmed). Experimental warming studies are lacking for many ecosystems. Even fewer night-time warming experiments have been conducted so far, with most being in shrublands[10] or grasslands and croplands[8]; warming experiments that truly impose asymmetry between day and night warming are rare[11]. There is a particularly urgent need for warming studies in forests, which dominate the global carbon cycle and climate feedbacks.

However, there are substantial technological challenges to conducting such experiments in large-statured ecosystems. Forest mesocosm experiments would require exceedingly complex and expensive facilities. Despite these limitations, Peng and colleagues' results argue strongly for an increased focus on the differing ecological impacts of night-time and daytime temperatures, to improve our ability to understand and predict how warming will affect Earth's ecosystems. ■

Christopher Still *is in the Department of Forest Ecosystems and Society, Oregon State University, Corvallis, Oregon 97331, USA.*
e-mail: chris.still@oregonstate.edu

1. Solomon, S. *et al.* (eds) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2007).
2. Peng, S. *et al. Nature* **501,** 88–92 (2013).
3. Chevallier, F. *et al. J. Geophys. Res.* **115,** D21307 (2010).
4. Clark, D. A., Piper, S. C., Keeling, C. D. & Clark, D. B. *Proc. Natl Acad. Sci. USA* **100,** 5852–5857 (2003).
5. Corlett, R. T. *Trends Ecol. Evol.* **26,** 606–613 (2011).
6. Doughty, C. E. & Goulden, M. L. *J. Geophys. Res.* **113,** G00B07 (2008).
7. Zhang, J.-L., Poorter, L., Hao, G.-Y. & Cao, K.-F. *Ann. Bot.* **110,** 1027–1033 (2012).
8. Wan, S., Xia, J., Liu, W. & Niu, S. *Ecology* **90,** 2700–2710 (2009).
9. Alward, R. D., Detling, J. K. & Milchunas, D. G. *Science* **283,** 229–231 (1999).
10. Beier, C. *et al. Sci. Total Environ.* **407,** 692–697 (2008).
11. Phillips, C. L., Gregg, J. W. & Wilson, J. K. *Glob. Change Biol.* **17,** 3263–3273 (2011).

CELL BIOLOGY

Recycling in sight

Vision requires the continuous recycling of photobleached pigments. An atypical form of a degradative pathway called autophagy seems to participate in this process in retinal pigment epithelial cells.

PATRICIA BOYA & PATRICE CODOGNO

Vision begins with the absorption of photons by light-sensitive photoreceptor cells in the retina. When photons arrive, chromophore molecules in the photoreceptors undergo conformational changes and trigger the phototransduction cascade, which converts light into electrical impulses that travel to the brain and are transformed into the images that we see. To sustain vision, the chromophore 11-*cis*-retinal must return to its original conformation through a process known as the visual cycle. Writing in *Cell*, Kim *et al.*[1] demonstrate that one process that contributes to the proper functioning of the visual cycle is an atypical form of autophagy — the 'self-eating' pathway through which cells recycle their components by degrading them in cellular organelles called lysosomes.

The visual cycle involves tight regulation of

the interaction between photoreceptors and the adjacent retinal pigment epithelium (RPE) cell layer; the latter nourishes photoreceptors and improves the quality of the optical system of the eye by absorbing scattered light. Another key role of the RPE is degradation and recycling of the photoreceptor outer segments (POSs), which are damaged by the impact of millions of photons every day. Each morning, the distal 10% of POSs are engulfed (phagocytosed) by the RPE and degraded inside lysosomes. Whereas this phagocytic process has been known for decades, the intimate molecular and cellular mechanisms of chromophore and photoreceptor recycling have remained unclear.

Kim *et al.* demonstrate that selected proteins implicated in autophagy are necessary for the degradation of POSs. Specifically, they show that POS phagocytosis coincides with an increase in the levels of the autophagy marker protein LC3-II, and that it involves an atypical autophagy pathway called LC3-associated phagocytosis (LAP). Moreover, in the RPE, LAP requires typical autophagy regulators such as the Atg5 protein (Fig. 1). However, it is independent of the proteins of the autophagy pre-initiation complex that in typical autophagy mediate the formation of autophagosomal vesicles to sequester cytoplasmic components[2], consistent with previous work[3]. Given that, in LAP, the phagosomal vesicles carrying the material to be degraded have already formed, it is perhaps not surprising that this process is independent of the autophagy pre-initiation complex.

Maturation of phagosomes usually involves their acidification and the acquisition of degradative enzymes to break down the ingested material. Why, then, are autophagy proteins required for LAP? It is thought that recruitment of autophagy proteins to the phagosomal membrane allows its rapid fusion with lysosomes, enhancing degradation[4]. In support of this view, Kim and co-workers observed no defects in phagocytic uptake in Atg5-deficient RPE cells. However, the phagosomes in these cells failed to migrate towards the basal side of the cells, where fusion with lysosomes occurs. Furthermore, maturation of the protease enzymes required for degradation was impaired, and the expression of lysosomal-membrane glycoproteins in these phagosomes was reduced, suggesting impaired phagosome–lysosome fusion. At least in this setting, therefore, LAP does seem to promote trafficking of phagosomes and their fusion with lysosomes, ensuring rapid recycling of POSs.

The visual cycle is a transcellular process by which RPE cells maintain the supply of chromophores for the regeneration of visual pigment in photoreceptors (Fig. 1). When photoreceptors absorb light, 11-*cis*-retinal is converted to all-*trans*-retinal through isomerization. All-*trans*-retinal is then released from the membrane-bound receptor opsin and reduced to all-*trans*-retinol (vitamin A), which diffuses to the intercellular space and enters the adjacent RPE cells, where it is transformed back into 11-*cis*-retinal[5]. Other sources of vitamin A are the blood and POS phagocytosis by the RPE[6].

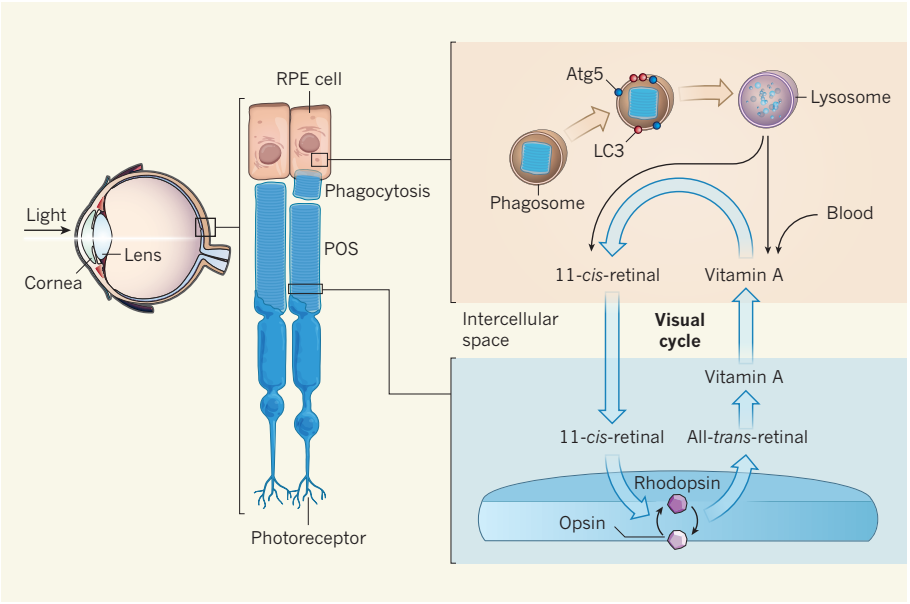


Figure 1 | Role of autophagy in vision. Light enters the eye and strikes the retina, where it activates rhodopsin pigments in the membranes of photoreceptor outer segments (POSs). This induces the release of both the bleached pigment (opsin) and a molecule of all-*trans*-retinal, which is converted to vitamin A (all-*trans*-retinol) before forming the chromophore 11-*cis*-retinal in the adjacent retinal pigment epithelium (RPE) cells through the visual cycle. The binding of 11-*cis*-retinal to opsin restores the photoactivable pigment rhodopsin, which can react again with light and start the phototransduction cascade. Light also damages POSs, which are subsequently phagocytosed and eliminated by RPE cells. Kim *et al.*[1] report that, after phagosome formation, an atypical autophagy pathway involving the proteins Atg5 and LC3 triggers phagosome fusion with the lysosome, resulting in POS degradation. The degradation products exit the lysosome and are recycled to sustain the visual cycle. Blood is another source of vitamin A.

Kim and co-authors show that autophagy proteins mediate 11-*cis*-retinal recycling following POS phagocytosis. Although many aspects of this process remain unclear, chromophore recovery from discarded POSs may represent a highly efficient mechanism for recycling. Each day during POS shedding, around 10% of the total ocular pool of retinoids passes through the phagolysosomal system of RPE cells. The present paper suggests that at least some of the 11-*cis*-retinal can be recovered and used to regenerate visual pigment without the involvement of intercellular transport systems or enzymes of the visual cycle, which are often mutated in retinal diseases[6].

The authors further show that mice specifically lacking Atg5 in their RPE have impaired vision and reduced chromophore levels. Interestingly, previous work[7] has found that retinal ageing is associated with decreased autophagic activity and a parallel reduction in night vision; these traits were mimicked by Atg5 deletion in retinal precursors. Also, in the lens, defects in autophagy-protein function result in age-related cataracts[8]. And retinal ganglion cells lacking Atg5 and Atg4, another autophagy-associated protein, show increased sensitivity to optic-nerve damage[9].

A common feature of these autophagy-impaired mutations is the accumulation of toxic products, suggesting a defect in intracellular quality control. For instance, accumulation of lipofuscin, a cellular waste product derived from the incomplete digestion of POSs, is frequently observed in human retinal diseases such as age-dependent macular degeneration[6]. This disease is caused by a primary malfunction of the RPE, which leads to photoreceptor death and subsequent blindness.

Kim *et al.* did not investigate whether there was an increase in lipofuscin levels in their Atg5-deficient mice. But photoreceptor numbers remained normal in these mutants, even after 7 months. And the authors could reverse the effects of Atg5 deletion on vision by giving the animals retinoid supplements. These results suggest that downregulation of autophagy primarily causes alterations in the visual cycle and not defects in intracellular quality control, as was previously thought.

Vitamin A and retinoid supplements have also been used to treat some retinal disorders associated with visual-cycle defects and ageing[6]. Intriguingly, retinoid derivatives increase the activity of chaperone-mediated autophagy[10] — a form of autophagy that is

upregulated in the retina as a compensatory response to age-associated decreases in macroautophagy, one of the more common autophagic pathways[7]. Regardless of the details, two points are becoming clear: autophagy proteins play a part in maintaining vision fitness, and retinoids show potential for the treatment of retinal diseases. ■

Patricia Boya *is in the Department of Cellular and Molecular Biology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, 28040 Madrid, Spain.* **Patrice Codogno** *is in INSERM U845, Necker Growth and Signaling Research Center, Université Paris Descartes, 75014 Paris, France.*
e-mail: pboya@cib.csic.es

1. Kim, J.-Y. *et al. Cell* **154,** 365–376 (2013).
2. Boya, P., Reggiori, F. & Codogno, P. *Nature Cell Biol.* **15,** 713–720 (2013).
3. Codogno, P., Mehrpour, M. & Proikas-Cezanne, T. *Nature Rev. Mol. Cell Biol.* **13,** 7–12 (2011).
4. Sanjuan, M. A. *et al. Nature* **450,** 1253–1257 (2007).
5. Rando, R. R. *Chem. Rev.* **101,** 1881–1896 (2001).
6. Travis, G. H., Golczak, M., Moise, A. R. & Palczewski, K. *Annu. Rev. Pharmacol. Toxicol.* **47,** 469–512 (2007).
7. Rodriguez-Muela, N. *et al. Aging Cell* **12,** 478–488 (2013).
8. Morishita, H. *et al. J. Biol. Chem.* **288,** 11436–11447 (2013).
9. Rodriguez-Muela, N., Germain, F., Mariño, G., Fitze, P. S. & Boya, P. *Cell Death Differ.* **19,** 162–169 (2012).
10. Anguiano, J. *et al. Nature Chem. Biol.* **9,** 374–382 (2013).

IMMUNOLOGY

B-cell development in the gut

B cells arise in the bone marrow and go on to produce antibodies that protect against microbial infection. Surprisingly, it seems that B-cell development also occurs in the gut, where it is stimulated by resident microbes. SEE LETTER P.112

MARK SCHLISSEL

Mammalian intestines contain vast numbers of immune cells. This is not surprising, given that the gut lumen is home to a microbial ecosystem with cell numbers[1] that exceed those of its human host by around 100 to 1. Microbial colonization of the gut begins at birth, and gut microbes make a crucial contribution to host metabolism. Appreciation is also growing of the role of the microbiome in shaping the host immune system[2]. For example, differentiation of a type of immune cell called T_H17 cells requires the presence of certain bacteria in the gut[3]. Conversely, mice that have no resident microorganisms have perturbed immune responses and a diminished amount of lymphoid tissue, which is an integral part of the immune system. A study by Wesemann *et al.*[4], reported on page 112 of this issue, extends the spectrum of these mutualistic interactions by reporting that B cells can undergo development in the mouse intestine, and that this is influenced by resident non-pathogenic bacteria*.

B cells produce antibodies. These bind specifically to a wide array of antigen molecules and flag them for elimination by other immune cells. Antibodies consist of a heterotetramer of proteins: two identical immunoglobulin heavy (IgH) chains and two identical immunoglobulin light (IgL) chains. The genes encoding these proteins exist as segments that are assembled into functional genes during B-cell development by a process known as V(D)J recombination[5]. Permutations in the combinations of these segments allow the genome to encode a vast repertoire of antigen-specific antibody molecules. The standard model of mammalian B-cell development holds that B cells arise from haematopoietic stem cells (which can give rise to all types of blood cell), in the liver during fetal life and in the bone marrow after birth. Once a B cell has produced a functional antibody molecule, this is transported to the cell membrane, where it serves a surveillance function as the cell's receptor for antigen.

The enormous diversity of antibodies generated by this process creates the problem of self tolerance. But several mechanisms exist to avoid the production of antibodies that attack normal host tissues, including clonal deletion, clonal anergy and receptor editing[6]. Receptor editing occurs in immature B cells (a late stage of B-cell development) and involves rearrangement of IgL-chain gene segments and IgL-chain replacement until a receptor is produced that does not recognize a self-antigen (Fig. 1a). In the bone marrow, which is sterile,

*This article and the paper under discussion[4] were published online on 21 August 2013.

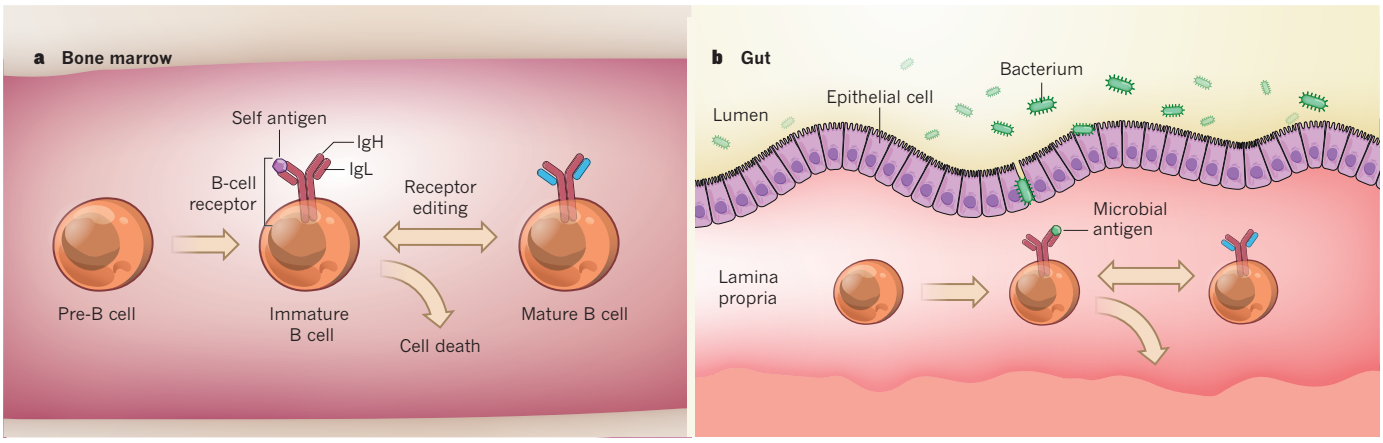


Figure 1 | Receptor editing. **a**, The standard model of B-cell development holds that it occurs in the bone marrow and is regulated by interactions with self antigens. As a pre-B cell develops into an immature B cell, it starts to express a B-cell receptor, comprising immunoglobulin heavy (IgH) chains and immunoglobulin light (IgL) chains, on its surface. Any cell expressing a receptor that binds to self-antigens will be marked for receptor editing — further rearrangement and replacement of the IgL-chain genes that occurs until non-self-recognizing receptors are produced or the cell dies. **b**, Wesemann *et al.*[4] suggest that B-cell-receptor editing also occurs in the lamina propria of the gut, where it is directed by interactions with non-pathogenic microorganisms in the gut lumen. This process may serve to induce B-cell tolerance to these resident microorganisms and to further diversify the repertoire of B-cell receptors.

receptor editing is stimulated by self-antigens. Wesemann and colleagues' results suggest that microbial antigens might also drive receptor editing in the lamina propria, an immune-cell-rich tissue layer in the mucous membranes that line body cavities, such as the gut and respiratory tract.

The authors studied mice that expressed a fluorescent-protein gene from within the *Rag2* gene. The RAG2 protein, in a complex with RAG1, comprises the recombinase enzyme that is necessary for V(D)J recombination, and is expressed almost exclusively in developing T and B cells[7]. They observed a small but significant fraction of RAG2-expressing developing B cells in the intestinal lamina propria. Some of these cells also expressed the marker enzyme TdT, which suggested that they were at the pro-B-cell stage of development. But, perhaps more interestingly, most of the cells were at either the pre-B or immature developmental stages, meaning that they were undergoing IgL-chain gene-segment rearrangement or receptor editing (Fig. 1b). In addition to expressing the recombinase, these cells expressed reaction intermediates that also indicated ongoing recombination[8]. The IgH-chain repertoire of the cells resembled that of B cells elsewhere in the body, but their IgL-chain repertoire was distinctive.

In germ-free mice, which have extremely low bacterial loads in their guts, the authors found that gut-developing B cells were present, but that their numbers peaked when the mice were about 3 weeks old, suggesting that the process peters out in the absence of ongoing microbial stimulation. However, when the mice were colonized with gut microorganisms or stimulated with inflammatory molecules, the numbers of gut-developing B cells increased significantly.

The most straightforward interpretation of these results is that B-cell-receptor editing occurs in immature B cells developing in the gut, and that this is driven by interactions with antigens from resident microorganisms. Although the lamina propria is a sterile location, it is easy to imagine how either local physical trauma or inflammation might allow small numbers of microbes to penetrate the mucosal surface and induce these interactions. It is also possible that specialized mechanisms exist whereby the developing B cells could 'sample' the gut microbiome without breaching the mucosal barrier; such mechanisms are used by M cells that overlie specialized lymphoid structures known as Peyer's patches[9].

That aspects of antibody-gene diversification occur outside the bone marrow is not without precedent. In rabbits, for example, V(D)J recombination during fetal life produces a repertoire of modest diversity that then undergoes significant diversification in the animals' appendix through a gene-conversion process that uses information stored in vast numbers of pseudogenes[10].

Wesemann and colleagues' work raises a number of intriguing questions, most importantly, how this mode of B-cell development contributes to immune homeostasis. One possibility is that receptor editing driven by the vast array of microbial antigens in the gut substantially enhances overall antibody diversity. Alternatively, receptor editing in the gut might serve to decrease B-cell reactivity to the antigens of the most prevalent gut bacteria and thus reduce the possibility of inflammation in the bowel. Finally, B-cell development at this site might facilitate tolerance against self-antigens that are predominantly expressed there.

Other key questions include whether haematopoietic stem cells reside in the lamina propria and, if not, which B-cell precursors seed the gut, and when. Do these gut-derived B cells contribute to the pool of antibodies that is present in the body's serum, or do they exist predominantly in the gut lumen? Furthermore, this research was conducted in mice, so it remains to be determined whether B-cell development and receptor editing in the lamina propria also occur in humans. If they do, this surprising observation will have several implications for our understanding of human health and disease. ■

Mark Schlissel *is in the Departments of Molecular Biology, Cell Biology, and Biochemistry and Molecular Microbiology and Immunology, Brown University, Providence, Rhode Island 02912, USA.*
e-mail: mssbu@brown.edu

1. Ley, R. E., Peterson, D. A. & Gordon, J. I. *Cell* **124,** 837–848 (2006).
2. Round, J. L. & Mazmanian, S. K. *Nature Rev. Immunol.* **9,** 313–323 (2009).
3. Ivanov, I. I. *et al. Cell* **139,** 485–498 (2009).
4. Wesemann, D. R. *et al. Nature* **501,** 112–115 (2013).
5. Jung, D. & Alt, F. W. *Cell* **116,** 299–311 (2004).
6. Nemazee, D. *Nature Rev. Immunol.* **6,** 728–740 (2006).
7. Kuo, T. C. & Schlissel, M. S. *Curr. Opin. Immunol.* **21,** 173–178 (2009).
8. Constantinescu, A. & Schlissel, M. S. *J. Exp. Med.* **185,** 609–620 (1997).
9. Wolf, J. L. & Bye, W. A. *Annu. Rev. Med.* **35,** 95–112 (1984).
10. Lanning, D., Zhu, X., Zhai, S.-K. & Knight, K. L. *Immunol. Rev.* **175,** 214–228 (2000).

NEUROSCIENCE

Bacteria get on your nerves

During infection, the inflammatory immune response can cause pain by activating nociceptor neurons. A bacterial pathogen also seems to stimulate pain directly, modulating the immune response in its favour. SEE ARTICLE P.52

VICTOR NIZET & TONY YAKSH

In his first-century-BC treatise *De Medicina*, Aulus Cornelius Celsus described the four cardinal signs of acute inflammation: *rubor* (redness), *tumor* (swelling), *calor* (warmth) and *dolor* (pain). In the context of innate immunity, inflammation reflects local vasodilation and the influx of leukocyte cells to injured or infected tissue locations, amid a flurry of lipids, enzymes, and cytokine and chemokine molecules. However, the fascinating work of Chiu *et al.*[1] reported on page 52 of this issue shows that acute pain accompanying infections with the bacterium *Staphylococcus aureus* is primarily caused by the direct activation of peripheral sensory neurons (nociceptors) by bacterial components and toxins, rather than by host-derived inflammatory mediators*. Disconcertingly, this may be to the pathogen's gain; in response to bacterial stimulation, the nociceptor terminals could release certain neurotransmitter molecules that impair the proper recruitment and activation of innate immune cells.

Consequences of local inflammation are continual pain and hyperalgesia — an exaggerated pain response to low-intensity stimuli[2]. Mechanistically, such enhanced responsiveness is triggered by molecules that are released into the local milieu of injured or infected tissue; these molecules are recognized by specific receptors on the peripheral terminals of afferent neurons, which reach out to every millimetre of the body's exterior and interior. The activation of these receptors induces a concentration-dependent depolarization of the terminals, triggering kinase enzymes that phosphorylate various terminal receptors and channels to produce continuing afferent activity and enhanced response to subsequent stimuli[3].

The concept that infectious microbes can directly activate pain receptors, rather than acting through an immune-cell intermediary, has emerged in recent years (Fig. 1). Supporting data include the discoveries that lipopolysaccharide (LPS) molecules of the outer membrane

*This article and the paper under discussion[1] were published online on 21 August 2013.

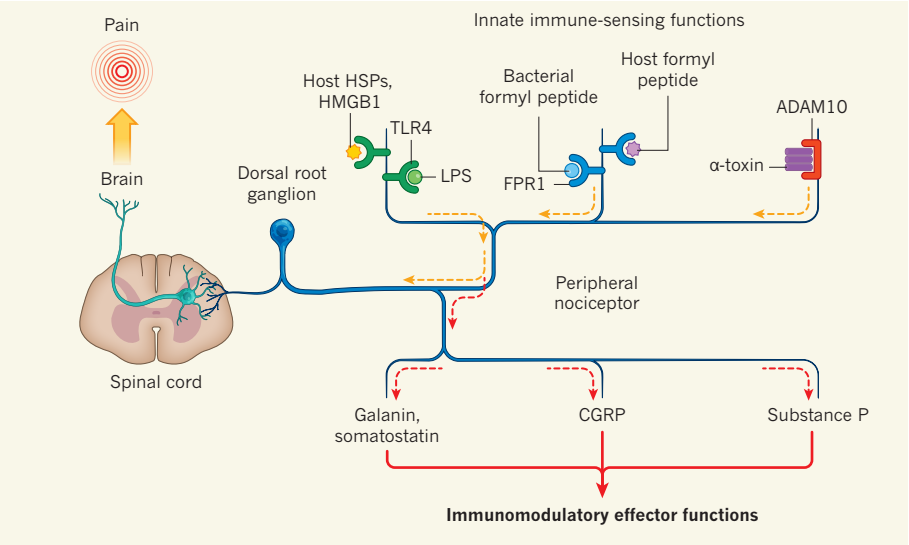


Figure 1 | Immune-related functions of peripheral nociceptors. Terminal fibres of peripheral afferent nociceptors express receptors that allow direct activation of these neurons by host-derived formyl peptides and possibly HSPs and HMGB1 released from injured tissues and by bacterial products (lipopolysaccharide (LPS) is detected by TLR4; formyl peptides are detected by FPR1; and staphylococcal α-toxin is detected by ADAM10). Chiu and colleagues' data suggest[1] that in addition to transducing a pain signal to the central nervous system through the dorsal root ganglion (yellow arrows), this stimulation elicits antidromic action potentials (red arrows) that evoke the release of bioactive peptides such as substance P, CGRP, galanin and somatostatin from the peripheral terminal. This antidromic signalling can thus modulate the local inflammatory response.

of Gram-negative bacteria can stimulate production of the vasodilator CGRP from dorsal root ganglion (DRG) neurons[4] and that the LPS co-receptors TLR4 and CD14, which are normally found on immune cells, are also expressed on trigeminal nociceptive neurons[5]. Moreover, exposure of the mouse urinary tract to live pathogenic *Escherichia coli* bacteria or purified LPS triggered pain by a mechanism that depended on TLR4 but not on the inflammatory response of the immune system's neutrophils or mast cells[6]. Furthermore, LPS binding to TLR4 and its co-receptors on DRG neurons prompted the release of nociceptin, an opioid-related peptide that is upregulated during peripheral inflammation and is associated with hyperalgesia[7].

Chiu and colleagues' results are surprising because they reveal that the activation of pain receptors by *S. aureus* involves neither TLR2, the key immune-system pattern-recognition receptor (PRR) for cell-wall components of Gram-positive bacteria, nor MyD88, a universal adaptor protein that is involved in transducing TLR signals. Instead, they detected two alternative receptor-mediated activation pathways in mouse nociceptors. These neurons expressed FPR1, a G-protein-coupled PRR that responds to formyl peptides on the *S. aureus* cell wall. Moreover, they express ADAM10, a cell-surface metalloprotease enzyme that binds to and facilitates the activity of the pore-forming staphylococcal α-toxin, thereby leading to rapid calcium fluxes within the nociceptors (Fig. 1). These pathways produced changes in pain-perception threshold that were proportional to the bacterial load but, strikingly, the changes were independent of the magnitude of the inflammatory responses.

At first glance, the expression of PRRs for microbial components on nociceptors could imply an evolutionary benefit for the host in return for experiencing acute pain due to infection. A classic study[8] noted that the stimulation of DRG neurons by signals originating from peripheral nociceptors triggers vasodilation. Moreover, antidromic (opposite direction) activities in the small peripheral nociceptor can promote the release of vasodilators at its peripheral terminal. Subsequent work[9,10] emphasized a key role of CGRP and the neurotransmitter substance P in initiating neuron-mediated inflammation, a collection of processes that aid pathogen clearance by the immune system.

Unexpectedly, however, Chiu *et al.* found that genetic ablation of all nociceptors in mice was associated with greater lymph-node swelling — a sign of immune activation — in response to *S. aureus* infection. The authors' further analysis revealed that CGRP and other nociceptive afferent-derived peptides (galanin and somatostatin) have previously unknown anti-inflammatory properties that limit the release of cytokines by macrophages, a key immune cell type.

Perhaps these results reflect yet another capacity of *S. aureus* to manipulate and thwart the innate immune-response pathways that are normally effective against 'lesser' pathogens; the large number of virulence factors that this bacterium releases subvert the normal function of phagocytes and the complement system of innate immunity[11,12]. Or one could propose an alternative role for nociceptor expression of TLR4 and FPR1, because these receptors respond to damage-associated molecular patterns such as HMGB1, HSPs or mitochondrial formyl peptides released from host cells after injury[13] (Fig. 1). In this sense, bacteria-induced pain could be an epiphenomenon in a broader selective advantage provided by pain-induced behavioural responses that limit traumatic tissue damage.

The current paper adds to the emerging view of the extensive and complex interaction between the peripheral nervous system and the innate immune system. Sensory afferent nociceptor neurons express receptors that detect bacteria and their toxins, leading to downstream signal transduction and the local release of vasoactive and immunomodulatory peptides; all of this is concurrent with the propagation of action potentials by the axonal processes of these cells and the subjective experience of pain.

There is evidence that this interplay is not limited to the body's peripheral nociceptors but extends to other sensory receptor systems. For instance, T2R38, the receptor for bitter taste, was recently found to detect molecules secreted by the bacterium *Pseudomonas aeruginosa*, stimulating effective clearance of this pathogen[14]. Could other sensory systems (visual, auditory and olfactory) receive direct molecular input from pathogens or the commensal microbiota? Better understanding of these processes could provide innovative targets and approaches to improve treatment outcome in infection-associated disorders. ■

Victor Nizet *is in the Department of Pediatrics and at the Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California 92093, USA, and* **Tony Yaksh** *is in the Department of Anesthesiology, School of Medicine, University of California, San Diego.*
e-mail: vnizet@ucsd.edu; tyaksh@ucsd.edu

1. Chiu, I. M. *et al. Nature* **501,** 52–57 (2013).
2. Kidd, B. L. & Urban, L. A. *Br. J. Anaesth.* **87,** 3–11 (2001).
3. Hucho, T. & Levine, J. D. *Neuron* **55,** 365–376 (2007).
4. Hou, L. & Wang, X. *J. Neurosci. Res.* **66,** 592–600 (2001).
5. Wadachi, R. & Hargreaves, K. M. *J. Dent. Res.* **85,** 49–53 (2006).
6. Rudick, C. N. *et al. J. Infect. Dis.* **201,** 1240–1249 (2010).
7. Acosta, C. & Davies, A. *J. Neurosci. Res.* **86,** 1077–1086 (2008).
8. Bayliss, W. M. *J. Physiol.* **26,** 173–209 (1901).
9. Richardson, J. D. & Vasko, M. R. *J. Pharmacol. Exp. Ther.* **302,** 839–845 (2002).
10. Chiu, I. M., von Hehn, C. A. & Woolf, C. J. *Nature Neurosci.* **15,** 1063–1067 (2012).
11. Rooijakkers, S. H., van Kessel, K. P. & van Strijp, J. A. *Trends Microbiol.* **13,** 596–601 (2005).
12. DeLeo, F. R., Diep, B. A. & Otto, M. *Infect. Dis. Clin. North Am.* **23,** 17–34 (2009).
13. Piccinini, A. M. & Midwood, K. S. *Mediators Inflamm.* **2010,** 672395 (2010).
14. Lee, R. J. *et al. J. Clin. Invest.* **122,** 4145–4159 (2012).

REVIEW

doi:10.1038/nature12481

Self-propagation of pathogenic protein aggregates in neurodegenerative diseases

Mathias Jucker[1,2] & Lary C. Walker[3,4]

For several decades scientists have speculated that the key to understanding age-related neurodegenerative disorders may be found in the unusual biology of the prion diseases. Recently, owing largely to the advent of new disease models, this hypothesis has gained experimental momentum. In a remarkable variety of diseases, specific proteins have been found to misfold and aggregate into seeds that structurally corrupt like proteins, causing them to aggregate and form pathogenic assemblies ranging from small oligomers to large masses of amyloid. Proteinaceous seeds can therefore serve as self-propagating agents for the instigation and progression of disease. Alzheimer's disease and other cerebral proteopathies seem to arise from the *de novo* misfolding and sustained corruption of endogenous proteins, whereas prion diseases can also be infectious in origin. However, the outcome in all cases is the functional compromise of the nervous system, because the aggregated proteins gain a toxic function and/or lose their normal function. As a unifying pathogenic principle, the prion paradigm suggests broadly relevant therapeutic directions for a large class of currently intractable diseases.

Proteins are essential to cellular metabolism and communication, and they form the framework on which cells and tissues are built. To undertake these roles, most proteins fold into a specific, three-dimensional architecture that is largely determined by their distinctive sequences of amino acids. Others have a degree of structural flexibility that enables them to tailor their shape to the task at hand[1,2]. For proteins, then, as for the rest of biology, structure governs function. Hence, it is critical for cells to maintain an efficient quality-control system that ensures the proper production, folding and elimination of proteins[3,4]. When a protein misfolds and evades normal clearance pathways, a pathogenic process can ensue in which the protein aggregates progressively into intracellular and/or extracellular deposits. The consequence is a diverse group of disorders, each of which entails the aggregation of particular proteins in characteristic patterns and locations (Fig. 1)[5–8]. New insights into the ontogeny of these proteopathies are beginning to emerge from the unusual properties of the prion, arguably one of the most provocative molecules in the annals of medicine.

The prion paradigm

Prions ('proteinaceous infectious particles') are unconventional infectious agents consisting of misfolded prion protein (PrP) molecules; in

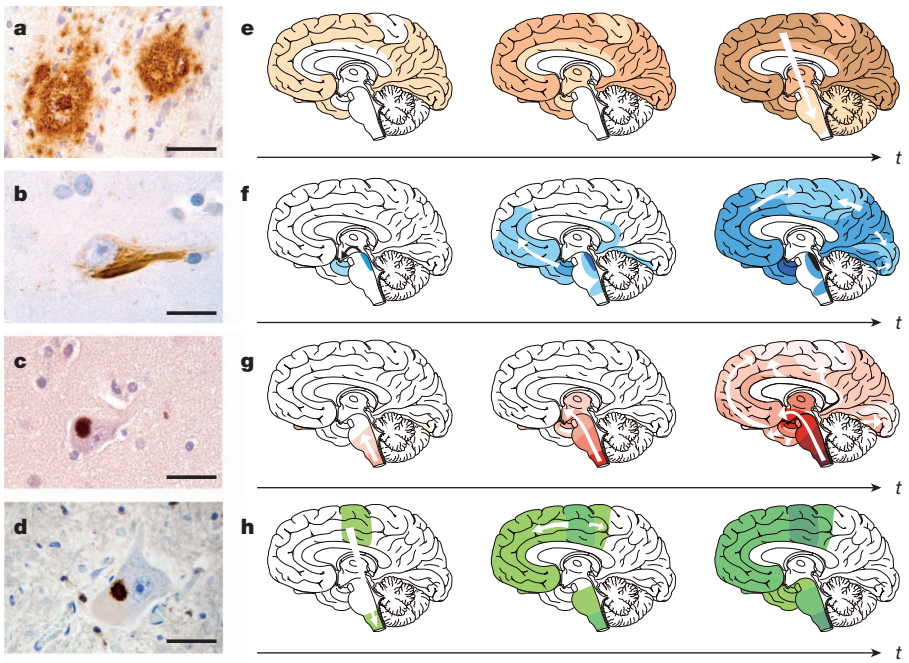


Figure 1 | Commonalities among age-related neurodegenerative diseases. The deposited proteins adopt an amyloid conformation and show prion-like self-propagation and spreading in experimental settings, consistent with the progressive appearance of the lesions in the human diseases. **a**, Amyloid-β deposits (senile plaques) in the neocortex of a patient with Alzheimer's disease. **b**, Tau inclusion as a neurofibrillary tangle in a neocortical neuron of a patient with Alzheimer's disease. **c**, α-Synuclein inclusion (Lewy body) in a neocortical neuron from a patient with Parkinson's disease/Lewy body dementia. **d**, TDP-43 inclusion in a motoneuron of the spinal cord from a patient with amyotrophic lateral sclerosis. Scale bars are 50 μm in **a** and 20 μm in **b**–**d**. **e**–**h**, Characteristic progression of specific proteinaceous lesions in neurodegenerative diseases over time (*t*, black arrows), inferred from post-mortem analyses of brains. Amyloid-β deposits and tau inclusions in brains of patients with Alzheimer's disease (**e** and **f**), α-synuclein inclusions in brains of patients with Parkinson's disease (**g**), and TDP-43 inclusions in brains of patients with amyotrophic lateral sclerosis (**h**). Three stages are shown for each disease, with white arrows indicating the putative spread of the lesions (for details see refs 5–8). Panels **e** and **f** are reproduced, with permission, from ref. 61.

[1]Department of Cellular Neurology, Hertie Institute for Clinical Brain Research, University of Tübingen, D-72076 Tübingen, Germany. [2]DZNE, German Center for Neurodegenerative Diseases, D-72076 Tübingen, Germany. [3]Yerkes National Primate Research Center, Emory University, Atlanta, Georgia 30329, USA. [4]Department of Neurology, Emory University, Atlanta, Georgia 30322, USA.

their misshapen state, the molecules aggregate with one another and impose their anomalous structure on benign PrP molecules[9–12]. Prions thus act as corruptive templates (seeds) that incite a chain-reaction of PrP misfolding and aggregation. As prions grow, fragment and spread, they perturb the function of the nervous system and ultimately cause the death of the affected individual.

The prion diseases in humans include Creutzfeldt–Jakob disease (CJD), Gerstmann–Sträussler–Scheinker disease, fatal insomnia and kuru; in nonhuman species they comprise scrapie, bovine spongiform encephalopathy, chronic wasting disease, transmissible mink encephalopathy and others[9–11,13]. Prion diseases are remarkable in that they can be genetic, infectious or sporadic in origin. Infectivity involves the transfer of seeds (prions) from one organism to another, whereas the genetic and idiopathic cases seem to develop endogenously, owing to the spontaneous misfolding and nucleation of PrP molecules into a self-propagating seed.

Pathologically, the prionotic brain is marked by spongiform degeneration, loss of neurons, gliosis and the accumulation of aggregated PrP[14,15]. Despite their similar molecular origins, the prion diseases can vary behaviourally and pathologically within and among species[14]. The diverse phenotypes are governed in part by the attributes of the affected organism, and in part by the strain-like functional diversity of prions, which in turn is thought to reflect distinct conformations of the misfolded PrP[9,16–18]. Prions exist in a range of sizes, but the most potent prions are relatively small, soluble assemblies[19]. Although molecular structure provides a framework for understanding the seeding of prions, there is evidence that prion propagation and prion toxicity are partly distinct[20].

The essence of prion disease is a crystallization-like chain reaction by which malformed PrP seeds force naive PrP molecules into a similar pathogenic architecture[21]. Recent findings now suggest that this 'prion paradigm'—the seeded corruption of otherwise harmless proteins—also underlies the ontogeny of a widening spectrum of maladies, including common age-related neurodegenerative diseases such as Alzheimer's disease and Parkinson's disease. In many of these disorders, as in prion disease, the aggregating proteins form characteristic lesions generically known as amyloid (Fig. 1).

Prion-like properties of amyloidogenic proteins

'Amyloid' refers to multimeric proteinaceous assemblies with distinctive histochemical features in tissues, and a cross-β quaternary structure as determined by X-ray fibre diffraction analysis (Box 1). In each manifestation of amyloidosis a specific protein is involved, and more than 30 different amyloidogenic proteins have been linked to disease[22,23]. Some types of amyloidosis, such as Alzheimer's disease and Parkinson's disease, are localized to the brain, but others affect organs such as the kidney, liver, heart and spleen[24].

A protein in the amyloid state (Box 1) is thermodynamically highly stable. Although all proteins have the potential to generate amyloid under the right conditions, protective mechanisms ensure that relatively few do so in living organisms[25,26]. In some cases, the amyloid state can be biologically useful in nature ('functional amyloids')[27–31]. However, in the amyloidoses the aggregation of proteins disrupts function either because the aggregates directly harm cells, tissues and organs (a gain of function), and/or because molecules sequestered within aggregates are unable to execute their required tasks[24,32]. Amyloidogenesis also involves the formation of intermediate assemblies known as oligomers and protofibrils, which can themselves be highly toxic[33].

The molecular structure responsible for the toxicity of amyloid and amyloid intermediates is not fully understood (Box 1). Contemporary work suggests that a deviation from the energetically favourable and stable parallel and in-register β-sheet amyloid state is linked to toxicity[34,35]. Consistent with this concept, assemblies consisting of out-of-register β-sheets have been identified as toxic amyloid entities[36]. There is also evidence that the process of amyloid formation itself can be harmful to cells[37]. Like the prion protein[9,17,28], other amyloidogenic proteins also can adopt conformationally distinct amyloid 'strains' (Box 1) with different biological activities and toxicities[32,38,39].

In vivo conditions that promote amyloid formation are: increased concentration of the culpable protein; mutations that destabilize the native form and allow the amyloid-prone segments of a protein to interact with each other; exposure and/or *de novo* generation of amyloid-prone segments through cleavage of the native protein or aberrant translation; thermodynamically destabilizing conditions such as conducive pH or temperature; and/or a deterioration of cellular protein quality control (proteostasis), as occurs with advancing age[3,32]. However, a highly effective way to stimulate amyloid formation is by seeding naive amyloidogenic protein molecules with β-sheet-rich aggregates of a cognate protein[21] (Box 1). Studies in experimental animals provide compelling evidence that some systemic amyloidoses can be induced in this way[40–44]. Moreover, recent research indicates that amyloidogenic proteins linked to various neurodegenerative diseases also exhibit prion-like seeded self-propagation *in vivo* and *in vitro*.

The prion-like properties of amyloid-β

The amyloid-β protein (Aβ) is a normal cleavage product of the Aβ precursor protein (APP)[45]. The physiological role of Aβ remains indeterminate, but it is now apparent that Aβ aggregation is a key player in Alzheimer's disease, the most common cause of dementia[46–48]. The self-propagation of aggregates of Aβ was predicted decades ago from *in vitro* studies[21] and from inoculation experiments with nonhuman primates[49]. However, it has only recently been established that Aβ can be induced to deposit in the living brain by a prion-like mechanism, as a result of experiments using genetically modified rodent models.

Rodent Aβ does not readily form amyloid *in vivo*; however, transgenic animal models expressing human APP generate Aβ plaques and cerebral Aβ-amyloid angiopathy (CAA) at predictable ages[50]. The intracerebral infusion of dilute Aβ-amyloid-rich brain extracts from Alzheimer's disease patients or from aged APP-transgenic mice stimulates the premature formation of plaques and CAA in these models[51,52]. The pathogenic process is also inducible in transgenic rodents that express human-sequence Aβ but do not otherwise develop Aβ deposits within their normal lifespans[53,54] (Box 2). The induction of Aβ deposition is dependent both on the concentration of Aβ seeds in the extract and on the production of human-sequence Aβ by the host brain[52]. Control brain extracts lacking aggregated Aβ do not induce Aβ deposition, and the denaturation of proteins or the selective removal of Aβ completely negates the ability of the extracts to seed Aβ aggregation[52]. Seeding of Aβ deposition also is abrogated by active or passive immunization of the mice against Aβ[52].

Like prions, Aβ seeds range in size from small, soluble, protease-sensitive aggregates to large, insoluble, protease-resistant fibrils[55]. Aβ lesions in APP transgenic mice are inducible by injections of pure, synthetic human Aβ fibrils[56], although (similar to recombinant prion protein[57]), the potency of synthetic Aβ seeds is less than that of Aβ aggregates formed within the living brain. The reasons for the relative ineffectiveness of synthetic Aβ seeds are not yet apparent; it may be that additional co-factors or chaperones must participate in their production (as has been shown for recombinant prions[58]), or that Aβ (like PrP) can fold into polymorphic and polyfunctional strains (Box 1)[32,38,52,59]. Indeed, strain-like variations in the morphology of Aβ lesions can be seeded in APP-transgenic mice by the injection of Aβ-laden brain extracts from different sources[60].

Injection of Aβ-rich brain extract into one brain region eventually triggers deposits in axonally coupled regions[61]. Over time, widespread areas of the brain are involved, including neocortical and subcortical regions[62], similar to areas affected in Alzheimer's disease (Fig. 1). In addition, Aβ deposition is inducible in the brain by the injection of extract into the peritoneal cavity[63]. Taken together, these studies reinforce the hypothesis that the aggregation of Aβ can be seeded in susceptible hosts by a mechanism that closely resembles the molecular templating of prions.

The prion-like properties of tau and α-synuclein

Accumulating experimental data indicate that the seeding principle also applies to other pathogenic proteins, many of which form amyloid-like inclusions within cells. Tau is a cytoplasmic protein that normally helps

BOX 1

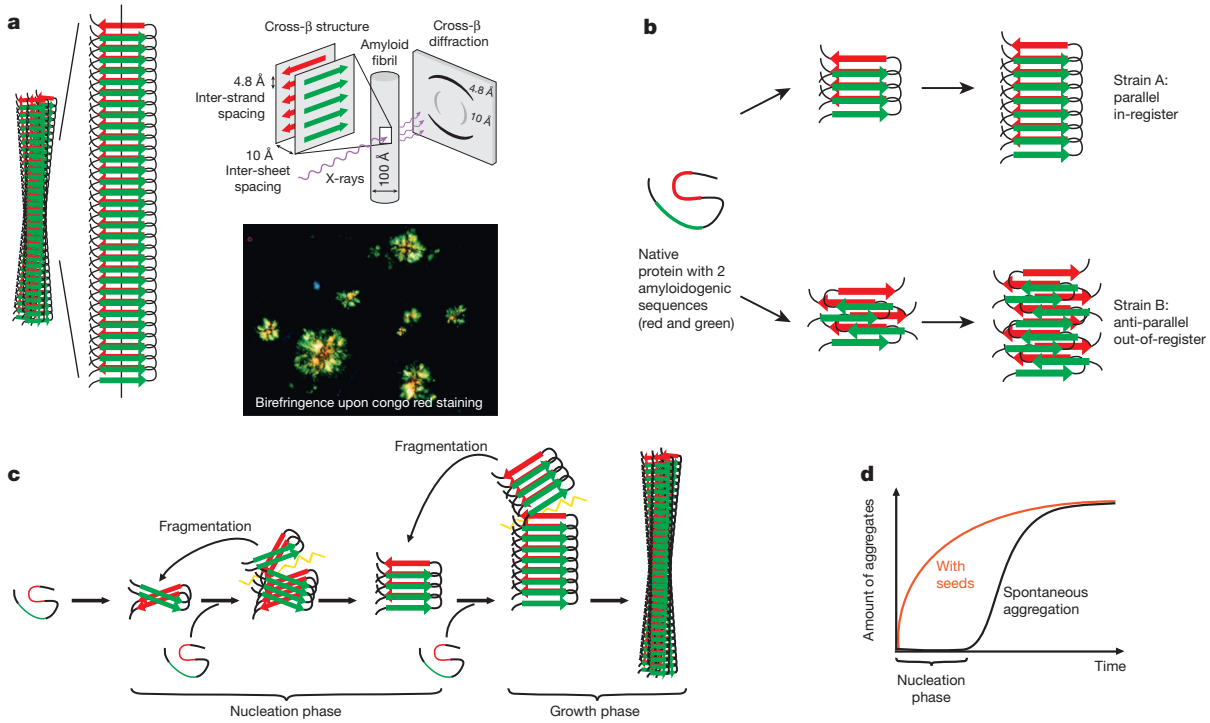
The amyloid state of proteins as a framework to explain prion-like seeding

A protein in the amyloid state forms bundles of twisted, unbranched filaments. Each filament is composed of sheets of β-strands (panel a). These β-sheets run parallel to the filament axis, and the strands are nearly perpendicular to the long axis. This structural arrangement produces a distinctive, cross-β X-ray diffraction pattern that reflects the characteristic spacing between the β-sheets and the β-strands[32]. Biophysicists classify amyloid based on the X-ray diffraction pattern, whereas pathologists define amyloid as deposits of fibrillar protein in cells or tissues that show reddish/green birefringence under cross-polarized light after staining with the dye Congo red[22] (panel a). Part of panel a is modified, with permission, from ref. 32.

In the most common amyloids, the β-sheets consist of parallel β-strands that are hydrogen-bonded by their backbones. The sheet is 'in-register' when identical side chains are on top of each other. The sheets are bonded to each other via amino acid side chains that are inter-digitated like a zipper. Amyloid 'steric zippers' can be formed from identical or different β-strands (homosteric versus heterosteric zippers). An amyloid-forming protein may contribute more than one β-strand segment to the cross-β amyloid backbone (spine)[32].

At the molecular level, amyloids can be highly polymorphic; that is, a given β-strand segment is able to form a variety of distinct cross-β amyloid spines and filamentous structures (panel b). Such conformational variants are suggested to be the molecular basis of amyloid 'strains', that is, amyloids formed from a particular protein but with different biological activities[16,32].

Amyloid formation (panel c) starts with a slow nucleation phase (the aggregation of the protein into a seed) that may go through a series of intermediate states until the initial segment of the amyloid spine is formed[21,124]. Monomers or oligomeric structures are then bonded to the ends of the initial amyloid seed by conformational conversion. With increasing length, and depending on the conformational stability of the amyloid spine, the growing fibril can eventually break, either spontaneously or actively through cellular processes. In this way, amyloid formation becomes self-propagating through the generation and spread of new amyloid seeds. The kinetics of amyloid fibril formation are a function of the rates of nucleation, growth, and fragmentation[21,125]. The lag time that precedes protein aggregation *in vitro* can be greatly shortened by the addition of pre-formed exogenous seeds (panel d).



to stabilize microtubules, but it becomes hyperphosphorylated and prone to aggregation (Fig. 1) in a variety of neurodegenerative conditions, including Alzheimer's disease, frontotemporal lobar degeneration, progressive supranuclear palsy, corticobasal degeneration, chronic traumatic encephalopathy and others[64,65]. The intracerebral injection of brain extracts containing aggregated tau induces tauopathy in tau-transgenic host mice, and the induced tau lesions propagate systematically from the injection site to axonally connected areas, consistent with neuronal uptake, transport and release of tau seeds[66]. Unlike Aβ lesions, tauopathy can be seeded by exogenous tau aggregates in non-transgenic (wild-type) mice[67,68], indicating that endogenous murine tau itself is amyloidogenic, even though wild-type mice do not develop tauopathy spontaneously (Box 2, panel b). Tau aggregation is also inducible by recombinant tau fibrils in cultured cells[69] and in tau-transgenic mice[70]. As in the case of Aβ and prions[19,55], tau seeds are of many sizes, and small, soluble assemblies are effective seeds[67]. Intracerebral injections of brain extracts from various human tauopathies have shown that tau lesions in mice can be induced to resemble those in the corresponding human diseases[68]. These results exemplify the conformation-dependent templated propagation of tau multimers, and evoke the hypothesis that specific tau conformers give rise to clinically distinct tauopathies, reminiscent of prion strains.

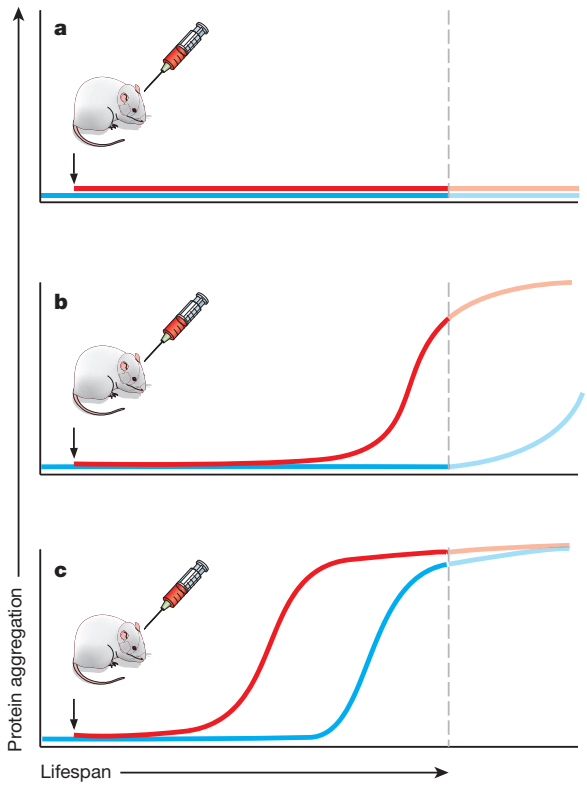
In α-synucleinopathies such as Parkinson's disease and dementia with Lewy bodies, misfolded α-synuclein assembles into intracellular fibrillar inclusions called Lewy bodies (Fig. 1) and Lewy neurites[71]. Clues to the induction of these lesions *in vivo* emerged from the surprising discovery that some foetal dopaminergic neurons transplanted into the brains of Parkinson's disease patients contain α-synuclein-positive Lewy bodies 11 to 16 years after transplant surgery[72,73]. Studies *in vitro*[74–76] and in experimental animals[74,75,77–79] suggest that the inductive agent probably consists of α-synuclein seeds formed within the host brain that transfer to the

grafted neurons to induce α-synuclein aggregation. Consistent with such a transfer, intracerebral injection of brain extracts containing aggregated α-synuclein into young, α-synuclein-transgenic mice stimulates the formation of α-synuclein lesions in the host[78,79]; the lesions then emerge in anatomically linked regions of the brain[79], suggesting a parallel with the apparent spread of α-synuclein deposits in the human brain (Fig. 1). Eventually, the mice develop progressive neurodegeneration, signs of Parkinson's-disease-like motor dysfunction, and premature death[78,79]. Importantly, intracerebral injections of synthetic (human or mouse) α-synuclein fibrils (as well as autopsy-derived brain extracts from a case of Lewy body disease) also induce Lewy-body-like pathology and neuronal degeneration in non-transgenic (wild-type) host mice[80,81]. Finally, synthetic α-synuclein fibrils are able to induce tauopathy, possibly by a cross-seeding mechanism[82]. When synthetic fibrils with distinct proteinase K cleavage sites are generated, these variant α-synuclein fibrils differ in their ability to stimulate tauopathy[82]. Differential proteinase K cleavage of α-synuclein aggregates is also observed in the brains of Parkinson's patients, indicative of alternative α-synuclein conformations[82].

BOX 2

Hypothetical model of seeded amyloid induction in mice

An endogenous protein (panel a, blue) is not amyloidogenic under physiological conditions. In this case, the application of a seed (panel a, red) will not induce protein aggregation. A protein is amyloidogenic but does not aggregate during the lifespan of the mouse (which ends at the dotted line), either because endogenously formed seeds are removed by an effective proteostasis network, or because seed formation is inefficient and therefore unlikely to occur during the mouse's lifetime (panel b, blue). However, with an appropriate seed (single inoculation), the onset of protein aggregation is advanced and occurs before the mouse reaches the end of its lifespan (panel b, red). A protein is highly amyloidogenic and typically aggregates with ageing of the mouse (panel c, blue). The addition of an appropriate seed advances the onset of protein aggregation (panel c, red).



The growing family of prion-like proteins

Amyotrophic lateral sclerosis (ALS) and frontotemporal lobar degeneration (FTLD) are two representatives of a cluster of related brain disorders[83,84]. Genetic factors play a conspicuous role in the ontogeny of these diseases, which can result from mutations involving genes coding for tau, superoxide dismutase-1 (SOD1), TAR DNA-binding protein 43 (TDP-43), fused in sarcoma (FUS), C9orf72, heterogeneous nuclear ribonucleoproteins (hnRNPs) and others[83–86]. A commonality among these proteins is their susceptibility to aggregation, particularly in the context of ribonucleoprotein granules that form in cells as a result of stress. Several of these proteins include aggregation-prone stretches of amino acids known as 'prion-like domains', which occur in a significant subgroup of RNA-binding proteins[87]. Under stressful conditions, the proteins form an amyloid-like cross-β sheet conformation that enables them to recruit and structurally convert like molecules, in which state their function is blocked[88]. In some instances, the aggregating proteins trap and thereby disable other proteins as well[89]. Importantly, ribonucleoprotein granules can be disassembled when conditions are more favourable for the cell[88]. However, the amyloid-like cross-β sheet conformations can also convert into stable amyloid conformations either naturally or by disease-linked mutations in the prion domains, as shown for hnRNPs[86] and probably also for TDP-43 (ref. 88). Indeed, TDP-43 inclusions in sporadic ALS appear in a sequential pattern consistent with the hypothesis of seeded propagation (Fig. 1). Although it has not been shown *in vivo*, a prion-like cell-to-cell propagation of TDP-43 and SOD1 aggregates has been demonstrated in cell cultures[90–92], supporting the hypothesis that the prion paradigm may also apply to diseases in the ALS–FTLD spectrum[93].

Finally, it is worth noting that diseases involving prion-like seeding mechanisms are likely to expand beyond the classical amyloidoses and neurodegenerative disorders. The tumour suppressor protein p53 acts mainly in the nucleus to prevent cells from dividing uncontrollably, and in many human cancers p53 function is diminished[94]. Interestingly, p53 can aggregate into amyloid-like assemblies that seed the further aggregation and mislocalization of the protein; in this ectopic state, it is thought that p53 loses its ability to suppress cell proliferation, thus promoting malignant growth[95,96].

The spectre of infectivity

A defining feature of prions is their infectivity[9]; that is, their capacity to cause disease when transferred from an affected individual to a naive recipient. The ease with which prion disease is transmissible from animal to animal under ordinary circumstances varies; natural infection by prions is facile in some nonhuman prionoses, such as chronic wasting disease in cervids, and scrapie in sheep[97,98]. However, in contemporary humans exogenous transmission of prion disease is rare, totalling approximately 2,700 documented cases of kuru among the Fore people of Papua New Guinea between 1957 and 2004 (ref. 99) and approximately 700 known cases of CJD as of 2012 (ref. 100). Most instances of human transmission have involved unusual routes of exposure such as ritual cannibalism (for kuru), treatment with contaminated growth hormone extracted from human pituitaries, or the use of contaminated dura mater in neurosurgical procedures. With the cessation of cannibalism among the Fore in the 1950s, recognition of the prion as the disease agent, and implementation of preventive measures, the transmission of prion disease to humans has virtually ceased[99,100].

We do not yet know why some proteopathies can move from animal to animal whereas others do not[61,101]. To become infectious under natural conditions, pathogenic seeds must exit the body and travel intact to another organism, where they resume replication. During this journey, the seeds must resist destruction and overcome a number of biological and physical barriers. Prions thus owe their infectivity, at least in part, to their durability, their replication over multiple serial passages from one host to another, and to attributes of the host that either promote or

restrain transmission[9–11]. The precise structure of the infectious PrP assembly has not yet been solved, although most data point to a β-sheet-rich amyloid-like conformation[34]. It is conceivable that the infectivity of prions results from a particular molecular architecture of PrP (Box 1). A comparison of prions with other self-propagating protein assemblies, including systemic amyloids with transmissible properties[40,43], could furnish clues to the features that determine whether or not a proteopathic agent is infectious under everyday circumstances.

Therapeutic implications

Proteins that aggregate in disparate neurodegenerative disorders share with prions the molecular properties of nucleation, templating, growth, multiplication and spread. Each of these phenomena presents potential therapeutic targets. For example, reducing the production or stimulating the removal of amyloidogenic proteins will lower the concentration of the protein and thereby impede nucleation and growth[32]. Another strategy is to stabilize the native conformation of an amyloidogenic protein, an approach that has been pioneered to treat transthyretin amyloidosis[102]. Stabilizing existing aggregates also could have therapeutic value, by hindering their fragmentation and thus the multiplication of seeds. Another option may be to exploit structural elements of the cross-β sheet to prevent the corruptive templating of cognate proteins. This strategy is particularly promising when the atomic structure of the seed is known, and interfering agents can therefore be designed to bind distinct amyloid conformations[103].

The progressive march of symptoms in neurodegenerative diseases has been proposed to involve the systematic advance of a pathogen along neuronal pathways[5–8,61,71,104,105]. Imaging studies confirm that the differential vulnerability of brain regions to neurodegenerative changes is correlated with the strength of neuronal connections among the affected areas[106–109]. These patterns could reflect the trafficking of proteopathic seeds among interconnected brain regions. The discharge of amyloid-β[110,111] and tau[112] into the extracellular space is regulated by neuronal activity. Accordingly, the progression of disease may be abrogated either by arresting seeds as they travel between cells[113], or by targeting the cellular processes of release, uptake and transport[77,92,114–121].

Finally, a hallmark of chronic diseases is a long silent phase of pathogenesis preceding the onset of symptoms[122]. In Alzheimer's disease, the pathological cascade is thought to begin 10 to 20 years or more before the first clinical symptoms appear[47,48]. Thus, in addition to being therapeutic targets, small, soluble proteinaceous seeds in bodily fluids could serve as early biomarkers for pre-symptomatic disease.

Perspectives

The theoretical and practical implications of the prion paradigm hinge on the resolution of a number of issues. Perhaps the most pressing need is a precise description of the molecular structure of the amyloidogenic seeds. The physicochemical and cellular conditions that promote the formation, growth and proliferation of seeds *in vivo* also are still ambiguous. Another question is whether the clinical and pathological diversity of neurodegenerative diseases reflects the strain-like structural diversity of the aggregates *in vivo*. In addition, different proteopathies often coincide in the ageing brain, and it remains uncertain whether these diseases result from independent pathologic processes, a coincidental response to a common instigator, or from the 'cross-seeding' of one type of aggregated protein by another. At present, it seems unlikely that non-prion neurodegenerative diseases are infectious under ordinary circumstances[61,101], but further epidemiological studies are warranted, particularly with regard to uncommon routes of exposure[61].

After decades of controversy, the prion now is widely accepted as an unorthodox, but formidable, agent of disease. On the margins of the prion debate has long been the notion that prion-like processes might drive the misfolding and aggregation of proteins involved in other diseases[9,123]. Recent experimental work supports this hypothesis. By establishing seeded protein aggregation as a cardinal pathogenic principle, the prion paradigm stands to consolidate and focus treatment strategies for a broad spectrum of diseases.

Received 14 May; accepted 17 July 2013.

1. Uversky, V. N. & Dunker, A. K. The case for intrinsically disordered proteins playing contributory roles in molecular recognition without a stable 3D structure. *F1000 Biol. Rep.* **5,** 1 (2013).
2. Malinovska, L., Kroschwald, S. & Alberti, S. Protein disorder, prion propensities, and self-organizing macromolecular collectives. *Biochim. Biophys. Acta* **1834,** 918–931 (2013).
3. Balch, W. E., Morimoto, R. I., Dillin, A. & Kelly, J. W. Adapting proteostasis for disease intervention. *Science* **319,** 916–919 (2008).
4. Hartl, F. U., Bracher, A. & Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature* **475,** 324–332 (2011).
5. Thal, D. R., Rub, U., Orantes, M. & Braak, H. Phases of amyloid-β-deposition in the human brain and its relevance for the development of AD. *Neurology* **58,** 1791–1800 (2002).
6. Braak, H. & Braak, E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* **82,** 239–259 (1991).
7. Braak, H. *et al.* Staging of brain pathology related to sporadic Parkinson's disease. *Neurobiol. Aging* **24,** 197–211 (2003).
8. Brettschneider, J. *et al.* Stages of pTDP-43 pathology in amyotrophic lateral sclerosis. *Ann. Neurol.* http://dx.doi.org/10.1002/ana.23937 (2013).
9. Prusiner, S. B. Prions. *Proc. Natl Acad. Sci. USA* **95,** 13363–13383 (1998).
10. Collinge, J. Prion diseases of humans and animals: their causes and molecular basis. *Annu. Rev. Neurosci.* **24,** 519–550 (2001).
11. Aguzzi, A. & Calella, A. M. Prions: protein aggregation and infectious diseases. *Physiol. Rev.* **89,** 1105–1152 (2009).
12. Caughey, B., Baron, G. S., Chesebro, B. & Jeffrey, M. Getting a grip on prions: oligomers, amyloids, and pathological membrane interactions. *Annu. Rev. Biochem.* **78,** 177–204 (2009).
13. Head, M. W. & Ironside, J. W. Review: Creutzfeldt–Jakob disease: prion protein type, disease phenotype and agent strain. *Neuropathol. Appl. Neurobiol.* **38,** 296–310 (2012).
14. DeArmond, S. J. & Prusiner, S. B. Etiology and pathogenesis of prion diseases. *Am. J. Pathol.* **146,** 785–811 (1995).
15. Wadsworth, J. D. & Collinge, J. Molecular pathology of human prion disease. *Acta Neuropathol.* **121,** 69–77 (2011).
16. Tanaka, M., Collins, S. R., Toyama, B. H. & Weissman, J. S. The physical basis of how prion conformations determine strain phenotypes. *Nature* **442,** 585–589 (2006).
17. Collinge, J. & Clarke, A. R. A general model of prion strains and their pathogenicity. *Science* **318,** 930–936 (2007).
18. Colby, D. W. & Prusiner, S. B. Prions. *Cold Spring Harb. Perspect. Biol.* **3,** a006833 (2011).
19. Silveira, J. R. *et al.* The most infectious prion protein particles. *Nature* **437,** 257–261 (2005).
20. Sandberg, M. K., Al-Doujaily, H., Sharps, B., Clarke, A. R. & Collinge, J. Prion propagation and toxicity *in vivo* occur in two distinct mechanistic phases. *Nature* **470,** 540–542 (2011).
21. Jarrett, J. T. & Lansbury, P. T., Jr. Seeding "one-dimensional crystallization" of amyloid: a pathogenic mechanism in Alzheimer's disease and scrapie? *Cell* **73,** 1055–1058 (1993).
One of the first papers to suggest a common pathogenic mechanism between prion diseases and Alzheimer's disease based on *in vitro* protein aggregation studies.
22. Sipe, J. D. *et al.* Amyloid fibril protein nomenclature: 2012 recommendations from the Nomenclature Committee of the International Society of Amyloidosis. *Amyloid* **19,** 167–170 (2012).
23. Buxbaum, J. N. & Linke, R. P. A molecular history of the amyloidoses. *J. Mol. Biol.* **421,** 142–159 (2012).
24. Blancas-Mejía, L. M. & Ramirez-Alvarado, M. Systemic amyloidoses. *Annu. Rev. Biochem.* **82,** 745–774 (2013).
25. Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci.* **24,** 329–332 (1999).
An early paper suggesting that all proteins have the intrinsic potential to form amyloid under suitable conditions.
26. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75,** 333–366 (2006).
27. Maji, S. K. *et al.* Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325,** 328–332 (2009).
28. Greenwald, J. & Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **18,** 1244–1260 (2010).
29. Hou, F. *et al.* MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell* **146,** 448–461 (2011).
30. Blanco, L. P., Evans, M. L., Smith, D. R., Badtke, M. P. & Chapman, M. R. Diversity, biogenesis and function of microbial amyloids. *Trends Microbiol.* **20,** 66–73 (2012).
31. Wickner, R. B. *et al.* Amyloids and yeast prion biology. *Biochemistry* **52,** 1514–1527 (2013).
32. Eisenberg, D. & Jucker, M. The amyloid state of proteins in human diseases. *Cell* **148,** 1188–1203 (2012).
33. Mucke, L. & Selkoe, D. J. Neurotoxicity of amyloid β-protein: synaptic and network dysfunction. Cold Spring Harbor Perspect. Med. **2,** a006338 (2012).

34. Tycko, R. & Wickner, R. B. Molecular structures of amyloid and prion fibrils: consensus versus controversy. *Acc. Chem. Res.* **46,** 1487–1496 (2013).
35. Liu, C. *et al.* Out-of-register β-sheets suggest a pathway to toxic amyloid aggregates. *Proc. Natl Acad. Sci. USA* **109,** 20913–20918 (2012).
36. Laganowsky, A. *et al.* Atomic view of a toxic amyloid small oligomer. *Science* **335,** 1228–1231 (2012).
37. Jan, A. *et al.* Aβ42 neurotoxicity is mediated by ongoing nucleated polymerization process rather than by discrete Aβ42 species. *J. Biol. Chem.* **286,** 8585–8596 (2011).
38. Petkova, A. T. *et al.* Self-propagating, molecular-level polymorphism in Alzheimer's β-amyloid fibrils. *Science* **307,** 262–265 (2005).
This study demonstrates that the morphology and toxicity of synthetic β-amyloid fibrils are linked to variations in the molecular structure of the protein, and that these properties can be propagated to subsequent generations of fibrils *in vitro* by a seeding mechanism.
39. Toyama, B. H. & Weissman, J. S. Amyloid structure: conformational diversity and consequences. *Annu. Rev. Biochem.* **80,** 557–585 (2011).
40. Westermark, G. T. & Westermark, P. Prion-like aggregates: infectious agents in human disease. *Trends Mol. Med.* **16,** 501–507 (2010).
An overview of the prion-like properties of systemic amyloids.
41. Yan, J. *et al.* Cross-seeding and cross-competition in mouse apolipoprotein A-II amyloid fibrils and protein A amyloid fibrils. *Am. J. Pathol.* **171,** 172–180 (2007).
42. Xing, Y. *et al.* Transmission of mouse senile amyloidosis. *Lab. Invest.* **81,** 493–499 (2001).
43. Zhang, B. *et al.* Fecal transmission of AA amyloidosis in the cheetah contributes to high incidence of disease. *Proc. Natl Acad. Sci. USA* **105,** 7263–7268 (2008).
44. Korenaga, T. *et al.* Transmission of amyloidosis in offspring of mice with AApoAII amyloidosis. *Am. J. Pathol.* **168,** 898–906 (2006).
45. De Strooper, B. Proteases and proteolysis in Alzheimer disease: a multifactorial view on the disease process. *Physiol. Rev.* **90,** 465–494 (2010).
46. Holtzman, D. M., Morris, J. C. & Goate, A. M. Alzheimer's disease: the challenge of the second century. *Sci. Transl. Med.* **3,** 77s71 (2011).
47. Bateman, R. J. *et al.* Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N. Engl. J. Med.* **367,** 795–804 (2012).
48. Villemagne, V. L. *et al.* Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: a prospective cohort study. *Lancet Neurol.* **12,** 357–367 (2013).
49. Baker, H. F., Ridley, R. M., Duchen, L. W., Crow, T. J. & Bruton, C. J. Evidence for the experimental transmission of cerebral β-amyloidosis to primates. *Int. J. Exp. Pathol.* **74,** 441–454 (1993).
Evidence that senile plaques in the brains of nonhuman primates are inducible by the intracerebral injection of Alzheimer's brain homogenates.
50. Jucker, M. The benefits and limitations of animal models for translational research in neurodegenerative diseases. *Nature Med.* **16,** 1210–1214 (2010).
51. Kane, M. D. *et al.* Evidence for seeding of β-amyloid by intracerebral infusion of Alzheimer brain extracts in β-amyloid precursor protein-transgenic mice. *J. Neurosci.* **20,** 3606–3611 (2000).
52. Meyer-Luehmann, M. *et al.* Exogenous induction of cerebral β-amyloidogenesis is governed by agent and host. *Science* **313,** 1781–1784 (2006).
The first conclusive demonstration that an aggregated form of Aβ is the β-amyloid-inducing agent in donor brain extracts, and that the transmission of cerebral β-amyloidosis is dependent on the nature of both the seed and the host.
53. Morales, R., Duran-Aniotz, C., Castilla, J., Estrada, L. D. & Soto, C. *De novo* induction of amyloid-β deposition *in vivo*. *Mol. Psychiatry* **17,** 1347–1353 (2012).
54. Rosen, R. F. *et al.* Exogenous seeding of cerebral β-amyloid deposition in βAPP-transgenic rats. *J. Neurochem.* **120,** 660–666 (2012).
55. Langer, F. *et al.* Soluble amyloid-β seeds are potent inducers of cerebral β-amyloid deposition. *J. Neurosci.* **31,** 14488–14495 (2011).
56. Stöhr, J. *et al.* Purified and synthetic Alzheimer's amyloid beta (amyloid-β) prions. *Proc. Natl Acad. Sci. USA* **109,** 11025–11030 (2012).
Demonstration that Aβ deposition can be seeded in the brain by synthetic Aβ seeds.
57. Legname, G. *et al.* Synthetic mammalian prions. *Science* **305,** 673–676 (2004).
58. Wang, F., Wang, X., Yuan, C. G. & Ma, J. Generating a prion with bacterially expressed recombinant prion protein. *Science* **327,** 1132–1135 (2010).
59. LeVine III, H. & Walker, L. C. Molecular polymorphism of amyloid-β in Alzheimer's disease. *Neurobiol. Aging* **31,** 542–548 (2010).
60. Heilbronner, G. *et al.* Seeded strain-like transmission of β-amyloid morphotypes in APP transgenic mice. *EMBO Rep.* http://dx.doi.org/10.1038/embor.2013.137 (3 September 2013).
61. Jucker, M. & Walker, L. C. Pathogenic protein seeding in Alzheimer disease and other neurodegenerative disorders. *Ann. Neurol.* **70,** 532–540 (2011).
62. Hamaguchi, T. *et al.* The presence of amyloid-β seeds, and not age per se, is critical to the initiation of amyloid-β deposition in the brain. *Acta Neuropathol.* **123,** 31–37 (2012).
63. Eisele, Y. S. *et al.* Peripherally applied amyloid-β-containing inoculates induce cerebral β-amyloidosis. *Science* **330,** 980–982 (2010).
Demonstration that cerebral Aβ deposition can be induced by the introduction of Aβ seeds into the peritoneal cavity (a site outside the brain).
64. Lee, V. M., Goedert, M. & Trojanowski, J. Q. Neurodegenerative tauopathies. *Annu. Rev. Neurosci.* **24,** 1121–1159 (2001).
65. Goldstein, L. E. *et al.* Chronic traumatic encephalopathy in blast-exposed military veterans and a blast neurotrauma mouse model. *Science Transl. Med.* **4,** 134ra160 (2012).
66. Clavaguera, F. *et al.* Transmission and spreading of tauopathy in transgenic mouse brain. *Nature Cell Biol.* **11,** 909–913 (2009).
The first report that tau lesions can be instigated in tau-transgenic mice by the intracerebral infusion of brain extracts containing aggregated tau.
67. Lasagna-Reeves, C. A. *et al.* Alzheimer brain-derived tau oligomers propagate pathology from endogenous tau. *Scientific Rep.* **2,** 700 (2012).
68. Clavaguera, F. *et al.* Brain homogenates from human tauopathies induce tau inclusions in mouse brain. *Proc. Natl Acad. Sci. USA* **110,** 9535–9540 (2013).
69. Guo, J. L. & Lee, V. M. Neurofibrillary tangle-like tau pathology induced by synthetic tau fibrils in primary neurons over-expressing mutant tau. *FEBS Lett.* **587,** 717–723 (2013).
70. Iba, M. *et al.* Synthetic tau fibrils mediate transmission of neurofibrillary tangles in a transgenic mouse model of Alzheimer's-like tauopathy. *J. Neurosci.* **33,** 1024–1037 (2013).
71. Goedert, M., Spillantini, M. G., Del Tredici, K. & Braak, H. 100 years of Lewy pathology. *Nature Rev. Neurology* **9,** 13–24 (2013).
72. Kordower, J. H., Chu, Y., Hauser, R. A., Freeman, T. B. & Olanow, C. W. Lewy body-like pathology in long-term embryonic nigral transplants in Parkinson's disease. *Nature Med.* **14,** 504–506 (2008).
73. Li, J. Y. *et al.* Lewy bodies in grafted neurons in subjects with Parkinson's disease suggest host-to-graft disease propagation. *Nature Med.* **14,** 501–503 (2008).
These two reports (refs 72, 73) suggest that endogenous α-synuclein seeds in the brains of humans with Parkinson's disease can induce the aggregation of α-synuclein in grafted neurons.
74. Hansen, C. *et al.* α-Synuclein propagates from mouse brain to grafted dopaminergic neurons and seeds aggregation in cultured human cells. *J. Clin. Invest.* **121,** 715–725 (2011).
75. Kordower, J. H. *et al.* Transfer of host-derived α synuclein to grafted dopaminergic neurons in rat. *Neurobiol. Dis.* **43,** 552–557 (2011).
76. Volpicelli-Daley, L. A. *et al.* Exogenous α-synuclein fibrils induce Lewy body pathology leading to synaptic dysfunction and neuron death. *Neuron* **72,** 57–71 (2011).
77. Desplats, P. *et al.* Inclusion formation and neuronal cell death through neuron-to-neuron transmission of α-synuclein. *Proc. Natl Acad. Sci. USA* **106,** 13010–13015 (2009).
78. Mougenot, A. L. *et al.* Prion-like acceleration of a synucleinopathy in a transgenic mouse model. *Neurobiol. Aging* **33,** 2225–2228 (2012).
79. Luk, K. C. *et al.* Intracerebral inoculation of pathological α-synuclein initiates a rapidly progressive neurodegenerative α-synucleinopathy in mice. *J. Exp. Med.* **209,** 975–986 (2012).
These two studies (refs 78, 79) first showed that α-synuclein lesions can be induced in α-synuclein-transgenic mice by the intracerebral inoculation of brain extracts rich in aggregated α-synuclein.
80. Luk, K. C. *et al.* Pathological α-synuclein transmission initiates Parkinson-like neurodegeneration in nontransgenic mice. *Science* **338,** 949–953 (2012).
Describes the instigation of α-synucleinopathy by the intracerebral injection of synthetic α-synuclein fibrils into non-transgenic mice.
81. Masuda-Suzukake, M. *et al.* Prion-like spreading of pathological α-synuclein in brain. *Brain* **136,** 1178–1138 (2013).
82. Guo, J. L. *et al.* Distinct α-synuclein strains differentially promote tau inclusions in neurons. *Cell* **154,** 103–117 (2013).
This study describes the strain-dependent induction of tau aggregation by aggregated α-synuclein *in vitro* and *in vivo*.
83. Van Langenhove, T., van der Zee, J. & Van Broeckhoven, C. The molecular basis of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum. *Ann. Med.* **44,** 817–828 (2012).
84. Rademakers, R., Neumann, M. & Mackenzie, I. R. Advances in understanding the molecular basis of frontotemporal dementia. *Nature Rev. Neurol.* **8,** 423–434 (2012).
85. Cruts, M., Gijselinck, I., Van Langenhove, T., van der Zee, J. & Van Broeckhoven, C. Current insights into the C9orf72 repeat expansion diseases of the FTLD/ALS spectrum. *Trends Neurosci.* **36,** 450–459 (2013).
86. Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495,** 467–473 (2013).
87. King, O. D., Gitler, A. D. & Shorter, J. The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. *Brain Res.* **1462,** 61–80 (2012).
88. Li, Y. R., King, O. D., Shorter, J. & Gitler, A. D. Stress granules as crucibles of ALS pathogenesis. *J. Cell Biol.* **201,** 361–372 (2013).
89. Olzscha, H. *et al.* Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. *Cell* **144,** 67–78 (2011).
90. Furukawa, Y., Kaneko, K., Watanabe, S., Yamanaka, K. & Nukina, N. A seeding reaction recapitulates intracellular formation of Sarkosyl-insoluble transactivation response element (TAR) DNA-binding protein-43 inclusions. *J. Biol. Chem.* **286,** 18664–18672 (2011).
91. Münch, C., O'Brien, J. & Bertolotti, A. Prion-like propagation of mutant superoxide dismutase-1 misfolding in neuronal cells. *Proc. Natl Acad. Sci. USA* **108,** 3548–3553 (2011).
92. Grad, L. I. *et al.* Intermolecular transmission of superoxide dismutase 1 misfolding in living cells. *Proc. Natl Acad. Sci. USA* **108,** 16398–16403 (2011).
93. Polymenidou, M. & Cleveland, D. W. The seeds of neurodegeneration: prion-like spreading in ALS. *Cell* **147,** 498–508 (2011).
94. Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. p53 mutations in human cancers. *Science* **253,** 49–53 (1991).
95. Xu, J. *et al.* Gain of function of mutant p53 by coaggregation with multiple tumor suppressors. *Nature Chem. Biol.* **7,** 285–295 (2011).
96. Ano Bom, A. P. *et al.* Mutant p53 aggregates into prion-like amyloid oligomers and fibrils: implications for cancer. *J. Biol. Chem.* **287,** 28152–28162 (2012).

These two studies (refs 95, 96) demonstrate the aggregation and prion-like characteristics of p53, a protein that normally regulates the cell-cycle and acts to inhibit tumour formation.
97. Sigurdson, C. J. & Aguzzi, A. Chronic wasting disease. *Biochim. Biophys. Acta* **1772,** 610–618 (2007).
98. Hoinville, L. J. A review of the epidemiology of scrapie in sheep. *Rev. Sci. Tech.* **15,** 827–852 (1996).
99. Collinge, J. *et al.* Kuru in the 21st century—an acquired human prion disease with very long incubation periods. *Lancet* **367,** 2068–2074 (2006).
100. Brown, P. *et al.* Iatrogenic Creutzfeldt–Jakob disease, final assessment. *Emerg. Infect. Dis.* **18,** 901–907 (2012).
101. Irwin, D. J. *et al.* Evaluation of potential infectivity of Alzheimer and Parkinson disease proteins in recipients of cadaver-derived human growth hormone. *JAMA Neurol.* **70,** 462–468 (2013).
102. Johnson, S. M., Connelly, S., Fearns, C., Powers, E. T. & Kelly, J. W. The transthyretin amyloidoses: from delineating the molecular mechanism of aggregation linked to pathology to a regulatory-agency-approved drug. *J. Mol. Biol.* **421,** 185–203 (2012).
103. Sievers, S. A. *et al.* Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **475,** 96–100 (2011).
104. Saper, C. B., Wainer, B. H. & German, D. C. Axonal and transneuronal transport in the transmission of neurological disease: potential role in system degenerations, including Alzheimer's disease. *Neuroscience* **23,** 389–398 (1987).
105. Ravits, J. M. & La Spada, A. R. ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration. *Neurology* **73,** 805–811 (2009).
106. Eidelberg, D. & Surmeier, D. J. Brain networks in Huntington disease. *J. Clin. Invest.* **121,** 484–492 (2011).
107. Zhou, J., Gennatas, E. D., Kramer, J. H., Miller, B. L. & Seeley, W. W. Predicting regional neurodegeneration from the healthy brain functional connectome. *Neuron* **73,** 1216–1227 (2012).
108. Raj, A., Kuceyeski, A. & Weiner, M. A network diffusion model of disease progression in dementia. *Neuron* **73,** 1204–1215 (2012).
109. Gardner, R. C. *et al.* Intrinsic connectivity network disruption in progressive supranuclear palsy. *Ann. Neurol.* **73,** 603–616 (2013).
110. Bero, A. W. *et al.* Neuronal activity regulates the regional vulnerability to amyloid-β deposition. *Nature Neurosci.* **14,** 750–756 (2011).
111. Dolev, I. *et al.* Spike bursts increase amyloid-β 40/42 ratio by inducing a presenilin-1 conformational change. *Nature Neurosci.* **16,** 587–595 (2013).
112. Pooler, A. M., Phillips, E. C., Lau, D. H., Noble, W. & Hanger, D. P. Physiological release of endogenous tau is stimulated by neuronal activity. *EMBO Rep.* **14,** 389–394 (2013).
113. Bae, E. J. *et al.* Antibody-aided clearance of extracellular α-synuclein prevents cell-to-cell aggregate transmission. *J. Neurosci.* **32,** 13454–13469 (2012).
114. Polymenidou, M. & Cleveland, D. W. Prion-like spread of protein aggregates in neurodegeneration. *J. Exp. Med.* **209,** 889–893 (2012).
115. Holmes, B. B. & Diamond, M. I. Cellular mechanisms of protein aggregate propagation. *Curr. Opin. Neurol.* **25,** 721–726 (2012).
116. Wu, J. W. *et al.* Small misfolded Tau species are internalized via bulk endocytosis and anterogradely and retrogradely transported in neurons. *J. Biol. Chem.* **288,** 1856–1870 (2013).
117. Aguzzi, A. & Rajendran, L. The transcellular spread of cytosolic amyloids, prions, and prionoids. *Neuron* **64,** 1856–790 (2009).
118. Nath, S. *et al.* Spreading of neurodegenerative pathology via neuron-to-neuron transmission of β-amyloid. *J. Neurosci.* **32,** 8767–8777 (2012).
119. Kfoury, N., Holmes, B. B., Jiang, H., Holtzman, D. M. & Diamond, M. I. Trans-cellular propagation of Tau aggregation by fibrillar species. *J. Biol. Chem.* **287,** 19440–19451 (2012).
120. Freundt, E. C. *et al.* Neuron-to-neuron transmission of α-synuclein fibrils through axonal transport. *Ann. Neurol.* **72,** 517–524 (2012).
121. Ren, P. H. *et al.* Cytoplasmic penetration and persistent infection of mammalian cells by polyglutamine aggregates. *Nature Cell Biol.* **11,** 219–225 (2009).
122. Selkoe, D. J. Resolving controversies on the path to Alzheimer's therapeutics. *Nature Med.* **17,** 1060–1065 (2011).
123. Gajdusek, D. C. Spontaneous generation of infectious nucleating amyloids in the transmissible and nontransmissible cerebral amyloidoses. *Mol. Neurobiol.* **8,** 1–13 (1994).
124. Lee, J., Culyba, E. K., Powers, E. T. & Kelly, J. W. Amyloid-β forms fibrils by nucleated conformational conversion of oligomers. *Nature Chem. Biol.* **7,** 602–609 (2011).
125. Knowles, T. P. & Buehler, M. J. Nanomechanics of functional and pathological amyloid materials. *Nature Nanotechnol.* **6,** 469–479 (2011).

Acknowledgements We thank D. Eisenberg, H. LeVine, A. Aguzzi, J. Collinge, R. Rosen, Y. Eisele, A. Mehta, M. Gearing, J. Manson, M. Neumann, and the members of our laboratories for critical discussions and comments. The help of H. Braak with Fig. 1, and the help of S. Eberle with the manuscript and figures is gratefully acknowledged. This work was supported by grants from the Competence Network on Degenerative Dementias (BMBF-01GI0705), ALZKULT (BMBF-031A198A), NGFN2 (BMBF-01GS08131), and anonymous foundations (to M.J.), and by National Institutes of Health grants R21AG040589, P51RR165, P51OD11132, and the CART Foundation (to L.C.W.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contributions M.J. and L.C.W. contributed to the writing of the review

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 should be addressed to M.J. (mathias.jucker@uni-tuebingen.de) or L.C.W. (lary.walker@emory.edu).

ARTICLE

doi:10.1038/nature12479

Bacteria activate sensory neurons that modulate pain and inflammation

Isaac M. Chiu[1], Balthasar A. Heesters[2,3], Nader Ghasemlou[1], Christian A. Von Hehn[1], Fan Zhao[4], Johnathan Tran[1], Brian Wainger[1], Amanda Strominger[1], Sriya Muralidharan[1], Alexander R. Horswill[5], Juliane Bubeck Wardenburg[6], Sun Wook Hwang[1,7], Michael C. Carroll[2] & Clifford J. Woolf[1]

Nociceptor sensory neurons are specialized to detect potentially damaging stimuli, protecting the organism by initiating the sensation of pain and eliciting defensive behaviours. Bacterial infections produce pain by unknown molecular mechanisms, although they are presumed to be secondary to immune activation. Here we demonstrate that bacteria directly activate nociceptors, and that the immune response mediated through TLR2, MyD88, T cells, B cells, and neutrophils and monocytes is not necessary for *Staphylococcus aureus*-induced pain in mice. Mechanical and thermal hyperalgesia in mice is correlated with live bacterial load rather than tissue swelling or immune activation. Bacteria induce calcium flux and action potentials in nociceptor neurons, in part via bacterial *N*-formylated peptides and the pore-forming toxin α-haemolysin, through distinct mechanisms. Specific ablation of Nav1.8-lineage neurons, which include nociceptors, abrogated pain during bacterial infection, but concurrently increased local immune infiltration and lymphadenopathy of the draining lymph node. Thus, bacterial pathogens produce pain by directly activating sensory neurons that modulate inflammation, an unsuspected role for the nervous system in host–pathogen interactions.

A dense network of low- and high-threshold sensory nerves innervate peripheral tissues including the skin, respiratory tract and gastrointestinal tract, which are often exposed to bacterial pathogens. Bacterial infection induces inflammation through immune cell recruitment[1]. Inflammatory pain during infection has been thought to be triggered by the action of immune-derived proteins (for example, cytokines and growth factors), lipids (for example, prostaglandins) and other mediators such as amines, potassium and protons on receptors expressed by nociceptors[2,3].

Staphylococcus aureus is a major cause of wound and surgical infections, leading to painful abscesses, cellulitis and necrotizing fascitis[4,5]. *S. aureus* releases toxins including haemolysins, Panton–Valentine leukocidin and phenol soluble modulins, which have roles in bacterial dissemination and tissue damage[6–8]. We have now investigated the molecular mechanisms of pain generation during *S. aureus* infection. Unexpectedly, key immune activation pathways were not necessary for hyperalgesia during acute infection. Rather, bacteria directly activated nociceptors through *N*-formyl peptides and the pore-forming toxin α-haemolysin (αHL). Moreover, we find that nociceptors release neuropeptides that modulate innate immune activation during infection.

Pain correlates with bacterial load

To study the nature of pain caused by bacterial pathogens, we established a *S. aureus* infection model of the mouse hindpaw. Subcutaneous injection of LAC/USA300 (5×10^6 colony-forming units (c.f.u.)), a community-associated methicillin-resistant *S. aureus* strain (CA-MRSA)[8,9], resulted in mechanical, heat and cold hypersensitivity within 1 h that lasted for 48–72 h (Fig. 1a). This hyperalgesia peaked at 6 h after infection, and began to decrease at 24 h.

We quantified the kinetics of tissue swelling, immune activation and bacterial clearance. Tissue swelling did not correlate with pain, but showed an immediate peak after bacterial injection, and a second peak at 48 h after infection (Fig. 1a). Using flow cytometry, we found increases in myeloid immune cells in infected tissues ($CD11b^+CD45^+$; Fig. 1b), constituted primarily of $Ly6G^+$ neutrophils and $Ly6C^{lo}$ monocytes, with fewer $Ly6C^{hi}$ monocytes (Supplementary Fig. 1 and Fig. 1b). This influx began at 6 h, with a peak at 48 h after infection and a reduction at 72 h (Fig. 1b). Levels of CXCL1 (also called KC) and MCP-1, chemokines that mediate neutrophil/monocyte recruitment, peaked early (Supplementary Fig. 2), whereas levels of TNF-α and IL-1β, pro-inflammatory cytokines which directly sensitize nociceptors[10,11], also increased in infected tissue but did not correlate with hyperalgesia (Fig. 1c).

We next analysed live bacterial load in infected tissues. *S. aureus* recovery peaked at 6 h, and then decreased over time, similar to the time course of pain hypersensitivity (Fig. 1d). This decrease was accompanied by myeloid cell ingestion of bacteria, as detected by flow cytometry (Supplementary Fig. 1). In infected Nav1.8-Cre/TdTomato reporter mice, *S. aureus* (GFP–USA300) were often found in close proximity to dermal but not epidermal nociceptor fibres (Fig. 1e and Supplementary Fig. 3), indicating potential direct nerve–bacteria interactions. Taken together, hyperalgesia mirrors the time course of bacterial expansion and not tissue swelling or immune activation (see diagram in Supplementary Fig. 4).

Host defences not necessary for *S. aureus*-induced pain

We determined whether key immune response pathways were necessary for *S. aureus*-induced pain. Innate immune cells recognize *S. aureus*-derived cell-wall components through TLR2 (ref. 12), leading to nuclear factor-κB (NF-κB) activation through adaptor myeloid differentiation factor 88 (MyD88). TLR2 and MyD88 protect mice against *S. aureus* skin infection[4,13]. TLR3, TLR4, TLR7 and TLR9 are also functionally expressed in sensory neurons[14,15]. Thus, we analysed pain responses in $Tlr2^{-/-}$ and $Myd88^{-/-}$ mice after *S. aureus* infection. Notably, mechanical and thermal hyperalgesia were not reduced in these mice (Fig. 2a and Supplementary Fig. 5). At 72 h, $Myd88^{-/-}$

[1]Kirby Neurobiology Center, Boston Children's Hospital, and Department of Neurobiology, Harvard Medical School, Boston, Massachusetts 02115, USA. [2]Boston Children's Hospital, Program in Cellular and Molecular Medicine, and Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Medical Microbiology, University Medical Center, Utrecht, Utrecht 3584 CX, The Netherlands. [4]Department of Chemistry, Quantitative Biology Program, Brandeis University, Waltham, Massachusetts 02454, USA. [5]Department of Microbiology, Roy J. and Lucille A. Carver College of Medicine, University of Iowa, Iowa City, Iowa 52242, USA. [6]Departments of Pediatrics and of Microbiology, University of Chicago, Chicago, Illinois 60637, USA. [7]Korea University Graduate School of Medicine, Seoul 136-705, South Korea.

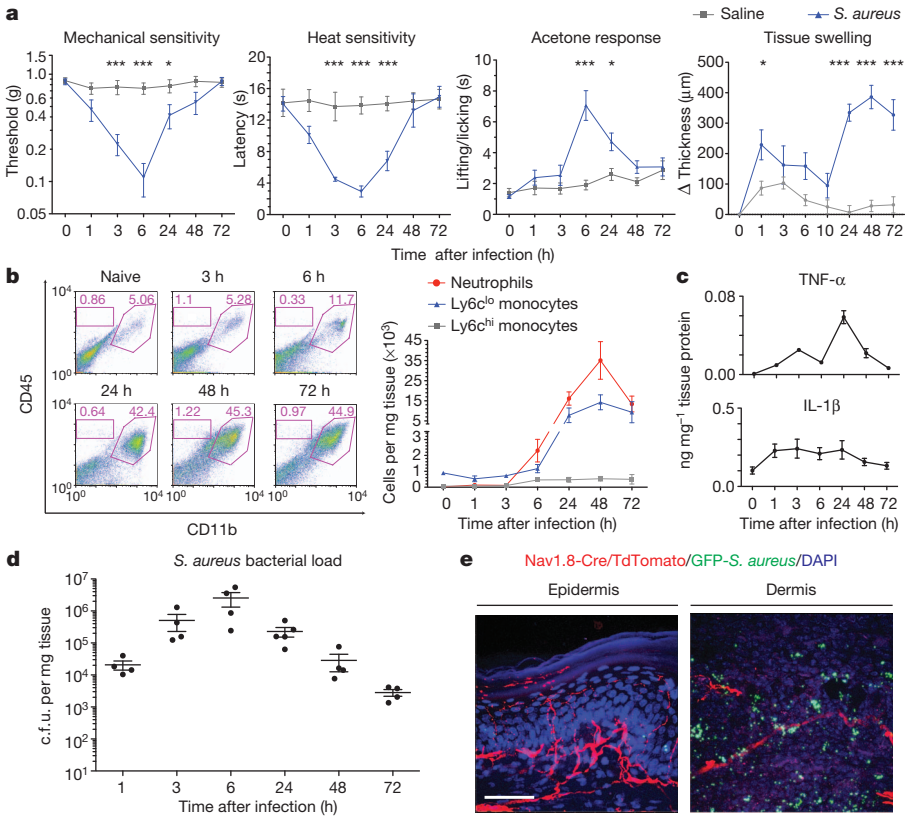


Figure 1 | *Staphylococcus aureus* infection induces pain hypersensitivity paralleling bacterial load but not immune activation. **a**, *S. aureus* infection induces mechanical hypersensitivity (Von Frey filaments, $P = 0.0021$, $n = 10$ per group), heat hypersensitivity (Hargreave's radiant heat test, $P < 0.0001$, $n = 10$ per group), acetone cold response (lifting/licking to acetone application, $P < 0.0001$, $n = 20$ per group) and tissue swelling ($P < 0.0001$, $n = 10$ per group). $^*P < 0.05$, $^{***}P < 0.001$. **b**, Left: flow cytometry shows myeloid ($CD11b^+CD45^+$) but not lymphoid ($CD11b^-CD45^+$) immune expansion in infected tissues. Right: quantification of infected tissue neutrophils ($CD11b^+Ly6G^+$), $Ly6c^{hi}$ monocytes ($CD11b^+Ly6G^-Ly6C^{hi}$), and $Ly6C^{lo}$ monocytes/macrophages ($CD11b^+Ly6G^-Ly6C^{lo}$). $n = 3$ per time point. **c**, TNF-α and IL-1β levels in infected tissues. $n = 4$ per time point. **d**, Bacterial load recovery. $n = 4$ per time point. **e**, GFP–*S. aureus* are in proximity with Nav1.8-Cre/$TdTomato^+$ dermal nerve fibres, 3 h after infection. Scale bar, 100 μm. Two-way analysis of variance (ANOVA) with Bonferroni post-tests for behavioural data. Error bars indicate mean ± s.e.m.

mice showed elevated pain-like hypersensitivity, which may reflect greater bacterial load due to reduced immune activation and bacterial clearance (Fig. 2a, b).

Neutrophils and monocytes are circulating leukocytes critical for innate immunity against *S. aureus*, rapidly infiltrating sites of infection to limit pathogen survival and spread[4,16]. We treated mice with GR1 antibody before infection, which eliminated blood-borne and splenic neutrophils and monocytes (Supplementary Fig. 6). Plantar tissue-infiltrating neutrophils and monocytes during *S. aureus* infection were completely depleted by GR1 (Supplementary Fig. 7; reduction of 97% of all $CD45^+$ immune cells in infected tissues). However, instead of decreasing hyperalgesia, GR1 depletion significantly increased mechanical and heat hypersensitivity (Fig. 2c). This was accompanied by higher bacterial load, reflecting the key role of these myelomonocytic cells in combating *S. aureus* (Fig. 2d). We repeated the experiment using injection of heat-killed *S. aureus* (10^8 c.f.u.), and found that GR1 treatment decreased tissue swelling, but did not affect pain-like hypersensitivity (Supplementary Fig. 8). The increased pain during *S. aureus* infection after GR1 depletion is probably linked to uncontrolled bacterial expansion (Fig. 2d), but may also reflect neutrophil analgesic factors[17].

To analyse the contribution of the immune system to pain further, we examined *S. aureus* infection in NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (also called NOD scid gamma) mice, which are deficient in natural killer, T and B cells[18]. Infection-induced mechanical and heat hyperalgesia did not differ between NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ and wild-type NOD mice (Supplementary Fig. 9). To control for strain-dependent differences, we also analysed C57BL/6 congenic *Rag1*$^{-/-}$ mice, which are deficient in mature T and B cells[19]. Pain-like hypersensitivity did not differ after *S. aureus* infection in C57BL/6 *Rag1*$^{-/-}$ mice compared to C57BL/6 wild-type mice (Supplementary Fig. 9). On the basis of these results, we conclude that adaptive immunity through T and B cells is not required for acute bacterial pain.

Formyl peptides and α-haemolysin activate nociceptors

The strong correlation between pain and bacterial load indicated that nociceptors directly interact with bacteria during pathogen invasion. To test this, we applied heat-killed bacteria on dorsal root ganglia (DRG) sensory neurons. Heat-killed *S. aureus* induced a robust calcium flux response in a subset of neurons that also responded to capsaicin, which activates transient receptor potential V1 (TRPV1) (Fig. 3a). Heat-killed *S. aureus* application also induced action-potential firing in capsaicin-responsive DRG neurons (Fig. 3b). Extending these results, we found that several other strains of heat-killed bacteria caused calcium flux in DRG neurons (Fig. 3c, responsive cells: heat-killed *S. aureus* = 152 of 1,046; heat-killed *Streptococcus pneumoniae* = 82 of 968; heat-killed *Listeria monocytogenes* = 67 of 852; heat-killed *Mycoplasma fermentans* = 9 of 339; heat-killed *Helicobacter pylori* = 85 of 1,365; heat-killed *Pseudomonas aeruginosa* = 14 of 269; heat-killed *Escherichia coli* = 3 of 233). Nav1.8-Cre/TdTomato reporter mice were used to genetically mark nociceptors[20] and all bacteria-evoked neuronal responses were within the Nav1.8-Cre/$TdTomato^+$ population (Supplementary Fig. 10). Patterns of nociceptor responsiveness to particular bacteria differed, indicating strain-specific ligands acting through disparate mechanisms (Supplementary Fig. 11). Intraplantar injection of different heat-killed bacterial strains (10^8 c.f.u.) induced acute pain responses similar to the relative efficacies of nociceptor activation *in vitro* (Fig. 3d). Heat-killed *S. aureus* and heat-killed *S. pneumoniae*, which caused the most acute pain after injection, also induced mechanical and heat hyperalgesia (Supplementary Fig. 12).

We next investigated which heat-stable molecular elements within bacteria mediate nociceptor activation. Application of *S. aureus* peptidoglycans and lipoteichoic acid did not produce calcium flux in DRG neurons (Supplementary Fig. 13), and together with *in vivo* data (Fig. 2), this indicates that TLR2 ligands do not contribute significantly to *S. aureus* pain. We therefore investigated the contribution of other bacterial molecules. *N*-formylated peptides, found in bacteria and mitochondria, are recognized by G-protein-coupled formyl peptide receptors (FPRs) on leukocytes to mediate immune chemotaxis during infection[21]. Mouse vomeronasal neurons express FPRs, so may detect formyl peptides during olfaction[22,23]. We found that fMLF, an *E. coli*-derived peptide, and fMIFL, a *S. aureus*-derived peptide[24], induced calcium flux in a subset of DRG neurons that also responded

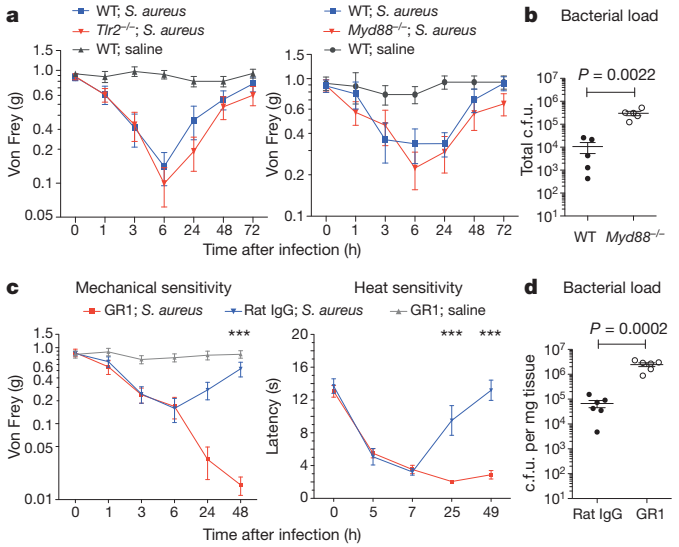


Figure 2 | Innate immunity through TLR2 and MyD88 and neutrophils and monocytes is not necessary for pain during *S. aureus* infection. **a**, Infection-induced mechanical hypersensitivity is similar in *Tlr2*$^{-/-}$ mice ($n = 10$ infected) compared to wild-type (WT) mice ($n = 10$ infected, $n = 10$ saline injected) ($P = 0.744$), and *Myd88*$^{-/-}$ mice ($n = 10$ infected) relative to wild-type mice ($n = 11$ infected, $n = 7$ saline injected) ($P = 0.533$). **b**, Bacterial load 3 days after infection ($n = 5$ each; P value, t-test). **c**, Infection-induced mechanical ($P < 0.0001$) and heat ($P < 0.0001$) hypersensitivity are increased in GR1-treated mice ($n = 10$ infected, $n = 10$ saline) compared to rat IgG-treated mice ($n = 10$ infected). ***$P < 0.001$. **d**, Bacterial load 2 days after infection ($n = 6$ each; P value, t-test). Two-way ANOVA with Bonferroni post-tests for behavioural data. Error bars indicate mean ± s.e.m.

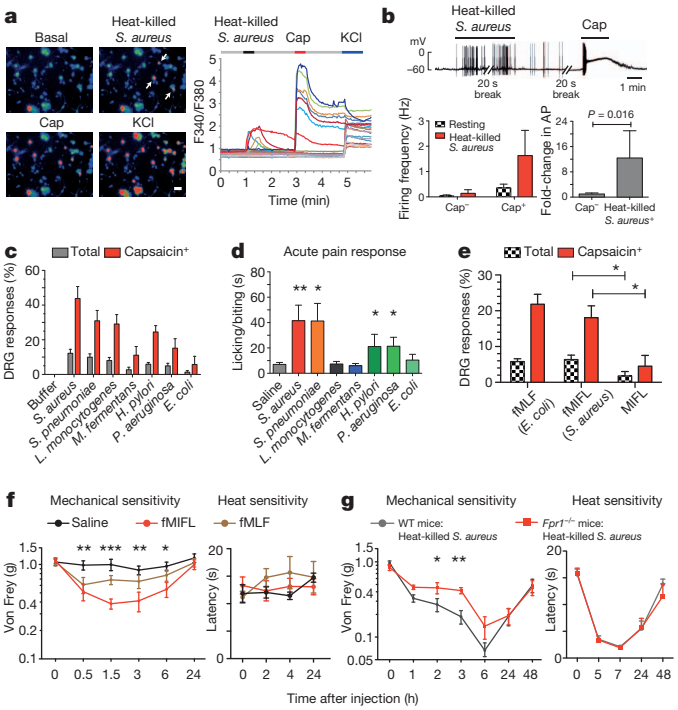


Figure 3 | Bacterial heat-stable components including *N*-formylated peptides activate nociceptors. **a**, Heat-killed *S. aureus* induces calcium flux in capsaicin (Cap) and KCl responsive DRG neurons (arrows, traces). **b**, Top row: representative recording; bottom row: firing frequency upon heat-killed *S. aureus* application (5 capsaicin-responsive cells, 9 unresponsive). AP, action potentials. **c**, DRG-responsive proportions to heat-killed bacteria ($n = 4$–26 fields per condition). **d**, Acute pain induction. Saline ($n = 13$), heat-killed *S. aureus* ($n = 12$), heat-killed *S. pneumoniae* ($n = 14$), heat-killed *L. monocytogenes* ($n = 5$), heat-killed *M. fermentans* ($n = 6$), heat-killed *H. pylori* ($n = 5$), heat-killed *P. aeruginosa* ($n = 8$), heat-killed *E. coli* ($n = 6$). **$P < 0.01$, *$P < 0.05$, t-test versus saline. **e**, DRG-responsive proportions to formyl peptides ($n = 3$–14 fields per condition; *$P < 0.05$, t-test). **f**, **g**, fMLF and fMIFL injection induces mechanical hypersensitivity. *Fpr1*$^{-/-}$ mice show reduced heat-killed *S. aureus* mechanical hypersensitivity ($P = 0.0089$). fMIFL versus saline and *Fpr1*$^{-/-}$ versus wild type: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$. Two-way ANOVA with Bonferroni post-tests for behavioural data. Error bars indicate mean ± s.e.m.

to capsaicin and to allyl isothiocyanate (AITC), a TRPA1 ligand (Fig. 3e and Supplementary Fig. 14). Moreover, unformylated MIFL activated fewer DRG neurons than fMIFL (Fig. 3e), similar to formyl group requirements for immune cell activation. When injected, fMLF and fMIFL induced mechanical but not heat hyperalgesia (Fig. 3f). This selective mechanical pain induction may be related to the restricted activation by formyl peptides of AITC-responsive neurons (Supplementary Fig. 14), which are mainly C-mechanosensitive fibres[25]. DRG and trigeminal ganglia express *Fpr1* and *Fpr2* (also known as *Fpr-rs2*), but not other *Fpr* genes (Supplementary Fig. 15). Microarray analysis of purified Nav1.8-Cre/TdTomato^{+} nociceptors confirmed the specific expression of *Fpr1* but not *Fpr-rs1*, *Fpr-rs3*, *Fpr-rs4*, or *Fpr-rs7* (Supplementary Fig. 16). These data indicate that nociceptor activation may be mediated via FPR1, the receptor that recognizes fMLF and fMIFL in immune cells[24]. The FPR1 antagonist Boc-MLF reduced neuron activation by fMLF and heat-killed bacteria (Supplementary Fig. 17). *Fpr1*$^{-/-}$ mouse DRG neurons showed decreased fMIFL calcium flux, and *Fpr1*$^{-/-}$ mice showed reduced mechanical hyperalgesia after fMIFL injection relative to wild-type mice (Supplementary Fig. 17). Whereas live *S. aureus* pain peaks at 6 h, time courses may differ for heat-killed bacteria and ligands, which are influenced by ligand diffusion and clearance kinetics. Ruthenium red significantly decreased fMIFL responses in amplitude and proportion, indicating that formyl peptides may gate a downstream large-pore calcium ion channel (Supplementary Fig. 18). Heat treatment did not affect fMIFL molecular composition or the ability of fMIFL to induce hyperalgesia after injection (Supplementary Fig. 19). Notably, *Fpr1*$^{-/-}$ mice showed a reduction in mechanical but not heat hypersensitivity after injection of heat-killed *S. aureus* (Fig. 3g). Thus, formyl peptides are heat-stable elements within bacteria that contribute to the activation of mechanical hypersensitivity.

Live bacteria actively release formyl peptides[26] and secrete a host of virulence factors, including pore-forming toxins (PFTs), to facilitate tissue dissemination[6–8,27–29]. We found that *S. aureus* culture supernatant induced calcium flux in DRG neurons (Supplementary Fig. 20). αHL is a PFT secreted by nearly all *S. aureus* strains, and is involved in tissue damage, bacterial spread and inflammation[27–29]. When flowed onto DRG neurons, αHL induced immediate calcium flux in nociceptors, which could be washed out by buffer (Supplementary Fig. 20). αHL inserts into cell membranes and assembles into heptameric pores that allow non-selective entry of cations[29], which may be sufficient to depolarize neurons. We found that prolonged bath application of αHL on DRG neurons elicited sustained bursts of calcium flux (Fig. 4a), selectively in capsaicin-responsive neurons (Supplementary Fig. 21). αHL induced a dose-dependent calcium flux in DRG neurons (half-maximum effective concentration (EC_{50}) 356 nM, Fig. 4b). αHL binds to cells via A disintegrin and metalloprotease 10 (ADAM10), leading to membrane pore assembly[27,30]. ADAM10 expression was detected by PCR with reverse transcription (RT–PCR) in DRG and trigeminal ganglia (Supplementary Fig. 15). Using an antibody that recognizes the ADAM10 ectodomain, we found that a subset of Nav1.8-Cre/TdTomato^{+} nociceptors (59.8%), but not TdTomato^{-} cells, expressed surface ADAM10, which may contribute to selective nociceptor activation (Supplementary Fig. 22). *In vivo*, αHL injection induced significant acute pain behaviour in a dose-dependent manner ($EC_{50} = 6.3$ pmol, Fig. 4c). Heat pre-treatment abolished the ability of αHL to induce pain (Fig. 4d), indicating that this mechanism of nociceptor activation is separate from heat-stable elements (Fig. 3). αHL also evoked action

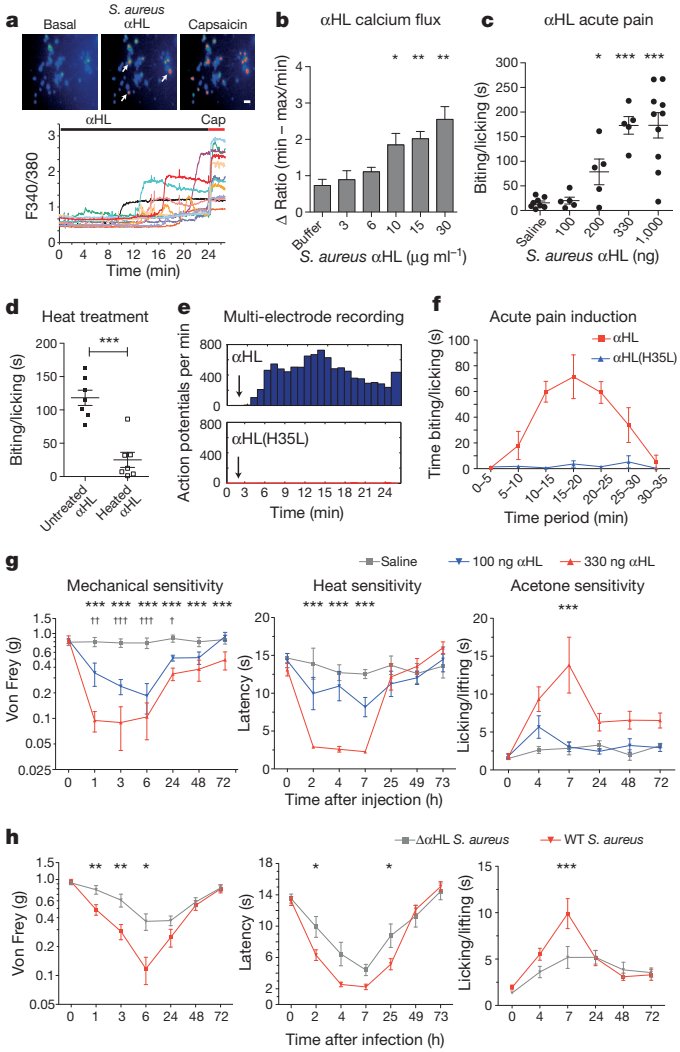


Figure 4 | Heat-sensitive *S. aureus* αHL activates nociceptors and contributes to infection-induced hyperalgesia. **a**–**c**, αHL application evoked DRG neuron calcium flux (**a**; arrows, traces), dose-dependent calcium flux (**b**; $n = 3$ per condition) and acute pain (**c**; $n = 5$–10 per group). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, *t*-test. **d**, Heat pre-treatment abolishes αHL-induced pain (1 μg, $n = 7$ per group, $^{***}P < 0.001$, *t*-test). **e**, αHL- and αHL(H35L)-evoked DRG neuron action potentials (arrow indicates αHL application, $n = 3$ per condition). **f**, αHL (1 μg, $n = 6$) but not αHL(H35L) (1 μg, $n = 5$) induced acute pain. **g**, Mechanical, heat and acetone hypersensitivity induced by injection of 100 ng αHL ($n = 8$), 330 ng αHL ($n = 8$), or saline ($n = 8$) is shown. 100 ng αHL versus saline: $^{***}P < 0.001$; 330 ng αHL versus saline: $^{\dagger}P < 0.05$, $^{\dagger\dagger}P < 0.01$, $^{\dagger\dagger\dagger}P < 0.001$. **h**, *S. aureus* lacking αHL ($n = 12$) produced less mechanical ($P = 0.0056$), heat ($P = 0.0193$) and acetone ($P = 0.0118$) hypersensitivity than wild-type *S. aureus* ($n = 13$). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Two-way ANOVA with Bonferroni post-tests for behavioural data. Error bars indicate mean ± s.e.m.

potential firing in DRG neurons (Fig. 4e). αHL neuronal activation did not involve voltage-gated calcium channels or large-pore cation channels (not blocked by $CdCl_2$ or ruthenium red, and not absent in *Trpv1*$^{-/-}$ neurons) but did require external calcium (Supplementary Fig. 23). Therefore, αHL nociceptor activation seems to be independent of most endogenous calcium ion channels, and its ability to assemble into membrane-perforating pores may be sufficient for neuronal depolarization. H35L mutant αHL (αHL(H35L)), which cannot form a stable oligomer or pore, did not evoke action potentials (Fig. 4e) or calcium flux in DRG neurons (Supplementary Fig. 23). αHL(H35L), unlike wild-type αHL, did not produce acute pain in mice (Fig. 4f). αHL was sufficient to induce mechanical, heat and cold hypersensitivity (Fig. 4g). Finally, we found that an isogenic *S. aureus* mutant devoid of αHL expression caused significantly less hyperalgesia than wild-type bacteria (Fig. 4h). Therefore, we conclude that αHL contributes to pain during *S. aureus* infection in a manner dependent on pore formation.

Nociceptor neuropeptides regulate inflammation

Nociceptor activation results in release of neuropeptides from peripheral terminals which can induce vasodilation and neurogenic inflammation[31]. To understand the role of nociceptors in modulating the immune response, we generated Nav1.8-Cre/diphtheria toxin A (DTA) mice to specifically ablate these cells[32]. Nav1.8-Cre/DTA DRG neurons did not show calcium flux upon stimulation with heat-killed bacteria (Fig. 5a). Mechanical and thermal hypersensitivity after *S. aureus* infection was abolished in Nav1.8-Cre/DTA mice, indicating that Nav1.8-lineage neurons are the major cell type mediating bacterial pain (Fig. 5b). Granulocytes were found 24 h after infection near Nav1.8-Cre/TdTomato^{+} nerve fibres (Supplementary Fig. 24), and when co-cultured, DRG neurons also formed close contacts with neutrophils and macrophages (Supplementary Fig. 25), raising the question of whether nociceptors may act on immune cells. After infection, Nav1.8-Cre/DTA mice displayed significantly increased tissue swelling relative to control littermates (Fig. 5c). Bacterial load did not differ significantly between the mice (Supplementary Fig. 26). Nav1.8-Cre/DTA mice showed increased infiltration of neutrophils/monocytes at infection sites (Fig. 5d). The popliteal lymph node which drains the footpad[33] was significantly larger, by weight and cellularity, in Nav1.8-Cre/DTA mice relative to control littermates following *S. aureus* infection (Fig. 5e and Supplementary Fig. 27). There was a difference in baseline lymph node size, but the increase was substantially greater in Nav1.8-Cre/DTA than control mice after infection. Nav1.8-Cre/DTA tissues also showed increased levels of

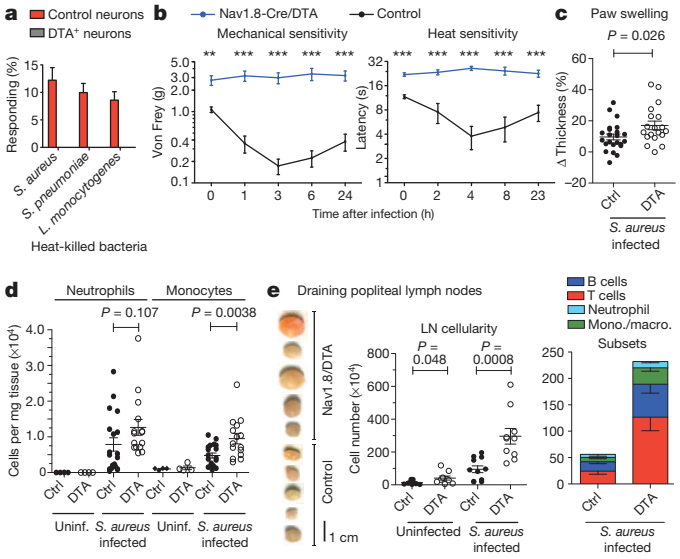


Figure 5 | Nociceptor ablation leads to increased local inflammation and lymphadenopathy after *S. aureus* infection. **a**, Nav1.8-Cre/DTA neurons lack heat-killed bacteria responses. **b**, Infection-induced mechanical ($P = 0.0027$) and heat hypersensitivity ($P = 0.0003$) in Nav1.8-Cre/DTA mice ($n = 10$ mechanical, $n = 6$ heat) and control littermates ($n = 12$ mechanical, $n = 6$ heat). $^{**}P < 0.01$, $^{***}P < 0.001$. **c**–**e**, Parameters analysed 24 h after infection. **c**, Tissue swelling in Nav1.8-Cre/DTA ($n = 23$) and control ($n = 19$) mice. *P* value, *t*-test. **d**, Plantar neutrophils and monocytes: Nav1.8-Cre/DTA mice, $n = 4$ uninfected, $n = 15$ infected; control mice, $n = 4$ uninfected, $n = 19$ infected. *P* values, *t*-test. **e**, Popliteal lymph node images (infected), lymph node cellularity (Nav1.8-Cre/DTA: $n = 9$ uninfected, $n = 10$ infected; control: $n = 9$ uninfected, $n = 11$ infected; *P* values, *t*-test) and lymph node monocyte/macrophage, neutrophil, T-cell and B-cell subsets ($n = 5$ each). Two-way ANOVA with Bonferroni post-tests for behavioural data. Error bars indicate mean ± s.e.m.

TNF-α (Supplementary Fig. 28), a cytokine that drives lymphadenopathy (hypertrophy) of the draining lymph node during bacterial infection[34]. This lymphadenopathy was localized, as downstream inguinal lymph nodes and spleens were not enlarged in infected Nav1.8-Cre/DTA mice (Supplementary Fig. 27). The popliteal lymphadenopathy was due to increased T cells, B cells and monocytes (Fig. 5e). Thus, nociceptor ablation led to increased local inflammation. To determine the molecular factors that mediate this immunomodulation, we performed microarray analysis of flow-cytometry-purified Nav1.8-Cre/TdTomato^{+} nociceptors from dorsal root, trigeminal and nodose ganglia (Fig. 6a). Neuropeptide expression levels were ranked based on a neuropeptide database (http://www.neuropeptides.nl; Fig. 6b, full data set in Supplementary Fig. 29). Microarray data from innate immune cell subsets[35] were analysed for neuropeptide receptor levels. CGRP, galanin and somatostatin receptors showed the highest expression in neutrophils, monocytes and macrophages; these neuropeptides were also highly expressed in purified nociceptors (Fig. 6b, full data set in Supplementary Fig. 30). *In vitro*, we found that CGRP, galanin and somatostatin all suppressed TNF-α release from macrophages stimulated with heat-killed *S. aureus* or lipoteichoic acid (Fig. 6c, complete analysis in Supplementary Fig. 31). Furthermore, *S. aureus* supernatant and αHL induced CGRP release from DRG neurons in a dose-dependent manner (Fig. 6d). CGRP injection during *S. aureus* infection did not alter inflammation at the infection site, but significantly suppressed lymphadenopathy of the draining lymph nodes (Fig. 6e and Supplementary Fig. 32). Therefore, upon infection, nociceptors may release neuropeptides that directly modulate innate immune activation.

Discussion

We analysed mechanisms responsible for nociceptor activation during *S. aureus* infection, which commonly causes pain. We found that bacterial-derived factors directly activate nociceptors and contribute to hyperalgesia *in vivo*. *N*-formylated peptides and the PFT αHL induced direct neuronal responses through distinct mechanisms: formyl peptides through FPR1 and αHL through pore assembly leading to ionic influx (diagram, Supplementary Fig. 33).

Direct activation of nociceptors by bacteria is probably a major mechanism leading to pain, especially early in *S. aureus* infection during active pathogen expansion (Figs 1, 3 and 4). After immune cell infiltration, bacteria are largely eliminated, concurrent with TNF-α production, and although pain is reduced it remains (Fig. 1). Immune-mediated mechanisms may have a role during these later time points. Live infection is complex and our blockade of host defences may also enhance direct pathogen-mediated mechanisms.

Pathogen virulence and immunogenicity probably contribute to the degree of direct nociceptor activation. USA300/LAC is a highly virulent strain, expressing several PFTs in addition to αHL including β-haemolysin, γ-haemolysin, Panton–Valentine leukocidin and phenol soluble modulins[7–9]. *S. aureus* also possesses effective immune evasion mechanisms, including toxins that lyse immune cells, staphylococcal protein A which impairs antibody function, and complement evasion strategies[8,36]. Thus, additional elements from the bacterial proteome may exist that activate nociceptors. *E. coli* and other Gram-negative bacteria also induce painful infections, and lipopolysaccharides have been found to sensitize TRPV1 (ref. 15). The balance of nociceptor activation by pathogenic and immune mechanisms may differ between dissimilar pathogenic bacteria species.

Although peripheral nociceptor activation contributes to neurogenic inflammation, inducing vasodilation and capillary permeability[31], we found that ablation led to increased immune influx and lymphadenopathy, implying pain-mediated immune suppression. Our data support a role for neuropeptides in regulating innate immune activation that occurs later than the acute vascular phase of inflammation. Receptors for CGRP, galanin and somatostatin are expressed by myeloid immune subsets, and these neuropeptides have been shown to have inhibitory functions on immunity[37–39]. In particular, CGRP dampens TNF-α transcription in dendritic cells through a cAMP-dependent repressor mechanism[37,40]. CGRP downregulates cytokine levels in endotoxic shock[41], and conversely, *Trpv1*$^{-/-}$ mice display increased inflammation during sepsis[42].

Potent immunomodulatory neural reflex circuits also exist that maintain immune homeostasis[43]. In *Caenorhabditis elegans*, a sensory neural circuit suppresses innate immunity and modulates survival during bacterial infection[44]. In mammals, vagal efferents suppress splenic macrophage activity, protecting against bacterial endotoxic shock[43,45], and activation of liver autonomic fibres modulates NK T-cell activity, leading to increased bacterial infection[34]. It has been proposed that sensory neurons initiate mammalian neural circuits[43], but a direct immunosuppressive role of nociceptive fibres, as revealed here, was not suspected. Lymph-node swelling during infection is often accompanied by pain, and nociceptor activation may limit immune influx into lymph nodes. Both nociceptive and autonomic fibres innervate lymph nodes[46]; therefore, lymph-node suppression may act through local neuropeptide release or initiation of autonomic reflex circuits. Highly pathogenic bacterial strains may have evolved the ability to exploit these neural-mediated immune regulation pathways for virulence and spread within infected tissues, by producing more nociceptor activation and greater immunosuppression.

Our data reveal an unsuspected mechanism for pain induction during bacterial infection: a direct pathogen-mediated activation of nociceptors. This neuron–pathogen interaction leads to a downregulation of the local inflammatory response. The nervous system therefore has direct sensory and modulatory roles in host–pathogen interactions during acute staphylococcal infection.

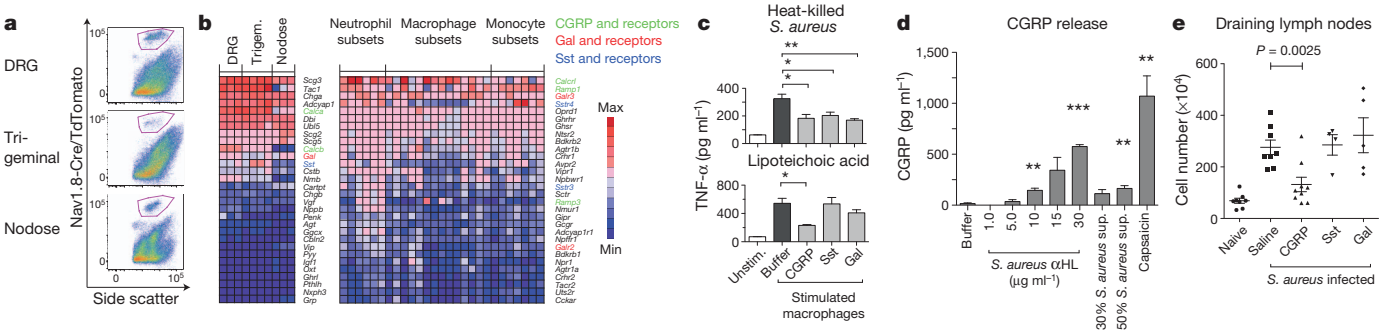


Figure 6 | Nociceptor-derived neuropeptides regulate innate immune activation. **a**, Nav1.8-Cre/TdTomato^{+} DRG, trigeminal and nodose ganglia neurons were purified by flow cytometry (gates shown). **b**, Top 30 nociceptor-expressed neuropeptides and myeloid immune-cell-expressed neuropeptide receptors, shown from maximum to minimum. Gal, galanin; Sst, somatostatin. **c**, TNF-α production by heat-killed *S. aureus* or lipoteichoic-acid-stimulated macrophages was suppressed by CGRP, somatostatin and galanin (neuropeptide concentrations, 1 μM; $*P < 0.05$, *t*-test). **d**, αHL, *S. aureus* supernatant and capsaicin (100 nM) induce DRG neuron CGRP release. $**P < 0.01$, $***P < 0.001$, *t*-test versus buffer. **e**, CGRP injection decreased lymphadenopathy 24 h after *S. aureus* infection. Error bars indicate mean ± s.e.m.

METHODS SUMMARY

All experiments were conducted according to institutional animal care guidelines. The following mouse strains were used: C57BL/6, C57BL/6 *Tlr2*$^{-/-}$, C57BL/6 *Myd88*$^{-/-}$, C57BL/6 *Rag1*$^{-/-}$, NOD wild type, NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ (Jackson Laboratories); C57BL/6 wild type, C57BL/6 *Fpr1*$^{-/-}$ (Taconic); Nav1.8-Cre/TdTomato, Nav1.8-Cre/DTA mice (Nav1.8-Cre (ref. 20) from R. Kuner, Heidelberg University). Bacterial strains: *S. aureus* LAC[9] and isogenic LAC deficient in αHL, described previously[6]. Recombinant αHL and αHL(H35L) are described previously[27,30]. For infection, 5×10^6 c.f.u. of *S. aureus* were injected into the plantar surface. For GFP–LAC/USA300 generation, statistical analysis, bacterial infection, formyl peptide synthesis, behaviour, microscopy, microarrays, neuronal cultures, electrophysiology, calcium imaging, flow cytometry, see Methods.

Full Methods and any associated references are available in the online version of the paper.

Received 7 November 2012; accepted 17 July 2013.
Published online 21 August 2013.

1. Medzhitov, R. Origin and physiological roles of inflammation. *Nature* **454,** 428–435 (2008).
2. White, R. J. Wound infection-associated pain. *J. Wound Care* **18,** 245–249 (2009).
3. Ren, K. & Dubner, R. Interactions between the immune and nervous systems in pain. *Nature Med.* **16,** 1267–1276 (2010).
4. Miller, L. S. & Cho, J. S. Immunity against *Staphylococcus aureus* cutaneous infections. *Nature Rev. Immunol.* **11,** 505–518 (2011).
5. Morgan, M. Treatment of MRSA soft tissue infections: an overview. *Injury* **42** (Suppl. 5), S11–S17 (2011).
6. Bubeck Wardenburg, J., Patel, R. J. & Schneewind, O. Surface proteins and exotoxins are required for the pathogenesis of *Staphylococcus aureus* pneumonia. *Infect. Immun.* **75,** 1040–1044 (2007).
7. Wang, R. *et al.* Identification of novel cytolytic peptides as key virulence determinants for community-associated MRSA. *Nature Med.* **13,** 1510–1514 (2007).
8. Gordon, R. J. & Lowy, F. D. Pathogenesis of methicillin-resistant *Staphylococcus aureus* infection. *Clin. Infect. Dis.* **46** (Suppl. 5), S350–S359 (2008).
9. Diep, B. A. *et al.* Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant *Staphylococcus aureus*. *Lancet* **367,** 731–739 (2006).
10. Binshtok, A. M. *et al.* Nociceptors are interleukin-1β sensors. *J. Neurosci.* **28,** 14062–14073 (2008).
11. Zhang, X. C., Kainz, V., Burstein, R. & Levy, D. Tumor necrosis factor-α induces sensitization of meningeal nociceptors mediated via local COX and p38 MAP kinase actions. *Pain* **152,** 140–149 (2011).
12. Müller-Anstett, M. A. *et al.* Staphylococcal peptidoglycan co-localizes with Nod2 and TLR2 and activates innate immune response via both receptors in primary murine keratinocytes. *PLoS ONE* **5,** e13153 (2010).
13. Miller, L. S. *et al.* MyD88 mediates neutrophil recruitment initiated by IL-1R but not TLR2 activation in immunity against *Staphylococcus aureus*. *Immunity* **24,** 79–91 (2006).
14. Liu, T. *et al.* Emerging roles of toll-like receptors in the control of pain and itch. *Neurosci. Bull.* **28,** 131–144 (2012).
15. Diogenes, A. *et al.* LPS sensitizes TRPV1 via activation of TLR4 in trigeminal sensory neurons. *J. Dent. Res.* **90,** 759–764 (2011).
16. Rigby, K. M. & DeLeo, F. R. Neutrophils in innate host defense against *Staphylococcus aureus* infections. *Semin. Immunopathol.* **34,** 237–259 (2011).
17. Rittner, H. L. *et al.* Mycobacteria attenuate nociceptive responses by formyl peptide receptor triggered opioid peptide release from neutrophils. *PLoS Pathog.* **5,** e1000362 (2009).
18. Shultz, L. D. *et al.* Human lymphoid and myeloid cell development in NOD/LtSz-scid IL2Rγ null mice engrafted with mobilized human hemopoietic stem cells. *J. Immunol.* **174,** 6477–6489 (2005).
19. Mombaerts, P. *et al.* RAG-1-deficient mice have no mature B and T lymphocytes. *Cell* **68,** 869–877 (1992).
20. Agarwal, N., Offermanns, S. & Kuner, R. Conditional gene deletion in primary nociceptive neurons of trigeminal ganglia and dorsal root ganglia. *Genesis* **38,** 122–129 (2004).
21. Le, Y., Murphy, P. M. & Wang, J. M. Formyl-peptide receptors revisited. *Trends Immunol.* **23,** 541–548 (2002).
22. Liberles, S. D. *et al.* Formyl peptide receptors are candidate chemosensory receptors in the vomeronasal organ. *Proc. Natl Acad. Sci. USA* **106,** 9842–9847 (2009).
23. Rivière, S., Challet, L., Fluegge, D., Spehr, M. & Rodriguez, I. Formyl peptide receptor-like proteins are a novel family of vomeronasal chemosensors. *Nature* **459,** 574–577 (2009).
24. Southgate, E. L. *et al.* Identification of formyl peptides from Listeria monocytogenes and *Staphylococcus aureus* as potent chemoattractants for mouse neutrophils. *J. Immunol.* **181,** 1429–1437 (2008).
25. Lennertz, R. C., Kossyreva, E. A., Smith, A. K. & Stucky, C. L. TRPA1 mediates mechanical sensitization in nociceptors during inflammation. *PLoS ONE* **7,** e43597 (2012).
26. Durr, M. C. *et al.* Neutrophil chemotaxis by pathogen-associated molecular patterns–formylated peptides are crucial but not the sole neutrophil attractants produced by *Staphylococcus aureus*. *Cell. Microbiol.* **8,** 207–217 (2006).
27. Inoshima, I. *et al.* A *Staphylococcus aureus* pore-forming toxin subverts the activity of ADAM10 to cause lethal infection in mice. *Nature Med.* **17,** 1310–1314 (2012).
28. Kennedy, A. D. *et al.* Targeting of alpha-hemolysin by active or passive immunization decreases severity of USA300 skin infection in a mouse model. *J. Infect. Dis.* **202,** 1050–1058 (2010).
29. Dinges, M. M., Orwin, P. M. & Schlievert, P. M. Exotoxins of *Staphylococcus aureus*. *Clin. Microbiol. Rev.* **13,** 16–34 (2000).
30. Wilke, G. A. & Bubeck Wardenburg, J. Role of a disintegrin and metalloprotease 10 in *Staphylococcus aureus* α-hemolysin-mediated cellular injury. *Proc. Natl Acad. Sci. USA* **107,** 13473–13478 (2010).
31. Chiu, I. M., von Hehn, C. A. & Woolf, C. J. Neurogenic inflammation and the peripheral nervous system in host defense and immunopathology. *Nature Neurosci.* **15,** 1063–1067 (2012).
32. Abrahamsen, B. *et al.* The cell and molecular basis of mechanical, cold, and inflammatory pain. *Science* **321,** 702–705 (2008).
33. McLachlan, J. B. *et al.* Mast cell-derived tumor necrosis factor induces hypertrophy of draining lymph nodes during infection. *Nature Immunol.* **4,** 1199–1205 (2003).
34. Wong, C. H., Jenne, C. N., Lee, W. Y., Leger, C. & Kubes, P. Functional innervation of hepatic iNKT cells is immunosuppressive following stroke. *Science* **334,** 101–105 (2011).
35. Gautier, E. L. *et al.* Gene-expression profiles and transcriptional regulatory pathways that underlie the identity and diversity of mouse tissue macrophages. *Nature Immunol.* **13,** 1118–1128 (2012).
36. Kim, H. K., Thammavongsa, V., Schneewind, O. & Missiakas, D. Recurrent infections and immune evasion strategies of *Staphylococcus aureus*. *Curr. Opin. Microbiol.* **15,** 92–99 (2011).
37. Holzmann, B. Modulation of immune responses by the neuropeptide CGRP. *Amino Acids* **45,** 1–7 (2011).
38. Lang, R. & Kofler, B. The galanin peptide family in inflammation. *Neuropeptides* **45,** 1–8 (2010).
39. Pintér, E., Helyes, Z. & Szolcsanyi, J. Inhibitory effect of somatostatin on inflammation and nociception. *Pharmacol. Ther.* **112,** 440–456 (2006).
40. Harzenetter, M. D. *et al.* Negative regulation of TLR responses by the neuropeptide CGRP is mediated by the transcriptional repressor ICER. *J. Immunol.* **179,** 607–615 (2007).
41. Gomes, R. N. *et al.* Calcitonin gene-related peptide inhibits local acute inflammation and protects mice against lethal endotoxemia. *Shock* **24,** 590–594 (2005).
42. Fernandes, E. S. *et al.* TRPV1 deletion enhances local inflammation and accelerates the onset of systemic inflammatory response syndrome. *J. Immunol.* **188,** 5741–5751 (2012).
43. Andersson, U. & Tracey, K. J. Reflex principles of immunological homeostasis. *Annu. Rev. Immunol.* **30,** 313–335 (2012).
44. Sun, J., Singh, V., Kajino-Sakamoto, R. & Aballay, A. Neuronal GPCR controls innate immunity by regulating noncanonical unfolded protein response genes. *Science* **332,** 729–732 (2011).
45. Rosas-Ballina, M. *et al.* Acetylcholine-synthesizing T cells relay neural signals in a vagus nerve circuit. *Science* **334,** 98–101 (2011).
46. Weihe, E. *et al.* Molecular anatomy of the neuro-immune connection. *Int. J. Neurosci.* **59,** 1–23 (1991).

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

Acknowledgements We thank L. Barrett, V. Wang, N. Andrews, C. Melin, Y. Wang, K. Duong, E. Cobos del Moral, O. Babanyi and G. Bryman for technical help. We thank Y.-C. Cheng and R. Becker for technical advice; J. Sprague and A. Yekkirala for developing whole-well imaging; I. Inoshima for recombinant αHL; R. Malley, J. Steen and Q. Ma for discussions; J. Chiu for moral support; S. Liberles, B. Xu and V. Kuchroo for mentoring. This work was supported by NIH PO1AI078897, 5RO1AI039246 (M.C.C.), R37NS039518, 5P01NS072040 (C.J.W.), 5F32NS076297 (I.M.C.), FACS, and microarrays at Boston Children's Hospital IDDRC facilities (NIH-P30-HD018655).

Author Contributions I.M.C. and C.J.W. designed the study. I.M.C. and B.A.H.: infection and immune analysis; I.M.C. and N.G.: behavioural analysis; N.G. and A.S.: cytokine profiling; I.M.C. and S.M.: microscopy; I.M.C., C.A.V.H. and J.T.: neuronal culture, calcium imaging; S.W.H.: electrophysiology; F.Z.: peptide synthesis and chemistry; B.W.: multielectrode arrays. J.B.W. and A.R.H.: generation of bacterial strains; J.B.W.: recombinant αHL; J.B.W., M.C.C. and C.J.W.: supervision and expertise. I.M.C. and C.J.W. wrote the manuscript.

Author Information Microarray data are deposited at the NCBI GEO database under accession number GSE46546. 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 C.W. (clifford.woolf@childrens.harvard.edu).

METHODS

Mice. C57BL/6, C57BL/6 *Tlr2*$^{-/-}$, C57BL/6 *Myd88*$^{-/-}$, C57BL/6 *Rag1*$^{-/-}$, C57BL/6 TdTomato reporter mice (ai14 line[47]), C57BL/6 Zsgreen mice, C57BL/6 DTA reporter mice[48], C57BL/6 *Trpv1*$^{-/-}$ mice, NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Wjl}/SzJ and NOD wild-type mice were purchased from Jackson Laboratories; C57BL/6 *Fpr1*$^{-/-}$, C57BL/6 wild-type controls were purchased from Taconic Farms. Nav1.8-Cre (SNS-Cre) mice[20] were a gift from R. Kuner (University of Heidelberg). Nav1.8-Cre mice were bred with C57BL/6 TdTomato, C57BL/6 Zsgreen mice to generate Nav1.8-Cre/TdTomato and Nav1.8-Cre/Zsgreen mice. Nav1.8-Cre$^{+/-}$ mice were bred with C57BL/6 DTA$^{+/+}$ mice to generate nociceptor-deficient Nav1.8-Cre$^{+/-}$/DTA$^{+/-}$ and control littermates (Nav1.8-Cre$^{-/-}$/DTA$^{+/-}$). For infection and behavioural experiments, adult, 7–14-week-old male mice were used, except for when age-matched, male and female Nav1.8-Cre/DTA with control littermates were used. All bacterial and animal experiments were conducted according to institutional animal care and safety guidelines and with IACUC approval at Boston Children's Hospital and Harvard Medical School.

Statistical analysis. Sample sizes for all experiments were chosen according to standard practice in the field. Bar and line graphs are plotted as mean ± standard error (s.e.m.), and in some cases, individual mice are plotted as dots. '*n*' represents the number of mice used in each group. Statistical analysis of behavioural data of mechanical, heat and cold, as well as tissue swelling time courses was conducted by two-way repeated measures ANOVA, with Bonferroni post-tests conducted for each time point tested. In these behavioural analyses, saline-injected mice or wild-type mice served as control groups for each statistical comparison.

For electrophysiology, significance was calculated using the Mann–Whitney *U*-test. Statistical comparisons of acute nocifensive behaviour (total time licking/biting in 20 min), tissue bacterial load, tissue swelling measurements, immune cell influx, neuronal responses by calcium imaging, CGRP and TNF-α levels were by unpaired, student's *t*-test. In acute pain analysis, saline-injected mice were compared to treatment groups. In CGRP and TNF analysis, buffer-treated cells were controls for comparisons. In calcium imaging, the proportion of neuronal responses from at least three fields was quantified. Data were plotted using Prism (Graphpad).

Behavioural analysis. All animals were acclimatized to the behavioural testing apparatus used on at least three habituation sessions. At least two baseline measures were obtained for each behavioural test before testing. To measure mechanical sensitivity, animals were placed on an elevated wire grid and the lateral plantar surface of the hindpaw stimulated with von Frey monofilaments (0.007–8 g). The withdrawal threshold was determined as the filament at which the animal withdrew its paw at least five in ten applications. To measure cold sensitivity, animals were placed on an elevated wire grid and a drop of acetone was applied to the plantar hindpaw by syringe. The duration of time that the animal elevated or licked the paw over a 90-s period immediately after acetone application was measured. To measure heat sensitivity, mice were plated on the glass plate of a Hargreave's apparatus set at 29 °C (IITC Life Science), and a radiant heat source applied to the plantar hindpaw. Latency to hindpaw flicking/licking was recorded (maximum of 30 s). Acute nocifensive behaviour was scored by observation of mice under a glass beaker after intraplantar injection. Time spent lifting/flinching/licking the hindpaw was recorded in 5-min intervals. Acute nocifensive behaviour was quantified as the total licking/lifting in the first 20 min after injection. In infection and compound injection experiments, observers were blinded to mouse genotype/strain. When different substances were compared within experiments, animals were randomized so similar group mean baseline thresholds were present.

Bacterial strains. *S. aureus* CA-MRSA strains LAC/USA300 and LAC/300 deficient in αHL were generated as described previously[6,27]. Strains were grown in tryptic soy broth (TSB) and tryptic soy agar (TSA) (BD Biosciences). GFP–LAC/USA300 strain AH1726 was generated by transforming GFP-expressing, chloramphenicol-resistant plasmid pCM29[49] into LAC. Plasmid DNA was electroporated into *S. aureus* as previously described[50]. The resulting strain produces GFP in a constitutive manner from the *sarA* P1 promoter, and maintained in TSB or TSA supplemented with 10 μg ml^{-1} chloramphenicol.

***S. aureus* infection model and tissue measurements.** *S. aureus* strains were grown overnight in TSB to log phase, pelleted, and re-suspended in 0.9% saline at different dilutions. OD_{650} was measured to estimate bacterial density, with confirmation by TSA plating. For infection, 20 μl of bacteria (5×10^6 c.f.u.) in 0.9% saline was injected subcutaneously into the hindpaw plantar surface using a 25 μl syringe fitted with a 26-gauge needle (Hamilton Co). Mice were monitored closely after injection, and assayed for pain behaviour at different intervals. For neutrophil/monocyte depletion, mice were injected intraperitoneally (i.p.) with GR1 antibody (clone RB6-8C5, Bio-XCell) at a dose of 400 μg per 200 μl PBS, 24 h and 1 h before bacterial injections. Control rat IgG (Jackson Immunoresearch) was injected at the same dosage. For neuropeptide experiments, CGRP, galanin and somatostatin (Tocris, doses of 1 pmol per 200 μl saline) were injected i.p. into mice at 24 h and 4 h before infection, and 4 h and 20 h after infection. As controls, saline was injected at the same time points. For tissue swelling analysis, a digital micrometer (Mitutoyo) was used to measure thickness of the plantar area before and at defined time points after infection. Thickness increase was calculated as differences from baseline measurements (or as % increase normalized to baseline). For bacterial load determination, total paw tissue from epidermis to the tendons was dissected onto ice and weighed, dissociated by dounce homogenizer (Wheaton) in 1 ml PBS, serial dilutions made, plated on TSA plates, and bacterial recovery determined by counting colonies after overnight incubation (normalized by dissected tissue weight). GFP–LAC strain was grown on TSA plates with 10 μg ml^{-1} chloramphenicol.

Neuronal cultures and calcium imaging. Dorsal root ganglia (DRG) from adult mice (7–12 weeks) were dissected into neurobasal-A medium (Life Technologies), dissociated in 1 mg ml^{-1} collagenase A plus 2.4 U ml^{-1} dispase II (enzymes, Roche Applied Sciences) in HEPES-buffered saline (Sigma) for 70 min at 37 °C. After trituration with glass Pasteur pipettes of decreasing size, DRG cells were centrifuged over a 10% BSA gradient, plated on laminin-coated cell culture dishes in B27 supplemented neurobasal-A medium plus 50 ng ml^{-1} nerve growth factor (NGF) plus penicillin/streptomycin (Life Technologies). For bacterial co-cultures, Nav1.8-Cre/TdTomato or Nav1.8-Cre/ZsGreen DRG neurons were plated in poly-lysine, laminin pre-coated, 8-well chamber slides (Lab-Tek) overnight at 37 °C. 10^6 c.f.u. GFP–*S. aureus* or CMTMR-labelled *S. pneumoniae* were added in neurobasal-A to DRG neurons for 2 h at 37 °C; co-cultures were fixed with 4% PFA and mounted in Vectashield (Vector Labs) for microscopy. DRG neurons were used for calcium imaging and electrophysiology 16–48 h after plating.

For calcium imaging, cells were loaded with 10 μM Fura-2-AM (Life Technologies) at 37 °C for 45 min in neurobasal-A medium, washed into Standard Extracellular Solution (SES, 145 mM NaCl, 5 mM KCl, 2 mM $CaCl_2$, 1 mM $MgCl_2$, 10 mM glucose, 10 mM HEPES, pH 7.5), and imaged at room temperature. Cells were illuminated by an ultraviolet light source (Xenon lamp, 75 W, Nikon), 340 nm and 380 nm excitation alternated by a LEP MAC 5000 filter wheel (Spectra services), and fluorescence emission captured by Cool SNAP ES camera (Princeton Instruments). 340/380 ratiometric images were processed, background corrected, and analysed with IPLab software (Scientific Analytics). Microsoft Excel was used for further analyses (Microsoft). Ligands were flowed directly onto neurons using perfusion barrels followed by buffer washout and further application, or applied to the culture bath at the beginning of imaging. In some experiments, 1 μM capsaicin (Tocris), 100 μM AITC (Sigma) or 40 mM KCl (Sigma) was applied after bacterial ligands.

Electrophysiology. Whole-cell transmembrane voltages of DRG neurons were recorded at room temperature (21 ± 1 °C) in the current-clamp mode using an Axopatch 200A amplifier (Molecular Devices). The internal pipette solution consisted of (in mM): 140 KCl, 5 NaCl, 2 $MgCl_2$, 0.5 $CaCl_2$, 5 EGTA, 10 HEPES, 3 Na_2ATP and 0.1 MgGTP (pH 7.4 with KOH). The extracellular solution consisted of the following: 145 NaCl, 5 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 10 HEPES and 10 glucose (pH 7.4 with NaOH). Data were sampled at 5–10 kHz and analysed using pCLAMP 10.2 (Molecular Devices). Resting membrane potentials of DRG neurons were typically −60 to −65 mV and cell body diameters were 15–25 μm.

For multi-electrode arrays, DRG neurons were isolated as described, plated on MED-P515A 64-electrode probes, and recorded using a MED64 device (Alpha Med Scientific). Recombinant αHL was applied at 30 μg ml^{-1} concentration. Spikes were identified using Mobius software (Alpha Med Scientific), and histogram analysis was performed in Matlab (Mathworks).

Bacterial ligands. For heat-killed bacteria experiments, to standardize titres used for differential comparisons we used bacteria purchased from Invivogen (10^{10} c.f.u., each heat-treated for 30 min at 120 °C.): heat-killed *S. aureus*, American-type tissue culture (ATCC) strain 6539 Rosenbach; heat-killed *S. pneumoniae*, NETC7466; heat-killed *H. pylori*, ATCC strain 43504, NTCTC 11637; heat-killed *L. monocytogenes*, strain 9668P; heat-killed *M. fermentans*, ATCC strain 19989; heat-killed *P. aeruginosa*, strain ATCC strain BAA-47. For calcium imaging, bacteria were dissolved in SES at 10^7 c.f.u. ml^{-1}. We note that bacterial addition did not affect pH, which remained at pH 7.5. For injections, heat-killed bacteria strains were dissolved in 0.9% saline, 20 μl heat-killed bacteria (10^8 c.f.u. total) was injected into the hindpaw using a Hamilton syringe fitted with a 26-gauge needle.

N-formal peptides were synthesized by Foci solid-phase peptide synthesis and the peptides were formylated by 2,2,2-trifluoroethyl formate[51]. HLPC resulted in the pure compounds fMLF, fMIFL and MIFL, which were confirmed by mass spectrometry and NMR. Peptides were stored at 10 mM concentration dissolved in DMSO (Thermo Fisher) at −20 °C. For calcium imaging, fMLF, fMIFL and MIFL were dissolved in SES to 1 μM concentration. For intraplantar injections, fMLF and fMIFL were dissolved in 0.9% saline, and 1.0 μg of fMLF (2.1 nmol) or 1.3 μg of fMIFL (2.04 nmol) injected in a 20 μl volume.

For *S. aureus* supernatant collection, LAC bacteria were cultured overnight in TSB, removed by centrifugation, and resulting supernatant diluted to 5% in SES buffer for application to DRG neurons. Recombinant *S. aureus* αHL and mutant H35L (αHL(H35L)) were generated and purified as previously described[30]. Commercial αHL (Sigma) induced comparable results *in vitro* and *in vivo*. In some imaging experiments, αHL was bath applied or flowed directly onto neurons at 10 μg ml^{-1} (Fig. 4a and Supplementary Fig. 20). For pain behavioural studies, αHL and αHL(H35L) dissolved in 0.9% saline was injected in 20 μl volume into the hindpaw at described doses. For heat inactivation, the same batch of fMIFL or recombinant αHL was split into two aliquots, one of which was treated at 100 °C for 30 min. Untreated or heat-treated substances were used for mass spectrometry and pain hypersensitivity studies.

Whole-well imaging of calcium flux. DRG neurons were seeded in B27 supplemented Neurobasal-A at 2,000 neurons per well in laminin-coated 384-well microplates (Greiner) at 37 °C for 24 h. Neurons were loaded with Fura-2 AM for 30 min at 37 °C, then washed twice with HBSS (Life Technologies). A Hamamatsu FDSS 7000EX kinetic reader was used to dispense ligands onto DRG neurons and calcium flux recorded every 1.9 s at room temperature for 30 min total. To evaluate the role of extracellular cations, cells were analysed in calcium/magnesium-free HBSS (Life Technologies). For related experiments, EDTA (5 mM) or HBSS was added to wells during recording to evaluate αHL flux.

Immune stimulation and neuroimmune co-cultures. For peritoneal macrophages, 0.5 ml Brewer's thioglycollate solution (2%) was injected i.p. into mice; 4 days later, animals were killed and peritoneal cavities flushed using 10 ml DMEM/10% fetal calf serum (FCS). Cells were plated in 96-well plates at 5,000 cells per well. For bone-marrow-derived macrophages, tibias and femurs of C57BL/6 mice were flushed using a 27 gauge needle, and bone marrow plated in 15-cm Petri dishes in DMEM/10%FCS/50 μM β-mercaptoethanol/20% L929 conditioned media for 7 days. Differentiated macrophages were dislodged from plates using non-enzymatic dissociation media in HBSS (Sigma), and plated into 96-well plates at 5,000 cells per well in DMEM/10% FCS/50 μM β-mercaptoethanol. Anti-CGRP (rat), somatostatin and galanin (Tocris) were used at different concentrations in macrophage stimulation assays. Peritoneal macrophages were stimulated with 10^7 c.f.u. per ml heat-killed *S. aureus* (Invivogen) and bone marrow macrophages stimulated with 1 μg ml^{-1} *S. aureus*-derived lipoteichoic acid (Invivogen) for 16 h in DMEM/10% FCS; TNF-α levels in culture supernatant were determined using an ELISA kit (Biolegend).

For neuron-macrophage co-cultures, DRG neurons were isolated and bone marrow macrophages derived as described above, re-suspended together in neurobasal-A/ 2% FCS/50 ng ml^{-1} NGF, and co-plated at a density of 2,000 neurons + 10,000 macrophages per well into laminin/poly-D-lysine-coated 8-well chamber slides (Lab-Tek); co-cultures were fixed with 4% PFA after 20 h. For neuron–neutrophil co-cultures, DRG neurons were cultured overnight in 8-well chamber slides at 2,000 cells per well. Neutrophils were isolated as described[52,53]: briefly, bone marrow from tibias and femurs was depleted of red blood cells with ammonium-chloride-potassium buffer (0.15 M NH_4Cl, 10 mM $KHCO_3$, 0.1 mM EDTA), run over a discontinuous 52%, 69%, 78% percoll gradient for 30 min at 1,500*g*. The 69%/78% interface and 78% layer containing neutrophils were collected, washed, re-suspended in neurobasal medium/2% FCS/50 ng ml^{-1} NGF and added at 10,000 cells per well to DRG neurons; co-cultures were washed and fixed with 4% PFA after 6 h. Immunostaining was performed with anti-βIII tubulin (Tuj1, 1:1,000), rat anti-CD11b (Biolegend, 1:100), followed by Alexa 568 goat anti-mouse IgG (Life Technologies, 1:500) or Alexa 488 goat anti-rat IgG (Life Technologies, 1:500), and imaged by epifluorescence microscopy.

RT–PCR and quantitative PCR. RNA was extracted from mouse bone marrow, spleen, kidneys, whole DRG, trigeminal ganglia, or DRG neurons cultured for 24 h using Qiazol reagent, followed by the RNeasy mini kit (Qiagen, MD). DNase I treatment (Qiagen) was used to remove genomic DNA, and complementary DNA reverse transcribed using Superscript III with random hexamers (Life Technologies). PCR was conducted for 32 cycles on a Mastercycler Pro thermocycler (Eppendorf), and products run on agarose gels. RT–PCR primers were synthesized by integrated DNA technologies. Primer sequences: *Adam10*, forward 5′-ATGGAG CAAACATGGCATAA-3′, reverse 5′-GCAACATCTGGGGACAAACT-3′; β-actin, forward 5′-TGTTACCAACTGGGACGACA-3′, reverse 5′-TCTCAGCTGTGG TGGTGAAG-3′; *Fpr1*, forward 5′-CAGCCTGTACTTTCGACTTCTCC-3′, reverse 5′-ATTGGTGCCTGTATCACTGGTCT-3′; *Fpr-rs1*, forward 5′-GGCAACTCT GTTGAGGAAAG-3′, reverse 5′-AATGAACTGGTTGGATTAAC-3′; *Fpr-rs2*, forward 5′-CTTTATCTGCTGGTTTCCCTTTC-3′, reverse 5′-CTGCTGCTTG AATCACTGGTTTG-3′; *Fpr-rs3*, forward 5′-TGACACCTTTAATTGCTCCT-3′, reverse 5′-GTTTCTTAATCACTCTCATTGC-3′; *Fpr-rs4*, forward 5′-CAA GAGGGGATGTGTACTGT-3′, reverse 5′-TGTTAAAGGAAGCCAAGGTA-3′; *Fpr-rs6*, forward 5′-CCCCTGAGGAGCAAGTAAAAGTAT-3′, reverse 5′-CAG GGCTGAGTCCTCCCTTA-3′; *Fpr-rs7*, forward 5′-CCTGAGGAGCAGGTAA ACATGT-3′, reverse 5′-GGGCTGAATCCTCCCTCA-3′.

For qPCR, cDNA was subjected to 2-step thermocycling using fast SYBR green master mix (Life Technologies), and data collection performed on an Applied Biosystems 7500 machine (Life Technologies). The following primers were used: *Fpr1*, forward 5′-CATTTGGTTGGTTCATGTGCAA-3′, reverse 5′-AATACAG CGGTCCAGTGCAAT-3′; *Fpr-rs2* (FPR2) was based on ref. 22, forward 5′-TGACTACTGTTAGAATTCCTG-3′, reverse 5′-GATAGCTGTGTTCAACTT TTCTTCAT-3′; *Gapdh*, forward 5′-TGGCAAAGTGGAGATTGTTGCC-3′, reverse 5′-AAGATGGTGATGGGCTTCCCG-3′. Expression levels were normalized to *Gapdh* using the ΔΔcT method.

Flow cytometry analysis of infected tissues and lymph nodes. Paw tissue was minced and digested in 1 mg ml^{-1} collagenase A/2.4 U ml^{-1} dispase II (Roche Applied Sciences) in HEPES buffered saline (Sigma) for 2 h at 37 °C. After digestion, cells were washed into HBSS (Life Technologies) with 0.5% bovine serum albumin (BSA, Sigma), filtered through a 70 μm mesh, re-suspended in HBSS/ 0.5% BSA, blocked with FcBlock (rat anti-CD16/CD32; 1:10 hybridoma supernatant) on ice for 5 min, and incubated with mixtures of the following antibodies: anti-CD11b-PE (Biolegend, 1:1,000), anti-CD45-PercP (Biolegend, 1:200), anti-Ly-6G-APC (Biolegend, 1:200), anti-Ly-6C-FITC (Biolegend, 1:200). Flow cytometry was conducted on a FACSCalibur machine (Bectin Dickinson) equipped with an argon and helium-neon laser.

Lymph nodes were dissected into 5 ml RPMI-1640 on ice. After dissections, RPMI-1640 was removed and replaced with 2 ml enzyme mix: RPMI-1640 containing 0.8 mg ml^{-1} dispase + 0.2 mg ml^{-1} collagenase P (both from Roche) plus 0.1 mg ml^{-1} DNase I (Life Technologies). Tubes were incubated at 37 °C in a water bath and inverted at 5-min intervals to ensure well mixing. After 20 min, lymph nodes were gently aspirated and expirated using a 1-ml pipette, which disrupts the capsule to release leukocytes. The mixture was replaced in the water bath and large fragments allowed to settle for 30 s, after which the enzyme mix was removed and added to 10 ml of ice-cold FACS buffer (0.5% BSA, 2 mM EDTA, 10 mM NaN3, 15 mM HEPES in PBS, pH 7.4). 2 ml of fresh enzyme mix was added to the digestion and the process was repeated until all lymph node fragments were completely digested. Cells were filtered through 80 μm nylon mesh and counted by haemocytometer. 5×10^6 cells were incubated with 50 μl diluted antibodies (TCRβ, CD11b, Ly6C, Ly6G, CD19, antibodies from Biolegend) for 20 min at 4 °C in FACS buffer before acquisition on a FACSCalibur or FACSAria II (BD Biosciences). Neutrophils were considered $CD11b^+Ly6G^+$, monocytes $CD11b^+Ly6G^-Ly6C^+$, B cells $CD19^+$, T cells $TCR\beta^+$. Flow cytometry data were analysed using FlowJo software (TreeStar).

Immunostaining and microscopy. Mice were transcardially perfused with PBS followed by 4% PFA/PBS (Sigma). Plantar tissue was dissected, post-fixed for 2 h, cryoprotected in 30% sucrose/PBS, embedded, and frozen in Optimal cutting temperature compound (OCT, Electron Microscopy Sciences). Tissues were stored at −80 °C until sectioning. Cryosections were cut at 50 μm thickness onto Superfrost plus slides (Thermo Fisher). For haematoxylin and eosin, sections were dehydrated by sequential alcohol steps, mounted in Permount medium (Thermo Fisher), and imaged by light microscopy on a BZ-II analyser at ×15 magnification (Keyence). For immunostaining, sections were stained with GR1-Alexa 647 antibody (Biolegend, 1:100) for 2 h at room temperature, mounted in Vectashield with DAPI (Vector Labs) and imaged using an LSM700 laser-scanning confocal microscope (Carl Zeiss). ×10 Zeiss EC plan-NEOFLUAR dry and ×63 Zeiss plan-Apochromat oil objectives were used, with z-stacks of 40 μm total imaged at 1 μm steps; maximum projection images were exported for presentation.

For DRG, tissues were dissected, post-fixed, and embedded in OCT as described above. 14-μm cryosections were cut and stained with rabbit anti-CGRP (Millipore, PC205L, 1:500) followed by Alexa 488-anti-rabbit IgG (Life Technologies, 1:1,000) or chicken anti-neurofilament (Millipore, AB5539, 1:500), followed by Alexa 488 anti-chicken IgG (Life Technologies, 1:1,000). Sections were mounted in Vectashield with DAPI (Vector Labs), and imaged by Eclipse 50i epifluorescence microscope (Nikon).

Flow-cytometric purification of nociceptors and ADAM10 staining. DRG, trigeminal, or nodose ganglia from Nav1.8-Cre/TdTomato mice were dissected, dissociated into single cells by enzymatic digestion (1 mg ml^{-1} collagenase A plus 2.4 U ml^{-1} dispase II (Roche Applied Sciences) in HEPES buffered saline (Sigma)) for 60 min at 37 °C, filtered through a 70-μm filter, and stained with DAPI (Sigma, 20 ng ml^{-1}) as a dead cell stain in HBSS/0.5% BSA. Nociceptors were sorted on a FACSaria II machine (Bectin Dickinson) using a yellow–green laser to detect TdTomato fluorescence (gates, Fig. 6, Supplementary Fig. 14). To determine purity, cells were sorted into neurobasal-A and plated onto laminin-coated glass slides for microscopy. For microarrays, neurons were sorted directly into Qiazol reagent (Qiagen). For ADAM10 surface staining, Nav1.8-Cre/ TdTomato DRG were dissociated and stained with rat anti-ADAM10 ectodomain antibody (clone 139712, R&D systems; reviewed, 1DegreeBio) or with rat IgG2A isotype control (clone RTK2758, leaf purified, Biolegend) for 1 h on ice

($5\ \mu g\ ml^{-1}$ each). After washing with HBSS, cells were incubated with Alexa 488 goat anti-rat IgG ($4\ \mu g\ ml^{-1}$, Life technologies) for 30 min on ice. After two washes with HBSS, cells were re-suspended in HBSS/0.5% BSA and analysed on a BD FACSaria II machine. Flow cytometry data was analysed using FlowJo (Treestar).

Microarray analysis. Total RNA was extracted by sequential Qiazol extraction and purification through the RNeasy micro kit with on column genomic DNA digestion (Qiagen). RNA quality was determined by an Agilent 2100 Bioanalyzer using the RNA Pico Chip (Agilent). RNA was amplified into cDNA using the Ambion wild-type expression kit for whole transcript expression arrays, with Poly-A controls from the Affymetrix Genechip Eukaryotic Poly-A RNA control kit. The Affymetrix Genechip WT terminal labelling kit was used for fragmentation, biotin labelling. Affymetrix GeneChip Hybridization control kit and the Affymetrix GeneChip Hybridization, wash, stain kit was used to hybridize samples to Affymetrix Mouse Gene ST 1.0 GeneChips, fluidics performed on the Affymetrix Genechip Fluidics Station 450, and scanned using Affymetrix Genechip Scanner 7G (Affymetrix). Microarray work was conducted at the Boston Children's Hospital IDDRC Molecular Genetics Core, which is supported by NIH-P30-HD 18655. Affymetrix CEL files were normalized with the robust multi-array average (RMA) algorithm with quantile normalization, background correction, and median scaling. The ImmGen data set was also analysed (GEO accession number GSE15907). Heat maps were generated using GenePattern platform (Broad institute, MIT). Nociceptor microarray data sets are deposited at the GEO database under accession number GSE46546.

47. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nature Neurosci.* **13,** 133–140 (2009).
48. Voehringer, D. *et al.* Homeostasis and effector function of lymphopenia-induced ''memory-like'' T cells in constitutively T cell-depleted mice. *J. Immunol.* **180,** 4742–4753 (2008).
49. Pang, Y. Y. *et al.* agr-Dependent interactions of *Staphylococcus aureus* USA300 with human polymorphonuclear neutrophils. *J. Innate Immun.* **2,** 546–559 (2010).
50. Schenk, S. & Laddaga, R. A. Improved method for electroporation of *Staphylococcus aureus*. *FEMS Microbiol. Lett.* **94,** 133–138 (1992).
51. Hill, D. R. *et al.* 2,2,2-Trifluoroethyl fromate: a versatile and selective reagent for the formylation of alcohols, amines, and N-hydroxylamines. *Org. Lett.* **4,** 111–113 (2002).
52. Boxio, R. *et al.* Mouse bone marrow contains large numbers of functionally competent neutrophils. *J. Leukoc. Biol.* **75,** 604–611 (2004).
53. Sai, J. *et al.* Parallel phosphatidylinositol 3-kinase (PI3K)-dependent and Src-dependent pathways lead to CXCL8-mediated Rac2 activation and chemotaxis. *J. Biol. Chem.* **283,** 26538–26547 (2008).

ARTICLE

doi:10.1038/nature12504

Topoisomerases facilitate transcription of long genes linked to autism

Ian F. King[1], Chandri N. Yandava[2], Angela M. Mabb[1], Jack S. Hsiao[3], Hsien-Sung Huang[1]†, Brandon L. Pearson[1], J. Mauro Calabrese[4], Joshua Starmer[4], Joel S. Parker[4,5], Terry Magnuson[4,5], Stormy J. Chamberlain[3], Benjamin D. Philpot[1,2,6] & Mark J. Zylka[1,2,6]

Topoisomerases are expressed throughout the developing and adult brain and are mutated in some individuals with autism spectrum disorder (ASD). However, how topoisomerases are mechanistically connected to ASD is unknown. Here we find that topotecan, a topoisomerase 1 (TOP1) inhibitor, dose-dependently reduces the expression of extremely long genes in mouse and human neurons, including nearly all genes that are longer than 200 kilobases. Expression of long genes is also reduced after knockdown of *Top1* or *Top2b* in neurons, highlighting that both enzymes are required for full expression of long genes. By mapping RNA polymerase II density genome-wide in neurons, we found that this length-dependent effect on gene expression was due to impaired transcription elongation. Interestingly, many high-confidence ASD candidate genes are exceptionally long and were reduced in expression after TOP1 inhibition. Our findings suggest that chemicals and genetic mutations that impair topoisomerases could commonly contribute to ASD and other neurodevelopmental disorders.

Autism is a neurodevelopmental disorder with symptoms that include repetitive behaviours and deficits in social interactions. Hundreds of genes are now associated with ASD[1,2], suggesting that there are diverse genetic risk factors for autism. Environmental factors, including chemicals that are ingested during critical periods of brain development[3], can also increase autism risk. Many ASD candidate genes regulate synapse function[4–6]; however, whether there are additional mechanisms that unite ASD patients or expression of ASD genes is unclear.

Recently, we found that topoisomerase inhibitors can transcriptionally un-silence the paternal allele of *Ube3a* in mouse cortical neurons[7]. *Ube3a* is located adjacent to a cluster of imprinted genes, is normally expressed only from the maternal allele in neurons, and regulates synaptic function[8]. In humans, *UBE3A* is associated with two distinct neurodevelopmental disorders. Specifically, deletion or mutation of maternal *UBE3A* causes Angelman syndrome, whereas duplication of the chromosomal region containing maternal *UBE3A* is frequently detected in individuals with autism[9,10].

Intriguingly, mutations in topoisomerases were recently identified in some individuals with ASD[11,12]. However, precisely how topoisomerases regulate the expression of *Ube3a* and possibly other genes associated with autism is unknown. Topoisomerases, including *Top1* and *Top2b*, are expressed throughout the developing and adult mouse brain[13,14]. Topoisomerases are integral to gene expression, as they resolve DNA supercoiling that is generated during transcription[15–18]. Here we sought to determine whether topoisomerases preferentially regulate the expression of additional imprinted genes in neurons, or whether they have broader effects on gene expression. Using genome-wide approaches, we unexpectedly found that topoisomerases facilitate the expression of long genes, including numerous long genes associated with synaptic function and ASD. In addition, our study uncovers a transcriptional mechanism that is particularly important for maintaining the expression of numerous ASD genes at normal levels.

Gene length effects

To determine whether topotecan, a TOP1 inhibitor, altered the expression of imprinted genes, we treated cultured cortical neurons from C57BL/6J (B6) × CASTEi/J (CAST) F_1 hybrid mice with vehicle or 300 nM topotecan, then used high-throughput transcriptome sequencing (RNA-seq) to survey changes in gene expression genome-wide. Single nucleotide polymorphisms were used to determine the parent-of-origin of sequence reads for autosomal genes[19]. We defined imprinted genes as those displaying statistically significant parent-of-origin expression bias in reciprocal B6 × CAST crosses (Fisher's exact test, $P < 0.05$ after adjustment for multiple comparisons). We found that cortical neurons expressed 49 known autosomal imprinted genes (Extended Data Table 1), but *Ube3a* was the only imprinted gene that showed a significant change in parental allele bias in reciprocal crosses upon topotecan treatment (Fisher's exact test, $P < 0.05$ after correction; Extended Data Table 1). Indeed, topotecan increased expression of the paternal allele of *Ube3a*, driving *Ube3a* levels significantly above wild-type levels (Extended Data Fig. 1a, b).

As we found previously[7], topotecan reduced the expression of an extremely long (>1 megabase), paternally expressed antisense transcript that overlaps *Ube3a* (*Ube3a-ATS*; Extended Data Fig. 1a, b). The *Ube3a-ATS* transcript is required for paternal *Ube3a* silencing[20,21]. Other imprinted genes in the same genomic region as *Ube3a* did not show changes in allelic expression following topotecan treatment (Extended Data Fig. 1b and Extended Data Table 1). Notably, topotecan also reduced the expression of *UBE3A-ATS* and increased the expression of *UBE3A* in induced pluripotent stem cell (iPSC)-derived neurons from an Angelman syndrome patient (Extended Data Fig. 1c). Topotecan thus had similar transcriptional effects at the *Ube3a* locus in mouse and human neurons.

Because *Ube3a-ATS* is extremely long and was strongly downregulated, we hypothesized that topotecan might reduce the expression of

[1]Department of Cell Biology and Physiology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [2]Carolina Institute for Developmental Disabilities, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [3]Department of Genetics and Developmental Biology, University of Connecticut Health Center, Farmington, Connecticut 06032, USA. [4]Department of Genetics, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [5]Lineberger Comprehensive Cancer Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. [6]UNC Neuroscience Center, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA. †Present address: Graduate Institute of Brain and Mind Sciences, College of Medicine, National Taiwan University, Taipei 10051, Taiwan.

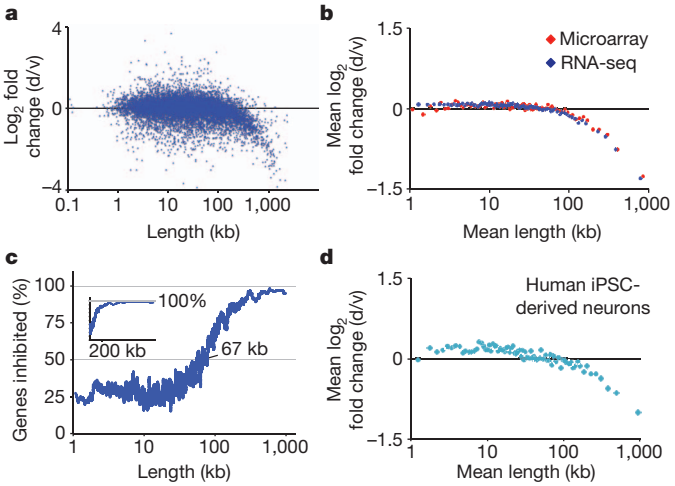


Figure 1 | TOP1 inhibition reduces expression of long genes in neurons. **a**, Mouse cortical neurons treated with vehicle (v) or 300 nM topotecan (drug; d) for 3 days ($n = 5$ biological replicates). RNA-seq gene expression versus gene length. **b**, Mean expression change in bins of 200 genes by length. **c**, Percentage of genes that were reduced in expression by topotecan; plotted as a sliding window of 100 genes by length, RNA-seq data (log scale). Inset, same data on linear scale. **d**, iPSC-derived human neurons treated with 1 μM topotecan for 6 days relative to vehicle, RNA-seq data in bins of 200 genes by length ($n = 2$ biological replicates).

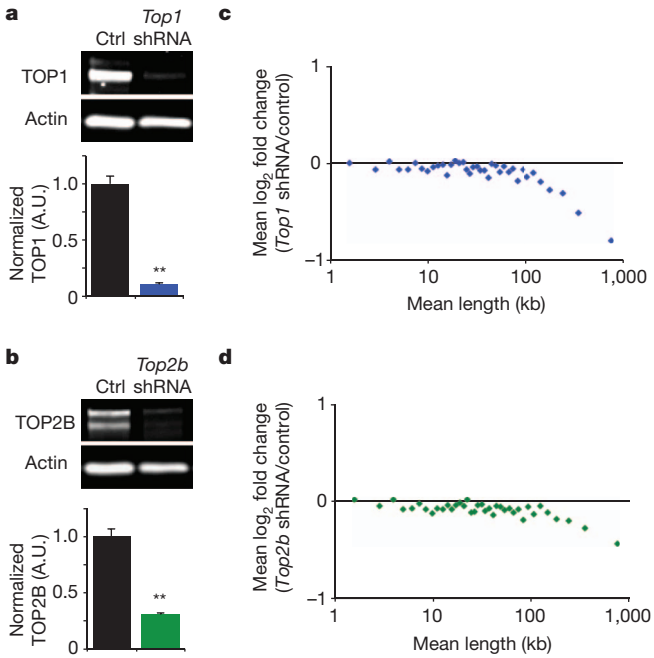


Figure 2 | Lentiviral shRNA knockdown of *Top1* or *Top2b* in mouse cortical neurons reduces the expression of long genes. **a**, **b**, Representative western blots and quantification of TOP1 (**a**) and TOP2B (**b**) 7 days post infection. Normalized to β-actin in arbitrary units (A.U.). $^{**}P < 0.01$ relative to scrambled control (Ctrl), Student's *t*-test. Error bars, s.e.m. $n = 3$ biological replicates. **c**, **d**, Gene expression from Affymetrix microarrays, relative to scrambled control shRNA. Plotted as mean expression change in bins of 200 genes by length.

other long genes. Notably, using RNA-seq and Affymetrix microarrays to quantify gene expression, we found that topotecan reduced the expression of nearly all extremely long genes in mouse cortical neurons (Fig. 1a–c), with a strong correlation between gene length and reduced expression (for genes longer than 67 kilobases (kb); Pearson's correlation coefficient (R) = −0.69). Topotecan also reduced the expression of long genes in iPSC-derived human neurons (Fig. 1d). Topotecan did not exclusively reduce the expression of extremely long genes, but instead acted over a continuum of gene lengths (Fig. 1c). Specifically, the percentage of genes that were inhibited (to any extent) by 300 nM topotecan increased from 50% for genes 67 kb in length to nearly 100% for genes ~200 kb and longer. In addition, inhibition of long genes by topotecan was highly dose dependent (Extended Data Fig. 2).

By contrast, topotecan increased the expression of most of the genes that were <67 kb in length (Fig. 1c), although the magnitude of this increase was very small for most genes (Fig. 1a, b). For some genes, this increase may reflect regulation by longer overlapping transcripts, like for *Ube3a*, or it might reflect other stimulatory effects of topoisomerase inhibitors[22,23].

The length-dependent effects on gene expression were not due to cell death or persistent DNA damage, as topotecan (300 nM for 3 days) did not kill neurons or damage DNA (Extended Data Fig. 3a, b). Moreover, agents that damage DNA in neurons (paraquat and H_2O_2) did not reduce the expression of long genes (Extended Data Fig. 3b–d). Notably, all length-dependent effects were fully reversible upon drug washout (Extended Data Fig. 3e), ruling out the possibility that gene expression changes were due to permanent effects (such as irreversible DNA damage and/or killing neurons).

A different TOP1 inhibitor, irinotecan, had a highly correlated length-dependent effect on gene expression in cortical neurons (Extended Data Fig. 4). In addition, we re-analysed published data from other laboratories and found that irinotecan and camptothecin (another TOP1 inhibitor) strongly reduced the expression of long genes and moderately increased the expression of shorter genes in several human cell lines (Extended Data Fig. 5a–e). Thus, the length-dependent effects we observed were not unique to postmitotic neurons and could be reproducibly detected in expression data acquired by other laboratories.

We found that lentiviral delivery of short hairpin RNA (shRNA) to knockdown *Top1* also reduced the expression of long genes in neurons (Fig. 2a, c), providing independent genetic support that TOP1 facilitates the expression of long genes. These gene knockdown results also rule out the possibility that TOP1–DNA covalent complexes, which form only in the presence of TOP1 inhibitors[18], block the expression of long genes. Unlike TOP1 inhibitors (Fig. 1a–c and Extended Data Figs 4a and 5a–e), *Top1* knockdown did not globally increase the expression of shorter genes (Fig. 2c). Thus TOP1 inhibitors probably increase the expression of shorter genes through a drug-specific effect that is unrelated to TOP1 depletion.

TOP2 enzymes (particularly TOP2B) also participate in gene transcription[15,16,24]. We next tested whether genetic or pharmacological inhibition of TOP2 enzymes could reduce the expression of long genes. Indeed, with new experiments and by re-analysing data from others[14,25], we found that the TOP2A/TOP2B inhibitor ICRF-193 reduced gene expression in a length-dependent manner in cultured mouse cortical neurons, embryonic stem (ES) cells and ES-cell-derived neurons (Extended Data Figs 6a and 7a, b). There was extensive overlap between genes affected by ICRF-193 and topotecan in cortical neurons, particularly for long genes, and the magnitudes of these effects were highly correlated (Extended Data Fig. 6b–e). Thus, TOP1 and TOP2 enzymes regulate the expression of many of the same genes.

Top2b is the predominant TOP2 expressed in neurons[25]. We next knocked down *Top2b* with shRNA (Fig. 2b, d) and found that this manipulation reduced the expression of long genes (Fig. 2d). Moreover, re-analysis of published data sets showed that expression of long genes was reduced in embryonic brain and ES-cell-derived neurons from *Top2b*$^{-/-}$ mice[14] (Extended Data Fig. 7c, f, g). By contrast, long genes were expressed normally in *Top2b*$^{-/-}$ ES cells and neuronal progenitors[25] (Extended Data Fig. 7d, e), presumably because these cell types express *Top2a* in addition to *Top2b*[25]. Last, two additional TOP2 inhibitors (doxorubicin and etoposide) reduced the expression of long genes in a human cancer cell line[26] (Extended Data Fig. 5f, g). Together, our data show that TOP1 and TOP2 enzymes are required for proper expression of long genes in mammalian cells. This stands in contrast to yeast, in

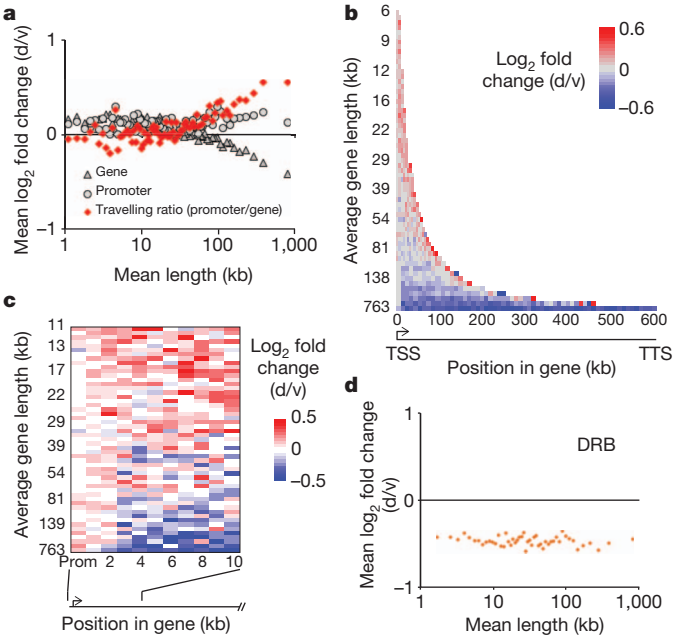


Figure 3 | Topotecan impairs transcription elongation of long genes. **a**, Mouse cortical neurons treated with vehicle or topotecan (300 nM for 3 days). Pol II density in gene bodies and promoter regions, averaged for bins of 200 genes by length. **b**, **c**, Change in Pol II density across genes (**b**) and across the first 10 kb (**c**), averaged for groups of 200 genes by length. Gene bins aligned relative to the transcription start site (TSS). TTS, transcription termination site. **d**, Mouse cortical neurons treated with 100 μM DRB for 3 days. Affymetrix microarray expression compared to controls in bins of 200 genes by length.

which length-dependent transcriptional effects specific to Top2 were observed[27].

Length-dependent effect on elongation

Topoisomerases interact directly with RNA polymerase II (Pol II) and are required for transcription elongation[17,28,29]. To study genome-wide effects of topotecan on transcription in neurons, we measured Pol II distribution by chromatin immunoprecipitation followed by high-throughput DNA sequencing (ChIP-seq). We calculated the travelling ratio for all genes bound by Pol II, defined as the ratio of Pol II density (read counts per million mapped reads) near transcription start sites to Pol II density in the body of the gene, a parameter that is increased when transcription elongation is inhibited[30–32]. We found that topotecan produced a distinct length-dependent increase in travelling ratio for very long genes, consistent with an impairment of transcription elongation (Fig. 3a and Extended Data Fig. 8).

The change in travelling ratio observed with topotecan treatment could reflect either a progressive inhibition of Pol II as it transcribes long genes, or it could reflect a block in the transition to productive elongation. To distinguish between these possibilities, we examined the change in Pol II density across the entire length of genes at high resolution (Fig. 3b). We found that Pol II density was slightly increased throughout the gene bodies of smaller genes, consistent with the modest increase in expression seen for smaller genes and with studies showing that TOP1 inhibitors can stimulate the transition to elongation of two genes (both shorter than 67 kb)[5,6]. However, for longer genes, topotecan strongly reduced Pol II density across the entire length of the gene body (Fig. 3b). These results are consistent with topotecan affecting the transition to productive elongation at long genes. On average, all genes showed a slight increase in Pol II density in the promoter-proximal region (Fig. 3c) and a stronger increase in Pol II density near the transcription termination site (Fig. 3b). However, these changes were independent of length, making it unlikely that they contributed to differential expression of short versus long genes.

TOP1 inhibitors can stimulate the release of the positive elongation factor P-TEFb[17]. Thus, we next tested whether P-TEFb inhibition would affect expression of longer genes differently from shorter genes by treating neurons with 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB; a P-TEFb inhibitor that can block the transition to elongation[17,30]). We found that DRB (100 μM) reduced gene expression equally across all gene lengths (Fig. 3d), suggesting that the requirement for P-TEFb is not influenced by length. (Note that 1 μM and 10 μM DRB had no length-dependent effects on expression (data not shown), ruling out the possibility that lower DRB concentrations preferentially impair expression of longer genes.)

Numerous long ASD genes affected

To investigate further the biological consequences of TOP1 inhibition in neurons, we defined a list of genes that were differentially expressed with high confidence. From our RNA-seq expression data, we found that topotecan significantly downregulated 155 genes and significantly upregulated 28 genes (Benjamini–Hochberg method, 5% false discovery rate) (Supplementary Data 1). The topotecan-downregulated genes were significantly longer (mean 591 kb, median 548 kb) than all expressed genes in cortical neurons (mean 59.3 kb, median 23.5 kb) and were significantly longer than topotecan-upregulated genes (mean 29.3 kb, median 16.4 kb) (one-way analysis of variance versus all expressed genes $P = 2.2 \times 10^{-16}$, versus upregulated genes $P = 3.7 \times 10^{-14}$), further indicating that topotecan has pronounced effects on long genes.

On the basis of Gene Ontology and functional annotation terms, we found that many topotecan-downregulated genes were involved in neuronal development and synaptic function (Supplementary Data 2). Because ASD is thought to be a neurodevelopmental disorder that affects synapses, we cross-referenced our list of downregulated genes with known ASD candidate genes, combining genes in the SFARI Gene database with candidates identified in recent exome sequencing studies[11,12,33–35] (Supplementary Data 3). Notably, 27% ($n = 49$) of the 183 differentially expressed genes are known ASD candidate genes (Table 1 and Supplementary Data 3), a proportion that is highly significant compared to chance ($P = 4.4 \times 10^{-8}$, Fisher's exact test). Independent microarray experiments showed that these ASD genes were dose-dependently downregulated by topotecan (Extended Data Fig. 9). Notably, ASD candidate genes are exceptionally long as a group; genes in the SFARI Gene database (as of June 20, 2013) are 3.7-fold longer on average than all genes expressed in cortical neurons (means of 217.3 kb versus 59.3 kb). Thus, mutations that alter topoisomerase activity might reduce expression of numerous long ASD genes and might contribute significantly to ASD. Consistent with this hypothesis, recent sequencing studies of autism patient cohorts uncovered rare *de novo* missense mutations in *TOP1*, *TOP3B*, *TOPORS* (a TOP1-SUMO ligase[36]) and several other genes that directly connect to TOP1[11,12].

Discussion

Our study shows that topoisomerases facilitate the expression of a large number of ASD candidate genes, including many that are long and that are thought to have large effects on ASD pathology in isolation[1,2,37]. Pharmacological inhibition of topoisomerases also reduced expression of long genes in other cell types, suggesting that this length-dependent transcriptional effect is fundamental to all mammalian cells. Our data rule out numerous possibilities as to why topoisomerase inhibitors reduce the expression of long genes (for example, cell death, DNA damage, formation of covalent complexes), and instead implicate a gene-length-dependent block in transcription elongation. Pol II and topoisomerases dynamically form and remodel large supercoiling domains[38], and the effects of topoisomerases on gene expression are strongly influenced by genomic structure and context[23,39]. Thus, we speculate that higher order structure differentially constrains shorter and longer genes, and that this creates distinct length-dependent requirements for topoisomerases in transcription elongation.

Table 1 | Topotecan reduces expression of numerous ASD candidate genes in neurons

Gene	Length in mouse (kb)	Length in human (kb)	Log_2 fold change (mouse)	*P* (adjusted)	SFARI Gene	Mutation
Cntnap2*	2241.3	2304.6	−4.5718	0.0055	Yes	H275R[34]
Csmd1	1642.8	2059.4	−3.4896	<0.0001	No	Q2254R[33]
Nrxn3	1612.1	1460.7	−2.9494	<0.0001	Yes	
Fhit	1611.9	1502.1	−2.1229	0.0490	Yes	
Rbfox1	1527.7	1694.2	−1.3178	0.0441	Yes	
Grid2	1409.4	1468.1	−2.4100	0.0204	Yes	
Il1rapl1	1368.4	1368.3	−3.0444	0.0213	Yes	
Il1rapl2	1275.4	1200.8	−2.6076	0.0034	Yes	
Park2	1223.0	1380.2	−2.8055	0.0003	Yes	
Cntn5	1243.9	1337.9	−2.3360	0.0009	Yes	
Ptprt	1133.5	1117.2	−2.0706	<0.0001	Yes	
Frmpd4	1105.9	586.1	−1.5332	0.0028	Yes	
Erbb4	1068.1	1162.9	−3.3890	<0.0001	Yes	
Nrxn1	1059.2	1114.0	−2.1914	<0.0001	Yes	Y587STOP[12]
Gpc6	1054.2	1181.2	−2.3530	<0.0001	Yes	
Astn2	1023.7	989.8	−2.7763	0.0001	Yes	
Cntn4	1021.6	959.1	−2.0717	0.0092	Yes	
Pard3b	1003.5	1074.4	−2.9457	<0.0001	Yes	
Epha6	952.0	934.3	−1.8366	0.0301	Yes	
Grm7	921.6	880.4	−2.2647	<0.0001	No	R622Q[33]
Nckap5	917.0	896.7	−1.3453	0.0089	Yes	
Nlgn1	900.1	884.9	−1.8339	<0.0001	Yes	H795Y[34]
Pcdh9	875.2	927.5	−1.8314	0.0002	Yes	
Pcdh15	827.9	998.5	−3.2033	<0.0001	Yes	
Grid1	761.0	766.9	−1.4922	0.0448	Yes	
Mdga2	756.5	835.3	−1.6942	0.0012	Yes	
Dpp10	713.4	1402.4	−1.7914	0.0007	Yes	
Kcnma1	705.4	768.2	−1.9021	0.0004	Yes	F372V[11]
Grik2	689.3	669.2	−1.5500	0.0028	Yes	
Plcb1	689.1	752.3	−1.4067	0.0075	Yes	
Ptprm	687.6	839.5	−1.7610	0.0057	No	A535T[12]
Rbms3	679.7	729.1	−1.6432	0.0191	Yes	
Vps13b	559.7	864.3	−1.0465	0.0317	Yes	
Nbea	558.5	730.5	−1.3664	0.0147	Yes	
Grm5	550.9	559.1	−1.3240	0.0343	Yes	Δaa679[12]
Kalrn	545.0	626.5	−1.3646	0.0186	Yes	
Ptprk	522.6	551.9	−1.6458	0.0005	No	R784H[34]
Atrnl1	522.3	855.3	−1.2194	0.0211	Yes	
Robo2	518.6	1712.5	−1.4921	0.0017	Yes	
Cacna1c	515.9	644.7	−1.1665	0.0236	Yes	
Lrrc7	477.8	363.3	−1.6330	0.0340	Yes	
Reln	460.3	517.7	−2.0687	<0.0001	Yes	Q417STOP[11] R2290C[12]
Exoc6b	451.0	646.7	−0.9691	0.0285	Yes	
Gabrb1	437.2	395.2	−1.4909	0.0041	Yes	
Nfia	341.3	385.5	−1.8218	0.0186	Yes	R30STOP[12]
Nxph1	298.6	319.0	−1.8182	0.0181	Yes	
Myb	36.1	37.9	−3.8779	0.0011	No	R45Q[33]
C3	24.1	42.8	−2.2892	<0.0001	No	I1569M[11]
C4b	15.5	1.7	−1.5240	0.0121	Yes	

* Isoform a (*uc009bst.2*).
aa, amino acid.

Some long genes were not strongly reduced in expression following topotecan treatment (Fig. 1a). In many cases this reflected ambiguity in gene annotation (data not shown). For example, a number of long genes also express shorter transcripts, making it difficult to distinguish expression of short isoforms from long isoforms. Alternatively, some long genes might be located within genomic regions that are more permissive for expression when TOP1 is inhibited.

Intriguingly, numerous genes associated with transcription are mutated in autism patients[11,40,41], although how these diverse transcriptional regulators contribute to autism is unclear. Our study highlights a mechanistic link between a critical step in transcription elongation and expression of numerous long ASD candidate genes. Our data suggest that chemicals or genetic mutations that impair topoisomerases, and possibly other components of the transcription elongation machinery that interface with topoisomerases, have the potential to profoundly affect the expression of long ASD candidate genes. Length-dependent impairment of gene transcription, particularly in neurons and during critical periods of brain development, may thus represent a unifying cause of pathology in many individuals with ASD and other neurodevelopmental disorders.

METHODS SUMMARY

All procedures using vertebrate animals were approved by the University of North Carolina Institutional Animal Care and Use Committee. Embryonic (E13.5–15.5) cortical neuron cultures from C57BL/6J × CASTEi/J mice (The Jackson Laboratory) were used for RNA-seq and ChIP-seq studies. C57BL/6J mice were used for all other mouse experiments. Cultures were prepared as previously described[7]. Affymetrix Mouse Genome 430 2.0 arrays and Mouse Gene 1.1 ST 24-array plates were used for microarray studies. ChIP-seq experiments were performed essentially as described[30,42] using RNA Pol II N20 antibody (sc-899, Santa Cruz Biotechnology).

Online Content Any additional Methods, 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 17 January; accepted 24 July 2013.
Published online 28 August 2013.

1. Abrahams, B. S. & Geschwind, D. H. Advances in autism genetics: on the threshold of a new neurobiology. *Nature Rev. Genet.* **9,** 341–355 (2008).
2. State, M. W. & Levitt, P. The conundrums of understanding genetic risks for autism spectrum disorders. *Nature Neurosci.* **14,** 1499–1506 (2011).
3. Christensen, J. G. T. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *J. Am. Med. Assoc.* **309,** 1696–1703 (2013).

4. Delorme, R. *et al.* Progress toward treatments for synaptic defects in autism. *Nature Med.* **19,** 685–694 (2013).
5. Betancur, C., Sakurai, T. & Buxbaum, J. D. The emerging role of synaptic cell-adhesion pathways in the pathogenesis of autism spectrum disorders. *Trends Neurosci.* **32,** 402–412 (2009).
6. Peça, J. & Feng, G. Cellular and synaptic network defects in autism. *Curr. Opin. Neurobiol.* **22,** 866–872 (2012).
7. Huang, H.-S. *et al.* Topoisomerase inhibitors unsilence the dormant allele of *Ube3a* in neurons. *Nature* **481,** 185–189 (2012).
8. Mabb, A. M., Judson, M. C., Zylka, M. J. & Philpot, B. D. Angelman syndrome: insights into genomic imprinting and neurodevelopmental phenotypes. *Trends Neurosci.* **34,** 293–303 (2011).
9. Cook, E. H. *et al.* Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *Am. J. Hum. Genet.* **60,** 928–934 (1997).
10. Moreno-De-Luca, D. *et al.* Using large clinical data sets to infer pathogenicity for rare copy number variants in autism cohorts. *Mol. Psychiatry* http://dx.doi.org/10.1038/mp.2012.138 (2012).
11. Neale, B. M. *et al.* Patterns and rates of exonic *de novo* mutations in autism spectrum disorders. *Nature* **485,** 242–245 (2012).
12. Iossifov, I. *et al. De novo* gene disruptions in children on the autistic spectrum. *Neuron* **74,** 285–299 (2012).
13. Plaschkes, I., Silverman, F. W. & Priel, E. DNA topoisomerase I in the mouse central nervous system: age and sex dependence. *J. Comp. Neurol.* **493,** 357–369 (2005).
14. Lyu, Y. L. *et al.* Role of topoisomerase IIβ in the expression of developmentally regulated genes. *Mol. Cell. Biol.* **26,** 7929–7941 (2006).
15. Vos, S. M., Tretter, E. M., Schmidt, B. H. & Berger, J. M. All tangled up: how cells direct, manage and exploit topoisomerase function. *Nature Rev. Mol. Cell Biol.* **12,** 827–841 (2011).
16. Baranello, L., Levens, D., Gupta, A. & Kouzine, F. The importance of being supercoiled: how DNA mechanics regulate dynamic processes. *Biochim. Biophys. Acta* **1819,** 632–638 (2012).
17. Capranico, G., Marinello, J. & Baranello, L. Dissecting the transcriptional functions of human DNA topoisomerase I by selective inhibitors: implications for physiological and therapeutic modulation of enzyme activity. *Biochim. Biophys. Acta.* **1806,** 240–250 (2010).
18. Pommier, Y. Drugging topoisomerases: lessons and challenges. *ACS Chem. Biol.* **8,** 82–95 (2013).
19. Babak, T. *et al.* Global survey of genomic imprinting by transcriptome sequencing. *Curr. Biol.* **18,** 1735–1741 (2008).
20. Meng, L., Person, R. E. & Beaudet, A. L. *Ube3a-ATS* is an atypical RNA polymerase II transcript that represses the paternal expression of *Ube3a*. *Hum. Mol. Genet.* **21,** 3001–3012 (2012).
21. Landers, M. *et al.* Regulation of the large (~1000 kb) imprinted murine *Ube3a* antisense transcript by alternative exons upstream of *Snurf/Snrpn*. *Nucleic Acids Res.* **32,** 3480–3492 (2004).
22. Ljungman, M. & Hanawalt, P. C. The anti-cancer drug camptothecin inhibits elongation but stimulates initiation of RNA polymerase II transcription. *Carcinogenesis* **17,** 31–36 (1996).
23. Collins, I., Weber, A. & Levens, D. Transcriptional consequences of topoisomerase inhibition. *Mol. Cell. Biol.* **21,** 8437–8451 (2001).
24. Liu, L. F. & Wang, J. C. Supercoiling of the DNA template during transcription. *Proc. Natl Acad. Sci. USA* **84,** 7024–7027 (1987).
25. Tiwari, V. K. *et al.* Target genes of topoisomerase IIβ regulate neuronal survival and are defined by their chromatin state. *Proc. Natl Acad. Sci. USA* **109,** E934–E943 (2012).
26. Troester, M. A., Hoadley, K. A., Parker, J. S. & Perou, C. M. Prediction of toxicant-specific gene expression signatures after chemotherapeutic treatment of breast cell lines. *Environ. Health Perspect.* **112,** 1607–1613 (2004).
27. Joshi, R. S., Piña, B. & Roca, J. Topoisomerase II is required for the production of long Pol II gene transcripts in yeast. *Nucleic Acids Res.* http://dx.doi.org/10.1093/nar/gks626 (2012).
28. Wu, H.-Y., Shyy, S., Wang, J. C. & Liu, L. F. Transcription generates positively and negatively supercoiled domains in the template. *Cell* **53,** 433–440 (1988).
29. Stewart, A. F., Herrera, R. E. & Nordheimt, A. Rapid induction of c-fos transcription reveals quantitative linkage of RNA polymerase II and DNA topoisomerase I enzyme activities. *Cell* **60,** 141–149 (1990).
30. Rahl, P. B. *et al.* c-Myc regulates transcriptional pause release. *Cell* **141,** 432–445 (2010).
31. Muse, G. W. *et al.* RNA polymerase is poised for activation across the genome. *Nature Genet.* **39,** 1507–1511 (2007).
32. Zeitlinger, J. *et al.* RNA polymerase stalling at developmental control genes in the *Drosophila melanogaster* embryo. *Nature Genet.* **39,** 1512–1516 (2007).
33. Sanders, S. J. *et al. De novo* mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485,** 237–241 (2012).
34. O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485,** 246–250 (2012).
35. Basu, S. N., Kollu, R. & Banerjee-Basu, S. AutDB: a gene reference resource for autism research. *Nucleic Acids Res.* **37,** D832–D836 (2009).
36. Hammer, E., Heilbronn, R. & Weger, S. The E3 ligase Topors induces the accumulation of polysumoylated forms of DNA topoisomerase I *in vitro* and *in vivo*. *FEBS Lett.* **581,** 5418–5424 (2007).
37. Zhao, X. *et al.* A unified genetic theory for sporadic and inherited autism. *Proc. Natl Acad. Sci. USA* **104,** 12831–12836 (2007).
38. Naughton, C. *et al.* Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nature Struct. Mol. Biol.* **20,** 387–395 (2013).
39. Sano, K., Miyaji-Yamaguchi, M., Tsutsui, K. M., Tsutsui, K. & Topoisomerase, I. I. Activates a subset of neuronal genes that are repressed in AT-rich genomic environment. *PLoS ONE* **3,** e4103 (2008).
40. Ben-David, E. & Shifman, S. Combined analysis of exome sequencing points toward a major role for transcription regulation during brain development in autism. *Mol. Psychiatry* http://dx.doi.org/10.1038/mp.2012.148 (2012).
41. Ronan, J. L., Wu, W. & Crabtree, G. R. From neural development to cognition: unexpected roles for chromatin. *Nature Rev. Genet.* **14,** 347–359 (2013).
42. Forsberg, E. C., Downs, K. M. & Bresnick, E. H. Direct interaction of NF-E2 with hypersensitive site 2 of the β-globin locus control region in living cells. *Blood* **96,** 334–339 (2000).

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

Acknowledgements We thank M. Vernon at the UNC Expression Profiling Core for assistance with microarray experiments and with data analysis, H. Kelkar for pilot bioinformatics support, P. Mieczkowski and A. Brandt at the UNC High Throughput Sequencing Facility for advice and assistance with Illumina library preparation and sequencing and T. Kafri and P. Zhang at the UNC Lenti-shRNA Core for assistance with preparation of lentiviral vectors. This work was supported by grants to M.J.Z. and B.D.P. from The Angelman Syndrome Foundation, The Simons Foundation (SFARI 10-3625) and The National Institute of Mental Health (R01MH093372). I.F.K. and A.M.M. were supported by Joseph E. Wagstaff Postdoctoral Research Fellowships from the Angelman Syndrome Foundation. J.M.C. was supported by a grant from the American Cancer Society (117571-PF-09-124-01-DDC). J.M.C., J.S. and T.M. were supported by a grant from the National Institutes of Health (NIH) (R01GM101974). B.L.P. was supported by a NIH postdoctoral training grant (T32HD040127). S.J.C. was supported by a grant from National Institute of Child Health and Human Development (NICHD) (R01HD068730). The expression profiling and bioinformatics cores were funded by grants from the National Institute of Neurological Disorders and Stroke (P30NS045892) and NICHD (P30HD03110).

Author Contributions I.F.K., H.-S.H., A.M.M., J.S.H., S.J.C., B.D.P. and M.J.Z. conceived and designed experiments. I.F.K. performed RNA-seq and ChIP-seq experiments with mouse neurons. I.F.K., C.N.Y., J.M.C., J.S. and J.S.P. analysed data from genome-wide experiments and from published datasets. J.M.C. performed SNP filtering of RNA-seq data and J.S. performed statistical analysis of RNA-seq data. A.M.M. performed lentiviral shRNA knockdown experiments. J.S.H. and S.J.C. performed all experiments with iPSC-derived human neurons. B.L.P. assessed propensity of compounds to kill neurons and damage DNA. T.M. provided bioinformatics support. I.F.K., H.-S.H. and A.M.M. performed microarray experiments. I.F.K. and M.J.Z. wrote the manuscript.

Author Information Data from Microarray, RNA-seq and ChIP-seq experiments have been deposited in the Gene Expression Omnibus under accession number (GSE43900). 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 B.D.P. (bphilpot@med.unc.edu) or M.J.Z. (zylka@med.unc.edu).

METHODS

Mouse cell culture. Cortical neurons were cultured from E13.5–E15.5 mouse embryos as described[7]. For RNA-seq and ChIP-seq, neurons were seeded on 10-cm diameter poly-D-lysine-treated culture plates at a density of $5–10 \times 10^6$ per plate. Microarray experiments used 6-well plates seeded at 1×10^6 cells per well. After 7 days in culture, drugs or an equivalent amount of vehicle were added and left in the culture medium for 24 or 72 h. For topotecan, irinotecan, DRB and paraquat, vehicle was 0.1% dimethylsulphoxide (DMSO). For ICRF-193, vehicle was 0.02% DMSO.

Lentiviral shRNA experiments used viruses from The RNAi Consortium (acquired from Sigma-Aldrich and from the UNC Lenti-shRNA core): *Top1* knockdown used clone TRCN0000011884 (5′-CCGGCCAGCGAAGATTCT ATCTTATCTCGAGATAAGATAGAATCTTCGCTGGTTTTT-3′) and *Top2b* knockdown used clone TRCN0000070988 (5′-CCGGCCTTGTGTTGTCCTTT GTCTTCTCGAGAAGACAAAGGACAACACAAGGTTTTTG-3′). Virus expressing non-targeting hairpin RNA (SHC002, Sigma; 5′-CCGGCGTGATCTTCACCG ACAAGATCTCGAGATCTTGTCGGTGAAGATCACGTTTTT-3′) was used as a control. Neurons were seeded on 24-well plates at 2×10^5 per well. After 3 days in culture, cells were treated with lentivirus at a multiplicity of infection of at least 1. Virus was removed after 24 h, and RNA collected after a further 6 days in culture. Western blotting to assess knockdown was performed with anti-TOP1 (Santa Cruz Biotechnology, H5) or anti-TOP2B (Santa Cruz, H286) antibodies, with signal normalized to β-actin (Millipore, C4).

Cell death was assayed using Sytox Green (Invitrogen Molecular Probes). γ-H2AX foci were measured by immunohistochemistry. Primary antibodies were anti-γ-H2AX (Millipore, 1:500 dilution) and anti-NeuN (Millipore, 1:500), used to mark neurons.

iPSC culture and neuronal differentiation. Human iPSC work was approved by the University of Connecticut Stem Cell Research Oversight Committee. iPSCs that carry a large deletion of maternal 15q11–q13 (AGdel1-0; see also ref. 43), this cell line was deemed exempt from IRB approval at the University of Connecticut due to its establishment in 1995 and lack of identifying information) were cultured on irradiated mouse embryonic fibroblasts and manually passaged as described[43]. iPSCs were differentiated into forebrain cortical neurons as described[43] with the following modifications: neural progenitors were generated by culturing iPSCs on feeders in N2B27 medium supplemented with noggin (500 ng ml^{-1}) for 8 days and then manually picking neural rosettes for two additional passages using trypsin and standard cell culture protocols. Topotecan was applied to mature neurons and RNA was collected by standard protocols 6 days after the addition of drug or vehicle.

qPCR. qRT–PCR was carried out as described[43] using Taqman (Life Technologies) gene expression assays for *UBE3A* (Hs00166580_m1) and *UBE3A-ATS* (Hs03454279_m1) according to the manufacturer's instructions. The Taqman assay for GAPDH was used as a control.

RNA-seq. Total RNA was collected using Trizol reagent (Invitrogen). Mouse polyA-selected messenger RNA libraries were then prepared using the Illumina True-Seq kit for RNA. For RNA-seq on human-iPSC-derived neuronal samples, stranded multiplexed mRNA libraries were prepared using Illumina kits. Cluster generation and sequencing were performed using the Illumina HiSeq 2000 platform. For allele-specific expression analysis, equal amounts of total RNA from 3–6 biological replicates were pooled before polyA mRNA purification and library preparation.

For non-allelic expression analysis, data from five biological replicates were included. mRNA was isolated and libraries prepared independently for each replicate sample. Library preparation incorporated barcoded adapters and all samples were sequenced in one lane, using 50-base-pair (bp) reads.

RNA-seq expression analysis. For allelic expression analysis, informative CAST/B6 SNPs were downloaded from http://www.sanger.ac.uk/resources/mouse/genomes/. CAST alleles were then substituted into their corresponding mouse reference genome (mm9) positions and sequence reads were aligned to the mm9 and the CAST version of mm9 using Bowtie, selecting for unique matches. Filtered read counts for autosomal genes were tested for allelic bias using Fisher's exact test against a background model derived from autosomes, and *P* values adjusted for multiple comparisons using the Benjamini–Hochberg procedure. Statistical analysis was performed using R. Genomic intervals were derived from UCSC known genes or created manually where annotation was absent, namely for *Ube3a-ATS*.

For non-allelic analysis, reads were aligned to mm9 using Bowtie. Read counts were obtained using DEGseq, and normalization and analysis of differential gene expression was performed using the R package, edgeR, using a negative binomial model.

ChIP-seq. ChIP-seq against RNA Pol II was performed as described previously[30]. Cultures ($n = 4$ per condition) totalling approximately 2×10^7 neurons were treated with vehicle or 300 nM topotecan as described above and pooled. Nuclear lysates were sheared to an average fragment size of approximately 200 bp. 2 μg anti-RNA Pol II N20 (Santa Cruz Biotechnology) was added, and the sample incubated at 4 °C for 16 h. ChIP was performed as described previously[30,42].

ChIP-seq libraries were prepared from immunoprecipitated samples and their corresponding inputs using the Illumina Tru-Seq kit for ChIP-seq. Ligation products were size-selected by purification on 2% PippinPrep gels (Sage Science). Samples from vehicle- and drug-treated cells and their inputs were sequenced using the HiSeq 2000 platform with single-end reads of 50 bp.

ChIP-seq analysis. Short read sequences were aligned to the mouse reference genome (mm9) with Bowtie. Duplicate reads were removed. The quality of the experiment and false discovery rate for enriched peaks was assessed using MACS 1.4.2. CoverageBED was used to obtain read counts covering the promoter region (from −30 to +300 bp, relative to TSS) and gene bodies (from +300 after TSS to 3,000 bp after the annotated TTS), and to count reads in intervals across genes. Read counts were normalized to the number of mapped unique reads per sample per base.

Affymetrix microarrays. For single-dose microarray experiments, cultured cortical neurons were treated with 300 nM or 1 μM topotecan, 10 μM irinotecan (Sigma), 3 μM ICRF-193 (Santa Cruz Biotechnology), 100 μM DRB (Sigma), 100 μM H_2O_2 (Fisher Scientific) or 10 μM paraquat (Sigma) for 24 or 72 h. For topotecan dose response, cells were treated with 3 nM, 30 nM, 150 nM, 300 nM, 500 nM and 1,000 nM topotecan or vehicle for 72 h. Total RNA was used for all Affymetrix microarray experiments. Comparative expression with topotecan and ICRF-193 was measured with Affymetrix mouse genome 430 2.0 arrays. All other microarray experiments used Affymetrix Mouse Gene 1.0 ST 24-array plates. Linear RMA background correction and normalization was used for all microarray data.

43. Chamberlain, S. J. *et al.* Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader–Willi syndromes. *Proc. Natl Acad. Sci. USA* **107,** 17668–17673 (2010).
44. Lamb, J. *et al.* The connectivity map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* **313,** 1929–1935 (2006).
45. Iorio, F. *et al.* Discovery of drug mode of action and drug repositioning from transcriptional responses. *Proc. Natl Acad. Sci. USA* **107,** 14621–14626 (2010).
46. Groschel, B. & Bushman, F. Cell cycle arrest in G2/M promotes early steps of infection by human immunodeficiency virus. *J. Virol.* **79,** 5695–5704 (2005).
47. Carson, J. P. *et al.* Pharmacogenomic identification of targets for adjuvant therapy with the topoisomerase poison camptothecin. *Cancer Res.* **64,** 2096–2104 (2004).
48. MacIsaac, J. L., Bogutz, A. B., Morrissy, A. S. & Lefebvre, L. Tissue-specific alternative polyadenylation at the imprinted gene *Mest* regulates allelic usage at *Copg2*. *Nucleic Acids Res.* **40,** 1523–1535 (2012).

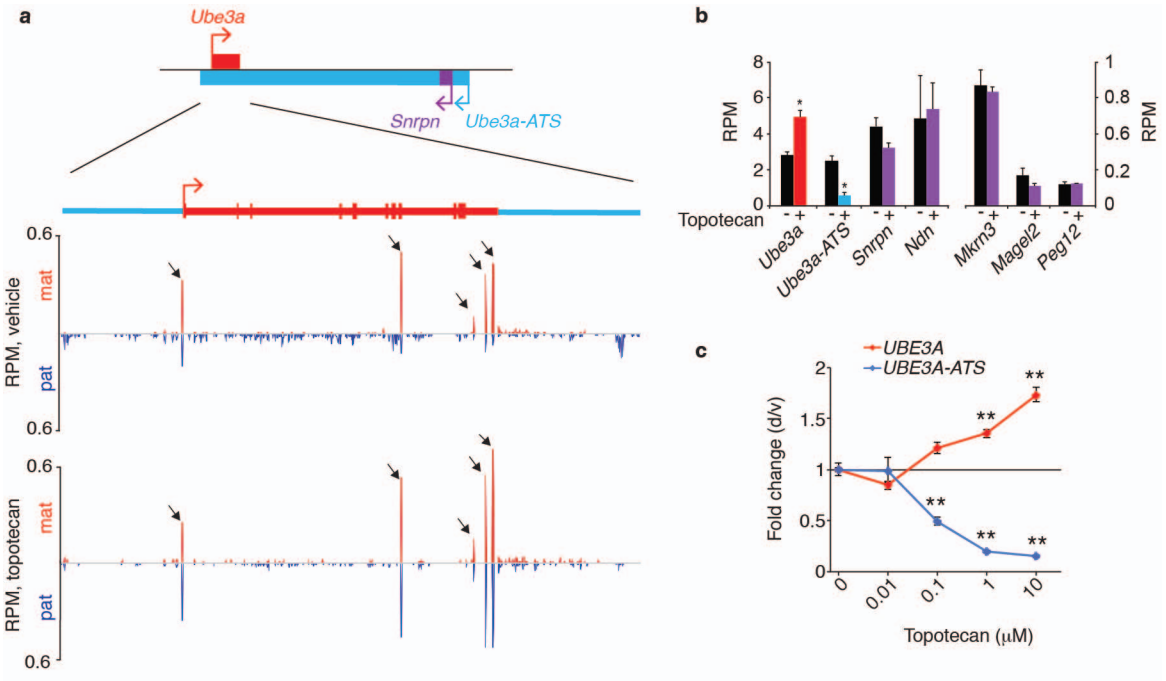
Extended Data Table 1 | Topotecan does not alter parent-of-origin bias of imprinted genes.

	Parent of origin bias (adjusted p value)		Change in bias, topotecan vs. vehicle (adjusted p value)	
	B6 x CAST	*CAST x B6*	*B6 x CAST*	*CAST x B6*
Maternal Bias				
*Ube3a**	$6.1x10^{-29}$	$1.2x10^{-20}$	***3.9x10^-7***	***3.3x10^-14***
Meg3	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Rian	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Kcnk9	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Mirg	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
H13	$3.0x10^{-56}$	$3.3x10^{-67}$	ns	ns
Eif2c2	$2.7x10^{-39}$	$6.3x10^{-28}$	ns	ns
H19	$1.3x10^{-29}$	$6.9x10^{-56}$	ns	ns
Rtl1	$1.2x10^{-14}$	$5.3x10^{-30}$	ns	ns
Cdkn1c	0.00035	$9.1x10^{-5}$	ns	ns
Igf2r	0.0056	0.00019	ns	ns
Paternal Bias				
*Ndn**	$<1x10^{-100}$	$7.1x10^{-61}$	ns	ns
*AK045535**	$<1x10^{-100}$	$<1x10^{-100}$	ns	$7.4x10^{-9}$
*Snurf**	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
*Snrpn**	$2.3x10^{-98}$	$<1x10^{-100}$	ns	ns
*A330076H08Rik**	$2.2x10^{-88}$	$<1x10^{-100}$	ns	ns
*A230057D06Rik**	$6.9x10^{-77}$	$<1x10^{-100}$	ns	ns
*Ube3a-ATS**	$2.1x10^{-59}$	$1.2x10^{-66}$	ns	0.00074
*Dokist4**	$1.3x10^{-46}$	$3.4x10^{-80}$	ns	ns
*Mkrn3**	$6.4x10^{-18}$	$1.3x10^{-19}$	ns	ns
*AK086712**	$4.2x10^{-13}$	$1.9x10^{-40}$	ns	ns
*U80893**	$1.7x10^{-9}$	$6.9x10^{-10}$	ns	ns
*Magel2**	$1.0x10^{-6}$	0.00065	ns	ns
*Peg12**	0.00032	0.0040	ns	ns
*AK039108**	0.0011	$1.0x10^{-10}$	ns	ns
Peg13	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Peg3	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Peg3as	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Plagl1	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Impact	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Rasgrf1	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Zim3	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Inpp5f	$<1x10^{-100}$	$1.3x10^{-58}$	ns	ns
Usp29	$4.4x10^{-99}$	$<1x10^{-100}$	ns	ns
Sgce	$4.2x10^{-95}$	$<1x10^{-100}$	ns	ns
Mest	$2.5x10^{-78}$	$1.2x10^{-36}$	ns	ns
Peg10	$6.2x10^{-77}$	$3.0x10^{-99}$	ns	ns
Zdbf2	$9.1x10^{-52}$	$1.8x10^{-43}$	ns	ns
Zrsr1	$3.2x10^{-48}$	$1.7x10^{-42}$	ns	ns
Nap1l5	$4.1x10^{-44}$	$9.2x10^{-28}$	ns	ns
Kcnq1ot1	$1.2x10^{-38}$	$1.0x10^{-73}$	ns	ns
Kcnq1	$2.1x10^{-36}$	$1.3x10^{-74}$	ns	ns
Adam23	$1.8x10^{-13}$	0.00010	ns	ns
Grb10	$4.0x10^{-13}$	$4.3x10^{-27}$	ns	ns
Dlk1	$6.8x10^{-5}$	$4.3x10^{-6}$	ns	ns
Copg2†	$<1x10^{-100}$	$<1x10^{-100}$	ns	0.014
AF217545†	$<1x10^{-100}$	$<1x10^{-100}$	ns	ns
Cdh15	$2.6x10^{-9}$	$3.7x10^{-12}$	ns	ns
Bcl2l1	0.028	0.0021	ns	ns

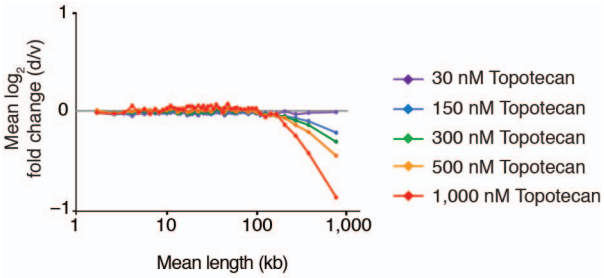
*Genes in the *Ube3a* imprinted cluster

†*AF217545* is contained within *Copg2*, which has also been reported to be maternally expressed[48]

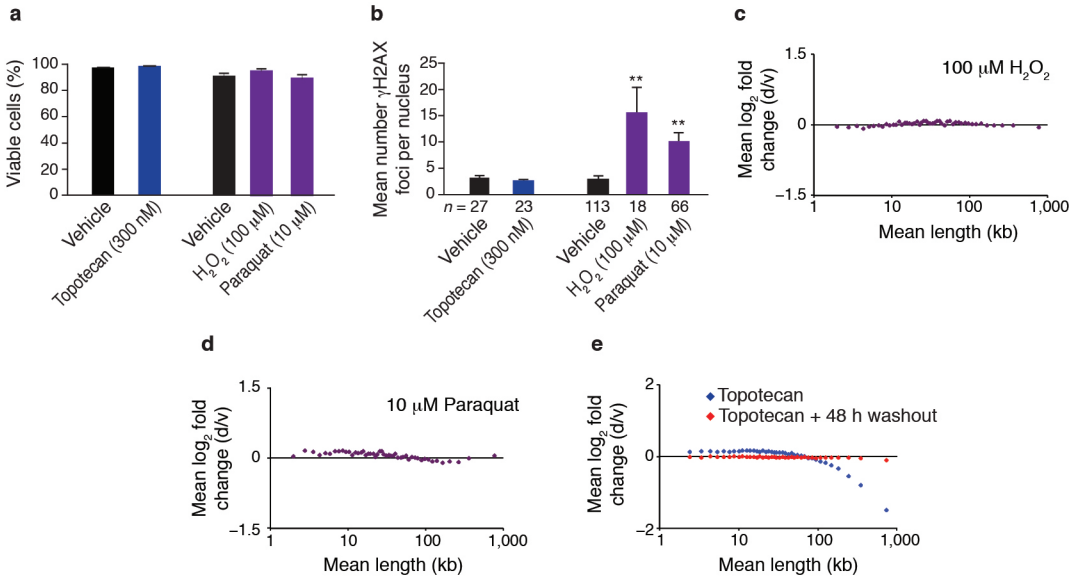
Genes with significant parent-of-origin expression bias in cultured cortical neurons. Shown are *P* values for allelic bias in baseline expression, and *P* values for change in expression bias with topotecan treatment (Fisher's exact test, adjusted for multiple comparisons). NS = $P > 0.05$. Embryos from CAST male × B6 female crosses ($n = 3$) were pooled and B6 male × CAST female crosses ($n = 6$) were pooled for sequencing.



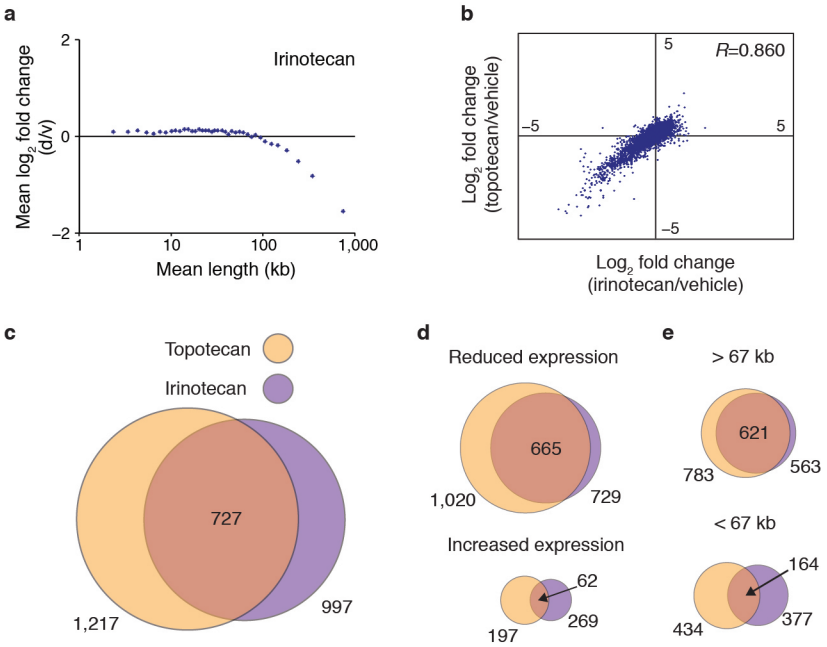
Extended Data Figure 1 | Topotecan affects allelic expression of *Ube3a* and *Ube3a-ATS* but not expression of nearby genes. **a**, Parent-of-origin-specific RNA-seq reads for *Ube3a* and *Ube3a-ATS* in vehicle- and topotecan-treated (300 nM for 3 days) mouse cortical neurons ($n = 5$ biological replicates). Shown are SNP-filtered read counts per million mapped reads (RPM) from the maternally (mat) and paternally (pat) inherited chromosomes. Reads from exons and 3′-untranslated region of *Ube3a* are indicated by arrows. **b**, Expression of *Ube3a* and imprinted genes near *Ube3a* in mouse cortical neurons ± 300 nM topotecan for 3 days. $^{*}P < 0.05$, Student's *t*-test. Error bars represent s.e.m. **c**, Expression of *UBE3A* and *UBE3A-ATS* in iPSC-derived neurons from an Angelman syndrome patient carrying a maternal deletion of the 15q11–q13 region. Differentiated neuronal cultures were treated with 10 nM–10 μM topotecan or vehicle for 6 days. Expression quantified by qPCR. $^{**}P < 0.01$, one-way analysis of variance with Dunnet's post-hoc test. $n = 4$. Error bars represent s.e.m.



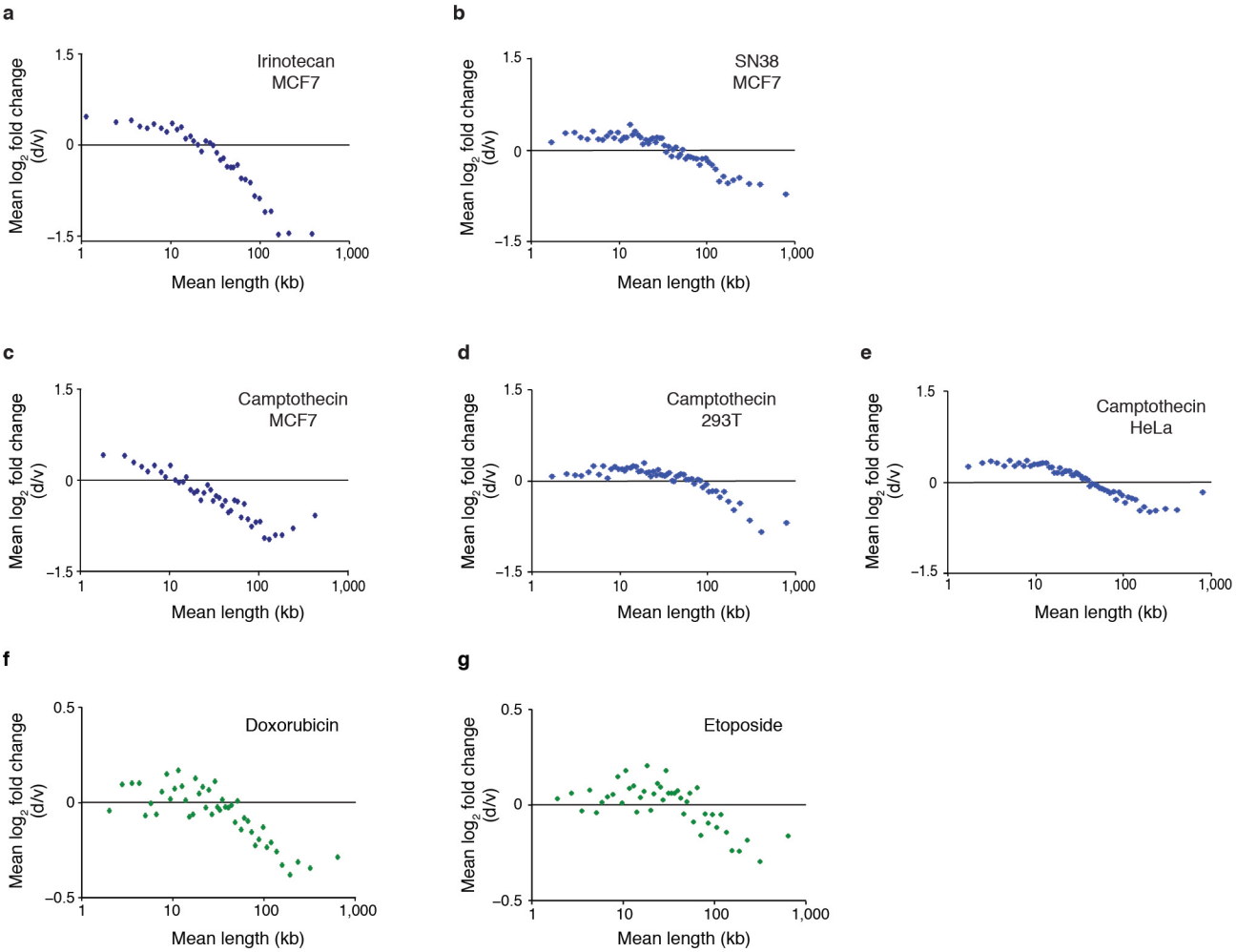
Extended Data Figure 2 | Topotecan dose–response. Mouse cortical neurons were treated with 3 nM, 30 nM, 150 nM, 300 nM and 1,000 nM topotecan for 3 days ($n = 3$ for 300 nM topotecan, all other doses $n = 1$). Gene expression was analysed by Affymetrix microarrays, plotted as mean expression change in bins of 200 genes.



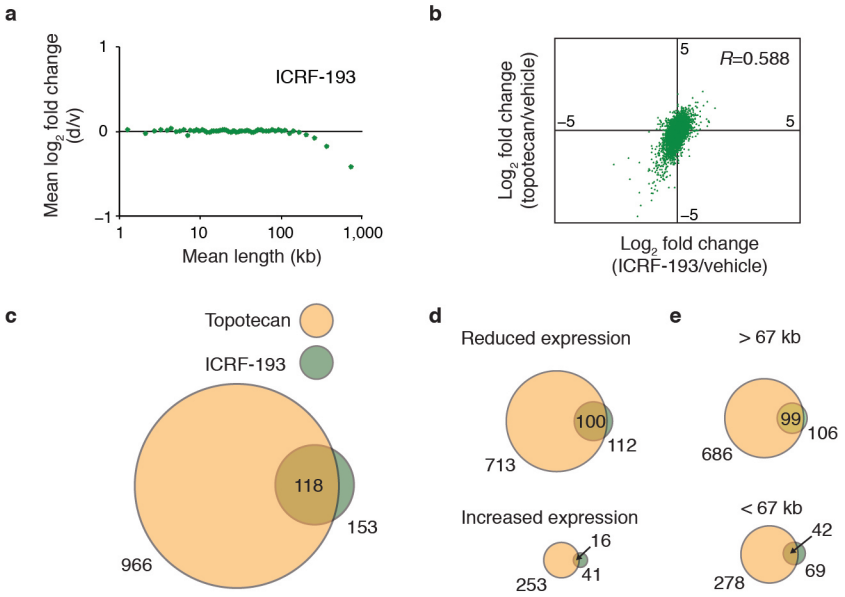
Extended Data Figure 3 | DNA damage does not inhibit gene expression in a length-dependent manner. **a**, Cultured mouse cortical neuron viability, assayed by Sytox Green staining after 72 h treatment with 300 nM topotecan, and after 24 h with 100 μM H_2O_2 or 10 μM paraquat, compared to vehicle-treated controls. Error bars represent s.e.m. $n = 4$. **b**, Average number of γ-H2AX foci per nucleus for cultured cortical neurons treated with 300 nM topotecan for 72 h, 100 μM H_2O_2 for 24 h and 10 μM paraquat for 24 h, compared to vehicle-treated controls. $^{**}P < 0.01$, Student's *t*-test. Number of cells counted is indicated for each sample. **c**, Gene expression compared to vehicle controls in bins of 200 genes by length, for cultured cortical neurons treated with 100 μM H_2O_2 for 24 h. **d**, Gene expression compared to vehicle controls in bins of 200 genes by length, for cultured cortical neurons treated with 10 μM paraquat for 24 h. **e**, Gene expression in cultured cortical neurons treated with 300 nM topotecan for 24 h, or treated for 24 h followed by 48 h without drug (washout). Average change in expression for bins of 200 genes by length.



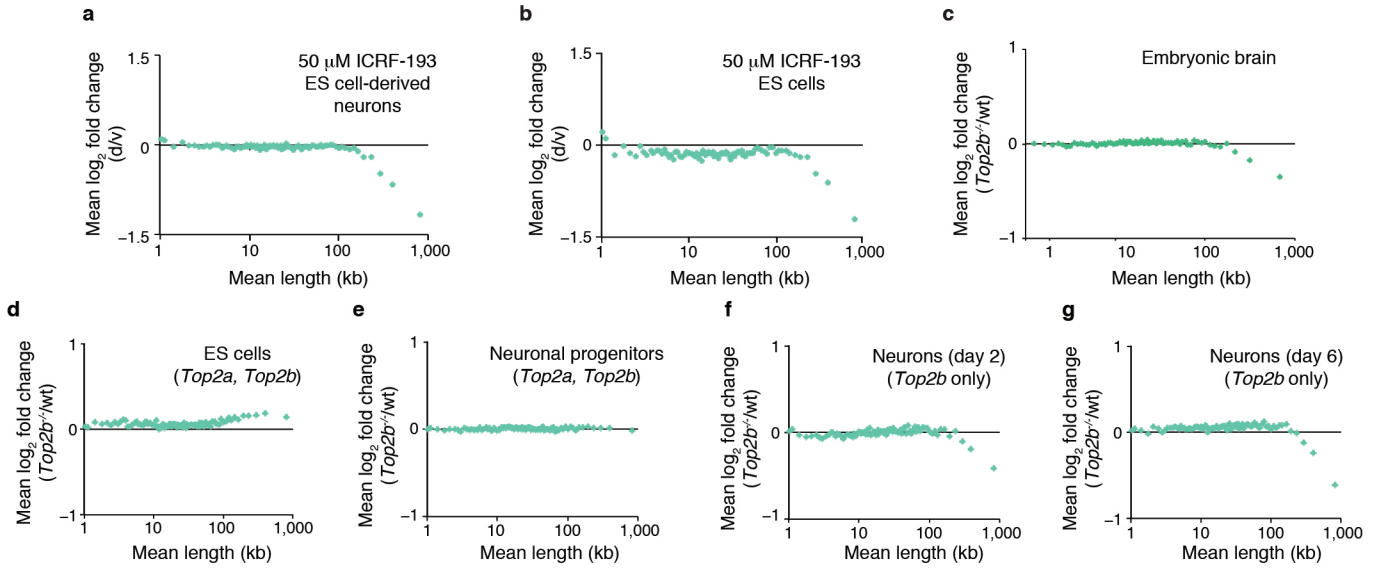
Extended Data Figure 4 | Topotecan and irinotecan have highly similar effects on gene expression. **a**, Affymetrix microarray analysis of gene expression in cultured mouse cortical neurons treated with vehicle or 10 μM irinotecan ($n = 3$ biological replicates), an inhibitor of TOP1, for 3 days. Mean expression fold change in bins of 200 genes, plotted by average gene length. **b**, Scatterplot of fold change with 1 μM topotecan ($n = 6$ biological replicates) versus fold change with 10 μM irinotecan for all expressed genes, measured by Affymetrix microarray. Pearson's $R = 0.860$. **c**, Overlap between genes showing positive or negative fold change of $\log_2 = 0.5$ or greater with topotecan and irinotecan treatment. **d**, Overlap between genes reduced or increased in expression. **e**, Overlap between differentially expressed genes that are greater or less than 67 kb.



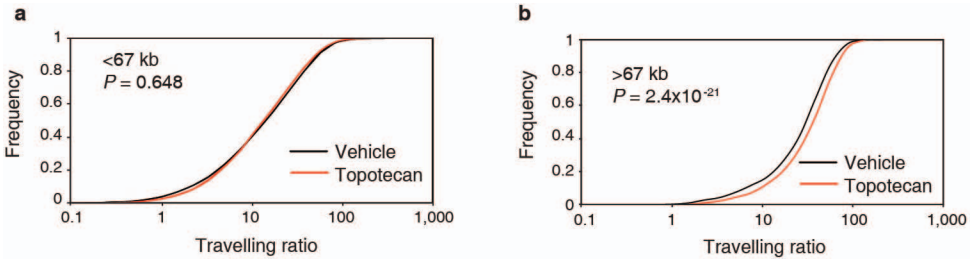
Extended Data Figure 5 | TOP1 and TOP2 inhibitors reduce expression of long genes in human cell lines. Re-analysis of microarray gene expression data sets from other laboratories. All plots are mean fold change in expression compared to vehicle controls in bins of 200 genes, plotted by average gene length. **a**, MCF7 cells treated with 10 μM irinotecan for 24 h, from the CMAP2 project[44]. **b**, MCF7 cells treated with 165 nM SN38, the active metabolite of irinotecan, for 6 h[45]. **c**–**e**, Gene expression in three human cell lines treated with camptothecin. **c**, MCF7 cells treated for 24 h with 10 μM camptothecin, from CMAP2. **d**, 293T cells treated with 2 μM camptothecin for 48 h[46]. **e**, HeLa cells treated with 10 μM camptothecin for 8 h[47]. **f, g**, Re-analysis of microarray data from ref. 26, comparing gene expression in TOP2 inhibitor- and vehicle-treated ME16C cells. **f**, ME16C cells treated with 0.5 μM doxorubicin for 36 h. **g**, ME16C cells treated with 50 μM etoposide for 36 h.



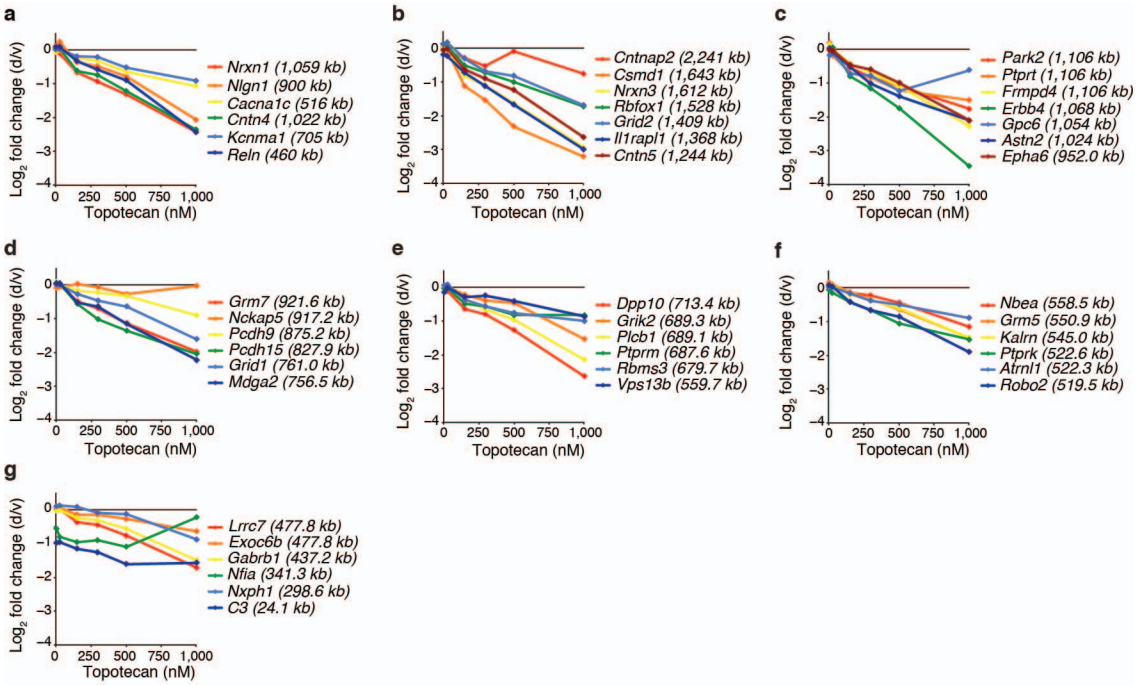
Extended Data Figure 6 | Topotecan and the TOP2 inhibitor ICRF-193 have similar effects on gene expression. **a**, Affymetrix microarray analysis of gene expression in cultured mouse cortical neurons treated with vehicle or 3 μM ICRF-193 ($n = 3$ biological replicates), an inhibitor of TOP2 enzymes, for 3 days. Mean expression fold change in bins of 200 genes, plotted by average gene length. **b**, Scatterplot of fold change with 300 nM topotecan ($n = 3$ biological replicates) versus fold change with 3 μM ICRF-193 for all expressed genes, measured by Affymetrix microarray. Pearson's $R = 0.588$. **c**, Overlap between genes showing positive or negative fold change of $\log_2 = 0.5$ or greater with topotecan and ICRF-193 treatment. **d**, Overlap between genes reduced or increased in expression. **e**, Overlap between differentially expressed genes that are greater or less than 67 kb in length.



Extended Data Figure 7 | Pharmacological inhibition of TOP2 or genetic deletion of *Top2b* reduces expression of long genes in ES-cell-derived neurons. **a**, **b**, Re-analysis of microarray expression data from ref. 25. Mean fold change in expression in bins of 200 genes, plotted by average gene length. **a**, Gene expression in ES-cell-derived neurons treated with vehicle or 50 μM ICRF-193 for 3 days. **b**, Gene expression in ES cells treated with vehicle or 50 μM ICRF-193 for 3 days. **c**, Re-analysis of microarray expression data from ref. 14, comparing gene expression in whole brain from wild-type (WT) and *Top2b*$^{-/-}$ embryonic mice. Expression data from three developmental time points (E16.5, E17.5, E18.5) were averaged for each gene then plotted as mean fold change in expression between wild-type and *Top2b*$^{-/-}$ mice (in bins of 100 genes, by average gene length). **d**–**g**, Re-analysis of microarray expression data from ref. 25. Mean fold change in gene expression between wild-type and *Top2b*$^{-/-}$ cells, for bins of 200 genes, plotted by average gene length. Expression data from ES cells (**d**), neuronal progenitors (**e**), ES-cell-derived neurons, 2 days after plating of neuronal progenitors (**f**) and ES-cell-derived neurons, 6 days after plating of neuronal progenitors (**g**). ES cells and neuronal progenitors express *Top2a* and *Top2b* (indicated within parentheses) and do not show reduced expression of long genes when *Top2b* is knocked out, suggesting that *Top2a* and *Top2b* redundantly regulate the expression of long genes.



Extended Data Figure 8 | Topotecan increases travelling ratio in genes larger than 67 kb. **a**, **b**, Travelling ratio was calculated for all genes bound by Pol II in vehicle- and topotecan-treated (300 nM for 3 days) cultured cortical neurons. 92.1% of all genes had travelling ratio values greater than 2, consistent with a previous report using mouse ES cells[30]. Frequency of travelling ratios for genes less than 67 kb in length (**a**) and for genes greater than 67 kb in length (**b**). Mean travelling ratio of vehicle- and topotecan-treated samples was significantly different for genes greater than 67 kb ($P = 2.4 \times 10^{-21}$, Student's *t*-test) but not for genes less than 67 kb ($P = 0.648$, Student's *t*-test).



Extended Data Figure 9 | Topotecan dose-dependently reduces expression of ASD candidate genes in cortical neurons. **a**, Topotecan dose–response for high-confidence autism candidate genes ($n = 3$ for 300 nM topotecan, all other doses $n = 1$). **b**–**g**, Topotecan dose–response for other ASD candidate genes organized by length from longest to shortest. Mouse cortical neurons were treated with vehicle or the indicated doses of topotecan for 3 days. Gene expression was quantified using Affymetrix microarrays. Dose–responses are from all topotecan-downregulated ASD candidate genes that were identified by RNA-seq (Table 1) and that were present on the Affymetrix microarrays.

ARTICLE

doi:10.1038/nature12510

Key tissue targets responsible for anthrax-toxin-induced lethality

Shihui Liu[1], Yi Zhang[1], Mahtab Moayeri[1], Jie Liu[2], Devorah Crown[1], Rasem J. Fattah[1], Alexander N. Wein[1], Zu-Xi Yu[3], Toren Finkel[2] & Stephen H. Leppla[1]

***Bacillus anthracis*, the causative agent of anthrax disease, is lethal owing to the actions of two exotoxins: anthrax lethal toxin (LT) and oedema toxin (ET). The key tissue targets responsible for the lethal effects of these toxins are unknown. Here we generated cell-type-specific anthrax toxin receptor capillary morphogenesis protein-2 (CMG2)-null mice and cell-type-specific CMG2-expressing mice and challenged them with the toxins. Our results show that lethality induced by LT and ET occurs through damage to distinct cell types; whereas targeting cardiomyocytes and vascular smooth muscle cells is required for LT-induced mortality, ET-induced lethality occurs mainly through its action in hepatocytes. Notably, and in contradiction to what has been previously postulated, targeting of endothelial cells by either toxin does not seem to contribute significantly to lethality. Our findings demonstrate that *B. anthracis* has evolved to use LT and ET to induce host lethality by coordinately damaging two distinct vital systems.**

Bacillus anthracis, the causative agent of anthrax, causes disease by growing to high numbers in the blood and secreting the anthrax exotoxins, consisting of three components: protective antigen (PA), lethal factor (LF) and oedema factor (EF)[1]. PA is the receptor-binding moiety that binds to either tumour endothelium marker-8 (TEM8, also called anthrax toxin receptor 1 (ANTXR1)) or CMG2 (also called anthrax toxin receptor 2 (ANTXR2)) on target cells[2–4]. LF and EF then bind to receptor-associated PA and are transported to the cytosol. The *in vivo* toxic effects of LT and ET are principally mediated through PA binding to CMG2 (ref. 4). EF, which with PA forms ET, is a calmodulin-dependent adenylate cyclase that elevates intracellular cAMP levels and has been shown to cause skin oedema and lethality in experimental animals[5,6]. LF, which forms LT with PA, is lethal to animals. LF is a Zn^{+2}-dependent metalloproteinase that cleaves and inactivates the mitogen-activated protein kinase kinases[7–9] and inflammasome sensor NLRP1 (refs 10, 11).

The toxins have essential roles in anthrax pathogenesis[12]. At the early stages of infection, LT and ET coordinately impair the immune system to establish infection[13,14]. At later stages, the toxins accumulate to high levels and cause death through mechanisms that are still not fully understood, despite studies in challenge models ranging from zebrafish to non-human primates[14–18]. The consequences of targeting specific tissues or cell types cannot be accurately assessed in systemic challenge models. In this study we generated various cell-type-specific CMG2-null mice as well as the corresponding cell-type-specific CMG2-expressing mice, and identified the key tissue targets of LT and ET.

LT targeting of endothelial cells is not lethal to mice

Anthrax toxins induce a vascular shock state in animal models[15], and this has been proposed to be due to their effects on the endothelium[14,17–21]. To assess the role of endothelial cells, we generated endothelial-cell-specific CMG2-null mice ($CMG2^{fl/fl}$/*Cdh-cre*, hereafter referred to as $CMG2(EC)^{-/-}$) (Table 1 and Extended Data Fig. 1a, b) in which the endothelial-cell-specific deletion of *CMG2* was verified. Endothelial cells isolated from the lungs of $CMG2(EC)^{-/-}$ mice were 300-fold more resistant than endothelial cells from control $CMG2^{fl/fl}$ mice to FP59 (refs 22, 23), an LF fusion toxin that kills cells in a PA-dependent manner (Extended Data Fig. 1c). $CMG2(EC)^{-/-}$ endothelial cells also became resistant to proliferation arrest induced by LT (Extended Data Fig. 1d). Interestingly, $CMG2(EC)^{-/-}$ endothelial cells were completely resistant to a PA mutant (PA(L687A)) that binds preferentially to CMG2 over TEM8 (Extended Data Fig. 1e, f), further confirming the complete deletion of CMG2 in $CMG2(EC)^{-/-}$ endothelial cells. In contrast, cells other than endothelial cells (non-endothelial cells) from both $CMG2(EC)^{-/-}$ and control $CMG2^{fl/fl}$ mice were similarly sensitive to both PA variants (Extended Data Fig. 1c, f). Thus, CMG2, the major anthrax toxin receptor in endothelial cells, was specifically and completely deleted from endothelial cells in $CMG2(EC)^{-/-}$ mice.

To determine the role of targeting endothelial cells in LT pathogenesis *in vivo*, we challenged the $CMG2(EC)^{-/-}$ mice with 100 μg LT. Notably, the $CMG2(EC)^{-/-}$ mice displayed similar sensitivity compared with their $CMG2(EC)^{+/-}$ and $CMG2^{+/+}$ littermates in a manner independent of challenge route, whereas whole-body $CMG2^{-/-}$ mice were completely resistant (Fig. 1a and Extended Data Fig. 1g, h). Therefore, LT targeting of endothelial cells seems not to be required for lethality.

To examine the possibility that damaging endothelial cells alone would induce symptoms or lethality, we also generated *CMG2* transgenic mice in which CMG2 was expressed only in endothelial cells. The *CMG2* transgenic vector contains a *loxP-stop-loxP* (LSL) cassette[24] with an enhanced green fluorescent protein (eGFP) coding sequence located between the ubiquitous CAG promoter and the *CMG2* transgene (Extended Data Fig. 2a). Thus, the resulting transgenic mice (*LSL-CMG2*) exhibited whole-body green fluorescence and did not express the transgene until bred with tissue-specific *cre*-transgenic mice, allowing loss of the LSL cassette and fluorescence in that particular tissue. Endothelial-cell-specific CMG2-expressing mice ($CMG2^{EC}$, Table 1) were obtained by breeding *LSL-CMG2* mice with *Cdh-cre* mice, and by subsequent breeding with whole-body $CMG2^{-/-}$ mice to eliminate expression of the endogenous CMG2. We confirmed the endothelial-cell-restricted expression of the *CMG2* transgene by demonstrating the

[1]Microbial Pathogenesis Section, Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA. [2]Center for Molecular Medicine, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, USA. [3]Pathology Core Facility, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

Table 1 | Nomenclature of gene-targeted and transgenic mice

Mouse description	Genotype	Nomenclature used
Whole-body CMG2-null	*CMG2*$^{-/-}$	*CMG2*$^{-/-}$
CMG2 floxed	*CMG2*$^{fl/fl}$	*CMG2*$^{fl/fl}$
EC-specific CMG2-null	*CMG2*$^{fl/fl}$/*Cdh-cre*	*CMG2(EC)*$^{-/-}$
CM-specific CMG2-null	*CMG2*$^{fl/fl}$/*Myh6-cre*	*CMG2(CM)*$^{-/-}$
SM/CM-specific CMG2-null	*CMG2*$^{fl/fl}$/*SM22-cre*	*CMG2(SM/CM)*$^{-/-}$
Hep-specific CMG2-null	*CMG2*$^{fl/fl}$/*Alb-cre*	*CMG2(Hep)*$^{-/-}$
EC/SM/CM-specific CMG2-null	*CMG2*$^{fl/fl}$/*SM22-cre*/*Cdh-cre*	*CMG2(SM/CM/EC)*$^{-/-}$
EC/SM/CM/IE-specific CMG2-null	*CMG2*$^{fl/fl}$/*SM22-cre*/*Cdh-cre*/*Vil-cre*	*CMG2(SM/CM/EC/IE)*$^{-/-}$
CMG2 transgenic, inactive	P_{CAG}-*loxPstoploxP-CMG2*	*LSL-CMG2*
Inactive *CMG2* transgene in whole-body CMG2-null	P_{CAG}-*loxPstoploxP-CMG2*/*CMG2*$^{-/-}$	*LSL-CMG2/CMG2*$^{-/-}$
EC-specific CMG2-expressing	P_{CAG}-*loxPstoploxP-CMG2*/*Cdh-cre*/*CMG2*$^{-/-}$	*CMG2*^{EC}
CM-specific CMG2-expressing	P_{CAG}-*loxPstoploxP-CMG2*/*Myh6-Cre*/*CMG2*$^{-/-}$	*CMG2*^{CM}
SM/CM-specific CMG2-expressing	P_{CAG}-*loxPstoploxP-CMG2*/*SM22-cre*/*CMG2*$^{-/-}$	*CMG2*$^{SM/CM}$
Hep-specific CMG2-expressing	P_{CAG}-*loxPstoploxP-CMG2*/*Alb-cre*/*CMG2*$^{-/-}$	*CMG2*^{Hep}
IE-specific CMG2-expressing	P_{CAG}-*loxPstoploxP-CMG2*/*Vil-cre*/*CMG2*$^{-/-}$	*CMG2*^{IE}
EC-specific *cre* transgenic	*Cdh-cre*	*Cdh-cre*
CM-specific *cre* transgenic	*Myh6-cre*	*Myh6-cre*
SM/CM-specific *cre* transgenic	*SM22-cre*	*SM22-cre*
Hep-specific *cre* transgenic	*Alb-cre*	*Alb-cre*
IE-specific *cre* transgenic	*Vil-cre*	*Vil-cre*

CM, cardiomyocytes; EC, endothelial cells; Hep, hepatocytes; IE, intestine epithelial cells; loxPstoploxP (LSL), loxP-stop-loxP cassette; P_{CAG}, CAG promoter; SM, vascular smooth muscle cells.

regained sensitivity of endothelial cells from *CMG2*^{EC} mice to PA plus FP59 and PA(L687A) plus FP59 (Extended Data Fig. 2b, c). As expected, non-endothelial cells from *CMG2*^{EC} and whole-body *CMG2*$^{-/-}$ mice, which only expressed the minor receptor TEM8, had similar intermediate sensitivities to PA plus FP59 and were fully resistant to PA(L687A) plus FP59 (Extended Data Fig. 2b, c). These results demonstrate that the *CMG2* transgene is specifically expressed in endothelial cells of *CMG2*^{EC} mice.

We then challenged the *CMG2*^{EC} mice with LT. Notably, all the *CMG2*^{EC} mice survived two doses of 100 μg LT, whereas all their littermate *CMG2*$^{+/+}$ and *CMG2*$^{+/-}$ control mice succumbed to the challenges (Fig. 1b). The above results unequivocally demonstrate that LT targeting of endothelial cells is insufficient to cause lethality *in vivo*.

LT targets cardiomyocytes and smooth muscle cells

We next examined the effects on LT pathogenesis of targeting the other two major cell types of the cardiovascular system by generating cardiomyocyte- and vascular-smooth-muscle-cell-specific CMG2-null mice (Table 1). The cardiomyocyte-specific CMG2-null mice (*CMG2*$^{fl/fl}$/*Myh6-cre*; hereafter referred to as *CMG2(CM)*$^{-/-}$, see Table 1) had CMG2 deleted only in heart tissue (Extended Data Fig. 3a, left panel). The *CMG2*$^{fl/fl}$/*SM22-cre* mice showed that the deletion occurred both in the aorta (enriched with vascular smooth muscle cells) and the heart (Extended Data Fig. 3a, right panel), and therefore these mice are referred to as *CMG2(SM/CM)*$^{-/-}$ mice (Table 1), reflecting the fact that they are actually smooth-muscle/cardiomyocyte-specific CMG2-null mice. This is consistent with a previous study showing that the SM22α promoter is active in both smooth muscle cells and cardiomyocytes[25].

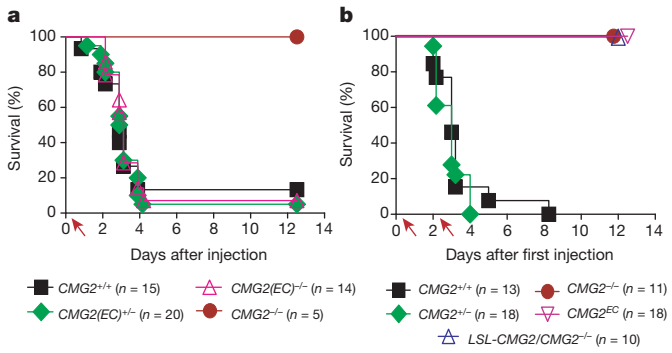


Figure 1 | LT targeting of endothelial cells is not lethal to mice. **a**, **b**, Susceptibility of *CMG2(EC)*$^{-/-}$ (**a**) and *CMG2*^{EC} (**b**) mice to LT. Mice were treated intraperitoneally with single (**a**) or two (**b**) doses (arrows) of 100 μg LT (100 μg PA + 100 μg LF), and monitored for survival.

We found that both *CMG2(CM)*$^{-/-}$ and *CMG2(SM/CM)*$^{-/-}$ mice were more resistant than their littermate controls to LT challenge, in that 52% of the *CMG2(CM)*$^{-/-}$ mice and nearly all of the *CMG2(SM/CM)*$^{-/-}$ mice survived (Fig. 2A, B and Extended Data Fig. 3b). Therefore, cardiomyocytes and smooth muscle cells seem to be major targets for LT-induced lethality.

To verify this finding further, we generated cardiomyocyte-specific (*CMG2*^{CM}, Table 1) or smooth-muscle- and cardiomyocyte-specific (*CMG2*$^{SM/CM}$, Table 1) CMG2-expressing mice using the strategy described for generation of *CMG2*^{EC} mice above (Extended Data Fig. 2a and Methods), where loss of GFP expression in cardiomyocytes or cardiomyocytes and smooth muscle cells, but not in other cell types, was verified (Fig. 2C, D and Extended Data Fig. 4). The activation of the *CMG2* transgene occurred selectively in vascular smooth muscle cells in aorta but not in the smooth muscle layer in small intestines and uterus of the *CMG2*$^{SM/CM}$ mice (Fig. 2D and Extended Data Fig. 4b). Notably, 100% of the *CMG2*^{CM} and *CMG2*$^{SM/CM}$ mice regained sensitivity and succumbed to two doses of LT (100 μg) whereas all the whole-body *CMG2*$^{-/-}$ and the *LSL-CMG2/CMG2*$^{-/-}$ mice survived without any sign of disease (Fig. 2E–H). Sensitivity of the *CMG2*$^{SM/CM}$ mice was significantly higher than that of the *CMG2*^{CM} mice in that 80% versus 30% of these mice, respectively, were killed by only one dose of LT ($P = 0.05$) (Fig. 2E–H). These results clearly demonstrate that cardiomyocytes as well as smooth muscle cells are key targets of LT.

Although nearly all mice with the CMG2 deletion in smooth muscle cells and cardiomyocytes survived one dose of LT challenge (Figs 2B and 3a), most of the mice were unable to survive two doses (Fig. 3b). This suggests that other target cells could contribute to lethality at higher doses of toxin. To evaluate whether the additional deletion of CMG2 in endothelial cells could augment the LT resistance of *CMG2(SM/CM)*$^{-/-}$ mice, we generated mice with the CMG2 deletion in cardiomyocytes, smooth muscle cells and endothelial cells (*CMG2*$^{fl/fl}$/*SM22-cre*/*Cdh-cre*, hereafter called *CMG2(SM/CM/EC)*$^{-/-}$, Table 1) and found that they were more resistant to LT than *CMG2(SM/CM)*$^{-/-}$ mice (Fig. 3a, b). These results demonstrate that the LT-induced lethality was largely due to targeting the three major cell types of the cardiovascular system through the CMG2 receptor, with LT targeting of endothelial cells contributing to the disease progression only at higher doses.

Serum levels of c-troponin-I, a cardiomyocyte-specific protein that is elevated during cardiac injury, were significantly increased in wild-type and *CMG2(EC)*$^{-/-}$ mice but remained low in both *CMG2(CM)*$^{-/-}$ and *CMG2(SM/CM)*$^{-/-}$ mice 48 h after LT challenge (Fig. 3c). Consistent with this, echocardiographic analyses showed that heart function was significantly compromised in LT-treated wild-type mice but not *CMG2(CM)*$^{-/-}$, *CMG2(SM/CM)*$^{-/-}$ and *CMG2*$^{-/-}$ mice (Fig. 3d).

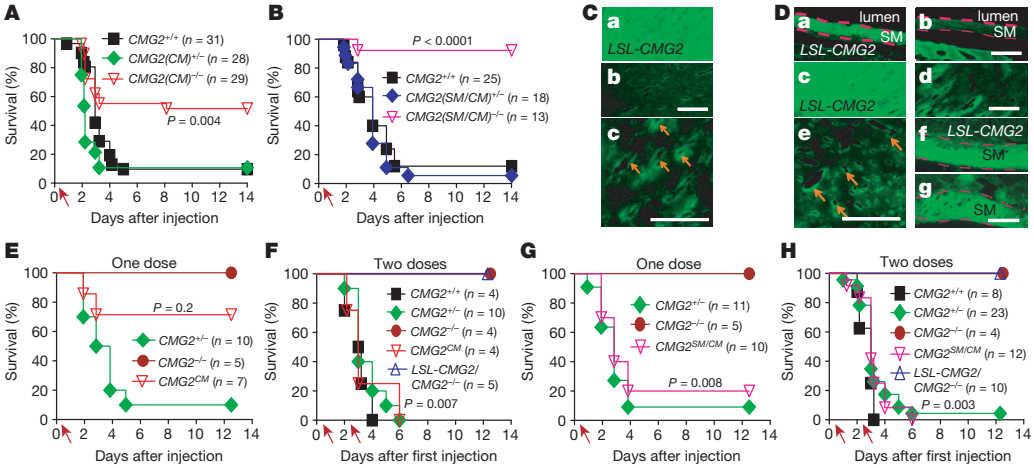


Figure 2 | Targeting of cardiomyocytes and smooth muscle cells by LT is sufficient for lethality. **A**, **B**, Increased resistance of cardiomyocyte- and smooth-muscle-cell-specific CMG2-null mice to LT. *CMG2(CM)*$^{-/-}$ (**A**), *CMG2(SM/CM)*$^{-/-}$ (**B**) and their littermate controls were challenged intraperitoneally with 100 μg LT. *CMG2(CM)*$^{-/-}$ versus *CMG2*$^{+/+}$ mice, $P = 0.004$; *CMG2(SM/CM)*$^{-/-}$ versus *CMG2*$^{+/+}$ mice, $P < 0.0001$. Log-rank test. **C**, Selective activation of the *CMG2* transgene in cardiomyocytes of *CMG2*^{CM} mice. Representative fluorescence microscopy of heart from *LSL-CMG2* mice (**a**) ($n = 2$) and *CMG2*^{CM} mice (**b**, **c**) ($n = 2$). Selective loss of GFP expression in cardiomyocytes (**b** and **c** compared to **a**) but not endothelial cells (arrows in **c**) is shown. Scale bar, 100 μm. **D**, Selective activation of *CMG2* transgene in smooth muscle cells and cardiomyocytes from *CMG2*$^{SM/CM}$ mice. Representative fluorescence microscopy of aorta (**b**), heart (**d** and **e**), small intestine (smooth muscle) (**g**) from *CMG2*$^{SM/CM}$ mice ($n = 3$), and aorta (**a**), heart (**c**), small intestine (**f**) from *LSL-CMG2* mice ($n = 2$). Selective loss of GFP expression in vascular smooth muscle cells (in aorta) and cardiomyocytes (**b**, **d**, **e** compared to **a** and **c**) but not smooth muscle cells in small intestine (**g** compared to **f**) and not endothelial cells in heart (arrows in **e**) is shown. Scale bar, 100 μm. **E**–**H**, Sensitivity of *CMG2*^{CM} mice (**E**, **F**) and *CMG2*$^{SM/CM}$ mice (**G**, **H**) to single (**E**, **G**) or two doses (**F**, **H**) of 100 μg LT (intraperitoneally). *CMG2*^{CM} versus *CMG2*$^{-/-}$ mice, $P = 0.2$ in **E**; $P = 0.007$ in **F**. *CMG2*$^{SM/CM}$ versus *CMG2*$^{-/-}$ mice, $P = 0.008$ in **G**; $P = 0.003$ in **H**. Log-rank test.

Histological analyses showed regions of cardiomyocyte degeneration in LT-treated wild-type but not *CMG2*$^{-/-}$ mice (Extended Data Fig. 5). Modest hepatocyte degeneration was also found in LT-treated wild-type but not *CMG2*$^{-/-}$ mice, perhaps reflecting previously reported LT-induced hypoxia-mediated liver damage (see below). No obvious abnormalities were found in the lungs, kidneys and spleens of the LT-treated mice (data not shown).

Resistance to *B. anthracis* infection

B. anthracis normally exists in spore form in the environment and germinates to toxin-producing vegetative bacteria at 37 °C in mammalian hosts. We next infected the *CMG2(SM/CM/EC)*$^{-/-}$ mice with toxigenic, unencapsulated *B. anthracis* Sterne strain spores. Notably, all the *CMG2(SM/CM/EC)*$^{-/-}$ mice survived the spore challenge whereas littermate heterozygous control mice (Fig. 3e) and *CMG2(EC)*$^{-/-}$ mice (Extended Data Fig. 6) were sensitive, indicating that toxin targeting endothelial cells alone has little role in this infection model. *CMG2(SM/CM/EC)*$^{-/-}$ mice were also more resistant than the littermate controls to infection with pre-germinated vegetative *B. anthracis* (75% versus 18% survival, $P < 0.006$) (Fig. 3f). Thus, CMG2 deficiency in the three major cell types of the cardiovascular system is sufficient to confer resistance to LT as well as to anthrax Sterne strain infection.

Liver is not a key target of LT

Early work in our laboratory showed hypoxia-mediated damage in the liver in response to LT[15]. In light of the above results which reveal that LT primarily induces lethality through targeting the cardiovascular system, any effects in the liver are probably secondary events. In support of this hypothesis, we found that hepatocyte-specific CMG2-null mice (*CMG2*$^{fl/fl}$*/Alb-cre*, hereafter called *CMG2(Hep)*$^{-/-}$, Table 1) remained sensitive to LT (Extended Data Fig. 7a), whereas hepatocyte-specific CMG2-expressing mice (*CMG2*^{Hep}, Table 1) were completely resistant to the toxin (Extended Data Fig. 7b, c). These results clearly demonstrate that liver is not a key target of LT.

ET causes intestinal fluid influx and liver oedema

Oedema toxin is another important anthrax toxin relevant to anthrax pathogenesis[1]. Subcutaneous injection of ET induces oedema in experimental animals. To determine whether ET also causes oedema in internal organs and to define its cell-type targets, we first measured the wet/dry ratios of various organs of mice injected with ET. ET caused remarkable fluid influx into intestinal lumen of wild-type as well as *CMG2(SM/CM/EC)*$^{-/-}$ mice, which did not occur in mice lacking CMG2 in the intestinal epithelial cells (*CMG2(IE)*$^{-/-}$ and *CMG2*$^{fl/fl}$*/SM22-cre/Cdh-cre/Vil-cre* mice, hereafter called *CMG2(SM/CM/EC/IE)*$^{-/-}$ mice, Table 1) (Fig. 4a). Additionally, ET induced much greater fluid accumulation in intestines in intestinal-epithelial-specific CMG2-expressing mice (*CMG2*^{IE}, Table 1) compared to *CMG2*$^{-/-}$ mice (Fig. 4b). Taken together, these results demonstrate that ET targeting of intestinal epithelial cells but not endothelial cells and smooth muscle cells is the cause of intestinal fluid accumulation.

Interestingly, ET also induced significant liver oedema in wild-type as well as *CMG2(IE)*$^{-/-}$, *CMG2(SM/CM/EC)*$^{-/-}$ and *CMG2(SM/CM/EC/IE)*$^{-/-}$ mice, but not in *CMG2(Hep)*$^{-/-}$ mice (Fig. 4c). Furthermore, this ET-induced liver oedema was not observed in *CMG2*$^{-/-}$ mice as well as *CMG2*^{EC}, *CMG2*^{IE}, or *CMG2*$^{SM/CM}$ transgenic mice, but occurred in the hepatocyte-specific CMG2-expressing *CMG2*^{Hep} mice at levels equal to and even exceeding those in *CMG2*$^{+/+}$ mice (Fig. 4d). Thus, ET induces liver oedema by acting directly on hepatocytes rather than endothelial cells and smooth muscle cells. Similarly, ET-induced skin oedema was also due to directly targeting cell types other than endothelial cells and smooth muscle cells (Extended Data Fig. 8a). We did not detect wet/dry ratio increases in the heart, spleen, kidney and lung of the ET-treated mice (Extended Data Fig. 8b–f), indicating that the oedema induced by ET was limited to certain tissues including skin and liver and not a general result of cAMP increases in all organs. Indeed, significant decreases in wet/dry ratios of lungs from ET-treated mice were detected (Extended Data Fig. 8f), probably reflecting the general dehydration status of the ET-treated mice resulting from fluid displacement to skin, liver and other tissues.

Liver is a key target of ET-induced lethality

To define the cell types that are responsible for ET-induced lethality, we next evaluated the sensitivities of the various cell-type-specific CMG2-null mice described above to ET. Surprisingly, mice with

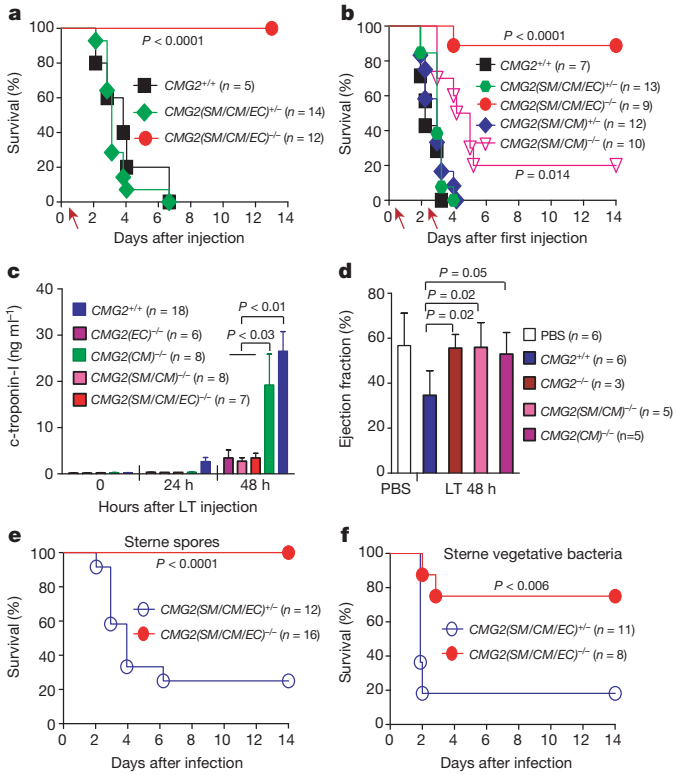


Figure 3 | Mice lacking CMG2 receptor in the three major cell types of the cardiovascular system are highly resistant to LT and *B. anthracis* infection. **a**, **b**, Resistance of $CMG2(SM/CM/EC)^{-/-}$ mice to challenge with one (**a**) or two doses (**b**) of 100 μg LT (intraperitoneally). In **a**, $CMG2(SM/CM/EC)^{-/-}$ versus $CMG2^{+/+}$ mice, $P < 0.0001$. In **b**, $CMG2(SM/CM/EC)^{-/-}$ versus $CMG2^{+/+}$ mice, $P < 0.0001$; $CMG2(SM/CM)^{-/-}$ versus $CMG2^{+/+}$ mice, $P = 0.014$. Log-rank test. **c**, Serum levels of cardiac biomarker c-troponin-I in mice treated intraperitoneally with 100 μg LT. Error bars indicate s.e. Two-tailed unpaired *t*-test. **d**, Echocardiography analyses of mice challenged intraperitoneally with 100 μg LT. Error bars indicate s.d. Two-tailed unpaired *t*-test. **e**, **f**, Resistance of the $CMG2(SM/CM/EC)^{-/-}$ mice to *B. anthracis* infection with 4×10^8 Sterne spores (**e**) or 2×10^6 vegetative bacteria (**f**). Log-rank test.

CMG2-specific deletion in cardiomyocytes or combined deletion in cardiomyocytes, vascular smooth muscle cells, endothelial cells, as well as intestinal epithelial cells ($CMG2(SM/CM/EC/IE)^{-/-}$) showed similar sensitivity to ET as their littermate controls (Extended Data Fig. 9). In contrast, mice with CMG2 deletion only in hepatocytes ($CMG2(Hep)^{-/-}$) displayed remarkable resistance to ET, with 82% surviving challenge (Fig. 5a). Notably, although all the ET-treated $CMG2(Hep)^{-/-}$ survivors did show initial signs of malaise, these mice recovered within 3 days of the challenge (Fig. 5a, right panel). We further tested the sensitivities of the various cell-type-specific CMG2-expressing mice to ET. Mice that expressed CMG2 only in endothelial cells, cardiomyocytes and vascular smooth muscle cells as well as the whole-body $CMG2^{-/-}$ mice remained resistant to ET and did not show significant signs of disease (Fig. 5b). In contrast, the $CMG2^{Hep}$ mice were fully sensitive to ET (Fig. 5b). Furthermore, the liver damage biomarkers aspartate aminotransferase and alanine aminotransferase were found to be significantly higher in ET-treated $CMG2^{+/+}$ and $CMG2^{Hep}$ mice than in $CMG2^{-/-}$ controls (Fig. 5c). Histological analyses readily identified necrotic areas in livers of ET-treated wild-type but not $CMG2^{-/-}$ mice (Extended Data Fig. 5b). Only scattered degenerated cardiomyocytes were found in ET-treated wild-type, but not $CMG2^{-/-}$, mice (Extended Data Fig. 5b). No obvious abnormalities were identified in the lungs, kidneys and spleens of ET-treated mice (data not shown). Together, the above results clearly demonstrate that hepatocytes (but not endothelial cells, cardiomyocytes or smooth muscle cells) are the key cell target of ET-induced lethality.

Discussion

Generation and analysis of the cell-type-specific CMG2-null mice, and the corresponding cell-type-specific CMG2-expressing mice, have allowed us to identify the target cell types responsible for lethality induced by anthrax toxins. Expression of the major toxin receptor CMG2 in cardiomyocytes and smooth muscle cells (but not in hepatocytes) is required for LT-induced lethality, whereas CMG2 expression in hepatocytes (but not in cardiomyocytes and smooth muscle cells) is critical for ET-mediated lethality. Therefore, *B. anthracis* has evolved to use the two distinct exotoxins to induce lethality by coordinately damaging two vital but distinct systems the cardiovascular system and liver to complete the pathogen's life cycle. Notably,

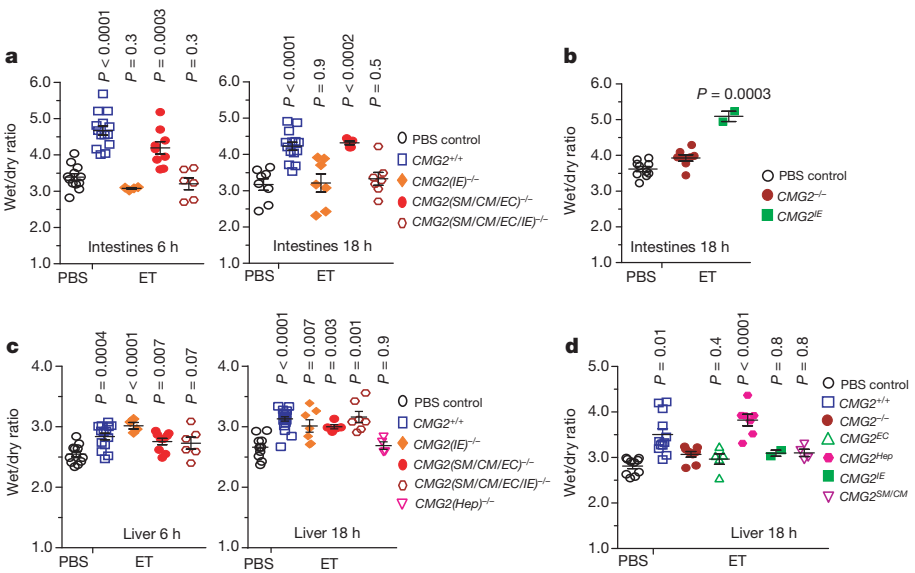


Figure 4 | ET directly targets intestinal epithelial cells and hepatocytes. **a**–**d**, Wet/dry ratios of intestines (**a**, **b**) or livers (**c**, **d**) of the cell-type-specific CMG2-null mice (**a**, **c**), and the cell-type-specific CMG2-expressing mice (**b**, **d**) after ET challenge (30 μg, intravenously). Because PBS-treated wild-type and mutant mice have a similar baseline of wet/dry ratio, the PBS control groups were pooled for each tissue. In **a** and **c**, the *P* values of the indicated groups versus PBS controls are given. In **b** and **d**, the *P* values of the indicated groups versus the $CMG2^{-/-}$ group are given. Horizontal lines indicate means. Error bars indicate s.e. Two-tailed unpaired *t*-test.

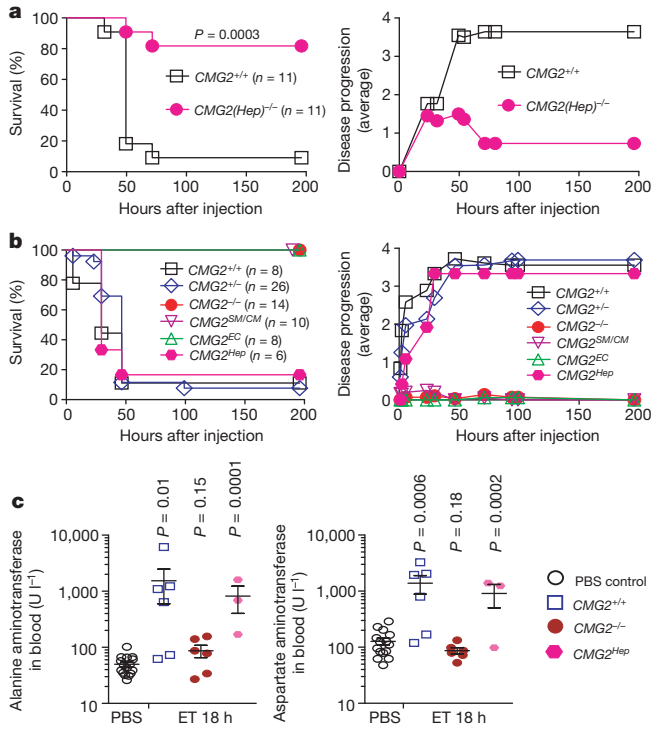


Figure 5 | Hepatocytes are a major target of ET-induced lethality. a, Survival of *CMG2(Hep)*$^{-/-}$ mice after ET (35 μg, intravenously) challenge. Right panel shows the disease progression of the challenged mice. See Methods for disease progression scoring criteria. Log-rank test. **b**, Sensitivity of *CMG2*^{Hep} mice to ET. The cell-type-specific CMG2-expressing mice and their respective littermate controls were challenged with ET (50 μg, intravenously). *CMG2*^{Hep} versus *CMG2*$^{-/-}$ mice, $P = 0.003$. Log-rank test. **c**, Serum levels of alanine aminotransferase and aspartate aminotransferase in ET-challenged mice (30 μg, intravenously). Horizontal lines indicate means. Error bars indicate s.e. Two-tailed unpaired *t*-test.

endothelial cells, the cells often considered to be a key target of LT-induced lethality[14,17–20], are not a primary target for either LT or ET.

Our studies also show that the previous hypoxia-mediated damage observed in the liver in response to LT is a secondary event not related to direct targeting of this organ by the toxin[15]. However, liver is the principal target of ET. In addition to subcutaneous oedema and fluid accumulation in the intestinal lumen, ET targeting of hepatocytes induces a unique liver oedema, which does not occur in other internal organs. Hepatocyte-specific CMG2-null mice were remarkably resistant to ET and, importantly, the mice with CMG2 expressed only in hepatocytes were susceptible to this toxin.

The relative importance of LT and ET in anthrax infections remains uncertain. LF and EF have been measured in the blood of infected animals. In nearly every successful measurement of toxin levels, LF concentrations exceeded those of EF. A recent analysis of *in vivo* toxin levels in anthrax-infected rabbits found LF at 10–35 μg ml^{-1} by 48 h and EF at about fivefold lower concentrations[26]. The 5:1 ratio of LF:EF has also been found in other anthrax infection models[27,28]. Thus, it is likely that lethal doses of ET may only be achieved in the very late stages of anthrax infection and that the majority of symptoms of infected animals may be caused by LT. Consistent with this view, we observed that the LT-resistant mice created in this study are much more likely to survive anthrax infections.

The findings reported here may have value in understanding the pathogenesis of human anthrax infections. Recognition that the anthrax toxins are targeting the cardiovascular system (LT) and liver (ET) may suggest specific supportive therapies that would limit tissue damage and increase survival. Re-examination of the clinical course, pathology and autopsy reports of the relatively few well documented human anthrax cases[29,30] in light of the data presented here may provide additional insights.

METHODS SUMMARY

In toxin challenge experiments, 8–10-week-old mice with various genotypes were injected with LT or ET either intraperitoneally or intravenously. *B. anthracis* spores were prepared from the non-encapsulated, toxigenic *B. anthracis* Sterne-like strain A35 (refs 31, 32). For infection studies, mice were injected with 4×10^8 A35 spores subcutaneously or 2×10^6 A35 vegetative bacteria via tail vein. For the tissue wet/dry ratio measurements, mice treated with PBS or ET were killed by CO_2 inhalation, and organs collected and weighed. Organs were dried by incubation in an oven at 45 °C overnight. Endothelial cells were isolated and cultured as described previously[33]. For cytotoxicity assays, endothelial cells were grown in 96-well plates and treated with serial dilutions of PA or PA(L687A) combined with 100 ng ml^{-1} FP59 for 48 h. Cell viabilities were then assayed by MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) as described previously[34]. Generation of the tissue-specific CMG2-null mice and the corresponding tissue-specific CMG2-expressing mice, anthrax toxin protein purification, endothelial cell isolation, and other methods are detailed in the Methods.

Online Content Any additional Methods, 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 18 April; accepted 31 July 2013.
Published online 28 August 2013.

1. Moayeri, M. & Leppla, S. H. Cellular and systemic effects of anthrax lethal toxin and edema toxin. *Mol. Aspects Med.* **30,** 439–455 (2009).
2. Bradley, K. A., Mogridge, J., Mourez, M., Collier, R. J. & Young, J. A. Identification of the cellular receptor for anthrax toxin. *Nature* **414,** 225–229 (2001).
3. Scobie, H. M., Rainey, G. J., Bradley, K. A. & Young, J. A. Human capillary morphogenesis protein 2 functions as an anthrax toxin receptor. *Proc. Natl Acad. Sci. USA* **100,** 5170–5174 (2003).
4. Liu, S. *et al.* Capillary morphogenesis protein-2 is the major receptor mediating lethality of anthrax toxin *in vivo*. *Proc. Natl Acad. Sci. USA* **106,** 12424–12429 (2009).
5. Leppla, S. H. Anthrax toxin edema factor: a bacterial adenylate cyclase that increases cyclic AMP concentrations of eukaryotic cells. *Proc. Natl Acad. Sci. USA* **79,** 3162–3166 (1982).
6. Firoved, A. M. *et al. Bacillus anthracis* edema toxin causes extensive tissue lesions and rapid lethality in mice. *Am. J. Pathol.* **167,** 1309–1320 (2005).
7. Duesbery, N. S. *et al.* Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science* **280,** 734–737 (1998).
8. Vitale, G. *et al.* Anthrax lethal factor cleaves the N-terminus of MAPKKs and induces tyrosine/threonine phosphorylation of MAPKs in cultured macrophages. *Biochem. Biophys. Res. Commun.* **248,** 706–711 (1998).
9. Vitale, G., Bernardi, L., Napolitani, G., Mock, M. & Montecucco, C. Susceptibility of mitogen-activated protein kinase kinase family members to proteolysis by anthrax lethal factor. *Biochem. J.* **352,** 739–745 (2000).
10. Newman, Z. L. *et al.* Susceptibility to anthrax lethal toxin-induced rat death is controlled by a single chromosome 10 locus that includes rNlrp1. *PLoS Pathog.* **6,** e1000906 (2010).
11. Levinsohn, J. L. *et al.* Anthrax lethal factor cleavage of Nlrp1 is required for activation of the inflammasome. *PLoS Pathog.* **8,** e1002638 (2012).
12. Pezard, C., Berche, P. & Mock, M. Contribution of individual toxin components to virulence of *Bacillus anthracis*. *Infect. Immun.* **59,** 3472–3477 (1991).
13. Liu, S. *et al.* Anthrax toxin targeting of myeloid cells through the CMG2 receptor is essential for establishment of *Bacillus anthracis* infections in mice. *Cell Host Microbe* **8,** 455–462 (2010).
14. Guichard, A., Nizet, V. & Bier, E. New insights into the biological effects of anthrax toxins: linking cellular to organismal responses. *Microbes Infect.* **5,** 48–61 (2012).
15. Moayeri, M., Haines, D., Young, H. A. & Leppla, S. H. *Bacillus anthracis* lethal toxin induces TNF-independent hypoxia-mediated toxicity in mice. *J. Clin. Invest.* **112,** 670–682 (2003).
16. Moayeri, M. *et al.* The heart is an early target of anthrax lethal toxin in mice: a protective role for neuronal nitric oxide synthase (nNOS). *PLoS Pathog.* **5,** e1000456 (2009).
17. Bolcome, R. E. III *et al.* Anthrax lethal toxin induces cell death-independent permeability in zebrafish vasculature. *Proc. Natl Acad. Sci. USA* **105,** 2439–2444 (2008).
18. Guichard, A. *et al.* Anthrax toxins cooperatively inhibit endocytic recycling by the Rab11/Sec15 exocyst. *Nature* **467,** 854–858 (2010).
19. Warfel, J. M., Steele, A. D. & D'Agnillo, F. Anthrax lethal toxin induces endothelial barrier dysfunction. *Am. J. Pathol.* **166,** 1871–1881 (2005).
20. Maddugoda, M. P. *et al.* cAMP signaling by anthrax edema toxin induces transendothelial cell tunnels, which are resealed by MIM via Arp2/3-driven actin polymerization. *Cell Host Microbe* **10,** 464–474 (2011).
21. Ghosh, C. C. *et al.* Impaired function of the Tie-2 receptor contributes to vascular leakage and lethality in anthrax. *Proc. Natl Acad. Sci. USA* **109,** 10024–10029 (2012).

22. Arora, N., Klimpel, K. R., Singh, Y. & Leppla, S. H. Fusions of anthrax toxin lethal factor to the ADP-ribosylation domain of Pseudomonas exotoxin A are potent cytotoxins which are translocated to the cytosol of mammalian cells. *J. Biol. Chem.* **267,** 15542–15548 (1992).
23. Liu, S. *et al.* Diphthamide modification on eukaryotic elongation factor 2 is needed to assure fidelity of mRNA translation and mouse development. *Proc. Natl Acad. Sci. USA* **109,** 13817–13822 (2012).
24. Bradley, S. V. *et al.* Degenerative phenotypes caused by the combined deficiency of murine HIP1 and HIP1r are rescued by human HIP1. *Hum. Mol. Genet.* **16,** 1279–1292 (2007).
25. Lepore, J. J. *et al.* High-efficiency somatic mutagenesis in smooth muscle cells and cardiac myocytes in SM22α-Cre transgenic mice. *Genesis* **41,** 179–184 (2005).
26. Dal Molin, F. *et al.* Ratio of lethal and edema factors in rabbit systemic anthrax. *Toxicon* **52,** 824–828 (2008).
27. Sirard, J. C., Mock, M. & Fouet, A. The three *Bacillus anthracis* toxin genes are coordinately regulated by bicarbonate and temperature. *J. Bacteriol.* **176,** 5188–5192 (1994).
28. Mabry, R. *et al.* Detection of anthrax toxin in the serum of animals infected with Bacillus anthracis by using engineered immunoassays. *Clin. Vaccine Immunol.* **13,** 671–677 (2006).
29. Jernigan, J. A. *et al.* Bioterrorism-related inhalational anthrax: The first 10 cases reported in the United States. *Emerg. Infect. Dis.* **7,** 933–944 (2001).
30. Guarner, J. *et al.* Pathology and pathogenesis of bioterrorism-related inhalational anthrax. *Am. J. Pathol.* **163,** 701–709 (2003).
31. Pomerantsev, A. P., Sitaraman, R., Galloway, C. R., Kivovich, V. & Leppla, S. H. Genome engineering in *Bacillus anthracis* using Cre recombinase. *Infect. Immun.* **74,** 682–693 (2006).
32. Hu, H., Sa, Q., Koehler, T. M., Aronson, A. I. & Zhou, D. Inactivation of *Bacillus anthracis* spores in murine primary macrophages. *Cell. Microbiol.* **8,** 1634–1642 (2006).
33. Reynolds, L. E. & Hodivala-Dilke, K. M. Primary mouse endothelial cell culture for assays of angiogenesis. *Methods Mol. Med.* **120,** 503–509 (2006).
34. Liu, S. & Leppla, S. H. Cell surface tumor endothelium marker 8 cytoplasmic tail-independent anthrax toxin binding, proteolytic processing, oligomer formation, and internalization. *J. Biol. Chem.* **278,** 5227–5234 (2003).

Acknowledgements This research was supported by the intramural research programs of the National Institute of Allergy and Infectious Diseases and the National Heart, Lung, and Blood Institute, National Institutes of Health. We thank L. Feigenbaum and the staff at SAIC/NCI Frederick for generation of the founder *CMG2* transgenic mice. We thank A. Kulkarni, B. Hall, B. Klaunberg, S. Anderson, I. Sastalla, C. Leysath and C. Bachran for discussions, and D. Despres for help with echocardiography.

Author Contributions Y.Z. maintained mouse colonies and performed animal experiments. M.M. and J.L. designed, performed experiments, analysed data and edited the paper. D.C. performed animal experiments. R.J.F. purified proteins. A.N.W. made the *CMG2* transgenic construct. Z.-X.Y. performed histological analyses. T.F. was involved in scientific discussions, providing reagents, and edited the paper. S.H.L. supervised research and edited the paper. S.L. conceived and supervised the project, designed and performed experiments, analysed data and 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 S.L. (shliu@niaid.nih.gov) or S.H.L. (sleppla@niaid.nih.gov).

METHODS

Generation of tissue-specific CMG2-null mice. Mice having the *CMG2* transmembrane-domain-encoding exon 12 flanked by *loxP* sites (a floxed allele), namely, the $CMG2^{fl/fl}$ mice (C57BL/6 background), were described previously[4]. To generate mice having *CMG2* deleted only in endothelial cells, the $CMG2^{fl/fl}$ mice were mated with *Cdh-cre* transgenic mice[35] (Cre-recombinase under the VE-cadherin promoter) (006137, the Jackson Laboratory). Endothelial-cell-specific CMG2-null mice ($CMG2(EC)^{-/-}$) were obtained by the subsequent intercrossing of the resulting $CMG2^{+/fl}$/*Cdh-cre* mice (see Table 1 for nomenclature). Similarly, to generate mice with *CMG2* deleted in cardiomyocytes, smooth muscle cells, hepatocytes and intestine epithelial cells, the $CMG2^{fl/fl}$ mice were mated with *Myh6-cre* (ref. 36) (Cre-recombinase under cardiomyocyte-specific α myosin heavy polypeptide 6 promoter) (011038, the Jackson Laboratory), *SM22-cre* (ref. 37) (Cre-recombinase under the vascular smooth-muscle-specific smooth muscle protein 22α promoter) (017491, the Jackson Laboratory), *Alb-cre* (ref. 38) (Cre-recombinase under mouse albumin promoter) (003574, the Jackson Laboratory), and *Vil-cre* (ref. 39) (Cre-recombinase under mouse intestinal-epithelial-cell-specific villin 1 promoter) (004586, the Jackson Laboratory) transgenic mice, respectively. All the tissue-specific Cre mice had been backcrossed to C57BL/6 for at least ten generations when purchased from the Jackson Laboratory, thus the resulting tissue-specific CMG2-null mice used in this study harbour LT-resistant *Nlrp1* alleles, and are resistant to LT-induced rapid myeloid cell death. Gene-targeted and transgenic mice used in this study are listed in Table 1.

Genotyping was performed by PCR using mouse ear DNA. In analysing *CMG2* expression, total RNA isolated from various mouse tissues using TRIZOL reagent (Invitrogen) was reverse transcribed using the SuperScript III Reverse Transcriptase (Invitrogen). The 451-bp transmembrane-containing *CMG2* cDNA fragment was amplified from wild-type tissues using a forward and reverse primer pair, as listed in Extended Data Table 1. These primers amplify a 355-bp transmembrane-deleted *CMG2* cDNA fragment from *CMG2* transmembrane-domain-deleted tissues. Two biological replicates were performed for reverse transcription (RT)–PCR analyses of tissue-specific deletion of CMG2.

Generation of tissue-specific *CMG2* transgenic mice. Mouse full-length *CMG2* cDNA was isolated by RT–PCR from mouse bone-marrow-derived macrophages and cloned into EcoRI/NheI sites of pCLE vector[24,40], resulting in *CMG2* transgenic vector pCLEmCMG2 containing mouse *CMG2* cDNA under the control of the CAG promoter (P_{CAG}, a combination of the cytomegalovirus early enhancer element and chicken β-actin promoter). A loxP-stop-loxP cassette (LSL, a DNA fragment containing eGFP coding sequence followed by a poly(A) terminator[24]) flanked by *loxP* sites was placed between P_{CAG} and *CMG2* cDNA (Extended Data Fig. 2a). After removing the non-relevant part by XhoI and DraIII digestions, the transgenic vector was microinjected into the pronuclei of fertilized eggs to generate transgenic mice (C57BL/6 background). The founder lines were genotyped by Southern blot analyses. Because P_{CAG} is a ubiquitous promoter, the resulting transgenic mice (*LSL-CMG2*) exhibit whole-body green fluorescence and thus could be identified with GFP visualization using GFP visualization goggles (Model FHS/F-00, BLS). The transgenic mice do not express the *CMG2* transgene until bred with cell-type-specific *cre*-transgenic mice to remove the LSL cassette and place P_{CAG} adjacent to the *CMG2* cDNA (Extended Data Fig. 2a). Thus, endothelial-cell-specific *CMG2*-expressing mice ($CMG2^{EC}$) were obtained by breeding the green *LSL-CMG2* mice with *Cdh-cre* mice to delete the LSL cassette only in endothelial cells, and by subsequent breeding with whole-body $CMG2^{-/-}$ mice to eliminate expression of the endogenous *CMG2* gene. Cardiomyocyte-, smooth-muscle/cardiomyocyte-, intestinal-epithelial- and hepatocyte-specific CMG2-expressing mice were generated similarly using the corresponding cell-type-specific Cre-expressing mice. The PCR primers for genotyping these mice are listed in Extended Data Table 1. *CMG2* transgenic mice were generated by SAIC/NCI-FCRDC-Frederick under contract with the National Institute of Allergy and Infectious Diseases.

Animal studies. All animal studies were carried out in accordance with protocols approved by the National Institute of Allergy and Infectious Diseases Animal Care and Use Committee. In LT challenge experiments, 8–10-week-old male and female mice with various genotypes were injected with one or two doses of 100 μg LT (100 μg PA plus 100 μg LF) in 0.5 ml PBS intraperitoneally or 50 μg LT in 0.2 ml PBS intravenously. In ET challenge experiments, mice described above were injected 25–50 μg ET in 0.2 ml PBS intravenously or in 0.5 ml PBS intraperitoneally. PA, LF, EF and FP59 proteins were purified from a non-virulent *B. anthracis* strain as previously described[41–43]. For infection studies, *B. anthracis* spores were prepared from the non-encapsulated, toxigenic *B. anthracis* Sterne-like strain A35 (ref. 31) as previously described[32]. Mice were injected with 4×10^8 A35 spores subcutaneously. For vegetative bacterial infections, mice were injected with 2×10^6 A35 bacteria via tail vein. Our previous studies demonstrate that 5–10 mice per treatment group are sufficient for statistical analyses in toxin challenge experiments[4,13,15]. Thus, we use 5–10 or more mice per group to ensure statistical power. At least two biological replicates were performed for each toxin challenge experiment. The mice were grouped based on genotypes. When the mice in the same genotype received different treatments, they were randomly assigned to treatment groups. We used the following criteria to score mouse disease progression induced by LT or ET: 0, healthy mouse; 1, slight ruffled coat but no problem in running around cage; 2, ruffled coat and decrease in activity; 3, ruffled coat, hunched posture and little movement; 3.5, moribund; 4, found dead. All toxin-challenged or infected mice were monitored twice daily for 2 weeks after challenge for signs of malaise or mortality by investigators and animal caretakers who were unaware of genotypes.

For the tissue wet/dry ratio measurements, mice treated with PBS or ET were killed by CO_2 inhalation, and organs collected and weighed. Organs were dried by incubation in an oven at 45 °C overnight. For the footpad skin oedema model, mice were injected intradermally in one foot pad with 0.25 μg ET in 20 μl PBS. Footpad oedema was monitored at 8 and 20 h after injection by measuring dorsal/planar and medial/lateral sizes using digital calipers.

Murine c-troponin-I ELISA kit (Life Diagnostics) was used to measure the cardiac biomarker c-troponin-I in mouse serum samples according to the manufacturer's protocol. Mice of various genotypes were treated intraperitoneally with 100 μg LT and bled at indicated time points after challenge for serum c-troponin-I measurements. For echocardiography studies, mice were kept under isoflurane anaesthesia on a heating pad with body temperature monitoring (by anal probe) throughout the procedure. Baseline left ventricular short and long axis views were obtained with the Vevo 770 system at the NIH Mouse Imaging Facility. All measurements were performed in blinded fashion.

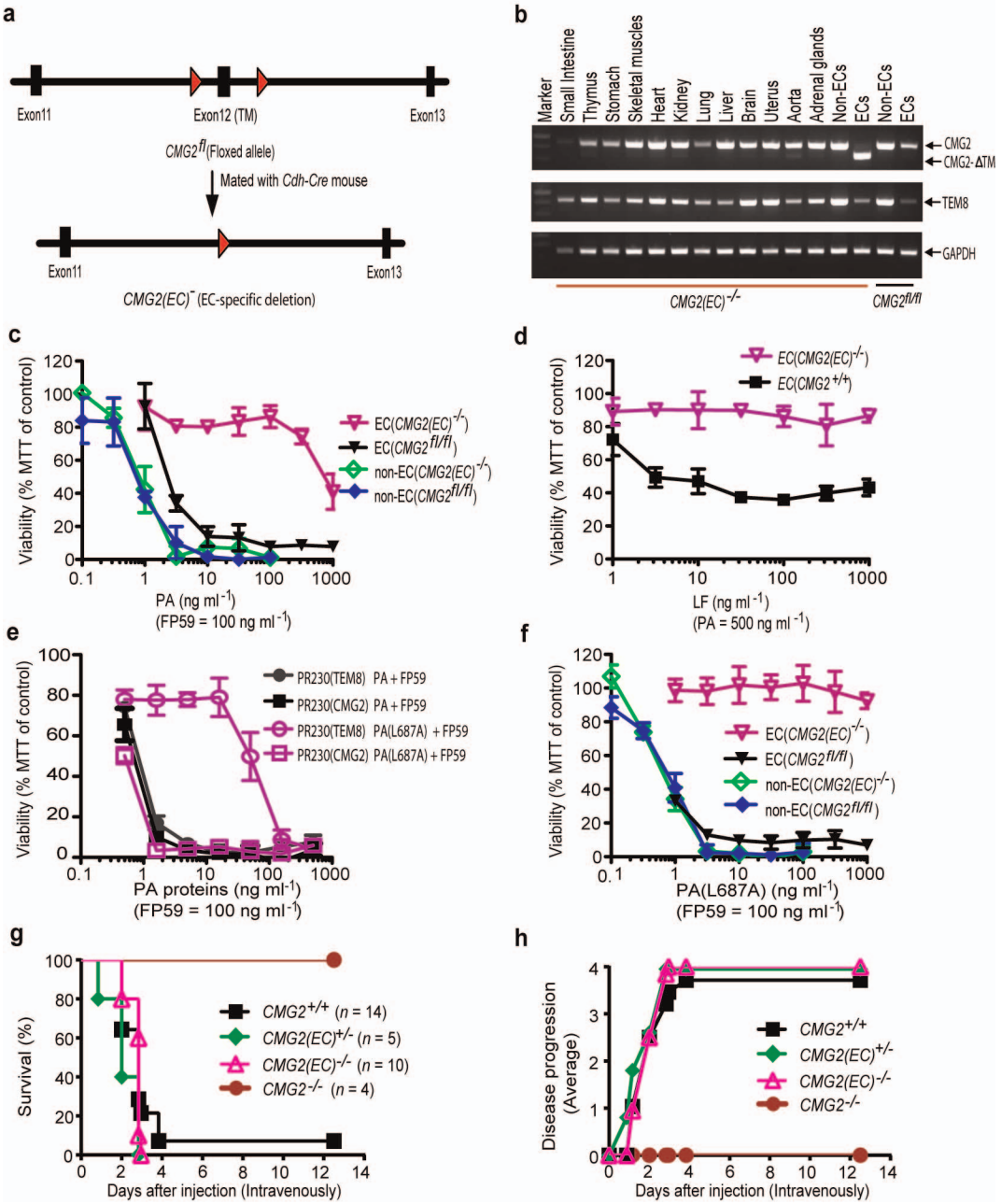
Mouse endothelial cell isolation and cytotoxicity assay. Endothelial cells were isolated and cultured as described previously[33]. Briefly, three mouse lungs were digested with type I collagenase and plated on gelatin- and collagen-coated flasks. The cells were then subjected to sequential negative sorting by magnetic beads coated with a sheep anti-rat antibody using a Fc blocker (rat anti-mouse CD16/CD32, catalogue number 553142, BD Pharmingen) to remove macrophages and positive sorting by magnetic beads using an anti- intermolecular adhesion molecule 2 (ICAM2 or CD102) antibody (catalogue 553326, rat anti-mouse CD102, BD Pharmingen) to isolate endothelial cells (ICAM2 positive cells). The cells other than endothelial cells (non-endothelial cells, ICAM2 negative cells) were also isolated simultaneously as controls. These primary cells were used within 5 passages after isolation. For cytotoxicity assays, endothelial cells, non-endothelial cells, and Chinese hamster ovary (CHO) cells were grown in 96-well plates and treated with serial dilutions of PA or PA(L687A) combined with 100 ng ml^{-1} FP59 for 48 h. PA(L687A) was first described in ref. 44 and found in this study to be a CMG2-preferable PA mutant. FP59 is a fusion protein of LF amino acids 1–254 and the catalytic domain of *Pseudomonas aeruginosa* exotoxin A that kills cells by ADP-ribosylation of eukaryotic elongation factor-2 after delivery to cytosol by PA[22,23]. Cell viabilities were then assayed by MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) as described previously[34], expressed as percentage of MTT signals of untreated cells. The CHO cells that express only TEM8 (PR230(TEM8)) or CMG2 (PR230(CMG2)) were described previously[42].

Histology. Mice treated with PBS, LT (100 μg, intraperitoneally), or ET (50 μg, intravenously) were killed by CO_2 inhalation, and hearts, livers, lungs, kidneys and spleens collected and fixed in 4% paraformaldehyde for 24 h, embedded in paraffin, sectioned, stained with haematoxylin/eosin and subjected to microscopic analysis.

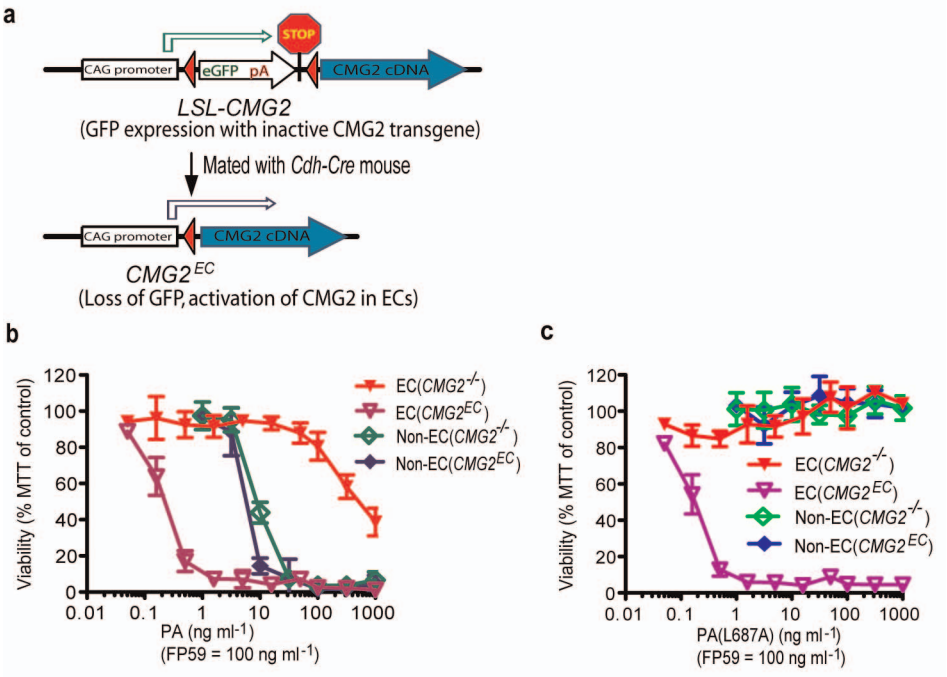
Statistical analyses. GraphPad Prism 6 was used for statistical analyses. For comparison of mouse survival curves, the log-rank (Mantel-Cox) test was used. For differences in tissue oedema between two groups, the two-tailed unpaired Student's *t*-test was used. $P < 0.05$ was considered statistically significant.

35. Alva, J. A. *et al.* VE-Cadherin-Cre-recombinase transgenic mouse: a tool for lineage analysis and gene deletion in endothelial cells. *Dev. Dyn.* **235,** 759–767 (2006).
36. Agah, R. *et al.* Gene recombination in postmitotic cells. Targeted expression of Cre recombinase provokes cardiac-restricted, site-specific rearrangement in adult ventricular muscle *in vivo*. *J. Clin. Invest.* **100,** 169–179 (1997).
37. Holtwick, R. *et al.* Smooth muscle-selective deletion of guanylyl cyclase-A prevents the acute but not chronic effects of ANP on blood pressure. *Proc. Natl Acad. Sci. USA* **99,** 7142–7147 (2002).
38. Postic, C. *et al.* Dual roles for glucokinase in glucose homeostasis as determined by liver and pancreatic beta cell-specific gene knock-outs using Cre recombinase. *J. Biol. Chem.* **274,** 305–315 (1999).
39. Braunstein, E. M. *et al.* Villin: A marker for development of the epithelial pyloric border. *Dev. Dyn.* **224,** 90–102 (2002).
40. Hall, B. E. *et al.* Conditional overexpression of TGF-β1 disrupts mouse salivary gland development and function. *Lab. Invest.* **90,** 543–555 (2010).
41. Pomerantsev, A. P. *et al.* A *Bacillus anthracis* strain deleted for six proteases serves as an effective host for production of recombinant proteins. *Protein Expr. Purif.* **80,** 80–90 (2011).

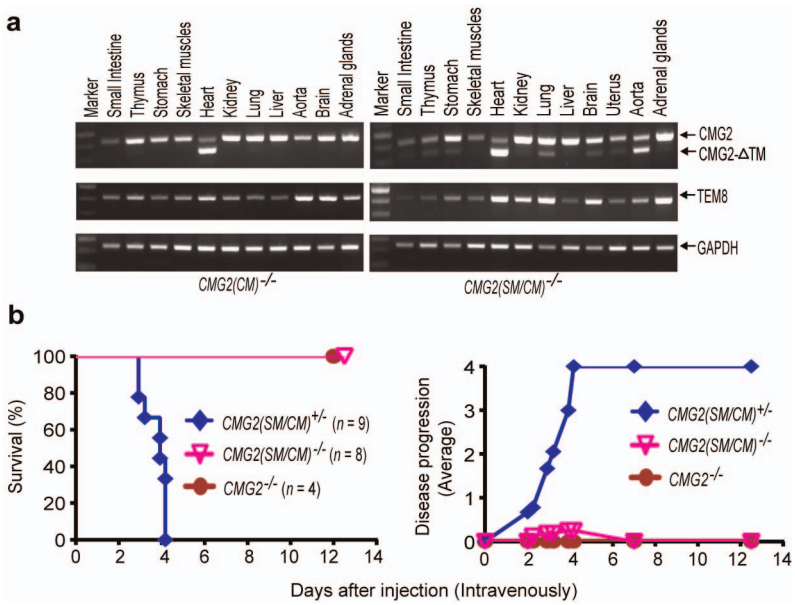
42. Liu, S., Leung, H. J. & Leppla, S. H. Characterization of the interaction between anthrax toxin and its cellular receptors. *Cell. Microbiol.* **9,** 977–987 (2007).
43. Gupta, P. K., Moayeri, M., Crown, D., Fattah, R. J. & Leppla, S. H. Role of N-terminal amino acids in the potency of anthrax lethal factor. *PLoS ONE* **3,** e3130 (2008).
44. Rosovitz, M. J. *et al.* Alanine scanning mutations in domain 4 of anthrax toxin protective antigen reveal residues important for binding to the cellular receptor and to a neutralizing monoclonal antibody. *J. Biol. Chem.* **278,** 30936–30944 (2003).



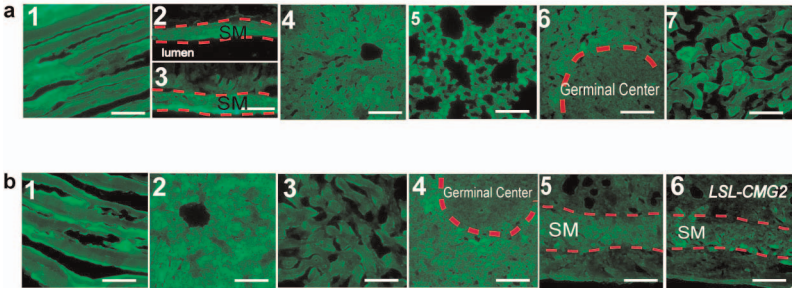
Extended Data Figure 1 | Generation of endothelial-cell-specific CMG2-null mice. **a**, Strategy for generation of endothelial-cell-specific CMG2-null mice. Diagram shows $CMG2^{fl}$ allele having exon 12 (encoding transmembrane domain, TM) flanked by *loxP* sites and the endothelial-cell-specific CMG2-null allele ($CMG2(EC)^{-}$). The red arrowheads indicate *loxP* sites. The homozygous endothelial-cell-specific CMG2-null mice ($CMG2(EC)^{-/-}$) were obtained by intercrossing of $CMG2(EC)^{+/-}$ mice. Other cell-type-specific CMG2-null mice were made similarly by using the corresponding cell-type-specific Cre transgenic mice. **b**, RT–PCR analyses of CMG2 transmembrane (TM) domain deletion in various tissues of $CMG2(EC)^{-/-}$ mice. Primers flanking the CMG2 transmembrane domain were used to amplify a *CMG2* cDNA fragment. Endothelial cells and non-endothelial cells were isolated simultaneously from lungs pooled from three $CMG2(EC)^{-/-}$ mice and three $CMG2^{fl/fl}$ control mice. Representative of two independent experiments is shown. Expression of TEM8 and GAPDH in these samples is also shown. **c**, Sensitivity of endothelial cells and non-endothelial cells from $CMG2^{fl/fl}$ and $CMG2(EC)^{-/-}$ mice to PA plus FP59. Cells were treated with various concentrations of PA and FP59 (100 ng ml^{-1}) for 48 h. Cell viability was evaluated by MTT assay, expressed as relative MTT signals to untreated cells. Error bars indicate s.d. **d**, Resistance of endothelial cells from $CMG2(EC)^{-/-}$ mice to LT. Endothelial cells from $CMG2(EC)^{-/-}$ and wild-type mice were treated with various concentrations of LF and PA (500 ng ml^{-1}) for 48 h. **e**, PA(L687A) preferentially kills CMG2-expressing cells. Cells were treated with various concentrations of PA or PA(L687A) and 100 ng ml^{-1} FP59 for 48 h. PR230(TEM8) and PR230(CMG2) are engineered CHO cells expressing only TEM8 or CMG2. Note that PR230(TEM8) cells are 100-fold more resistant than PR230(CMG2) cells to PA(L687A) plus FP59. **f**, Sensitivity of endothelial cells and non-endothelial cells from $CMG2^{fl/fl}$ and $CMG2(EC)^{-/-}$ mice to PA(L687A) plus FP59. Cells were incubated for 48 h with various concentrations of PA(L687A) and 100 ng ml^{-1} FP59. Error bars indicate s.d. **g**, Susceptibility of $CMG2(EC)^{-/-}$ mice to LT. $CMG2(EC)^{-/-}$ mice and their littermate controls were injected intravenously with 50 μg LT (50 μg PA plus 50 μg LF), and monitored for survival. Whole-body $CMG2^{-/-}$ mice were included as additional controls. **h**, Disease progression of the LT-challenged mice in panel **g**. Please see Methods for disease progression scoring criteria.



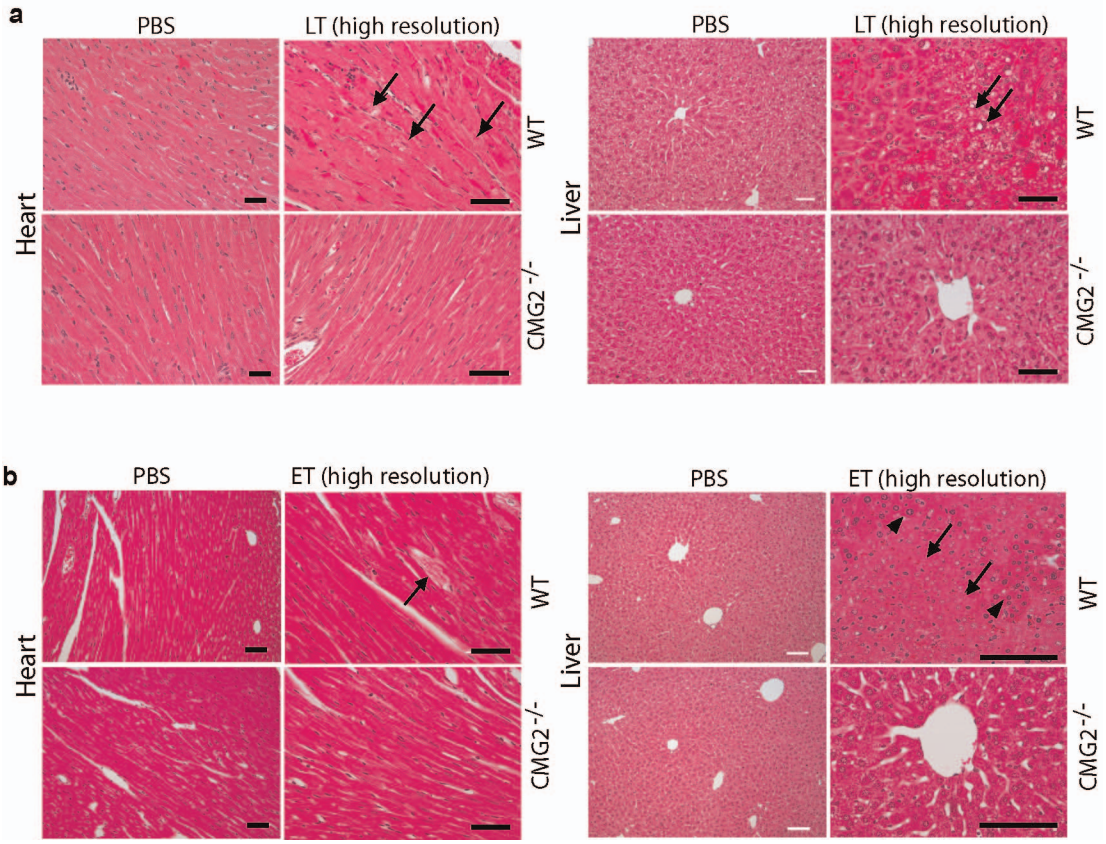
Extended Data Figure 2 | Generation of endothelial-cell-specific CMG2-expressing mice. **a**, Strategy for generation of endothelial-cell-specific CMG2-expressing mice. In the *CMG2* transgenic vector (*LSL-CMG2*), a *loxP-stop-loxP* cassette containing a promoterless eGFP and a poly(A) stop signal flanked by *loxP* sites was placed between the CAG promoter and *CMG2* cDNA. Activation of *CMG2* transgene in endothelial cells ($CMG2^{EC}$) was achieved by breeding *LSL-CMG2* mice with *Cdh-cre* mice to specifically remove the *loxP-stop-loxP* cassette in endothelial cells. Other cell-type-specific CMG2-expressing mice were made similarly by using the corresponding cell-type-specific Cre transgenic mice. **b**, **c**, Regained toxin sensitivity of endothelial cells from $CMG2^{EC}$ mice. Endothelial cells and non-endothelial cells from $CMG2^{EC}$ and whole-body $CMG2^{-/-}$ mice were incubated for 48 h with various concentrations of PA (**b**) or PA(L687A) (**c**) and 100 ng ml^{-1} FP59. Error bars indicate s.d.



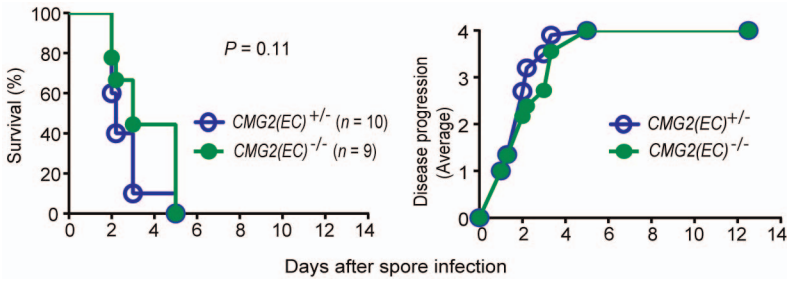
Extended Data Figure 3 | Tissue-specific deletion of CMG2 in *CMG2(CM)*$^{-/-}$ and *CMG2(SM/CM)*$^{-/-}$ mice. **a**, RT–PCR analyses of CMG2 deletion in tissues of *CMG2(CM)*$^{-/-}$ and *CMG2(SM/CM)*$^{-/-}$ mice. CMG2 deletion was detected in the hearts of *CMG2(CM)*$^{-/-}$ mice and in the hearts and aorta of *CMG2(SM/CM)*$^{-/-}$ mice. The small fraction of CMG2 deletion that occurred in other tissues of the *CMG2(SM/CM)*$^{-/-}$ mice was due to the existence of varying amounts of vascular smooth muscle cells in those tissues. Representative of two independent experiments is shown. **b**, Resistance of smooth-muscle/cardiomyocyte-specific CMG2-null mice to LT. *CMG2(SM/CM)*$^{-/-}$ mice and their littermate *CMG2(SM/CM)*$^{+/-}$ controls were injected intravenously with 50 μg LT, and monitored for survival. Whole-body *CMG2*$^{-/-}$ mice were included as additional controls. Right panel shows the disease progression of the challenged mice. *CMG2(SM/CM)*$^{-/-}$ versus *CMG2*$^{+/+}$ mice, $P = 0.0002$. Log-rank test.



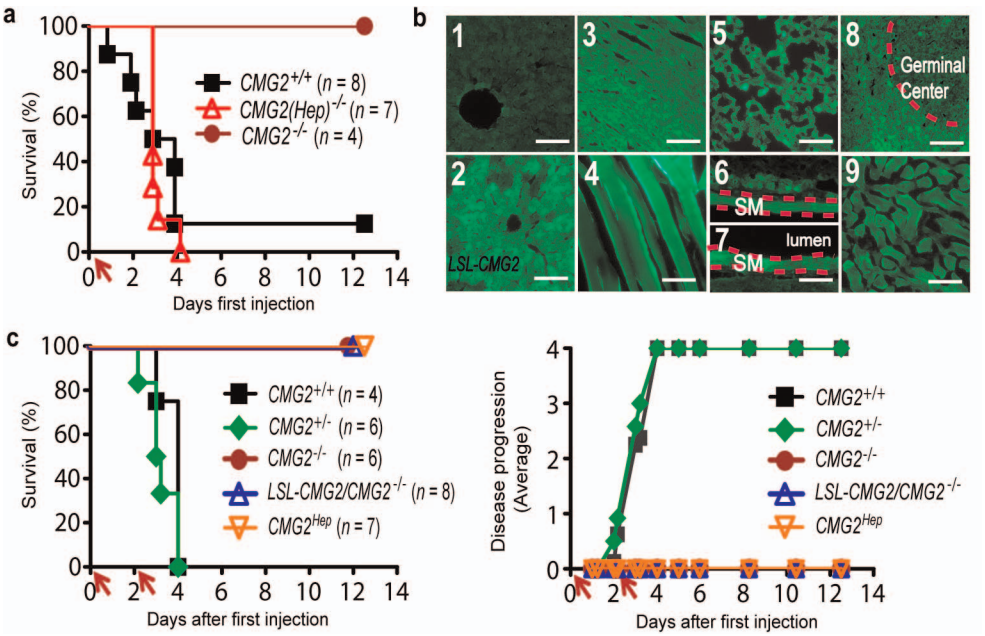
Extended Data Figure 4 | Fluorescence microscopic analyses of GFP expression in mouse tissues. **a**, Representative fluorescence microscopy of skeletal muscle (1), aorta (vascular smooth muscle) (2), small intestine (smooth muscle) (3), liver (4), lung (5), spleen (6) and kidney (cortex) (**7**) from $CMG2^{CM}$ mice ($n = 2$). Scale bar, 100 μm. **b**, Representative fluorescence microscopy of skeletal muscle (1), liver (2), kidney (cortex) (3), spleen (4) and uterus (5) from $CMG2^{SM/CM}$ mice ($n = 3$), and uterus (6) from *LSL-CMG2* mice ($n = 2$). Scale bar, 100 μm.



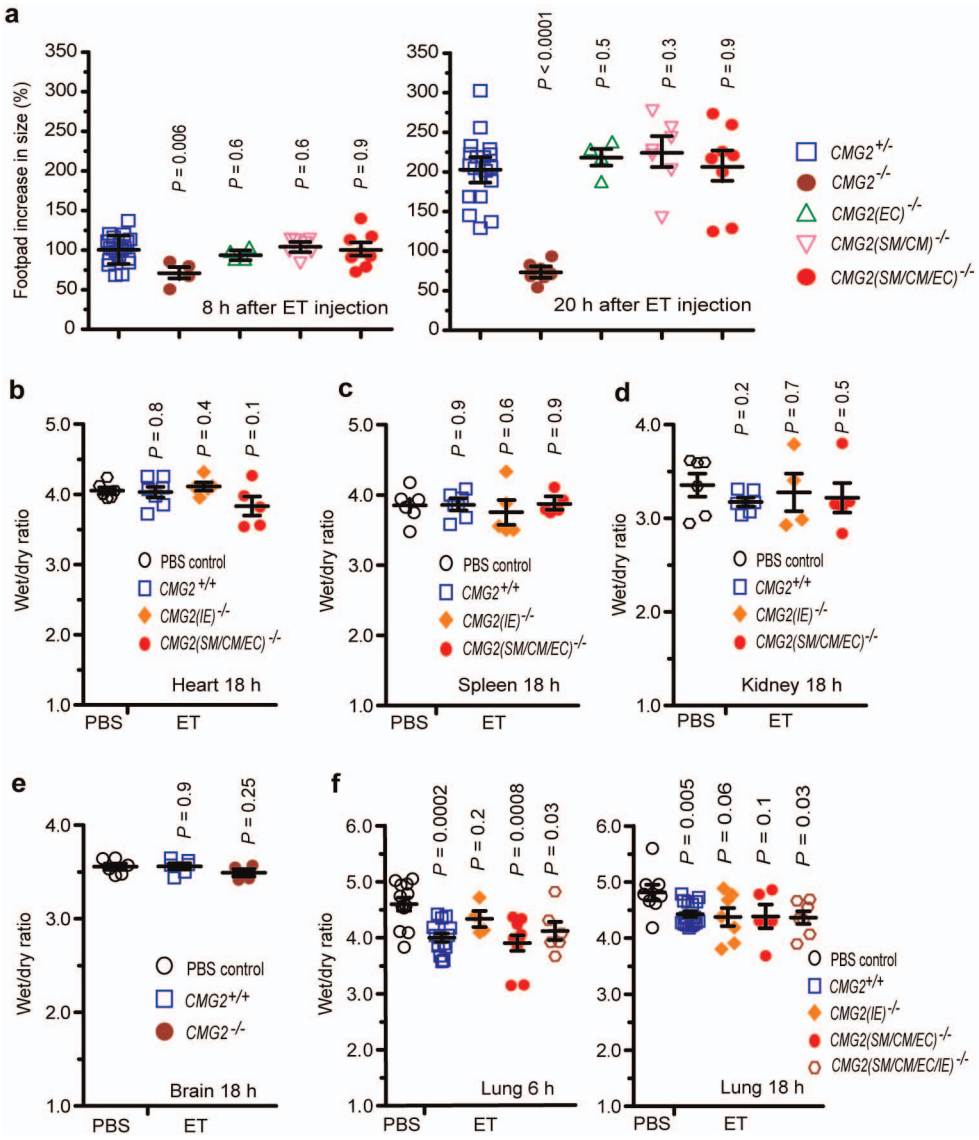
Extended Data Figure 5 | Histology of heart and liver of LT- and ET-treated mice. **a**, Haematoxylin and eosin staining of heart and liver from wild-type ($n = 3$) and $CMG2^{-/-}$ ($n = 3$) mice challenged intraperitoneally with 100 μg LT for 48 h. In heart, regions with cardiomyocyte degeneration were found in LT-treated wild-type but not $CMG2^{-/-}$ mice. Arrows show examples of degenerated cardiomyocytes. In liver, regions with mild to modest hepatocyte degeneration were identified in LT-treated wild-type but not $CMG2^{-/-}$ mice. Arrows show examples of degenerated hepatocytes with cytosol vacuolization changes. Scale bar, 50 μm. **b**, Haematoxylin and eosin staining of heart and liver from wild-type ($n = 4$) and $CMG2^{-/-}$ ($n = 3$) mice 18 h after 50 μg ET injection (intravenously). In liver, regions with hepatocyte necrotic changes were identified in ET-treated wild-type mice but not $CMG2^{-/-}$ mice. Arrows show necrotic regions, arrowheads indicate examples of intact hepatocytes remaining in the necrotic regions. In heart, only scattered degenerated cardiomyocytes (arrow) were found in ET-treated wild-type but not $CMG2^{-/-}$ mice. Scale bar, 50 μm.



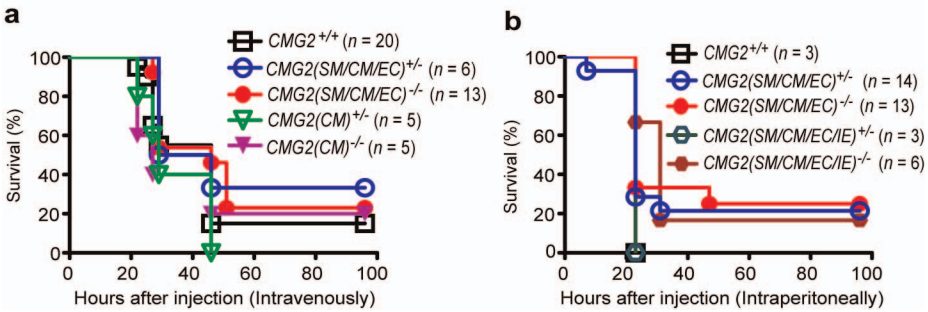
Extended Data Figure 6 | Endothelial-cell-specific CMG2-null mice are sensitive to *B. anthracis* infection. *CMG2(EC)*$^{-/-}$ mice and their littermate heterozygous mice were subcutaneously infected with 4×10^8 Sterne spores and monitored for survival. The right panel shows the disease progression of the challenged mice. Log-rank test.



Extended Data Figure 7 | LT targeting of liver does not contribute to lethality. **a**, Susceptibility of the hepatocyte-specific CMG2-null mice to LT. *CMG2(Hep)*$^{-/-}$ mice and their littermate *CMG2*$^{+/+}$ control mice were challenged intraperitoneally with 100 μg LT and monitored for survival. Whole-body *CMG2*$^{-/-}$ mice were included as additional controls. **b**, Selective activation of *CMG2* transgene in liver of *CMG2*^{Hep} mice. Representative fluorescence microscopy of the liver (1), heart (3), skeletal muscles (4), lung (5), small intestines (smooth muscle) (6), aorta (7), spleen (8) and kidney (cortex) (9) from *CMG2*^{Hep} mice ($n = 2$), and liver (2) from *LSL-CMG2* mice ($n = 2$). Selective loss of GFP expression in liver from *CMG2*^{Hep} mice but not *LSL-CMG2* mice (1 and 2) is shown. Scale bar, 100 μm. **c**, Susceptibility of the hepatocyte-specific CMG2-expressing mice to LT. *CMG2*^{Hep} mice and various control mice as indicated were intraperitoneally challenged with two doses of 100 μg LT and monitored for survival or signs of malaise. Right panel: disease progression of the challenged mice.



Extended Data Figure 8 | Oedema in ET-treated mice. **a**, ET-induced footpad skin oedema in mice. Mice with various genotypes were injected with 0.25 μg ET (in 20 μl PBS) and the thicknesses of footpads were measured at 0, 8 and 20 h after injection. ET only induced modest oedema in $CMG2^{-/-}$ mice, but caused much higher levels of oedema in $CMG2^{+/-}$, $CMG2(EC)^{-/-}$, $CMG2(SM/CM)^{-/-}$ and $CMG2(SM/CM/EC)^{-/-}$ mice. The *P* values of the indicated groups versus $CMG2^{+/-}$ control group are shown. Each symbol represents one mouse. **b**–**f**, ET does not cause oedema in heart, spleen, kidney, lung and brain. Mice were challenged intravenously with 30 μg ET or PBS, and hearts (**b**), spleens (**c**), kidneys (**d**), brains (**e**) and lungs (**f**) were collected at 6 h or 18 h for tissue wet/dry ratio measurements. The *P* values of the indicated groups versus the PBS control group are shown. No significant differences were detected among the groups in **b**–**e**. In **f**, decreases in wet/dry ratio of lungs (dehydration) from ET-treated mice were observed. Each symbol represents one mouse. In **a**–**e**, error bars indicate s.e.; two-tailed unpaired *t*-test.



Extended Data Figure 9 | Mice with CMG2 deletion in cardiovascular system and intestines remain sensitive to ET. a, **b**, Sensitivity of *CMG2(CM)*$^{-/-}$, *CMG2(SM/CM/EC)*$^{-/-}$, *CMG2(SM/CM/EC/IE)*$^{-/-}$ and their littermate control mice to ET. Mice were challenged intravenously with 25 μg ET (**a**) or intraperitoneally with 50 μg ET (**b**) and survival monitored after challenge.

Extended Data Table 1 | Primers used for PCR genotyping and cloning

Use		Primer sequence	Size of PCR product (bp)
CMG2 genotyping	Forward1	5'GACTCTTAGGAAGGGTTCCTACTGG3'	WT=350 KO=500 Floxed=550
	Forward2	5'CCAATTTGGAGCTCAGGTTGGTGGA3'	
	Reverse	5'TGTAAGTCATATGGGTAGTGACCTAT3'	
General Cre genotyping	Forward	5'ATGTCCAATTTACTGACCGTACACC3'	600
	Reverse	5'CACCGTCAGTACGTGAGATATC3'	
Cdh-Cre genotyping	Forward	5'CTAGAATTGAGGTATGAGTTGAATACC3'	700
	Reverse	5'CACCGTCAGTACGTGAGATATC3'	
Myh6-Cre genotyping	Forward	5'ATGACAGACAGATCCCTCCTATCTCC3'	300
	Reverse	5'CTCATCACTCGTTGCATCATCGAC3'	
SM22-Cre genotyping	Forward	5'TGGTGAGCCAAGCAGACTTCCATGG3'	650
	Reverse	5'CACCGTCAGTACGTGAGATATC3'	
mCMG2 cDNA cloning	Forward	5'AAAAGAATTCGCCACCATGGTGGCCGGTCGGTCCCGGGCGCGCAGCCC TGGGAGCT3'	
	Reverse	5'AAAAGCTAGCTTAATTAATTATTGATGTGGAACCCGGGAGAAGTTTATGC3'	
CMG2 expression	Forward	5'GGAAGAGCAGTCACGTCGATCAGTCA3'	WT=451 KO=355
	Reverse	5'GACCTCCGTAGTAGGAAGCGT3'	
TEM8 expression	Forward	5'TGGCATGAAAGCTGCACTGCAGGTCAGCAT3'	410
	Reverse	5'CATATTCTTGCTCTGGCATCTTGACTCGTG3'	

LETTER

doi:10.1038/nature12493

A quantum access network

Bernd Fröhlich[1,2], James F. Dynes[1,2], Marco Lucamarini[1,2], Andrew W. Sharpe[1], Zhiliang Yuan[1,2] & Andrew J. Shields[1,2]

The theoretically proven security of quantum key distribution (QKD) could revolutionize the way in which information exchange is protected in the future[1,2]. Several field tests of QKD have proven it to be a reliable technology for cryptographic key exchange and have demonstrated nodal networks of point-to-point links[3–5]. However, until now no convincing answer has been given to the question of how to extend the scope of QKD beyond niche applications in dedicated high security networks. Here we introduce and experimentally demonstrate the concept of a 'quantum access network': based on simple and cost-effective telecommunication technologies, the scheme can greatly expand the number of users in quantum networks and therefore vastly broaden their appeal. We show that a high-speed single-photon detector positioned at a network node can be shared between up to 64 users for exchanging secret keys with the node, thereby significantly reducing the hardware requirements for each user added to the network. This point-to-multipoint architecture removes one of the main obstacles restricting the widespread application of QKD. It presents a viable method for realizing multi-user QKD networks with efficient use of resources, and brings QKD closer to becoming a widespread technology.

In a nodal QKD network, multiple trusted repeaters are connected by means of point-to-point links between a quantum transmitter ('Alice') and a quantum receiver ('Bob'). These point-to-point links can be realized with long-distance optical fibres, and in the future might even use ground-to-satellite communication[6–8]. Although point-to-point connections are suitable to form a backbone quantum core network to bridge long distances, they are less suitable to provide the last-mile service needed to give a multitude of users access to this QKD infrastructure. Reconfigurable optical networks based on optical switches or wavelength-division multiplexing have been suggested to achieve more flexible network structures[3,9–12]; however, they also require the installation of a full QKD system for each user, which is prohibitively expensive for many applications.

Giving a multitude of users access to the nodal QKD network requires point-to-multipoint connections. In modern fibre-optic networks point-to-multipoint connections are often realized passively by using components such as optical power splitters[13]. Single-photon QKD with the sender positioned at the network node and the receiver at the user premises[14] lends itself naturally to a passive multi-user network (see Fig. 1a). However, this downstream implementation has two major shortcomings. First, every user in the network requires a single-photon detector; these are often expensive and difficult to operate. Second, it is not possible to deterministically address a user. All detectors therefore have to operate at the same speed as the transmitter so as not to miss photons, which means that most of the detector bandwidth is unused.

Here we show that both problems associated with a downstream implementation can be overcome with a conceptual advancement: the most valuable resource should be shared by all users and should operate at full capacity. We propose and demonstrate an upstream quantum access network, in which the transmitters are placed at the end user

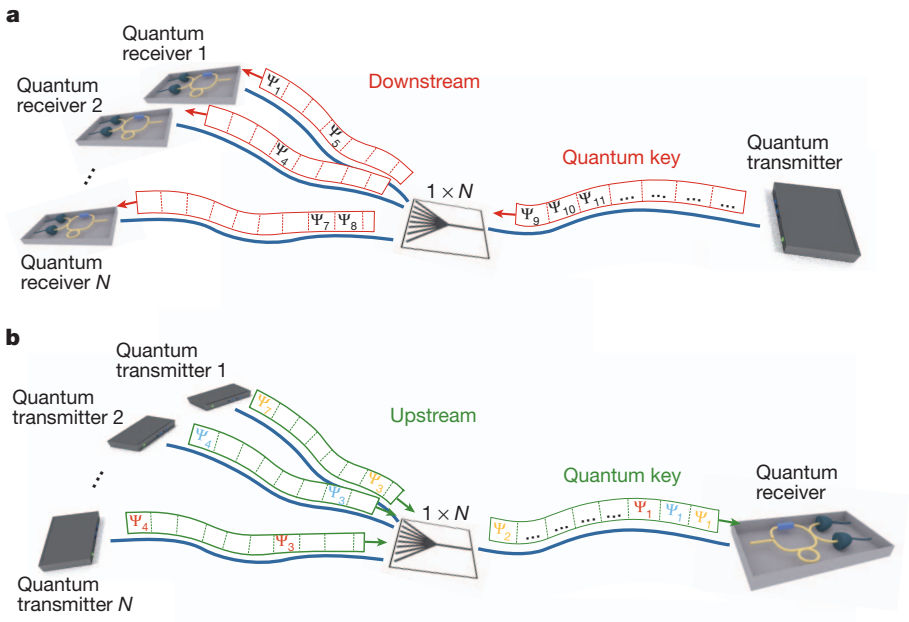


Figure 1 | Downstream and upstream quantum access network. **a**, In a downstream configuration the quantum transmitter is positioned at the network node. The transmitted quantum key is randomly directed to one of the quantum receivers by a passive optical splitter. Each user needs a single-photon detector, and the key is not distributed deterministically. **b**, The upstream configuration requires only a single detector at the network node. The quantum transmitters share this detector by ensuring that only photons from one transmitter at a time reach the receiver.

[1]Toshiba Research Europe Ltd, 208 Cambridge Science Park, Cambridge CB4 0GZ, UK. [2]Corporate Research and Development Center, Toshiba Corporation, 1 Komukai-Toshiba-Cho, Saiwai-ku, Kawasaki 212-8582, Japan.

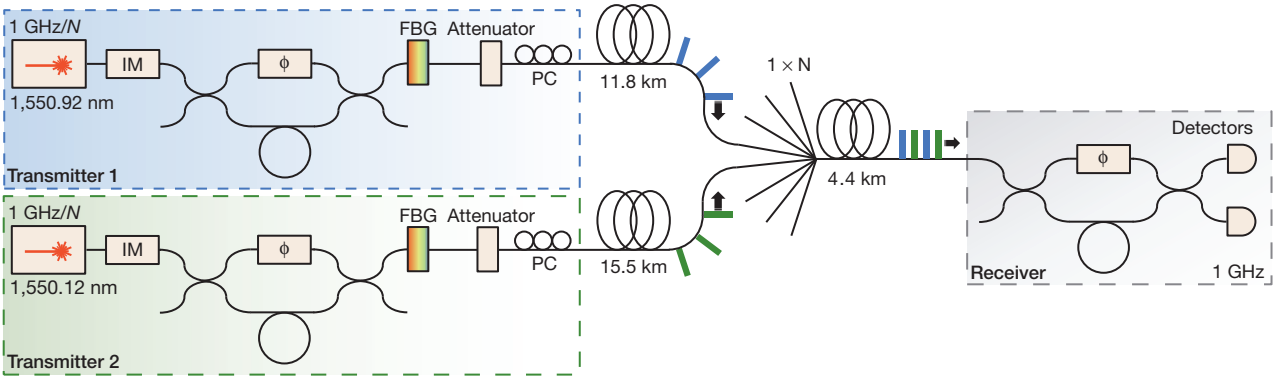


Figure 2 | Experimental set-up. Two quantum transmitters are connected to a single quantum receiver by means of a passive optical splitter and fibre spools. Each transmitter encodes bit and basis information on short laser pulses with an asymmetric Mach–Zehnder interferometer. The intensity of the pulses is modulated with an intensity modulator (IM) and attenuated to the single-photon level. A polarization controller (PC) pre-compensates the polarization and a fibre Bragg grating (FBG) compensates for pulse broadening in the fibre. The receiver decodes the phase information with a matching interferometer. The two outputs of this interferometer are connected to single-photon detectors operating at 1 GHz.

location and a common receiver is placed at the network node as shown in Fig. 1b. A careful study of the cross-talk between senders arising from the shared receiver topology shows that operation with up to 64 users is feasible, which we demonstrate by performing multi-user QKD over a 1×64 passive optical splitter. The results presented here highlight a practical and viable approach to extending the scope of QKD applications to many more users. Our approach would also be advantageous in a fully quantum network in which a quantum relay or repeater is located at the common node.

One of the main challenges for realizing an upstream quantum access network is to develop independently operating quantum transmitters that exchange secure keys efficiently with the receiver in parallel. For example, active stabilization in QKD systems is typically implemented at the receiver side[15]. In our scheme, however, the receiver is the reference for multiple transmitters and therefore each user has to pre-compensate system fluctuations individually. Figure 2 shows a diagram of our experimental set-up (see also Methods). We developed two flexible quantum transmitters that can operate at varying repetition rates and contain all the stabilization components necessary for continuous operation. They also include additional polarization control elements to achieve higher key rates (see Methods). At the centre of the quantum network is a passive optical splitter, which connects multiple transmitters to the receiver. The fibre distance between transmitters and receiver was chosen to be close to the maximum distance defined for gigabit passive optical networks in the International Telecommunication Union (ITU) standardization document[13]. We use a phase-encoding BB84 QKD protocol with decoy states[16–19] implemented with asymmetric Mach–Zehnder interferometers and intensity modulators. The quantum receiver decodes the phase information with a matching interferometer and uses two high-speed detectors based on avalanche photodiodes to detect the single photons with a rate of 1 GHz (refs 20, 21). We implement phase encoding because it is robust against fluctuations on the transmission channel and permits a simple stabilization mechanism.

To share a single-photon detector between multiple transmitters we adopt a novel and efficient scheme that allows the continuous and stable exchange of keys necessary to reduce the detrimental effect of finite-size samples[22–26]. Large fluctuations during a key session will reduce the number of sifted bits that are transmitted and therefore reduce the number of secure bits that can be distilled after privacy amplification. In our scheme all quantum transmitters operate continuously in parallel, permitting uninterrupted key sessions. We operate each transmitter at a fraction of the speed of the receiver, for example 1 GHz/$8 = 125$ MHz in an eight-user network. The transmitters are synchronized such that their pulses fall into subsequent detection time slots and can be clearly assigned to each user, as shown in Fig. 1b. This scheme has two main advantages: first, polarization, phase and synchronization tracking is done continuously by each QKD transmitter against the common quantum receiver, thus allowing stable operation of the quantum network; and, second, the transmitter can be realized with simpler electronics and optics because of the lower operational speed.

In a first experiment we demonstrate stable operation, over 12 h, of a 1×8 quantum access network populated by two users. We operate both transmitters at 125 MHz and use a passive 1×8 splitter to combine their signals; the total transmission losses including the splitter are 13.6 and 14 dB for transmitters 1 and 2, respectively. Counts in each detection gate are either allocated to one of the transmitters or identified as an empty gate, depending on their timing information. In a 20-min key session we record almost 300 Mbits of counts per transmitter. Figure 3a shows the quantum bit error rate (QBER) and secure bit rate for each key session. Although assigned to a specific transmitter, both the transmitter and the receiver subsystem contribute to the QBER. As a result of the low QBERs (transmitter 1, 1.28%; transmitter 2, 1.53%) we can exchange secure bits very efficiently, with average secure bit rates of 47.5 and 43.1 kbits s^{-1} for transmitters 1 and 2, respectively. Continuous operation over a month would allow unconditionally secure one-time-pad encryption of more than 10 Gbytes of data for each user, which is enough to protect more than 10^5 emails, for example.

The key rate can be increased further by the use of wavelength-division multiplexing optics instead of passive splitters because of the lower insertion loss of these devices. We demonstrate this in a second experiment by replacing the 1×8 splitter with an eight-channel thin-film dense-wavelength-division multiplexing module. The specified loss of the multiplexing module of about 2.5 dB is one-fifth of that for the 1×8 splitter and accordingly leads to a proportional increase in the count rate. In this experiment the quantum transmitters do not need to be modified because we designed the emission wavelengths of transmitters 1 and 2 to coincide with channels 33 (1,550.92 nm) and 34 (1,550.12 nm) of the ITU grid, respectively. Figure 3b shows the resulting QBER and secure bit rate for each key session. In addition to the higher transmission rate, the ratio of dark counts to photon counts decreases, leading to a decrease in the QBER (transmitter 1, 1.06%; transmitter 2, 1.17%) and therefore to an even higher increase in the secure bit rate (sixfold), corresponding to almost 100 Gbytes of key material per month (transmitter 1, 303 kbits s^{-1}; transmitter 2, 259 kbits s^{-1}).

We can extrapolate the performance of a network with more users by studying the cross-talk between two transmitters in detail. To determine the cross-talk we measured the average count rate of one transmitter with the other transmitter either on or off (C_{on} and C_{off}, respectively). As shown in the inset of Fig. 4a, the transmitters are operated at 125 MHz, allowing us to vary the gate separation between them. From these data we extract how many spurious detection events the second transmitter (green) causes in the detection gates allocated to the first transmitter

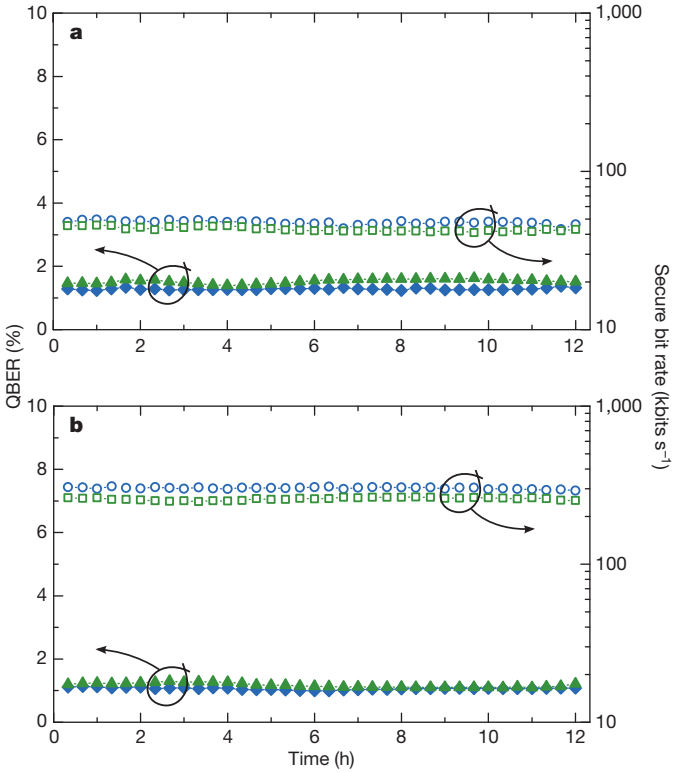


Figure 3 | Stable operation of quantum access network. **a**, Quantum bit error rate (QBER) (filled blue diamonds, transmitter 1; filled green triangles, transmitter 2) and secure bit rate (open blue circles, transmitter 1; open green squares, transmitter 2) for each 20-min key session in a network supporting eight users with a 1×8 passive optical splitter. **b**, QBER and secure bit rate for each 20-min key session in an eight-user network using dense-wavelength-division multiplexing optics.

(blue). Figure 4a displays the relative increase in count rate $(C_{on} - C_{off})/C_{off}$ for various gate separations of transmitters 1 and 2. Two effects contribute to the increase. First, the detection of pulses sent from the second transmitter causes after-pulses in the gate periods of the first transmitter. This leads to an increase in the count rate for each added user, independently of the gate separation by p_A/N as indicated by the dashed line, where p_A is the after-pulse probability of the detector and N is the number of users that the network supports. Second, smaller gate separation between transmitters 1 and 2 increases the cross-talk, most probably as a result of the late arrival of photons or ringing of the detector electronics.

Using the cross-talk data we simulate how the key rate of a single transmitter changes when more users are added to the network (see also Methods). Figure 4b shows logarithmic colour-scale plots of the secure bit rate per user as a function of fibre distance from the transmitters to the node and number of active users in the network. We simulate the rate for various network capacities N, which are given by the splitting ratio of the passive optical splitter installed in the system. For the simulation we assume that each active transmitter operates with 1 GHz/N and that the fibre distance to the node is the same for all transmitters. The data show clearly that even for a 64-user network, which inherently has a loss of about 20 dB from the 1×64 splitter, secure transmission is possible up to the maximum distance for gigabit passive optical networks of 20 km with all users active. In networks supporting fewer users there is a margin to allow longer fibre distances or, correspondingly, higher loss in the system.

To verify this result experimentally we switch to 500-MHz operation of the transmitters and vary the splitting ratio from 1×8 to 1×64. Operation at 500 MHz with two transmitters allows us to emulate a fully occupied network because photon detections are possible in all detector time slots. Figure 4c displays the estimated key rate per user based on the measured secure bit rates of transmitters 1 and 2, and shows how it compares with the expected value from the simulation. For the key rate estimation we add the key rate of the two transmitters and divide it by the network capacity. The data confirm

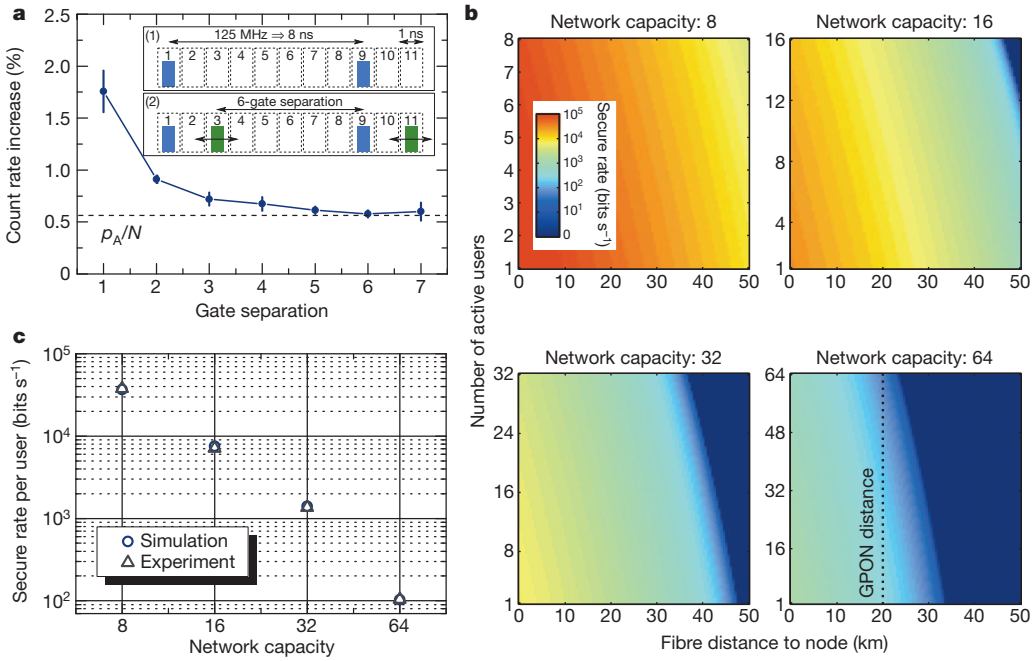


Figure 4 | Quantum access network with varying capacity. **a**, Relative increase of count rate due to cross-talk between the transmitters as a function of gate separation. The dashed line indicates the limit given by detector after-pulses. Error bars correspond to one standard deviation of three consecutive measurements. Inset: measurement principle (see the text). **b**, Simulation of the secure bit rate per user as a function of fibre distance and number of active users in the network for various network capacities (see the text and Methods). **c**, Secure bit rate per user in a quantum network with varying capacity estimated from a two-user measurement at 500 MHz.

the result obtained from the simulation: it demonstrates that a 64-user network is feasible with our scheme.

We have demonstrated that passive optical networks have the potential to scale up the number of users in a nodal QKD network. We have shown that the decrease in the secure key rate accompanying a time-division multiplexing approach can be greatly mitigated by using a high-speed single-photon detector. The network node in our scheme acts as a receiver and has to be trusted intrinsically; however, techniques such as classical secret sharing[27] or measurement-device-independent QKD[28] might be used to relax this requirement in the future. It might also be possible to combine classical data transport on the same fibre in quantum-secured access networks[29]. Quantum access networks could initially find application in protecting smart community or smart grid networks, for example, allowing authenticated data collection from multiple locations in a critical infrastructure network[30].

METHODS SUMMARY

Experimental set-up. We implement the standard BB84 protocol with decoy states[16–19], with photon fluxes of 0.5, 0.1 and 0.0002 photons per pulse sent with probabilities of 98.83%, 0.78% and 0.39%, respectively. Detection events are sorted into N time bins, where N is the number of users that the network supports. Each transmitter is aligned relative to a master clock from the receiver such that their photons fall into a specific time bin. We determine the secure bit rate for each user individually by estimating single-photon parameters from decoy states. Our security analysis[26] takes finite-size effects into account[23–25] and achieves bit rates close to the asymptotic limit for key sessions of 20 min.

Simulation. We first estimate the probability of obtaining an error count for a single transmitter in an otherwise empty network η_{err} on the basis of measured values for encoding imperfections and detector imperfections (after-pulses and dark counts). The QBER including cross-talk counts is then given by

$$e=\frac{\eta_{\mathrm{err}}+\frac{1}{2}p_{\mathrm{X}}(n-1)\eta}{\eta(1+p_{\mathrm{A}}/N)+p_{\mathrm{D}}+p_{\mathrm{X}}(n-1)\eta}$$

where n is the number of active users populating the network, η is the system detection probability, p_{A} is the after-pulse probability and p_{D} is the dark count probability. The average count rate increase per added user p_{X} is extracted from the data shown in Fig. 4a under the assumption of equal losses for all users. The secure bit rate is determined from a refined analysis calculating the count rates and QBER for all three signal levels of the decoy-state protocol. For the 32-user and 64-user networks we increased the key session length from 20 min to 2 h and 12 h, respectively, to compensate for the decreasing sample size due to the slower operation of the transmitter. Our measurements shown in Fig. 3 indicate that longer key sessions are feasible. For the comparison of simulation and experiment shown in Fig. 4c we adopt the simulation to account for the unequal fibre distances of the two transmitters, as well as the higher operational speed leading to different sample sizes compared with Fig. 4b.

Full Methods and any associated references are available in the online version of the paper.

Received 1 May; accepted 18 July 2013.

1. Lütkenhaus, N. & Shields, A. J. Focus on quantum cryptography: theory and practice. *New J. Phys.* **11,** 045005 (2009).
2. Scarani, V. *et al.* The security of practical quantum key distribution. *Rev. Mod. Phys.* **81,** 1301–1350 (2009).
3. Elliott, C. *et al.* in *Quantum Information and Computation III* (*Proc. SPIE*, vol. 5815) (eds Donkor, E. J., Pirich, A. R. & Brandt, H. E.) 138–149 (SPIE, 2005).
4. Peev, M. *et al.* The SECOQC quantum key distribution network in Vienna. *New J. Phys.* **11,** 075001 (2009).
5. Sasaki, M. *et al.* Field test of quantum key distribution in the Tokyo QKD Network. *Opt. Express* **19,** 10387–10409 (2011).
6. Ursin, R. *et al.* Entanglement-based quantum communication over 144 km. *Nature Phys.* **3,** 481–486 (2007).
7. Nauerth, S. *et al.* Air-to-ground quantum communication. *Nature Photon.* **7,** 382–386 (2013).
8. Wang, J.-Y. *et al.* Direct and full-scale experimental verifications towards ground-satellite quantum key distribution. *Nature Photon.* **7,** 387–393 (2013).
9. Toliver, P. *et al.* Experimental investigation of quantum key distribution through transparent optical switch elements. *IEEE Photon. Technol. Lett.* **15,** 1669–1671 (2003).
10. Chapuran, T. E. *et al.* Optical networking for quantum key distribution and quantum communications. *New J. Phys.* **11,** 105001 (2009).
11. Chen, T.-Y. *et al.* Metropolitan all-pass and inter-city quantum communication network. *Opt. Express* **18,** 27217–27225 (2010).
12. Wang, S. *et al.* Field test of the wavelength-saving quantum key distribution network. *Opt. Lett.* **35,** 2454–2456 (2010).
13. International Telecommunication Union. G.984.1: Gigabit-capable passive optical networks (GPON): general characteristics. http://www.itu.int/rec/T-REC-G.984.1-200803-I/en (2008).
14. Townsend, P. D. Quantum cryptography on multiuser optical fibre networks. *Nature* **385,** 47–49 (1997).
15. Dixon, A. R., Yuan, Z. L., Dynes, J. F., Sharpe, A. W. & Shields, A. J. Continuous operation of high bit rate quantum key distribution. *Appl. Phys. Lett.* **96,** 161102 (2010).
16. Hwang, W.-Y. Quantum key distribution with high loss: toward global secure communication. *Phys. Rev. Lett.* **91,** 057901 (2003).
17. Wang, X.-B. Beating the photon-number-splitting attack in practical quantum cryptography. *Phys. Rev. Lett.* **94,** 230503 (2005).
18. Lo, H.-K., Ma, X. & Chen, K. Decoy state quantum key distribution. *Phys. Rev. Lett.* **94,** 230504 (2005).
19. Ma, X., Qi, B., Zhao, Y. & Lo, H.-K. Practical decoy state for quantum key distribution. *Phys. Rev. A* **72,** 012326 (2005).
20. Yuan, Z. L., Kardynal, B. E., Sharpe, A. W. & Shields, A. J. High speed single photon detection in the near infrared. *Appl. Phys. Lett.* **91,** 041114 (2007).
21. Yuan, Z. L. *et al.* Gigahertz quantum key distribution with InGaAs avalanche photodiodes. *Appl. Phys. Lett.* **92,** 201104 (2008).
22. Hayashi, M. Upper bounds of eavesdropper's performances in finite-length code with the decoy method. *Phys. Rev. A* **76,** 012329 (2007).
23. Scarani, V. & Renner, R. Quantum cryptography with finite resources: unconditional security bound for discrete-variable protocols with one-way postprocessing. *Phys. Rev. Lett.* **100,** 200501 (2008).
24. Scarani, V. & Renner, R. Security bounds for quantum cryptography with finite resources. Preprint at http://arxiv.org/abs/0806.0120 (2008).
25. Cai, R. Y. Q. & Scarani, V. Finite-key analysis for practical implementations of quantum key distribution. *New J. Phys.* **11,** 045024 (2009).
26. Lucamarini, M., Dynes, J. F., Yuan, Z. L. & Shields, A. J. in *Electro-Optical Remote Sensing, Photonic Technologies, and Applications VI* (*Proc. SPIE*, vol. 8542) (eds Kamerman, G. W. *et al.*) 85421K (SPIE, 2012).
27. Barnett, S. & Phoenix, S. J. D. in *GCC Conference and Exhibition (GCC), 2011 IEEE* 143–145, http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=5746659 (IEEE, 2011).
28. Lo, H.-K., Curty, M. & Qi, B. Measurement-device-independent quantum key distribution. *Phys. Rev. Lett.* **108,** 130503 (2012).
29. Patel, K. A. *et al.* Coexistence of high-bit-rate quantum key distribution and data on optical fiber. *Phys. Rev. X* **2,** 041010 (2012).
30. Hughes, R. J. *et al.* Network-centric quantum communications with application to critical infrastructure protection. Preprint at http://arxiv.org/abs/1305.0305 (2013).

Acknowledgements This research is partly supported by Research and Development of Secure Photonic Network Technologies, the Commissioned Research of the National Institute of Information and Communications Technology (NICT), Japan.

Author Contributions B.F. performed the measurements and simulations. B.F., J.F.D. and A.W.S. developed the system. M.L. performed calculations for the security analysis. Z.Y. and A.J.S. conceived the experiment and guided the work. B.F. wrote the manuscript with contributions from the other authors. All authors discussed experiments, results and the interpretation of results.

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.J.S. (andrew.shields@crl.toshiba.co.uk), Z.Y. (zhiliang.yuan@crl.toshiba.co.uk) or B.F. (bernd.frohlich@crl.toshiba.co.uk).

METHODS

Experimental set-up. Each quantum transmitter consists of a source of short laser pulses, an intensity modulator, an asymmetric Mach–Zehnder interferometer including a phase modulator in one arm, a fibre Bragg grating, an attenuator and a polarization controller (see Fig. 2). The laser source is a distributed feedback laser generating laser pulses shorter than 50 ps with a selectable repetition rate of up to 1 GHz. The wavelengths of transmitters 1 and 2 are tuned to coincide with channels 33 and 34, respectively, of the grid defined by the International Telecommunication Union (ITU). The intensity modulator in combination with the attenuator at the output sets the power of each pulse to one of three power levels necessary for the decoy protocol[16–19]: 0.5 photons per pulse for signal pulses, 0.1 photons per pulse for decoy pulses and 0.0002 photons per pulse for vacuum pulses. We send signal, decoy and vacuum pulses with probabilities of 98.83%, 0.78% and 0.39%, respectively.

The asymmetric Mach–Zehnder interferometer is made of off-the-shelf fibre optic components, and the length difference between the long and short arms is matched to the receiver interferometer with a tunable optical delay. The components do not require temperature stabilization or vibration isolation. Drifts of the relative phase between the long and short arms of the interferometer with respect to the receiver interferometer are compensated for by applying a bias voltage across the phase modulator. Each transmitter compensates for the phase difference individually and independently of other users in the network. No compensation is necessary in the receiver interferometer, which acts as a phase reference. The output beam splitter of the transmitter interferometer and the input beam splitter of the receiver interferometer are polarizing to direct the photons into the correct arm of the receiver interferometer, thus avoiding a 3 dB penalty when using polarization-insensitive interferometers. We therefore additionally pre-compensate the polarization of the transmitted pulses with a polarization controller to achieve maximum count rate at the receiver. The fibre Bragg grating compensates for pulse broadening across the fibre link to avoid a further 1.5 dB decrease in the count rate.

The quantum receiver consists of the reference Mach–Zehnder interferometer that decodes the phase information of the pulses sent from the transmitters and detects the photons with InGaAs single-photon avalanche photodiodes with a rate of 1 GHz using a self-differencing technique[20,21]. The avalanche photodiodes are operated at a temperature of $-30\,^{\circ}\mathrm{C}$. A typical detection efficiency is 15%, with 8×10^{-6} dark counts per gate and an after-pulse probability of 4.5%. Detection events are sorted into N time bins, where N is the number of users that the network supports. Each transmitter is aligned relative to a master clock from Bob such that their photons coincide with a specific time bin and can be clearly assigned to a user. A variable time delay implemented in each transmitter permits control of the alignment continuously with a feedback signal generated from the detector count rate.

Secure key rate. We implement the standard BB84 protocol with decoy states in our set-up. The quantum transmitter prepares one of four phase states (0, π/2, π and 3π/2) with equal probabilities, and the receiver chooses either phase 0 or π/2. All events with non-matching basis are discarded in the sifting process. On the basis of the individual error rates of signal, decoy and vacuum states we estimate single-photon parameters for each user individually. Our security analysis[26] takes finite-size effects into account[23–25] and achieves bit rates close to the asymptotic limit for key sessions of 20 min. The secure key rate is bounded from below by

$$R=\{Q_1[1-H(e_1)]-Qf_{\mathrm{EC}}(e)H(e)+Q_0-\Delta\}/t$$

Here Q_1 is the estimated number of sifted bits from single-photon states, $H(e_1)$ is the binary entropy function of the estimated error rate of those bits, Q is the total number of sifted bits, f_{EC} is the error correction efficiency, which is set to 1.1, e is the QBER of sifted bits, Q_0 is the estimated number of sifted bits originating from vacuum pulses and t is the key session time. Finite-size effects are included by subtracting Δ, which is proportional to $\sqrt{Q}$ and to $\log_2(\varepsilon^{-1})$, where ε, equal to 10^{-10} in our system, is related to the overall security of the system[26].

Parameters of the protocol such as decoy level and decoy probability have to be chosen carefully to achieve optimal secure key rates. For example, the estimation of Q_1 depends directly on the chosen decoy photon flux and probability. We simulate the achievable secure key rates in advance to select suitable parameters for the experiment. For convenience we use one set of parameters for all measurements presented here, which we found to lead to stable results in all configurations considered.

Simulation. We simulate the secure key rate per user in a network that is populated by more than two users based on measured experimental parameters. For the simulation we assume that each active transmitter operates with 1 GHz/N and that the fibre distance to the node is the same for all transmitters. The starting point is calculating the probability to obtain an error count for a single transmitter in an otherwise empty network, using $\eta_{\mathrm{err}}=\eta(e_{\mathrm{opt}}+p_{\mathrm{A}}/2N)+p_{\mathrm{D}}/2$. Here e_{opt} is the optical error due to encoding imperfections, p_{A} is the after-pulse probability and p_{D} is the dark count probability of the detector, and the detection probability η is given by $\eta=\mu\times10^{-0.2L/10}\,l_{\mathrm{spl}}\eta_{\mathrm{Bob}}$, with μ being the photon flux, L the total fibre length in kilometres, l_{spl} the splitter loss and η_{Bob} the system detection efficiency of the receiver. We use the following parameters in the simulation: $e_{\mathrm{opt}}=0.5\%$, $p_{\mathrm{A}}=4.5\%$, $p_{\mathrm{D}}=2\times8\times10^{-6}$ and $\eta_{\mathrm{Bob}}=9.04\%$.

Adding more users to the network will increase the error rate as a result of cross-talk between the users. The QBER including cross-talk counts is then given by

$$e=\frac{\eta_{\mathrm{err}}+\frac{1}{2}p_{\mathrm{X}}(n-1)\eta}{\eta(1+p_{\mathrm{A}}/N)+p_{\mathrm{D}}+p_{\mathrm{X}}(n-1)\eta}$$

with n the number of active users in the network. The average increase in count rate per added user $p_{\mathrm{X}}=1.9\%/(N-1)+p_{\mathrm{A}}/N$ is extracted from the data shown in Fig. 4a by taking both the base increase by p_{A}/N and the enhanced cross-talk at short gate separation into account. For the splitter loss l_{spl} we use our measured values of 9.7, 13, 16.1 and 19.5 dB for 1×8, 1×16, 1×32 and 1×64 splitting ratios, respectively.

The secure key rate is determined from a refined analysis calculating the count rates and QBER for all three signal levels of the decoy-state protocol. We use the same routine as for the experimental data to determine the secure key rate from these values. For the 32-user and 64-user networks we increased the key session length from 20 min to 2 h and 12 h, respectively, to compensate for the decreasing sample size due to the slower operation of the transmitter. Our measurements shown in Fig. 3 indicate that longer key sessions are feasible. For the comparison of simulation and experiment shown in Fig. 4c we adopt the simulation to account for the unequal fibre distances of the two transmitters, as well as the higher operational speed leading to different sample sizes compared with Fig. 4b.

LETTER

doi:10.1038/nature12421

Microscopic origin of the '0.7-anomaly' in quantum point contacts

Florian Bauer[1,2]*, Jan Heyder[1,2]*, Enrico Schubert[1], David Borowsky[1], Daniela Taubert[1], Benedikt Bruognolo[1,2], Dieter Schuh[3], Werner Wegscheider[4], Jan von Delft[1,2] & Stefan Ludwig[1]

Quantum point contacts are narrow, one-dimensional constrictions usually patterned in a two-dimensional electron system, for example by applying voltages to local gates. The linear conductance of a point contact, when measured as function of its channel width, is quantized[1–3] in units of $G_Q = 2e^2/h$, where e is the electron charge and h is Planck's constant. However, the conductance also has an unexpected shoulder at $\sim 0.7 G_Q$, known as the '0.7-anomaly'[4–12], whose origin is still subject to debate[11–21]. Proposed theoretical explanations have invoked spontaneous spin polarization[4,17], ferromagnetic spin coupling[19], the formation of a quasi-bound state leading to the Kondo effect[13,14], Wigner crystallization[16,20] and various treatments of inelastic scattering[18,21]. However, explicit calculations that fully reproduce the various experimental observations in the regime of the 0.7-anomaly, including the zero-bias peak that typically accompanies it[6,9–11], are still lacking. Here we offer a detailed microscopic explanation for both the 0.7-anomaly and the zero-bias peak: their common origin is a smeared van Hove singularity in the local density of states at the bottom of the lowest one-dimensional subband of the point contact, which causes an anomalous enhancement in the Hartree potential barrier, the magnetic spin susceptibility and the inelastic scattering rate. We find good qualitative agreement between theoretical calculations and experimental results on the dependence of the conductance on gate voltage, magnetic field, temperature, source–drain voltage (including the zero-bias peak) and interaction strength. We also clarify how the low-energy scale governing the 0.7-anomaly depends on gate voltage and interactions. For low energies, we predict and observe Fermi-liquid behaviour similar to that associated with the Kondo effect in quantum dots[22]. At high energies, however, the similarities between the 0.7-anomaly and the Kondo effect end.

In our measurements, we use the multigate layout on the surface of a GaAs/AlGaAs heterostructure shown in Fig. 1a. By suitably tuning the central- and side-gate voltages, V_c and V_s, at a fixed top-gate voltage, V_t, we can use the device to define a short, one-dimensional (1D) channel, containing a smooth, symmetric barrier, in the two-dimensional electron system (2DES) buried in the heterostructure. To describe such a quantum point contact (QPC), we adopt a 1D model with local interactions and a smooth potential barrier. We treat interactions perturbatively, using either second-order perturbation theory[23] (SOPT) or the functional renormalization group[24–26] (FRG) approach (Supplementary Information, sections 7 and 6, respectively). The lowest 1D subband of the device is modelled by

$$\hat{H} = \sum_{j\sigma} \left[E_{j\sigma} \hat{n}_{j\sigma} - \tau_j \left(d^\dagger_{j+1\,\sigma} d_{j\sigma} + \text{h.c.} \right) \right] + \sum_j U_j \hat{n}_{j\uparrow} \hat{n}_{j\downarrow} \qquad (1)$$

Here $\hat{n}_{j\sigma} = d^\dagger_{j\sigma} d_{j\sigma}$ counts the number of electrons with spin σ (spin up, $\sigma = \uparrow$ or $+$; spin down, $\sigma = \downarrow$ or $-$) at site j of an infinite, tight-binding chain with hopping amplitude τ_j, on-site interaction U_j and potential energy $E_{j\sigma} = E_j - \sigma\tilde{B}/2$ (Supplementary Fig. 8), and 'h.c.' denotes Hermitian conjugate. The Zeeman energy, $\tilde{B} = |g_{el}|\mu_B B$, describes the effect of a uniform external parallel magnetic field B, where μ_B is the Bohr magneton and g_{el} is the effective g factor (<0 in GaAs). (When similar symbols are used for model parameters and experimental parameters, we add tildes to the former to distinguish them from the latter.) We neglect spin–orbit interactions and other orbital effects. The parameters E_j, U_j and τ_j vary smoothly with j and differ from their bulk values, $E_{bulk} = U_{bulk} = 0$ and $\tau_{bulk} = \tau$ (taken as the unit of energy), only within a central constriction region (CCR) of N sites around $j = 0$, representing the QPC. Sites $j < -N/2$ and $j > N/2$ represent two non-interacting leads, each with bandwidth 4τ, chemical potential μ and bulk Fermi energy $\varepsilon_F = 2\tau + \mu$; we choose $\mu = 0$, implying half-filled leads (Fig. 1b). We set U_j to a fixed value, U, for all but the outermost sites of the CCR, where it drops smoothly to zero.

Within the CCR, we define the QPC barrier by specifying the shape of the 'band bottom' as $\omega_j^{min} = E_j - (\tau_{j-1} + \tau_j) - \mu$ (Fig. 1b, solid black line). We choose ω_j^{min} to define a smooth, symmetric barrier within the CCR, parabolic near the top[3], where we parameterize it as $\omega_j^{min} \approx \tilde{V}_c - \Omega_x^2 j^2 / 4\tau_0$ (Supplementary Information, section 4D). Here $\tilde{V}_c$ sets the barrier height with respect to μ (Fig. 1b, dashed black line), and $\Omega_x \ll \tau$ characterizes its curvature. We first consider the theoretical case of zero temperature, $\tilde{T} = k_B T$ (k_B, Boltzmann's constant), source–drain voltage, $\tilde{V}_{sd} = |e| V_{sd}$, and field, $\tilde{B}$: $\tilde{T} = \tilde{V}_{sd} = \tilde{B} = 0$. As $\tilde{V}_c$ is decreased below 0, the conductance, $g = G/G_Q$, increases from 0 to 1, showing a step of width $\sim\Omega_x$ (about 1.5 meV in our experiment), whose shape depends on U (Fig. 1k). In the upper part of the step, say $0.5 \lesssim g \lesssim 0.9$, we say that the QPC is 'sub-open'; the sub-open regime is of special interest because for measured $g(V_c)$ curves it contains the 0.7-anomaly.

The bare local density of states (LDOS), $A_j^0(\omega)$, for equation (1) has a strong maximum just above the band bottom[18], seen as a yellow–red ridge-like structure in Fig. 1b. In a semiclassical picture, $A_j^0(\omega) \propto 1/v_j(\omega)$, where $v_j(\omega)$ is the velocity at site j of an electron with energy ω with respect to μ. The ridge-like maximum of $A_j^0(\omega)$ above the barrier reflects the fact that electrons move slowest there. In the CCR's outer flanks, this ridge develops smoothly into the van Hove singularity, $A_{bulk}^0 \propto \left[(\omega - \omega_{bulk}^{min})\tau \right]^{-1/2}$, in the bulk LDOS at the bulk band bottom in the leads, $\omega_{bulk}^{min} = -\varepsilon_F$. We therefore call this LDOS structure a 'van Hove ridge'. Near the barrier's centre, its curvature causes the singularity to be smeared out on a scale set by Ω_x. This limits the amplitude of the van Hove ridge to $\max\left[A_j^0(\omega)\right] \propto \mathcal{O}(\Omega_x \tau_0)^{-1/2}$ and shifts it upwards in frequency relative to the band by $\mathcal{O}(\Omega_x)$ (Fig. 1f–h).

The van Hove ridge has a strong, $\tilde{V}_c$-dependent effect on numerous QPC properties. Near those spatial locations where the ridge intersects the chemical potential ($\omega = 0$), the LDOS is enhanced, thus amplifying the effects of interactions by $\mathcal{O}(\Omega_x \tau_0)^{-1/2}$ (which grows with QPC

[1]Center for NanoScience and Fakultät für Physik, Ludwig-Maximilians-Universität München, Geschwister-Scholl-Platz 1, 80539 München, Germany. [2]Arnold Sommerfeld Center for Theoretical Physics, Ludwig-Maximilians-Universität München, Theresienstrasse 37, D-80333 München, Germany. [3]Institut für Angewandte Physik, Universität Regensburg, D-93040 Regensburg, Germany. [4]Laboratory for Solid State Physics, ETH Zürich, CH-8093 Zürich, Switzerland.
*These authors contributed equally to this work.

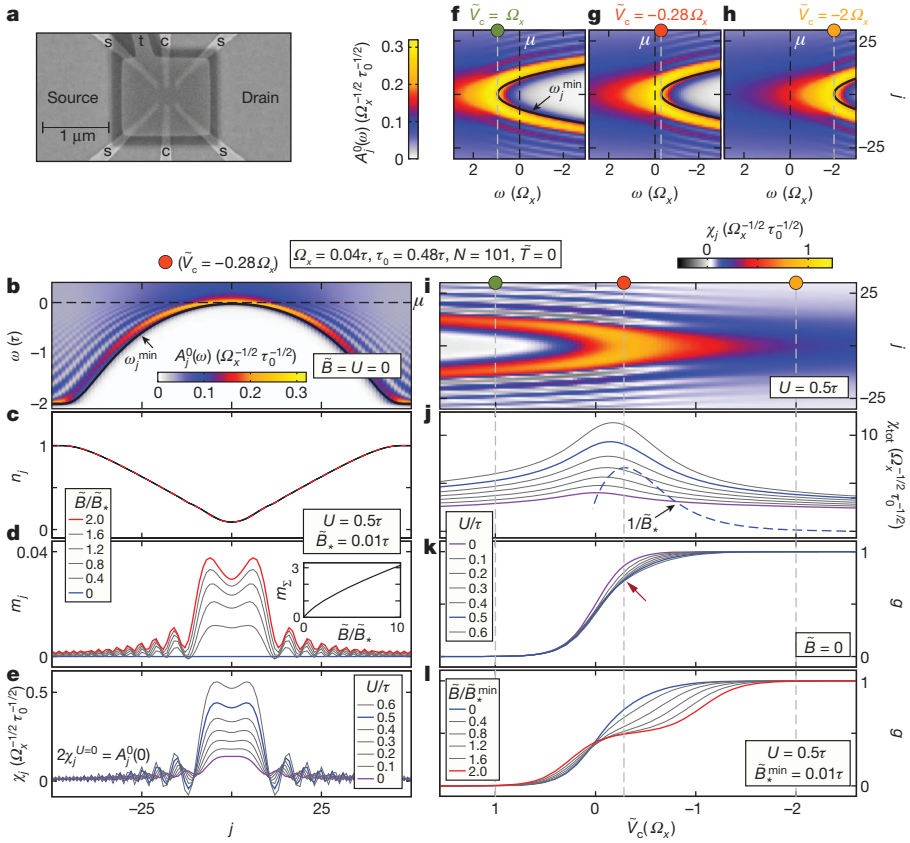


Figure 1 | Experimental set-up and model. **a**, Scanning electron microscope picture of the gate layout, featuring a top gate (t) at voltage V_t, two central gates (c) at voltage V_c and four side gates (s) at voltage V_s. Negative voltages V_c and V_s locally deplete the 2DES, which is 85 nm beneath the sample surface. Together with V_t, they induce a tunable electrostatic potential landscape in the 2DES. **b**, Barrier shape and LDOS. The bare ($U_j = 0$, $\tilde{B} = 0$) 1D LDOS per spin species, $A_j^0(\omega)$ (colour scale), as a function of energy, ω, and site index, j, for $\tilde{V}_c = -0.28\Omega_x$. The barrier shape is defined by the solid black line, showing the band bottom, ω_j^{min}. The LDOS vanishes exponentially rapidly below ω_j^{min} (Supplementary Fig. 11), and has a van Hove ridge (yellow–red) just above it, followed by Friedel oscillations (white fringes) at higher energies (up to $\omega \lesssim \tilde{V}_c$). **c**–**e**, Local properties of a sub-open QPC: FRG results for the sub-open barrier shown in **b**. **c**, **d**, The local density, n_j (**c**), and the magnetization, m_j (**d**), for several values of magnetic field, $\tilde{B}$. Inset of **d**, $m_\Sigma = \sum_{|j|\le 10} m_j$ as a function of $\tilde{B}$. **e**, The local spin susceptibility, χ_j, for several values of interaction strength, U. The shapes of m_j and χ_j are modulated by Friedel oscillations inherited from the bare LDOS (**b**), with locally varying wavelength, $\lambda \approx 1/n_j$. **f**–**l**, Changing barrier height. **f**–**h**, The bare LDOS, $A_j^0(\omega)$, for three successively lower barrier heights, $\tilde{V}_c/\Omega_x = 1$ (**f**), -0.28 (**g**) and -2 (**h**). The LDOS pattern is fixed with respect to $\tilde{V}_c$ (grey dashed lines) but shifts with respect to μ (black dashed lines). **i**–**l**, FRG results for the $\tilde{V}_c$ dependence of the local spin susceptibility, χ_j (colour scale), at fixed $U = 0.5\tau$ (**i**); the total spin susceptibility, $\chi_{tot} = \sum_j^{CCR} \chi_j$, for several U values (solid lines), and the inverse low-energy scale, $1/\tilde{B}_*$, for $U = 0.5\tau$ (dashed line) (**j**); the zero-temperature linear-response ($V_{sd} = 0$) conductance, $g = G/G_Q$, for several U values (at fixed $\tilde{B} = 0$) (**k**) and for several $\tilde{B}$ values (at fixed $U = 0.5\tau$) (**l**). For a large enough interaction, $U = 0.5\tau$, even for $\tilde{B} = \tilde{T} = \tilde{V}_{sd} = 0$ (blue lines in **k** and **l**), $g(\tilde{V}_c)$ has a shoulder (red arrow) at $g \approx 0.7$, the 0.7-anomaly. Three vertical dashed lines in **i**–**l** mark the three $\tilde{V}_c$ values used in **f**–**h**, as indicated by dots of matching colours.

length). In semiclassical terms, slow electrons feel interactions particularly strongly. When lowering the barrier top, $\tilde{V}_c$, to open the QPC, the van Hove ridge sweeps downwards (Fig. 1f–h); its interaction-amplifying effects are strongest in the $\tilde{V}_c$ regime where its apex, which has most weight, crosses μ. This happens for $0 \gtrsim \tilde{V}_c \gtrsim -\mathcal{O}(\Omega_x)$ (Fig. 1g), which, very importantly, encompasses the sub-open regime containing the 0.7-anomaly. Below, we show that the 0.7-anomaly and the zero-bias peak (ZBP) stem precisely from the amplification of interaction effects where the van Hove ridge intersects μ. The relevant implications are enhancements in the effective Hartree barrier governing elastic transmission, the spin susceptibility and the inelastic scattering rate, all of which lead to an anomalous reduction of g in the sub-open regime, especially for $T, B, V_{sd} > 0$.

Figure 1c–e illustrates several local properties, calculated at $\tilde{T} = 0$ using FRG, for the sub-open QPC barrier shown in Fig. 1b. We note four salient features, all intuitively expected. First, the local density, $n_j = \langle \hat{n}_{j\uparrow} + \hat{n}_{j\downarrow} \rangle$, is minimal at the barrier centre (Fig. 1c). Second, the local magnetization, $m_j = \langle \hat{n}_{j\uparrow} - \hat{n}_{j\downarrow} \rangle / 2$, vanishes at $B = 0$ (Fig. 1d, blue line); this reflects a physical assumption entailed in our calculations (Supplementary Information, section 6), namely that no spontaneous magnetization occurs, in contrast to the spontaneous spin splitting scenario advocated in refs 4, 8, 17. Third, m_j increases without saturation when $\tilde{B}$ becomes large (Fig. 1d, inset), indicating a smooth redistribution of spin, as expected for an open structure. Fourth, the local spin susceptibility, $\chi_j = (\partial m_j / \partial \tilde{B})_{\tilde{B}=0}$, is strongly enhanced with increasing U (Fig. 1e), because interactions amplify any field-induced spin imbalance.

The j dependence of χ_j is governed by that of $A_j^0(0)$ (in fact, $\chi_j^{U=0} = A_j^0(0)/2$), which is maximal near those sites where the van Hove ridge intersects μ. When $\tilde{V}_c$ is decreased through 0 (Figs 1f–h), these intersection points sweep out a parabolic arch in the $\tilde{V}_c - j$ plane, along which $\chi_j(\tilde{V}_c)$ (Fig. 1i, colour scale) is peaked, with most weight near the arch's apex. This leads to a corresponding peak in the total spin susceptibility, $\chi_{tot} = \sum_j^{CCR} \chi_j$, as a function of $\tilde{V}_c$ (Fig. 1j). This peak is strongly enhanced by increasing U (in accordance with the fourth feature above) and is located near the $\tilde{V}_c$ value where $g \approx 0.7$ (Fig. 1k). We will see further below that this peak strongly affects the $\tilde{B}$ dependence of the conductance (Fig. 1l).

Note that the spatial structure for $\chi_j(\tilde{V}_c)$ in Fig. 1i, namely two peaks merging into one as $\tilde{V}_c$ is lowered, is consistent with that, shown

in fig. 2b of ref. 14, for the density of spin-up electrons calculated using spin-density-functional theory, initialized in a small applied field to break spin symmetry. In ref. 14, the local maximum in the spin-up density was interpreted as evidence for a 'quasi-bound state' that was argued to host a spin-1/2 local moment; in contrast, features one and, especially, three above imply that our model yields no local moment.

Next we discuss the effect of the van Hove ridge on the conductance, $g(\tilde{V}_c)$, starting with its U dependence at $\tilde{B}=\tilde{T}=0$ (Fig. 1k). Increasing U skews the shape of the step in $g(\tilde{V}_c)$, which eventually develops a shoulder near $g\approx 0.7$ (red arrow). This shoulder develops because the increase in local density with decreasing $\tilde{V}_c$ is slightly nonlinear when the apex of the van Hove ridge drops past μ, causing a corresponding nonlinear upward shift in the effective Hartree barrier. For a parabolic barrier top, this occurs for $g\approx 0.7$. If the shape of the barrier top is changed to be non-parabolic, both the shape of the bare conductance step and the energy distance between the van Hove ridge apex and μ will change, which can cause the interaction-induced shoulder in g to shift away from 0.7. This explains the experimentally observed spread[6,12] of shoulders (that is, plateau values of the 0.7-anomaly) for $0.5 \lesssim g \lesssim 1$.

On increasing $\tilde{B}$ for fixed U and $\tilde{T}=0$ (Figs 1l and 2a), the shoulder in $g(\tilde{V}_c)$ becomes more pronounced, eventually developing into a spin-split plateau. Comparison of Fig. 2a with Fig. 2e shows that this development qualitatively agrees with experiment; the agreement was optimized by using U as fit parameter. Inspecting how the corresponding spin-resolved conductances, $g_\uparrow$ and $g_\downarrow$, change with $\tilde{B}$ (Fig. 2b), we note a strong asymmetry: although the bare barrier heights for spins $\uparrow$ and $\downarrow$ are shifted symmetrically by $-\tilde{B}/2$ and $\tilde{B}/2$, respectively, $g_\downarrow$ is decreased much more strongly than $g_\uparrow$ is increased. This is due to exchange interactions: increasing the spin-up density near the CCR centre (Fig. 1d) strongly raises the Hartree barrier, and more so for spin-down electrons than spin-up, owing to Pauli's exclusion principle. The consequences are most pronounced in the sub-open regime, owing to the van-Hove-ridge-induced peak in χ_{tot} there (Fig. 1j). We note, however, that $g_\uparrow = g_\downarrow$ at $\tilde{B}=0$, reflecting our above-mentioned assumption that no spontaneous spin splitting occurs.

Our FRG approach is limited to the case of zero temperature and zero source–drain voltage, for which no inelastic scattering occurs. To access qualitatively the effects of the latter at fixed U, we have instead used SOPT (Supplementary Information, section 7). Figure 2c–h shows a comparison of our SOPT results for the linear conductance, $g(\tilde{V}_c)$, calculated for several values of magnetic field, $\tilde{B}$, and temperature, $\tilde{T}=k_BT$, and our experimental data for $g(V_c)$. The measured conductance step shows a shoulder (Fig. 2e, f, red arrows) that becomes increasingly more pronounced with both increasing field, B (Fig. 2e), and increasing temperature, T (Fig. 2f), which is the hallmark of the 0.7-anomaly. Our perturbative calculations qualitatively reproduce both trends remarkably well. The only caveat is that the experimental curves in Fig. 2e, f show more pronounced shoulders than do the respective SOPT curves in Fig. 2c, d. This failure of SOPT to

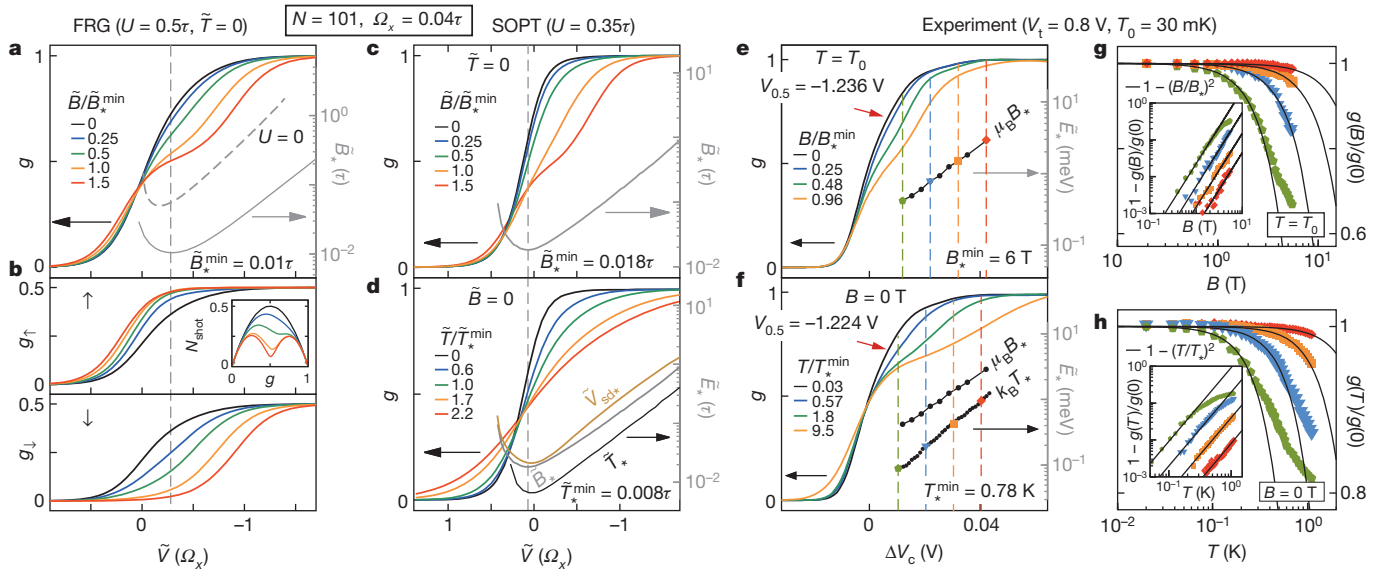


Figure 2 | Conductance: theory versus experiment. **a**, **b**, FRG results: the linear response conductance, $g(\tilde{V}_c, \tilde{B})$, of a QPC (**a**), and its spin-resolved components, $g_\uparrow$ and $g_\downarrow$ (**b**), plotted as functions of $\tilde{V}_c/\Omega_x$ for several values of $\tilde{B}$ at $\tilde{T}=0$ (but finite interaction U). The grey dashed and solid lines in **a** show the low-energy scale $\tilde{B}_*(\tilde{V}_c)$ for $U=0$ and $U=0.5$, respectively, plotted on the log-linear scale indicated on the right-hand axis (as also done in **c**–**f**). The small-field magnetoresponse in **a** is strongest when $\tilde{B}_*$ takes its smallest value, $\tilde{B}_*^{min}$ (vertical dashed lines). Inset of **b**, the shot noise factor, $N_{shot} = \sum_\sigma g_\sigma(1-g_\sigma)/2$, plotted as function of g. Its asymmetric development with $\tilde{B}$, which reflects that of $g_\uparrow$ and $g_\downarrow$, agrees qualitatively with experiment (see fig. 4d of ref. 7). **c**, **d**, SOPT results: $g(\tilde{V}_c, \tilde{B})$ at $\tilde{T}=0$ for several values of $\tilde{B}$ (**c**) and $g(\tilde{V}_c, \tilde{T})$ at $\tilde{B}=0$ for several values of $\tilde{T}$ (**d**), both plotted as functions of $\tilde{V}_c/\Omega_x$. The low-energy scale $\tilde{B}_*(\tilde{V}_c)$ is shown as a thin grey line in **c** and repeated in **d**; $\tilde{T}_*(\tilde{V}_c)$ and $\tilde{V}_{sd*}(\tilde{V}_c)$ are respectively shown as thin black and brown lines in **d**. The vertical dashed line indicates where $\tilde{B}_*$ takes its minimal value, $\tilde{B}_*^{min}$. For $\tilde{V}_c$ values below this dashed line, the lines for $\tilde{B}_*$, $\tilde{T}_*$ and $\tilde{V}_{sd*}$ in **d** are nearly straight on the log-linear scale, implying the behaviour summarized by equation (3), and are nearly parallel to each other, implying that the ratios $\tilde{B}_*/\tilde{T}_*$ and $\tilde{V}_{sd*}/\tilde{T}_*$ are essentially independent of $\tilde{V}_c$ there. **e**, **f**, Experiments—pinch-off curves. **e**, $g(V_c)$ measured at a low 2DES temperature, T_0, for various magnetic fields parallel to the 2DES, plotted as a function of $\Delta V_c = V_c - V_{0.5}$, where $V_{0.5}$ is the gate voltage for which the conductance at $B=0$ and $T=T_0$ is $g(V_{0.5})=0.5$. **f**, Analogous to **e**, but for $B=0$ and various temperatures T. Colours in **e** and **f** are chosen to provide comparability with theory curves in **a**, **c** and **d** (with the correspondence $|e|\Delta V_c \propto -\tilde{V}_c$). **g**, **h**, Experiments—Fermi-liquid behaviour: $g(B)/g(0)$ as function of B at temperature T_0 (**g**), and $g(T)/g(T_0)$ as function of T at $B=0$ (**h**), shown on log-linear scales (insets show their differences from unity on log-log scales) to emphasize small values of B and T. Coloured symbols distinguish data taken at different fixed V_c values, indicated by dashed lines of corresponding colour in **e** and **f**. The quadratic B and T dependences observed in **g** and **h** for each fixed V_c value confirm equation (2) and were used to determine the corresponding scales $B_*(V_c)$ and $T_*(V_c)$. (Black lines in **g** and **h** show $1-(B/B_*)^2$ and $1-(T/T_*)^2$, respectively.) The resulting energies, $E_* = \mu_B B_*(V_c)$ and $E_* = k_B T_*(V_c)$, are shown as functions of V_c in **e** (for B_*) and **f** (for both B_* and T_*) on a log-linear scale. The shape of these measured functions agrees qualitatively with the SOPT predictions in **c** and **d**, confirming the nearly exponential $\tilde{V}_c$ dependences and the nearly V_c-independent B_*/T_* ratio, discussed above. (For additional data, similar to that in **g** and **h**, see Supplementary Information, section 2B.)

produce real shoulders is present both in the low-field dependence at low temperature (compare Fig. 2e with Fig. 2c; the former, but not the latter, shows a weak shoulder even at zero field) and in the temperature dependence at zero field (compare Fig. 2f and Fig. 2d). In contrast, the more powerful FRG approach does reproduce the weak shoulder even for $\tilde{B}=\tilde{T}=0$, as discussed above; compare the black $g(\tilde{V}_c)$ curves in Fig. 2a (FRG) and Fig. 2c (SOPT). (That the latter curve, in contrast to the former, lies above its non-interacting version, $g^0(\tilde{V}_c)$, is an artefact of SOPT; see Supplementary Information, section 7D.)

We next focus on the limit of small energies $\tilde{B}$, $\tilde{T}$ and $\tilde{V}_{sd}$. Here our SOPT calculations yield three predictions, enumerated below, that are all consistent with our measurements. First, for fixed $\tilde{V}_c$, the leading dependence of the nonlinear conductance, $g_{nl}=(dI/d\tilde{V}_{sd})/G_Q$, on $\tilde{B}$, $\tilde{T}$ and $\tilde{V}_{sd}$ is predicted to be quadratic, as confirmed by the measured data in Figs 2g, h and 3a. This implies an expansion of the form

$$\frac{g_{nl}(\tilde{B},\,\tilde{T},\,\tilde{V}_{sd})}{g_{nl}(0,\,0,\,0)} \approx 1-\frac{\tilde{B}^2}{\tilde{B}_*^2}-\frac{\tilde{T}^2}{\tilde{T}_*^2}-\frac{\tilde{V}_{sd}^2}{\tilde{V}_{sd*}^2} \qquad (2)$$

for $\tilde{B}/\tilde{B}_*$, $\tilde{T}/\tilde{T}_*$, $\tilde{V}_{sd}/\tilde{V}_{sd*}\ll 1$, where $\tilde{B}_*$, $\tilde{T}_*$ and $\tilde{V}_{sd*}$ are $\tilde{V}_c$-dependent crossover scales that govern the 'strength' of the 0.7-anomaly for $U\neq 0$: the smaller these scales, the stronger the dependence on $\tilde{B}$, $\tilde{T}$ and $\tilde{V}_{sd}$ for a given $\tilde{V}_c$. Our SOPT results for these crossover scales are shown as thin lines on log-linear scales in Fig. 2c and Fig. 2d, respectively. Second, in that part of the sub-open regime where $g_{nl}(0,\,0,\,0)\approx 1$, they all depend exponentially on $\tilde{V}_c$:

$$\tilde{B}_*,\ \tilde{T}_*,\ \tilde{V}_{sd*} \propto \exp(-\pi\tilde{V}_c/\Omega_x) \qquad (3)$$

Third, and again for $g_{nl}(0,\,0,\,0)\approx 1$, the ratios $\tilde{B}_*/\tilde{T}_*$ and $\tilde{V}_{sd*}/\tilde{T}_*$ are essentially independent of $\tilde{V}_c$ (Supplementary Fig. 4). Remarkably, both the second and third predictions are confirmed by our experimental results (Fig. 2e for B_*, Fig. 2f for T_* and Supplementary Fig. 3 for V_{sd*}). The behaviour predicted by equation (3) for $\tilde{T}_*$ is also in accord with previous experiments[6] and with a perturbative treatment of interactions using Wentzel–Kramers–Brillouin wavefunctions[21]. Remarkably, the exponential $\tilde{V}_c$ dependence of the crossover scales stated in equation (3) can be understood from a non-interacting ($U=0$) theory, by using the bare transmission probability[3]

$$T_\sigma^0(\omega)=\left[e^{-2\pi(\omega-\tilde{V}_c+\sigma\tilde{B}/2)/\Omega_x}+1\right]^{-1} \qquad (4)$$

in the Landauer–Büttiker formula. A detailed analysis (Supplementary Information, section 5) shows that the crossover scales experience a further exponential reduction with increasing effective interaction strength, $U/\sqrt{\Omega_x\tau_0}$.

When plotted as a function of $\tilde{V}_c$, $1/\tilde{B}_*$ has a peak in the sub-open regime just before the onset of the exponential dependence of equation (3) (Fig. 1j). This peak is roughly similar in shape and position to that in $\chi_{tot}(\tilde{V}_c)$ (compare dashed and solid blue lines in Fig. 1j), except that the latter has a finite offset, reflecting the non-zero spin susceptibility of an open QPC. Thus, we predict, fourth, that $1/\tilde{B}_*$, which characterizes the strength of the low-field magnetoconductance, is roughly proportional to the spin susceptibility, χ_{tot}, of the CCR.

Next we address the remarkable experimental fact[6] that many low-energy properties of the 0.7-anomaly (including our first and third predictions) are similar to those seen in transport through a Kondo quantum dot (KQD). This led to the proposal[13,14] that a QPC harbours

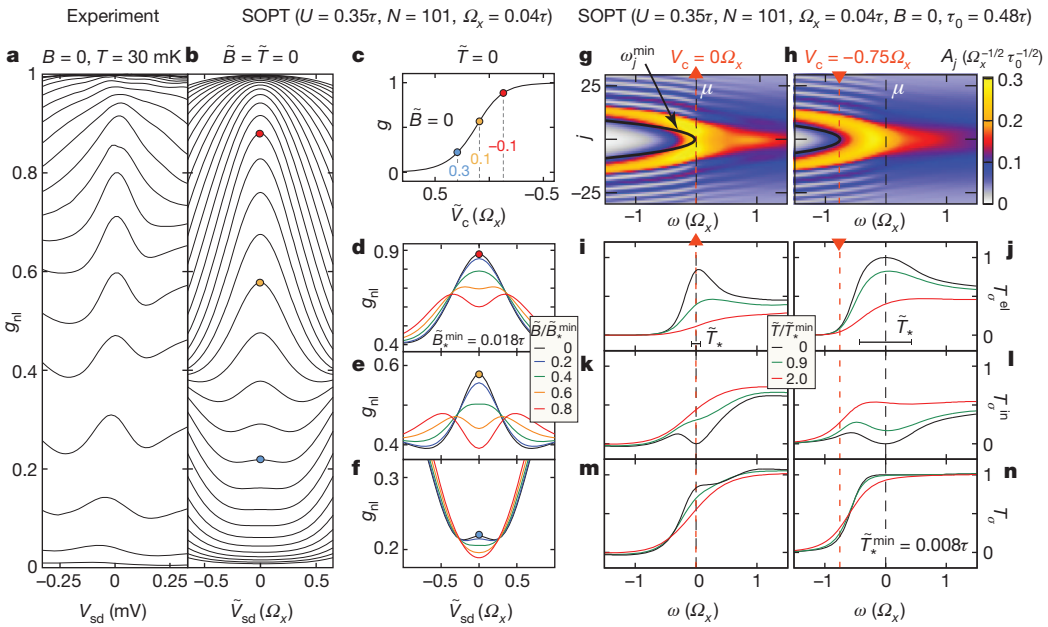


Figure 3 | Finite excitation energies. **a**–**f**, Zero-bias peak. **a**, Experimental data for the nonlinear conductance, g_{nl}, as a function of source–drain voltage, measured for several V_c values at a fixed low temperature and zero field. **b**, Keldysh SOPT results for $g_{nl}(\tilde{V}_{sd})$ for several $\tilde{V}_c$ values at $\tilde{T}=\tilde{B}=0$, showing qualitative agreement with **a**. **c**, The linear-response conductance, $g=g_{nl}(\tilde{V}_{sd}=0)$, as a function of $\tilde{V}_c$. **d**–**f**, $g_{nl}(\tilde{V}_{sd})$ as in **b**, but for three different $\tilde{V}_c$ values (compare colour-matched dots in **c** and **b**) and five different magnetic field values in each panel. Increasing $\tilde{B}$ causes the ZBP to split into two subpeaks once $\tilde{B}\gtrsim\tilde{B}_*$; the splitting is therefore most pronounced in **e**, for which $\tilde{B}_*$ is smallest. A detailed discussion of the ZBP, including its T dependence, will be published elsewhere. Here we would like to point out the qualitative agreement of **d**–**f** with published data; see, for example, fig. 2d of ref. 6. **g**, **h**, Interacting LDOS: $A_j(\omega)$, calculated using SOPT, shown for two fixed gate voltage values, $\tilde{V}_c/\Omega_x=0$ (**g**) and -0.75 (**h**) (red dashed lines). **i**–**n**, Equilibrium transmission probabilities: the corresponding elastic, inelastic and total transmission probabilities, T_σ^{el} (**i**, **j**), T_σ^{in} (**k**, **l**) and T_σ (**m**, **n**), calculated using SOPT and shown as functions of energy, ω, for three different temperatures. At $\tilde{T}=0$ (black curves) $T_\sigma^{in}(\omega)$ vanishes at $\omega=0$, where there is no phase space for inelastic scattering. However, it increases as ω changes from zero, causing a corresponding reduction in the elastic transmission for $\omega\neq 0$, such that $T_\sigma^{el}(\omega)$ has a narrow 'low-energy peak' around $\omega=0$. On increasing the temperature, the probability of inelastic scattering increases, causing the minimum in $T_\sigma^{in}(\omega)$ and the peak in $T_\sigma^{el}(\omega)$ to be smeared out. This leads to a net $\tilde{T}$-induced reduction in the total transmission, $T_\sigma(\omega)$ near $\omega=0$, causing a corresponding reduction in the conductance (Fig. 2d, f). This reduction is stronger for $\tilde{V}_c/\Omega_x=0$ (**m**) than for $\tilde{V}_c/\Omega_x=-0.75$ (**n**), because the probability of electron–hole pair creation during inelastic scattering is largest when apex of the van Hove ridge lies closest to μ (compare **g** and **h**).

a quasi-bound state, whose local moment gives rise to the Kondo effect. In contrast, our van-Hove-ridge scenario fully explains the 0.7-anomaly without invoking the Kondo effect. In particular, we find no indications that a smooth parabolic barrier hosts a discrete, localized spin (compare with the third feature above), and no Kondo effect/0.7-anomaly similarities (experimentally or theoretically) at high energies ($\gtrsim \tilde{B}_*$), where the Kondo effect is governed by an unscreened local moment. Nevertheless, the two phenomena do have similar low-energy behaviour. This is because both involve a spin-singlet ground state featuring spatially confined spin fluctuations. For a KQD they result from screening of the localized spin, whereas for a QPC they result from the extended structure of the van Hove ridge (Fig. 1i); but this distinction, which is important on short length scales (high energies), does not matter on long ones (low energies). These spin fluctuations are characterized by exponentially small energy scales, the Kondo temperature for a KQD, and $\tilde{T}_*$ for a QPC, both scaling inversely with the local spin susceptibility (for a QPC, this follows from prediction four). For a KQD, the local spin fluctuations can be described by Nozières–Fermi-liquid theory[27,28] in terms of scattering phase shifts, which determine its low-energy properties. Because a QPC, like a KQD, harbours spatially confined spin fluctuations, a similar Nozières–Fermi-liquid framework applies, explaining why its low-energy transport properties are similar to those of a KQD.

We next study finite excitation energies ($\tilde{T}$, $\tilde{V}_{sd} > 0$), where inelastic scattering becomes important (Fig. 3). We begin by considering the nonlinear differential conductance, g_{nl}, as a function of source–drain voltage, V_{sd}. Experimentally, g_{nl} shows a narrow peak at $V_{sd} = 0$ (Fig. 3a; see also refs 6, 9, 10). This ZBP appears strongest in the sub-open regime, but remains visible even very close to pinch off[10] ($g \to 0$). It splits with increasing field once B exceeds a V_c-dependent crossover value that is smallest when $g \approx 0.7$ (see fig. 2d of ref. 6). Remarkably, our model, treated using Keldysh SOPT (Supplementary Information, section 7B), yields a ZBP (Fig. 3b, d–f) that qualitatively reproduces this behaviour. In the sub-open regime ($0.5 \lesssim g \lesssim 0.9$), a ZBP arises even without interaction (this follows from equation (4)), but interactions modify it in two ways (Supplementary Information, section 7C): a finite V_{sd} causes a net charge enhancement at the barrier, resulting in a reduction of transmission due to Coulomb repulsion; and opens up a finite phase space for inelastic backscattering. Both effects strongly depend on the LDOS near μ (Fig. 3g, h), and are thus strongest when the apex of the van Hove ridge lies near μ (as in Figs 3g and 1g). However, the van Hove ridge intersects μ also for $g < 0.5$ (as in Fig. 1f), which explains why a ZBP is experimentally observed even close to pinch off[10].

The two modification mechanisms just discussed also apply to the case of increasing temperature. To highlight the role of inelastic scattering, we now discuss (for $\tilde{B} = \tilde{V}_{sd} = 0$) the transmission probability $T_\sigma(\omega) = T_\sigma^{el}(\omega) + T_\sigma^{in}(\omega)$, written as the sum of elastic and inelastic contributions corresponding respectively to transmission without or with the creation of electron–hole pairs (see Supplementary Information, section 7A, for their precise definition). Figure 3i–n shows examples of these quantities. With increasing temperature, the probability for inelastic scattering increases, causing $T_\sigma^{in}(\omega)$ to increase (Fig. 3k, l) and $T_\sigma^{el}(\omega)$ to decrease (Fig. 3i, j). This leads to a net temperature-induced reduction in the total transmission, $T_\sigma(\omega)$ (Fig. 3m, n), near $\omega = 0$, causing a corresponding reduction in the conductance (Fig. 2d, f). Importantly, this reduction is V_c dependent: it is strongest when the apex of the van Hove ridge lies near μ (as in Fig. 3m) and decreases away from this point (as in Fig. 3n), because the probability for electron–hole pair creation during inelastic scattering increases with the LDOS near μ. The fact that $T_\sigma(\omega)$ acquires a non-trivial, interaction-induced dependence on $\tilde{T}$ in the sub-open regime is consistent with the fact that near $g \approx 0.7$ the measured thermopower violates the Mott relation[5], which is based on the assumption of non-interacting electrons.

Finally, we note that we have studied the magnetic field dependence of the transconductance, dG/dV_c, both experimentally and by using FRG. We obtain excellent qualitative agreement between experiment and theory, showing that such measurements can be understood without invoking spontaneous spin polarization, as is often advocated to explain them[4,8,17]. A detailed analysis (Supplementary Information, section 2C, and Supplementary Fig. 5) establishes that the g factor is enhanced significantly by interactions, and that interaction strength can be tuned experimentally using a top gate.

We have presented detailed microscopic calculations that qualitatively reproduce the full phenomenology of the 0.7-anomaly. We argued that a van Hove ridge in the LDOS, combined with interactions, provides a natural explanation for the anomalous behaviour of the conductance of a sub-open ($g \gtrsim 0.5$) QPC. The experimentally observed[6] similarities between the 0.7-anomaly and the Kondo effect at low energies arise because both phenomena involve spatially localized spin fluctuations; at high energies, the similarities cease. We verified our Fermi-liquid predictions for the QPC conductance by systematic measurement of the conductance as a function of V_c, B and T. Strikingly, we demonstrated that the zero-bias peak in a QPC arises from the interplay of interactions and geometry. By implication, anomalous zero-bias behaviour might also arise in other systems involving interacting electrons traversing 1D low-density regions with slowly varying spatial inhomogeneities, such as the gated nanowires being studied in the search for Majorana fermions[29].

METHODS SUMMARY

The nanostructure is laterally defined in a 2DES located 85 nm beneath the surface of a GaAs/AlGaAs heterostructure. The low-temperature carrier density and mobility are $1.9 \times 10^{11}\,cm^{-2}$ and $1.2 \times 10^6\,cm^2\,V^{-1}\,s^{-1}$, respectively. Electron-beam lithography was used to create the Ti/Au gates. The top gate is electrically insulated from the others by cross-linked poly(methyl methacrylate). Perfect alignment of magnetic fields parallel to the 2DES and the 1D channel defining the QPC was ensured by using a two-axis magnet and was controlled by magnetotransport measurements. We used a dilution refrigerator and reached electron temperatures as low as $T_{2DES} \approx 30$ mK.

Our most accurate theoretical results were obtained by using FRG[24–26] to calculate $T = 0$ properties. FRG amounts to doing renormalization-group-enhanced perturbation theory in the interaction U. In setting up our FRG flow equations, we made two approximations, both exact to second order in U: we truncated the infinite hierarchy of flow equations by neglecting the flow of the three-particle vertex; and we set to zero all components of the two-particle vertex that are not already generated to second order in the interaction (coupled-ladder approximation).

To access the effects of inelastic scattering for $\tilde{T} > 0$ or $\tilde{V}_{sd} > 0$ at fixed U, we used SOPT: we dressed bare Green's functions by evaluating the self-energy perturbatively to second order in the interaction. For $\tilde{V}_{sd} = 0$, we calculated the linear conductance following the strategy in ref. 23, generalized to $\tilde{B} \neq 0$ and broken electron–hole symmetry. For $\tilde{V}_{sd} > 0$, we calculated the nonlinear conductance, $g_{nl} = (dI/d\tilde{V}_{sd})/G_Q$, using the Meir–Wingreen formula for the current (equation (6) of ref. 30).

Received 16 November 2012; accepted 26 June 2013.
Published online 28 August 2013.

1. van Wees, B. J. *et al.* Quantized conductance of point contacts in a two-dimensional electron gas. *Phys. Rev. Lett.* **60,** 848–850 (1988).
2. Wharam, D. A. *et al.* One-dimensional transport and the quantisation of the ballistic resistance. *J. Phys. C* **21,** L209–L214 (1988).
3. Büttiker, M. Quantized transmission of a saddle-point constriction. *Phys. Rev. B* **41,** 7906(R) (1990).
4. Thomas, K. J. *et al.* Possible spin polarization in a one-dimensional electron gas. *Phys. Rev. Lett.* **77,** 135–138 (1996).
5. Appleyard, N. J. *et al.* Direction-resolved transport and possible many-body effects in one-dimensional thermopower. *Phys. Rev. B* **62,** R16275–R16278 (2000).
6. Cronenwett, S. M. *et al.* Low-temperature fate of the 0.7 structure in a point contact: a Kondo-like correlated state in an open system. *Phys. Rev. Lett.* **88,** 226805 (2002).
7. DiCarlo, L. *et al.* Shot-noise signatures of 0.7 structure and spin in a quantum point contact. *Phys. Rev. Lett.* **97,** 036810 (2006).
8. Koop, E. J. *et al.* The influence of device geometry on many-body effects in quantum point contacts: signatures of the 0.7-anomaly, exchange and Kondo. *J. Supercond. Nov. Magn.* **20,** 433–441 (2007).
9. Sarkozy, S. *et al.* Zero-bias anomaly in quantum wires. *Phys. Rev. B* **79,** 161307R (2009).
10. Ren, Y. *et al.* Zero-bias anomaly of quantum point contacts in the low-conductance limit. *Phys. Rev. B* **82,** 045313 (2010).

11. Micolich, A. P. What lurks below the last plateau: experimental studies of the $0.7 \times 2e^2/h$ conductance anomaly in one-dimensional systems. *J. Phys. Condens. Matter* **23,** 443201 (2011).
12. Burke, A. *et al.* Extreme sensitivity of the spin-splitting and 0.7 anomaly to confining potential in one-dimensional nanoelectronic devices. *Nano Lett.* **12,** 4495–4502 (2012).
13. Meir, Y., Hirose, K. & Wingreen, N. S. Kondo model for the 0.7 anomaly in transport through a quantum point contact. *Phys. Rev. Lett.* **89,** 196802 (2002).
14. Rejec, T. & Meir, Y. Magnetic impurity formation in quantum point contacts. *Nature* **442,** 900–903 (2006).
15. Ihnatsenka, S. & Zozoulenko, I. V. Conductance of a quantum point contact based on spin density-functional theory. *Phys. Rev. B* **76,** 045338 (2007).
16. Matveev, K. A. Conductance of a quantum wire at low electron density. *Phys. Rev. B* **70,** 245319 (2004).
17. Reilly, D. J. Phenomenological model for the 0.7 conductance feature in quantum wires. *Phys. Rev. B* **72,** 033309 (2005).
18. Sloggett, C. Milstein, A. I. & Sushkov, O. P. Correlated electron current and temperature dependence of the conductance of a quantum point contact. *Eur. Phys. J. B* **61,** 427–432 (2008).
19. Aryanpour, K. & Han, J. E. Ferromagnetic spin coupling as the origin of 0.7 anomaly in quantum point contacts. *Phys. Rev. Lett.* **102,** 056805 (2009).
20. Güçlü, A. D. *et al.* Localization in an inhomogeneous quantum wire. *Phys. Rev. B* **80,** 201302(R) (2009).
21. Lunde, A. M. *et al.* Electron-electron interaction effects in quantum point contacts. *New J. Phys.* **11,** 023031 (2009).
22. Goldhaber-Gordon, D. *et al.* Kondo effect in a single-electron transistor. *Nature* **391,** 156–159 (1998).
23. Oguri, A. Transmission probability for interacting electrons connected to reservoirs. *J. Phys. Soc. Jpn* **70,** 2666–2681 (2001).
24. Andergassen, S. *et al.* Renormalization-group analysis of the one-dimensional extended Hubbard model with a single impurity. *Phys. Rev. B* **73,** 045125 (2006).
25. Karrasch, C., Enss, T. & Meden, V. Functional renormalization group approach to transport through correlated quantum dots. *Phys. Rev. B* **73,** 235337 (2006).
26. Metzner, W. *et al.* Functional renormalization group approach to correlated fermion systems. *Rev. Mod. Phys.* **84,** 299–352 (2012).
27. Nozières, P. A ''fermi-liquid'' description of the Kondo problem at low temperatures. *J. Low Temp. Phys.* **17,** 31–42 (1974).
28. Glazman, L&. Pustilnik, M. in *Nanophysics: Coherence and Transport* (eds Bouchiat, H. *et al.*) 427–478 (Elsevier, 2005).
29. Mourik, V. *et al.* Signatures of Majorana fermions in hybrid superconductor-semiconductor nanowire devices. *Science* **336,** 1003–1007 (2012).
30. Meir, Y. & Wingreen, N. S. Landauer formula for the current through an interacting electron region. *Phys. Rev. Lett.* **68,** 2512–2515 (1992).

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

Acknowledgements We thank B. Altshuler, P. Brouwer, R. Egger, J. Folk, L. Glazman, V. Golovach, A. Hamilton, A. Högele, Y. Imry, M. Kiselev, J. Kotthaus, D. Logan, D. Loss, C. Marcus, Y. Meir, H. Monien, M. Pepper, M. Pustilnik, A. Rosch, K. Schönhammer, B. Spivak and A. Yacoby for discussions, and, in particular, S. Andergassen, C. Honerkamp, S. Jakobs, C. Karrasch, V. Meden, M. Pletyukhov and H. Schoeller for FRG-related help and advice. We acknowledge support from the DFG through SFB-631, SFB-TR12, De730/3-2, De730/4-1, De730/4-2, De730/4-3, HO 4687/1-3, LU819/4-1 and the Cluster of Excellence Nanosystems Initiative Munich; from the Center for NanoScience; and from the US National Science Foundation under grant no. NSF PHY05-51164. S.L. acknowledges support through a Heisenberg fellowship of the DFG.

Author Contributions J.v.D. and S.L. coordinated the project: J.v.D. initiated and supervised the theoretical work, and S.L. planned and supervised the experiments and their analysis. F.B. and J.H. carried out the calculations using FRG, and J.H., F.B. and B.B. carried out the calculations using perturbation theory. D.S. and W.W. provided the wafer material, and D.B. fabricated the nanostructure. E.S., D.B., D.T. and S.L. carried out the measurements, and E.S., D.B., F.B. and J.H. carried out the experimental data analysis. J.H. and F.B. prepared the figures, and J.v.D., S.L., F.B., J.H. and E.S. 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 J.v.D. (vondelft@lmu.de) or S.L. (ludwig@lmu.de).

LETTER

doi:10.1038/nature12491

Odd and even Kondo effects from emergent localization in quantum point contacts

M. J. Iqbal[1], Roi Levy[2], E. J. Koop[1], J. B. Dekker[1], J. P. de Jong[1], J. H. M. van der Velde[1], D. Reuter[3], A. D. Wieck[3], Ramón Aguado[4], Yigal Meir[2,5] & C. H. van der Wal[1]

A quantum point contact (QPC) is a basic nanometre-scale electronic device: a short and narrow transport channel between two electron reservoirs. In clean channels, electron transport is ballistic and the conductance is then quantized as a function of channel width[1,2] with plateaux at integer multiples of $2e^2/h$ (where e is the electron charge and h is Planck's constant). This can be understood in a picture where the electron states are propagating waves, without the need to account for electron–electron interactions. Quantized conductance could thus be the signature of ultimate control over nanoscale electron transport. However, even studies with the cleanest QPCs generically show significant anomalies in the quantized conductance traces, and there is consensus that these result from electron many-body effects[3,4]. Despite extensive experimental and theoretical studies[4–11], understanding these anomalies is an open problem. Here we report that the many-body effects have their origin in one or more spontaneously localized states that emerge from Friedel oscillations in the electron charge density within the QPC channel. These localized states will have electron spins associated with them, and the Kondo effect—related to electron transport through such localized electron spins—contributes to the formation of the many-body state[5–7]. We present evidence for such localization, with Kondo effects of odd or even character, directly reflecting the parity of the number of localized states; the evidence is obtained from experiments with length-tunable QPCs that show a periodic modulation of the many-body properties with Kondo signatures that alternate between odd and even Kondo effects. Our results are of importance for assessing the role of QPCs in more complex hybrid devices[12,13] and for proposals for spintronic and quantum information applications[14,15]. In addition, our results show that tunable QPCs offer a versatile platform for investigating many-body effects in nanoscale systems, with the ability to probe such physics at the level of a single site.

There are two signatures of many-body physics that are generically observed for a wide variety of QPCs, including systems in GaAs (refs 3, 9), Si (ref. 16) and graphene[17], and that are found for transport of both electrons and holes[9,16]. First, the quantized conductance traces often show the so-called 0.7 anomaly: an additional small plateau at about $0.7(2e^2/h)$. Second, as a function of bias voltage across the channel, conductance G typically shows a peak around zero bias (named the zero-bias anomaly, ZBA), mostly below the first quantized plateau. The experimental observation[5] that the ZBA and the 0.7 anomaly had similarities with the Kondo effect for quantum dots (transport through a single localized electron state[18,19]) inspired theoretical work[6,7] that proposed that electron many-body physics could lead to localized electrons in the centre of the QPC. This is a remarkable phenomenon, because a QPC is a fully open quantum system. To avoid confusion with localization by an atomic impurity or disordered potential, these many-body states are termed self-consistent or emergent localized states (ELSs).

This theoretical work[6,7] developed the picture that the many-body effects in QPC channels are intimately related to the occurrence of a Friedel oscillation—an oscillation in the electron charge density that occurs when electron waves get reflected in a partially open QPC channel—which gets enhanced into an ELS with the charge of about one electron owing to Coulomb repulsion and exchange interactions between electrons. This reduces the conductance and can explain the 0.7 anomaly[6]. However, transport through such a state can be enhanced by the Kondo effect at temperatures below a typical Kondo temperature, T_K. This appears as a ZBA and also moves the 0.7 plateau towards unity (in units of $2e^2/h$), consistent with experiments. This theoretical work[6,7] also predicted that, depending on parameters, a pair of such ELSs may emerge in the channel, resulting in a double-peak ZBA (as observed in double quantum dots[20] owing to the two-impurity Kondo effect[21–26]). So far no such double-peak ZBAs have been reported for QPCs (Supplementary Information section 3).

We report here the observation of such double-peak ZBAs in a large number of conventional QPCs (with two gate fingers as in Fig. 1a, denoted QPC_{2F}). We also introduce a new type of QPC which has a tunable channel length (with six gate fingers as in Fig. 2a, denoted QPC_{6F}). In these devices, the 0.7 anomaly and the ZBA show a periodic modulation as a function of QPC length, which we attribute to an increasing number of ELSs. Thus, as the number of ELSs increases with QPC length, its parity alternates, giving rise to modulation between Kondo effects for an odd or even number of localized states (known as odd- and even-impurity Kondo effects), and, as a result, between single- and double-peak ZBAs. In addition, the 0.7 anomaly shows a periodic modulation because the enhancement of the 0.7 feature towards unit conductance depends on both the parity and the parameters of the Kondo system such as T_K, and these are both modulated as a function of QPC length.

The signatures of a pair of ELSs are more likely to be observed on shorter QPCs[7]. We thus focused on QPC_{2F} of lithographic length $L = 200$ nm (and width $W = 350$ nm), which is shorter than most QPCs reported in the literature. We searched for double-peak ZBAs in a set of 80 QPC_{2F} (realized in two different wafer materials, different fabrication runs, different cool-downs, and with or without gate-biased cool-down; see Methods) and found them in about half the studied devices. The ubiquity of the phenomenon, and the fact that such double-peak ZBAs were persistently observed in the same devices over different cool-downs, implies that it is a generic effect and not due to a fortuitous impurity nearby. Figure 1 presents data from two QPC_{2F} to illustrate that the signatures of many-body physics show qualitatively similar features, though with significant device-to-device variation (whereas there is no strong variation in the manifestation of non-interacting electron physics, such as the quantized conductance[8,10]). Figure 1b, c presents measurements of the linear conductance (Methods). In addition to the quantized conductance plateaux at integer multiples of $2e^2/h$, the trace in Fig. 1c shows an additional shoulder at $G \approx 0.7(2e^2/h)$ (the 0.7 anomaly, also observed for the device of Fig. 1b at higher temperatures). Results for the nonlinear conductance (Methods) of these same devices are presented in Fig. 1d, e. Most traces between 0 and $1(2e^2/h)$ show a

[1]Zernike Institute for Advanced Materials, University of Groningen, NL-9747AG Groningen, The Netherlands. [2]Department of Physics, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel. [3]Angewandte Festkörperphysik, Ruhr-Universität Bochum, D-44780 Bochum, Germany. [4]Instituto de Ciencia de Materiales de Madrid (ICMM), Consejo Superior de Investigaciones Científicas (CSIC), Sor Juana Ines de la Cruz 3, 28049 Madrid, Spain. [5]Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel.

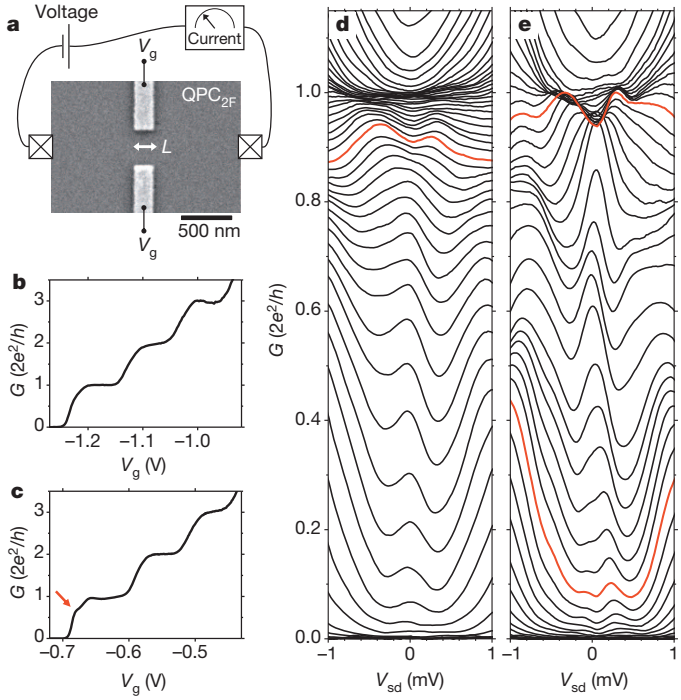


Figure 1 | Conductance of quantum point contacts. **a**, Electron microscope image of a conventional quantum point contact (QPC) with two gate fingers (QPC_{2F}). These gates are on the surface of a wafer with a 2DEG at 110 nm or 114 nm depth. Applying voltages V_g to these gates induces a narrow transport channel between source and drain regions of the 2DEG. The conductance of the QPCs is studied by applying and measuring voltage and current signals on contacts to the 2DEG. For QPC_{2F} the length of the transport channel is fixed by the lithographic length L. **b**, **c**, Linear conductance G measured on two different QPC_{2F} with $L = 200$ nm. The traces show clear quantized conductance plateaux at integer multiples of $2e^2/h$ (the shift in dependence on V_g for **c** as compared to **b** is due to a different cool-down procedure; see Methods). The plateaux and transitions between plateaux show small deviations from clean quantized conductance behaviour, such as an additional shoulder at $G \approx 0.7(2e^2/h)$ in **c** (red arrow). **d**, Nonlinear conductance G as a function of source–drain voltage V_{sd} at various V_g settings, for the device of **b**. The zero-bias anomaly (ZBA, enhanced conductance around $V_{sd} = 0$) has mostly single-peak character, but has double-peak character for $G \approx 0.9(2e^2/h)$ (for example the red trace). **e**, Similar results as **d** for the device of **c**. In this device, ZBAs with double-peak character appear at $G \approx 0.1(2e^2/h)$ and $G \approx 0.95(2e^2/h)$.

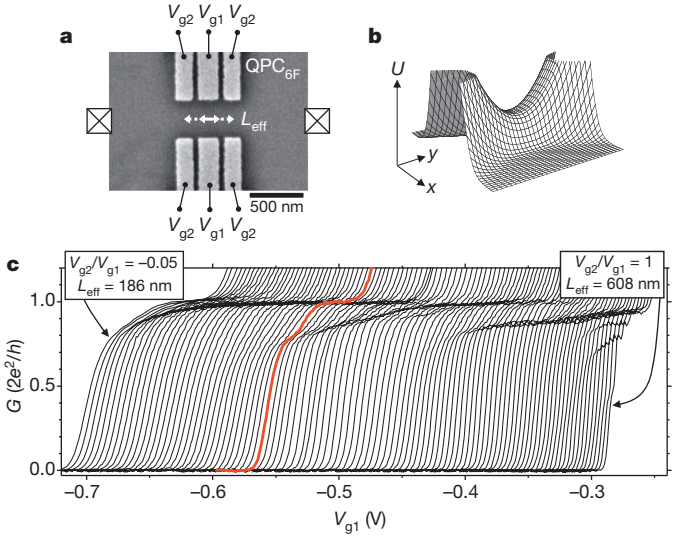


Figure 2 | Length-tunable QPCs. **a**, Electron microscope image of a QPC with six gate fingers (QPC_{6F}). It has a tunable effective length L_{eff} that is set by operating at a fixed ratio V_{g2}/V_{g1}. **b**, Saddle-point potential that illustrates the electron potential energy U (without many-body interactions) in the 2DEG plane in a QPC region. **c**, Linear conductance G as a function of V_{g1} (while co-sweeping V_{g2} at fixed V_{g2}/V_{g1}) measured on a QPC_{6F} for L_{eff} tuned from 186 nm to 608 nm (traces are not offset). In addition to the quantized conductance plateau at $1(2e^2/h)$, most traces show a smaller plateau in the range $0.7(2e^2/h)$ to $0.9(2e^2/h)$ due to many-body effects. For this signature of many-body effects, three periods of modulation can be observed in its dependence on V_{g2}/V_{g1} (that is, L_{eff}).

single-peak ZBA (the enhancement of conductance within ±0.5 mV around $V_{sd} = 0$ mV, where V_{sd} is the bias voltage across the channel). However, the red traces mark examples where the ZBA appears with double-peak character (the asymmetric character of these nonlinear conductance traces will be discussed below). For double-peak ZBAs just below $1(2e^2/h)$, we typically observe a peak splitting that increases with conductance.

Theoretical work[7] predicted that the Friedel oscillation resulting from the screening of the QPC potential would create two electron 'puddles', one on each side of the QPC. For most QPC geometries, it showed that lowering the potential by means of the gate voltage would give a single ELS in the centre of the QPC because the Friedel oscillations from both sides connect. However, for short QPCs (for L similar to the Fermi wavelength) the two ELSs remain intact as the potential is lowered. This gives the physics of a two-impurity Kondo system, and as the coupling between them gets stronger with increasing conductance, the ZBA splits[22–26]. At lower conductance values, the ZBA may appear as a single or a double peak, depending on the ratio between the Kondo temperature of the two ELSs and the strength of interaction between them. Although this is fully consistent with our observations (Fig. 1d, e), the coupling between these ELSs and the resulting splitting of the ZBA depend very sensitively on the device dimensions and the always-present remote imperfections (Supplementary Information section 4), and this leads to significant variability among devices. Thus, a more stringent test of this picture is possible with QPCs whose parameters can be modified continuously.

We addressed this by investigating a set of eight QPC_{6F} devices, for which we could gradually change the effective length L_{eff} of the saddle-point potential. These devices showed qualitatively identical behaviour, with an oscillatory signature of the 0.7 anomaly and regular modulation between single- and double-peak ZBAs with increasing L_{eff} (Fig. 3b, c). These reproducible observations on eight QPC_{6F} (and for different cool-downs of one QPC_{6F}; Supplementary Information section 10) are consistent with the emergence of an increasing number of ELSs due to many-body physics that generically occurs, even in ultra-clean QPC channels.

Figure 2a depicts the QPC_{6F} devices, for which the effective channel length L_{eff} could be tuned continuously. These were operated with the central gate voltage V_{g1} more negative than the side gate voltage V_{g2}, to avoid quantum dot formation. Our analysis shows that in this regime the gates induce a smooth saddle-point potential, despite the narrow gaps between the gate fingers. The effective channel length L_{eff} is set by V_{g2}/V_{g1} (short for V_{g2}/V_{g1} near 0, long for V_{g2}/V_{g1} near 1; see Supplementary Information section 5 for details). Our devices could thus be controlled to have L_{eff} in a range from about 186 nm to 608 nm. Making V_{g1} less negative, at a fixed ratio V_{g2}/V_{g1}, opens the QPC while keeping the length unchanged.

All QPC_{6F} showed clear quantized conductance plateaux. Figure 2c shows for one device how the 0.7 anomaly appears as an additional plateau (smaller than the quantized conductance plateaux) in the range $0.7(2e^2/h)$ to $0.9(2e^2/h)$, and that this plateau shows a dependence on L_{eff}, with three periods of modulation for the range $L_{eff} = 186$ nm to 608 nm. The nonlinear conductance measurements from this device (Fig. 3a) show how the ZBAs appear for $L_{eff} = 286$ nm. At a fixed length, the ZBAs alternate between single-peak and double-peak character when opening the QPC, again with increased splitting for the double-peak ZBA as the conductance approaches $1(2e^2/h)$. The overall

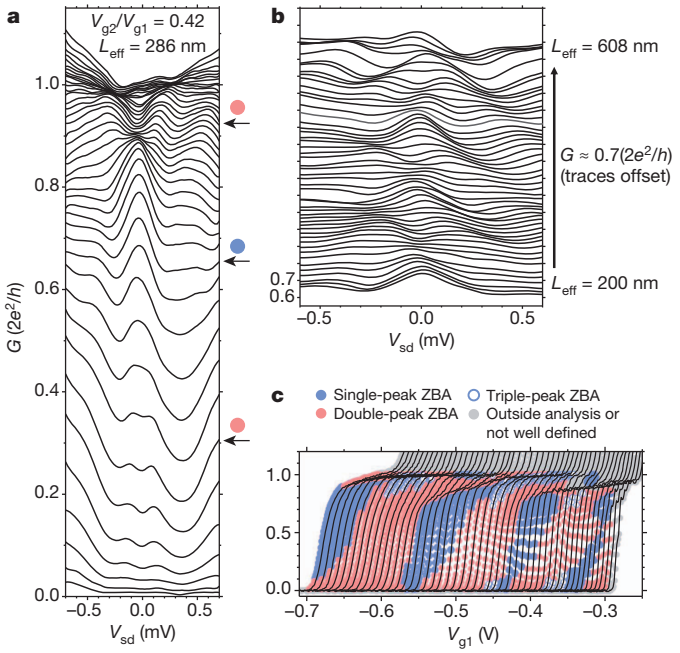


Figure 3 | ZBAs in the nonlinear conductance of a QPC_{6F}. **a**, Nonlinear conductance G as a function of source–drain voltage V_{sd} at various V_{g1} settings, for operation at L_{eff} = 286 nm. The ZBA appears alternately with single- or double-peak character. **b**, Evolution of the ZBA in the nonlinear conductance at fixed conductance level of ~$0.7(2e^2/h)$ as a function of L_{eff} (traces offset). The ZBA has alternately single- or double-peak character. **c**, The character of the ZBA (single-, double- or triple-peak, as labelled) mapped out on the linear conductance data of Fig. 2c.

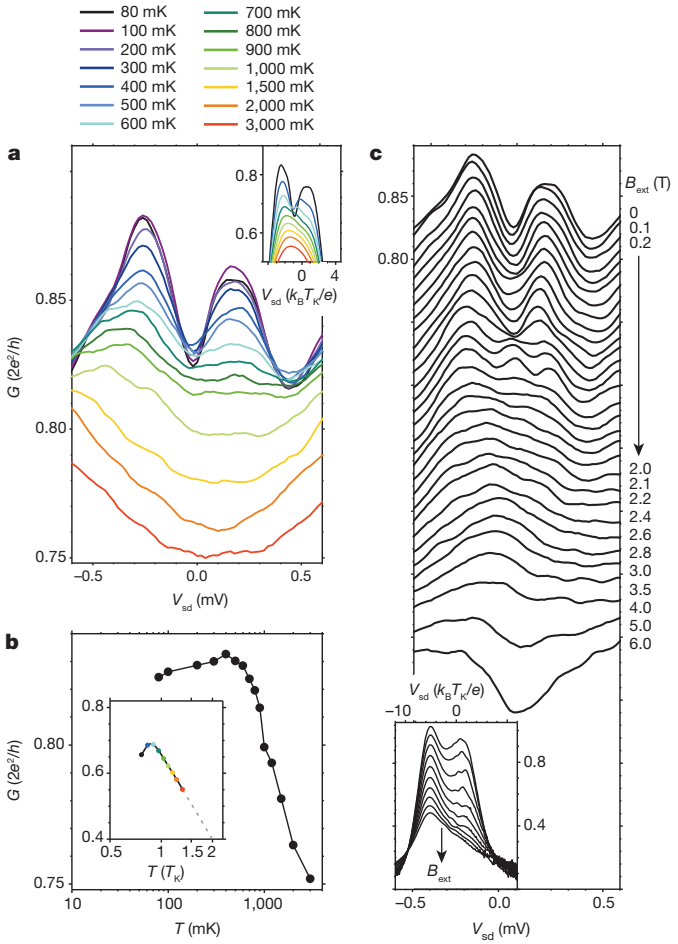


Figure 4 | Comparison between experiments and theory of the Anderson model for a two-impurity Kondo system. **a**, Evolution of the nonlinear conductance (with double-peak ZBA) as a function of temperature for a QPC_{6F} operated with fixed $V_{g1} = -0.528$ V and fixed $V_{g2}/V_{g1} = 0.3$ (traces not offset). Inset, calculated nonlinear-conductance traces from a two-impurity Kondo model (see main text). **b**, The conductance level of the traces in panel **a** near zero bias (between the two ZBA peaks, evaluated at $V_{sd} = -0.02$ mV) as a function of temperature (logarithmic axis). Inset, the same analysis applied to the theoretical results in the inset of **a**. **c**, Evolution of the nonlinear conductance (same device and operation point as **a**) as a function of applied in-plane magnetic field B_{ext} (traces offset by $-0.01(2e^2/h)$). Inset, calculated nonlinear-conductance traces from a two-impurity Kondo model.

appearance of the ZBAs is very similar to that of fixed-length QPC_{2F} (Fig. 1d, e). Figure 3b shows that there is also a modulation between single- and double-peak character when increasing L_{eff} at a fixed conductance level (as evaluated immediately next to the ZBA). Figure 3c plots the data of Fig. 2c again, with coloured data symbols on the traces that mark whether the ZBA at that point has single- or double-peak character (in some cases, we find ZBAs that are best described as triple-peak). The modulation between single-peak and double-peak ZBA as a function of L_{eff} also shows about three periods, and is clearly correlated with the modulation of the 0.7 anomaly.

To critically check the relevance of two-impurity Kondo physics to our observations, we measured the temperature and magnetic field dependence of double-peak ZBAs and compared these with theory for this system. We obtained the theoretical results by calculating the current through a two-impurity Anderson model. For this model, one expects[22–26] a current that gets enhanced by the Kondo effect when the temperature is lowered from above the Kondo temperature to below it. Depending on the strength of the effective coupling between the impurities relative to the Kondo temperature, the associated ZBA has either single- or double-peak character.

Figure 4a depicts how the nonlinear conductance develops from a background conductance of $0.75(2e^2/h)$ at 3,000 mK into a double-peak ZBA with peak values up to about $0.90(2e^2/h)$ as the temperature is decreased (we used the device of Fig. 3; Supplementary Fig. 9 shows the temperature dependence of the linear conductance). The conductance between the peaks (the linear conductance) has a non-monotonic temperature dependence (Fig. 4b) that is characteristic of two-impurity Anderson physics[27]. For comparison, insets in Fig. 4a and b depict the results of the theoretical calculations, and show good qualitative agreement. (We note that this description calculates only the Kondo contribution to the current, which yields zero current at high temperatures owing to Coulomb blockade; in contrast, the ELSs in QPCs are not expected to show strong Coulomb blockade at high temperatures but rather a finite background conductance, as observed.) The theoretical traces are calculated for two impurity sites with unequal coupling strength Γ to a neighbouring electron reservoir, and accordingly an unequal Kondo temperature T_K (the plots are presented on an energy scale given by the highest Kondo temperature of the two sites). The asymmetry between the two Γ parameters gives asymmetric double-peak ZBAs, with the minimum conductance between the peaks not exactly at $V_{sd} = 0$, very similar to the experimentally observed double-peak ZBAs. We should expect such asymmetries between the two Γ parameters, because the Γ values depend exponentially on the coupling between the ELSs and the electron reservoirs (Supplementary Information section 4).

Figure 4c presents the measured magnetic field dependence of this double-peak ZBA. Theoretically, one expects different evolutions with magnetic field depending on the relative magnitude of parameters—such as the Kondo temperature of each impurity, the coupling between the impurities, the asymmetry between the impurities and the temperature. The possibilities include each peak splitting into two sub-peaks, the peaks getting closer and merging, and the peaks moving closer and crossing each other. We see such diverse dependence on magnetic field for the different operation points of a given device, and for different

devices. Figure 4c shows one example, with a comparison to the theoretical expectations (other examples are given in Supplementary Information section 8). We see again good qualitative agreement between the theory of the two-impurity model and the experimental observations.

From the width of the peaks in Fig. 4a we can estimate the Kondo temperature T_K, and the splitting between the peaks is equal to twice the effective coupling between the two ELSs. This coupling is here ~0.2 meV (a typical value for the larger data set plotted in Fig. 3c). This should be consistent with the temperature at which the double-peak character of the ZBA is no longer resolved (and the temperature where the non-monotonic temperature dependence of conductance has a maximum[27]). In Fig. 4a, this occurs for ~800 mK (that is, ~0.1 meV). This is in reasonable agreement with the distance between the peaks. The fact that it is on the low side is probably because T_K is here at the same energy scale as the splitting (~0.15 meV when estimated as half the peak widths in Fig. 4a). We also analysed ZBAs with pronounced single-peak character in the same manner as in ref. 5, and found agreement with the single-impurity Kondo model to the same extent.

The increase of L_{eff} that induces one period of modulation for the 0.7 anomaly and the split-peak character of the ZBA in Fig. 3 (~100 nm to ~150 nm) matches the Fermi wavelength in the QPC channel (Supplementary Information section 6). This supports the hypothesis that the periodicity is linked to Friedel oscillations in the channel which gradually develop additional periods as it gets longer. (That such Friedel oscillations occur on scattering in a two-dimensional electron gas, 2DEG, has been observed directly in a different setting[28].) The increasing number of ELSs with channel length leads to alternation between odd- and even-impurity Kondo effects, and to modulation between single-peak and double-peak ZBAs, respectively (the three-impurity case was studied in ref. 29 and can show ZBAs with triple- or predominantly single-peak character).

To support this picture, we carried out numerical spin density functional theory (SDFT) calculations, generalizing earlier work[7] (Supplementary Information section 2). The length of the channel was determined by a gate of variable length. We studied whether SDFT yields localized states with the charge of about one electron as the state with lowest energy. The results (an example is in Supplementary Fig. 1) show an increasing number of ELSs as the channel gets longer, and also as the QPC is opened by lowering the QPC saddle-point potential at fixed length. These calculations support the picture we have presented here: the number of ELSs increases by one each time the QPC length increases over a range that allows for one additional period of the Friedel oscillation in the QPC channel. The SDFT results suggest that, depending on the overlap of the ELSs, higher spin states ($S = 1$ or $S = 3/2$) may develop in the QPC, making it possible to study Kondo effects for higher spin, and enabling investigation of transport through such exotic states.

The emergence of a ZBA for a hybrid device with a semiconductor channel was recently reported as a signature of Majorana fermions[12]. The fact that similar ZBAs occur with rich behaviour in plain semiconductor QPCs suggests that one should be cautious when ruling out alternative explanations for these Majorana signatures[13]. Evidently, basic understanding of the physics in QPCs is a crucial step in understanding more complex hybrid structures. Tunable QPCs offer an excellent new test ground for such studies, and they are also suited for detailed studies of Friedel oscillations[30] and strongly correlated electrons in low dimensions, at the level of a single site.

METHODS SUMMARY

Experimental methods. QPCs were realized by locally depleting the 2DEG below the surface of a GaAs/AlGaAs heterostructure (for details, see Methods). In the standard approach, applying negative voltage to two metallic gate fingers on the surface of such material (Fig. 1a) induces an electrostatic potential barrier between a source and drain reservoirs in the 2DEG, with a small tunable opening in the form of a saddle-point potential (Fig. 2b). Such devices with two gate fingers are denoted QPC_{2F} and these have a fixed channel length L. We also studied novel devices with six gate fingers (Fig. 2a), denoted QPC_{6F}, which have a channel with tunable length L_{eff}. We focus on the differential conductance G, which is obtained by applying and measuring voltage and current signals as in the simplified scheme in Fig. 1a (for details, see Methods). Results for G at zero bias voltage V_{sd} are called here linear conductance, whereas results for G as a function of V_{sd} (bias spectroscopy) are called here nonlinear conductance. Unless stated otherwise, we present data taken at a temperature T of 80 mK and zero magnetic field.

Kondo transport calculations. We obtained the theoretical results by calculating the current through a two-impurity Anderson model within the slave-boson non-crossing approximation[27] (detailed in Supplementary Information section 1).

SDFT calculations. We obtained the SDFT results by extending the work of ref. 7 (detailed in Supplementary Information section 2).

Full Methods and any associated references are available in the online version of the paper.

Received 25 May; accepted 19 July 2013.
Published online 28 August 2013.

1. van Wees, B. J. *et al.* Quantized conductance of point contacts in a two-dimensional electron gas. *Phys. Rev. Lett.* **60,** 848–850 (1988).
2. Wharam, D. A. *et al.* One-dimensional transport and the quantisation of the ballistic resistance. *J. Phys. Chem.* **21,** L209–L214 (1988).
3. Thomas, K. J. *et al.* Possible spin polarization in a one-dimensional electron gas. *Phys. Rev. Lett.* **77,** 135–138 (1996).
4. Micolich, A. P. What lurks below the last plateau: experimental studies of the $0.7\times2e^2/h$ conductance anomaly in one-dimensional systems. *J. Phys. Condens. Matter* **23,** 443201 (2011).
5. Cronenwett, S. M. *et al.* Low-temperature fate of the 0.7 structure in a point contact: a Kondo-like correlated state in an open system. *Phys. Rev. Lett.* **88,** 226805 (2002).
6. Meir, Y., Hirose, K. & Wingreen, N. S. Kondo model for the "0.7 anomaly" in transport through a quantum point contact. *Phys. Rev. Lett.* **89,** 196802 (2002).
7. Rejec, T. & Meir, Y. Magnetic impurity formation in quantum point contacts. *Nature* **442,** 900–903 (2006).
8. Koop, E. J. *et al.* The influence of device geometry on many-body effects in quantum point contacts: signatures of the 0.7 anomaly, exchange and Kondo. *J. Supercond. Nov. Magn.* **20,** 433–441 (2007).
9. Komijani, Y. *et al.* Evidence for localization and 0.7 anomaly in hole quantum point contacts. *Europhys. Lett.* **91,** 67010 (2010).
10. Burke, A. M. *et al.* Extreme sensitivity of the spin-splitting and 0.7 anomaly to confining potential in one-dimensional nanoelectronic devices. *Nano Lett.* **12,** 4495–4502 (2012).
11. Wu, P. M., Li, P., Zhang, H. & Chang, A. M. Evidence for the formation of quasibound states in an asymmetrical quantum point contact. *Phys. Rev. B* **85,** 085305 (2012).
12. Mourik, V. *et al.* Signatures of Majorana fermions in hybrid superconductor-semiconductor nanowire devices. *Science* **336,** 1003–1007 (2012).
13. Churchill, H. O. H. *et al.* Superconductor-nanowire devices from tunneling to the multichannel regime: zero-bias oscillations and magnetoconductance crossover. *Phys. Rev. B* **87,** 241401(R) (2013).
14. Bertoni, A., Bordone, P., Brunetti, R., Jacoboni, C. & Reggiani, S. Quantum logic gates based on coherent electron transport in quantum wires. *Phys. Rev. Lett.* **84,** 5912–5915 (2000).
15. Blaauboer, M. & DiVincenzo, D. P. Detecting entanglement using a double-quantum-dot turnstile. *Phys. Rev. Lett.* **95,** 160402 (2005).
16. Bagraev, N. T., Shelykh, I. A., Ivanov, V. K. & Klyachkin, L. E. Spin depolarization in quantum wires polarized spontaneously in zero magnetic field. *Phys. Rev. B* **70,** 155315 (2004).
17. Tombros, N. *et al.* Quantized conductance of a suspended graphene nanoconstriction. *Nature Phys.* **7,** 697–700 (2011).
18. Goldhaber-Gordon, D. *et al.* Kondo effect in a single-electron transistor. *Nature* **391,** 156–159 (1998).
19. Cronenwett, S. M., Oosterkamp, T. H. & Kouwenhoven, L. P. A tunable Kondo effect in quantum dots. *Science* **281,** 540–544 (1998).
20. Jeong, H., Chang, A. M. & Melloch, M. R. The Kondo effect in an artificial quantum dot molecule. *Science* **293,** 2221–2223 (2001).
21. Jones, B. A., Kotliar, B. G. & Millis, A. J. Mean-field analysis of 2 antiferromagnetically coupled Anderson impurities. *Phys. Rev. B* **39,** 3415–3418 (1989).
22. Ivanov, T. The nonlinear conductance of a double quantum dot in the Kondo regime. *Europhys. Lett.* **40,** 183–188 (1997).
23. Pohjola, T. *et al.* Resonant tunneling through a two-level dot and double quantum dots. *Europhys. Lett.* **40,** 189–194 (1997).
24. Aono, T., Eto, M. & Kawamura, K. Conductance through quantum dot dimer below the Kondo temperature. *J. Phys. Soc. Jpn* **67,** 1860–1863 (1998).
25. Georges, A. & Meir, Y. Electronic correlations in transport through coupled quantum dots. *Phys. Rev. Lett.* **82,** 3508–3511 (1999).
26. Aguado, R. & Langreth, D. C. Out-of-equilibrium Kondo effect in double quantum dots. *Phys. Rev. Lett.* **85,** 1946–1949 (2000).
27. Aguado, R. & Langreth, D. C. Kondo effect in coupled quantum dots: a noncrossing approximation study. *Phys. Rev. B* **67,** 245307 (2003).
28. Kanisawa, K., Butcher, M. J., Yamaguchi, H. & Hirayama, Y. Imaging of Friedel oscillation patterns of two-dimensionally accumulated electrons at epitaxially grown InAs(111)A surfaces. *Phys. Rev. Lett.* **86,** 3384–3387 (2001).
29. Vernek, E. *et al.* Kondo regime in triangular arrangements of quantum dots: molecular orbitals, interference, and contact effects. *Phys. Rev. B* **80,** 035119 (2009).

30. Simion, G. E. & Giuliani, G. F. Friedel oscillations in a Fermi liquid. *Phys. Rev. B* **72,** 045127 (2005).

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

Acknowledgements We thank B. J. van Wees, A. Aqeel, S. Ludwig, J. von Delft and Y. Komijani for discussions and B. Wolfs, J. Holstein and M. de Roosz for technical assistance. We acknowledge financial support from the German programmes DFG-SPP 1285, Research School Ruhr-Universität Bochum and BMBF QuaHL-Rep 16BQ1035, and grants FIS2009-08744 and FIS2012-33521 from the Spanish Ministry of Economy and Innovation. M.J.I. acknowledges a scholarship from the Higher Education Commission of Pakistan. Y.M. and R.L. acknowledge support from the ISF.

Author Contributions M.J.I. was the lead researcher for experiments, with C.H.v.d.W. as supervisor, experimental contributions from E.J.K., J.B.D., J.P.d.J. and J.H.M.v.d.V., and design contributions from Y.M. The devices were fabricated from wafer material that was grown by D.R. and A.D.W. The calculations of electron transport in Kondo systems were carried out by R.A. The SDFT contribution was worked out by R.L. with Y.M. as supervisor. M.J.I., C.H.v.d.W. and Y.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 M.J.I. (m.j.iqbal@rug.nl) or C.H.v.d.W. (c.h.van.der.wal@rug.nl).

METHODS

Materials and device fabrication. QPC devices were fabricated with two different GaAs/AlGaAs heterostructures containing a 2DEG in a heterojunction quantum well. They were grown by molecular-beam epitaxy under similar conditions, and the properties of the 2DEGs were similar. We obtained very similar results with both materials. Most of the results presented in this Letter come from material 2, only the data in Fig. 1b, d comes from material 1.

Material 1 was a GaAs/$Al_{0.32}Ga_{0.68}As$ heterostructure with a 2DEG 114 nm below the surface. The electrons of the 2DEG were supplied by modulation doping with Si at about $1 \times 10^{24}\,m^{-3}$. At 4.2 K, the mobility of the 2DEG was $\mu = 159\,m^2\,V^{-1}\,s^{-1}$, and the electron density $n_s = 1.5 \times 10^{15}\,m^{-2}$ after cooling down in the dark. The layer with modulation doping started at 37 nm distance from the 2DEG position towards the wafer surface (this material was uniquely used in the related results presented in ref. 8).

Material 2 was a GaAs/$Al_{0.35}Ga_{0.65}As$ heterostructure with a 2DEG 110 nm below the surface. This material also had modulation doping with Si at about $1 \times 10^{24}\,m^{-3}$. At 4.2 K, the mobility of the 2DEG was $\mu = 118\,m^2\,V^{-1}\,s^{-1}$, and the electron density $n_s = 1.60 \times 10^{15}\,m^{-2}$. Here the layer with modulation doping started at 45 nm distance from the 2DEG position.

QPCs were realized by locally depleting the 2DEG below the surface of the GaAs/AlGaAs heterostructures. In the standard approach, applying negative voltage V_g to two metallic gate fingers on the surface of such material (Fig. 1a) induces an electrostatic potential barrier between source and drain reservoirs in the 2DEG, with a small tunable opening in the form of a saddle-point potential (Fig. 2b). Such devices with two gate fingers are denoted QPC_{2F} and these have a fixed channel length L. We also studied novel devices with six gate fingers (Fig. 2a), denoted QPC_{6F}, which have a channel with tunable length L_{eff}. Tuning of V_g on the gates allows for controlling the effective QPC shape (for details, see Supplementary Information section 5).

The depletion gates were defined with standard electron-beam lithography and lift-off techniques and by depositing 15 nm of Au with a Ti sticking layer. The reservoirs were connected to macroscopic leads via ohmic contacts, which were realized by annealing a thin Au/Ge/Ni layer that was deposited on the surface. Part of our data (including all the results presented in the main text, except for the data in Fig. 1b, d) was obtained after cooling down with about +0.3 V on the gates to suppress $1/f$ and telegraph noise in the conductance signals due to charge instabilities in the doping layer (Supplementary Information section 4). We obtained (apart from the change in noise properties) similar results for the cases with and without biased cool-down.

Measurement techniques and set-up. The measurements focus on the differential conductance G, which is obtained by applying and measuring voltage and current signals as in the simplified scheme in Fig. 1a. Results for G at zero bias voltage V_{sd} are called linear conductance, while results for G as a function of V_{sd} (bias spectroscopy) are called nonlinear conductance. Unless stated otherwise, the presented data were taken at a temperature T of 80 mK and zero magnetic field.

The presented results of linear and nonlinear conductance measurements all concern the differential conductance $G = dI/dV_{sd}$ (where I is the measured current). For linear conductance measurements, we used standard lock-in techniques (typically at 387 Hz), with an a.c. voltage bias $V_{sd} = V_{ac} = 10\,\mu V$. For the nonlinear conductance measurements, we superimposed an a.c. and a d.c. voltage bias, $V_{sd} = V_{dc} + V_{ac}$. We used an effective four-terminal measurement where we locally measured the source–drain voltage drop V_{sd} across the QPC, such that we can present results without significant contributions from series resistance. Only one of the source–drain contacts was connected to the grounded shielding of our set-up, and all gate voltages were applied with respect to this ground.

Measurements were performed in a dilution refrigerator with the sample at temperatures from ~5 mK to 4.2 K. For all our data, the temperature dependence saturated when cooling below ~80 mK. This is consistent with independent measurements of the lowest effective electron temperature that could be achieved with this set-up. The electron temperature of 80 mK allows for probing peak structures in nonlinear conductance traces as narrow as $4k_BT/e = 0.03$ mV (k_B is Boltzmann's constant).

The in-plane magnetic field was applied both parallel and perpendicular to the current direction and we measured devices with the current along the [110] direction and also along the [−110] direction of the crystal, but the results did not depend significantly on these orientations. Alignment of the sample with the magnetic field was within 1°, as determined from Hall voltage measurements on the 2DEG.

Kondo transport calculations. We obtained the theoretical results by calculating the current through a two-impurity Anderson model within the slave-boson noncrossing approximation[27] (detailed in Supplementary Information section 1).

SDFT calculations. We obtained the SDFT results by extending the work of ref. 7 (detailed in Supplementary Information section 2).

LETTER

doi:10.1038/nature12435

Catalytic conversion of nitrogen to ammonia by an iron model complex

John S. Anderson[1], Jonathan Rittle[1] & Jonas C. Peters[1]

The reduction of nitrogen (N_2) to ammonia (NH_3) is a requisite transformation for life[1]. Although it is widely appreciated that the iron-rich cofactors of nitrogenase enzymes facilitate this transformation[2–5], how they do so remains poorly understood. A central element of debate has been the exact site or sites of N_2 coordination and reduction[6,7]. In synthetic inorganic chemistry, an early emphasis was placed on molybdenum[8] because it was thought to be an essential element of nitrogenases[3] and because it had been established that well-defined molybdenum model complexes could mediate the stoichiometric conversion of N_2 to NH_3 (ref. 9). This chemical transformation can be performed in a catalytic fashion by two well-defined molecular systems that feature molybdenum centres[10,11]. However, it is now thought that iron is the only transition metal essential to all nitrogenases[3], and recent biochemical and spectroscopic data have implicated iron instead of molybdenum as the site of N_2 binding in the FeMo-cofactor[12]. Here we describe a tris(phosphine)borane-supported iron complex that catalyses the reduction of N_2 to NH_3 under mild conditions, and in which more than 40 per cent of the proton and reducing equivalents are delivered to N_2. Our results indicate that a single iron site may be capable of stabilizing the various N_xH_y intermediates generated during catalytic NH_3 formation. Geometric tunability at iron imparted by a flexible iron–boron interaction in our model system seems to be important for efficient catalysis[13–15]. We propose that the interstitial carbon atom recently assigned in the nitrogenase cofactor may have a similar role[16,17], perhaps by enabling a single iron site to mediate the enzymatic catalysis through a flexible iron–carbon interaction[18].

Nitrogen is fixed on a vast scale by the industrial Haber–Bosch process using a solid-state Fe catalyst at very high temperatures and pressures, and in nature by nitrogenase enzymes under ambient conditions[1]. These enzymes feature active-site cofactors rich in S and Fe (FeFe-cofactor), and can additionally contain a Mo (FeMo-cofactor; Fig. 1) or V (FeV-cofactor) site[2,3].

The intimate mechanism of biological nitrogen fixation is a topic that has fascinated chemists, biochemists and biologists[4–7]. Synthetic chemists have searched for decades for well-defined complexes that can catalyse N_2 reduction to NH_3[19–22]. Chatt's early work with low-valent Mo complexes provided a mechanistic outline for approaching this problem now commonly called the Chatt, or distal, mechanism, wherein a terminal nitride intermediate is generated on liberation of the first NH_3 equiv. (ref. 9; Fig. 1, bottom). Other mechanisms, including an 'alternating' scenario (Fig. 1, top), have also received attention[6]. So far, even modest catalysis of NH_3 production directly from N_2 by a well-defined model complex is still limited to the original tri(amido)amine Mo systems[10] and the more recently discovered phosphine-pincer Mo system[11]. Earlier work had established the electrochemical feasibility of an NH_3 production cycle with W (ref. 19).

Synthetic efforts to establish whether one or more Fe sites can catalyse N_2 reduction to NH_3 in a well-defined model complex have progressed more slowly. For example, previous Fe–N_2 model complexes have not afforded more than ~10% of NH_3 per Fe equiv. on treatment with proton sources[7,23,24]. Despite important advances[24], which have most recently included reductive N_2 cleavage at Fe (ref. 22) and the catalytic reductive silylation of N_2 facilitated by unknown Fe species derived from Fe precursors such as ferrocene and iron pentacarbonyl[25], the delivery of protons and electrons to N_2 to generate NH_3 catalytically at a synthetic Fe complex has remained an unsolved challenge. Here we show that a recently reported mononuclear Fe complex[13,14], [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] (TPB, tris(phosphine)borane; Fig. 2a), can catalyse the reduction of N_2 to NH_3 by protons and electrons in solution at −78 °C under one atmosphere of N_2.

The Fe centre of the '(TPB)Fe' fragment readily binds N_2 as evidenced by the featured five-coordinate complex [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] and the previously reported neutral N_2 adduct, (TPB)Fe(N_2) (ref. 14). The same scaffold also accommodates a variety of other nitrogenous ligands relevant to NH_3 generation, including terminally bonded

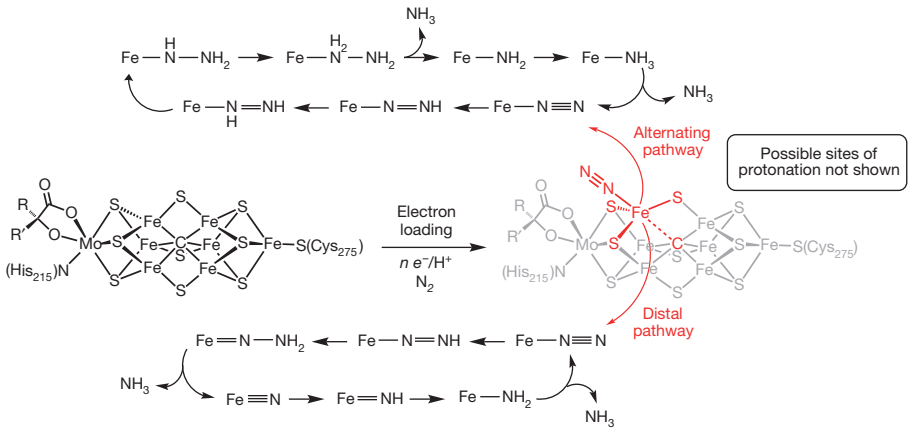


Figure 1 | Chemical line representations of the FeMo-cofactor of nitrogenase. A schematic depiction of postulated N_2 binding and reduction at an Fe site by limiting alternating (top) and distal (bottom) mechanisms. The drawing emphasizes a possible hemilabile role for the interstitial C atom with respect to an Fe–N_2 binding site.

[1]Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125, USA.

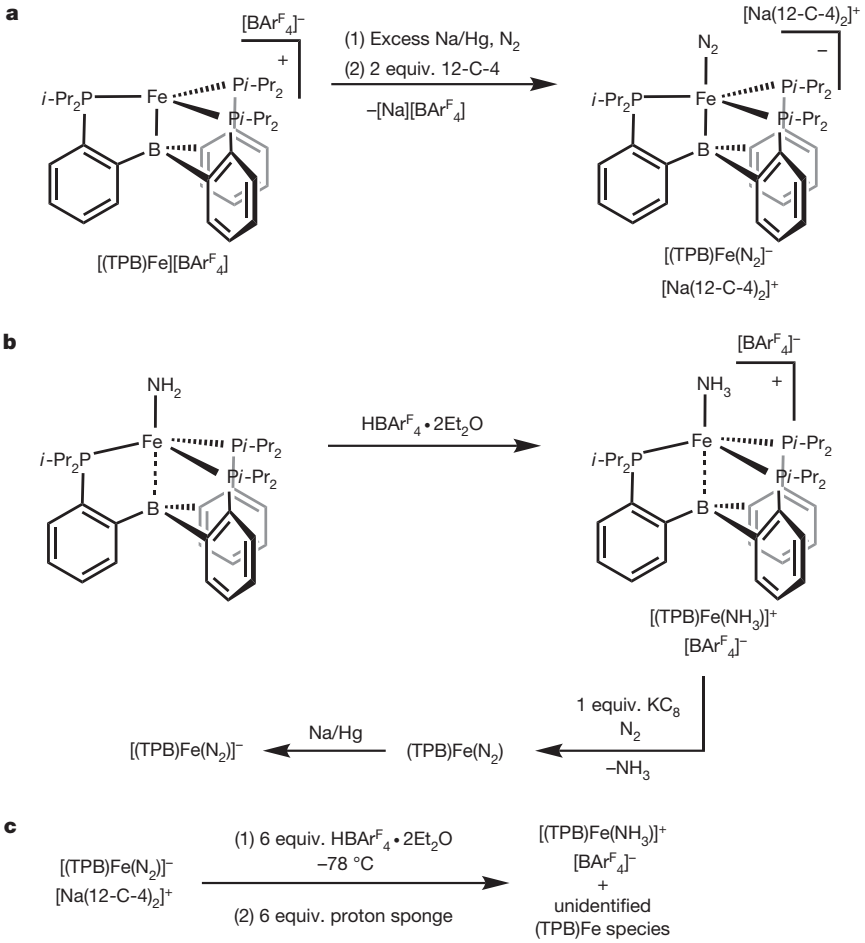


Figure 2 | Stoichiometric (TPB)Fe–N_2 model reactions. **a**, N_2 binding to [(TPB)Fe][BAr$^F_4$] under electron loading to generate [(TPB)Fe(N_2)][Na(12-crown-4)$_2$]. **b**, Reductive protonolysis of (TPB)Fe(NH_2) to release NH_3 with concomitant N_2 uptake. **c**, Generation of [(TPB)Fe(NH_3)][BAr$^F_4$] and other (TPB)Fe species on addition of acid to [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] at low temperature, followed by warming and then addition of base. 12-C-4, 12-crown-4.

NH_2, NH_3 and N_2H_4 ligands[15]. Both the substrate-free complex, [(TPB)Fe][BAr$^F_4$] (where [BAr$^F_4$]$^-$ represents the weakly coordinating anion [B(3,5-(CF$_3$)$_2$-C$_6$H$_3$)$_4$]$^-$; ref. 15), and (TPB)Fe(N_2) may be reduced to [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] by Na/Hg reduction under 1 atm N_2 followed by the addition of 2 equiv. of 12-crown-4 to encapsulate the sodium cation (Fig. 2a). Model reactions with silyl electrophiles have also shown that the Nβ of the coordinated N_2 ligand can be mono- or difunctionalized with concomitant increase in the Fe–B distance[13]. Furthermore, starting from (TPB)Fe(NH_2), a reductive protonation sequence has been established (Fig. 2b) that liberates NH_3 and affords (TPB)Fe(N_2) (ref. 15). This reaction sequence demonstrates the synthetic viability of reductive release of NH_3 from an Fe–NH_2 intermediate with simultaneous uptake of N_2. In sum, the rich reaction chemistry of the (TPB)Fe scaffold with nitrogenous ligands motivated us to examine more closely the possibility that it might catalyse N_2 fixation.

The addition of excess acid to [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] at −78 °C was investigated (Fig. 2c). When [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] was dissolved in tetrahydrofuran (THF), cooled to −78 °C and exposed to 6 equiv. of H^+ in the form of HBAr$^F_4$·2Et$_2$O, a previously unobserved yellow solution resulted that, on warming followed by addition of proton sponge (1,8-bis(dimethylamino)naphthalene), was shown by ^{1}H NMR analysis to contain the previously characterized complex[15] [(TPB)Fe(NH_3)][BAr$^F_4$] (~30–35% of the total Fe), along with resonances consistent with [(TPB)Fe][BAr$^F_4$] (~40–45% of the total Fe) and two other minor and as yet unidentified paramagnetic (TPB)Fe-species (Supplementary Information). An independent electron paramagnetic resonance (EPR) study of this low-temperature protonation reaction in 2-methyltetrahydrofuran revealed a new rhombic $S = 1/2$ signal (Supplementary Information) that we speculate may be (TPB)Fe(=N–NH_2)$^+$ or an alternative structural isomer such as (TPB)Fe(NH=NH)$^+$. Spin quantification of this species showed it to represent >85% of the Fe species in solution, and its rhombic EPR spectrum is highly similar to the rhombic EPR signature that is obtained on oxidation of (TPB)Fe=N(*p*-tolyl) to generate (TPB)Fe=N(*p*-tolyl)$^+$ (Supplementary Information). Subsequent low-temperature reduction of a similarly prepared mixture regenerates [(TPB)Fe(N_2)][Na(12-crown-4)$_2$], as determined by infrared spectroscopy, suggesting the possibility of cycling protonation and reduction with this Fe system.

To explore the possibility of N_2 reduction catalysis using this (TPB)Fe system, we tested several reductants (for example Na[naphthalenide], Mg(THF)$_3$(anthracene), Na/Hg and KC_8) and acids (for example HCl, trifluoromethanesulphonic acid and HBAr$^F_4$·2Et$_2$O) in a variety of solvents and solvent mixtures (for example THF, dimethoxyethane, diethyl ether and toluene). When carried out at −78 °C numerous reaction conditions showed unusually high yields of NH_3 relative to the number of Fe equivalents in the reaction vessel, and the combination of KC_8, HBAr$^F_4$·2Et$_2$O and Et$_2$O solvent enabled the catalytic generation of NH_3.

In a representative catalytic run, red [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] was suspended in diethyl ether in a reaction vessel at −78 °C, followed by the sequential addition of excess acid and then excess reductant. Ammonia analysis followed the indophenol protocol (Methods and Supplementary Information) and the independent identification of ammonium salts by ^{1}H NMR spectroscopy in dimethylsulphoxide

(DMSO) by comparison with an authentic sample of $[NH_4][Cl]$ (Fig. 3a). An experiment performed using the $[(TPB)Fe(^{14}N_2)][Na(12\text{-crown-}4)_2]$ catalyst under an $^{15}N_2$ atmosphere, followed by ^1H NMR analysis (Fig. 3a) of the volatile reaction products, confirmed the production of $[^{15}NH_4][Cl]$, as expected, with only trace $[^{14}NH_4][Cl]$. This overall procedure has been repeated many times, and Table 1 includes data averaged from 16 independent runs (entry 1) in which yields were, on average, 7.0 equiv. NH_3 per Fe equiv. Using 7.0 equiv. NH_3 as the product stoichiometry, 44% of the added protons are reliably delivered to N_2 to produce NH_3. Our individual runs reached a maximum of 8.5 equiv. NH_3 per Fe equiv. under these standard conditions. $[(TPB)Fe][BAr^F{}_4]$ is also an effective catalyst and afforded 6.2 ± 0.7 equiv. NH_3 per added Fe equiv. (Table 1, entry 2). For comparison, the Mo systems of refs 10, 11 have afforded between 7 and 12 equiv. NH_3 per Mo equiv. The current Fe system seems to be active at an unusually low temperature (−78 °C) but benefits from a strong reductant (KC_8). We do not yet know whether conditions can be found that will tolerate a milder reductant, for example by circumventing the need to generate the $(TPB)Fe(N_2)^-$ anion during catalysis.

Table 1 lists several sets of conditions (entries 10–15) other than the standard conditions described above that were studied. Several of these alternative conditions showed NH_3 generation, although not in catalytic or even in high yields. N_2H_4 was not detected as an additional product when using the standard catalytic protocol for NH_3 generation with $[(TPB)Fe(N_2)][Na(12\text{-crown-}4)_2]$ (Supplementary Information). If 2 equiv. N_2H_4 (per Fe) were added to $[(TPB)Fe(N_2)][Na(12\text{-crown-}4)_2]$ in diethyl ether and the mixture was then subjected to the standard catalytic conditions and work-up, only trace N_2H_4 remained (Supplementary Information). This result suggests that if N_2H_4 is generated as an intermediate during catalysis, it would not likely be detectable on work-up and analysis. It is worth noting that $HBAr^F{}_4 \cdot 2Et_2O$ and KC_8 react in the absence of an Fe precursor, under the standard catalytic conditions at −78 °C, to generate H_2 but not NH_3 (>75% yield of H_2 after 40 min). That H_2 generation is kinetically feasible without the addition of an Fe precursor, and yet NH_3 can nonetheless be generated on the addition of $[(TPB)Fe(N_2)][Na(12\text{-crown-}4)_2]$ or $[(TPB)Fe][BAr^F{}_4]$,

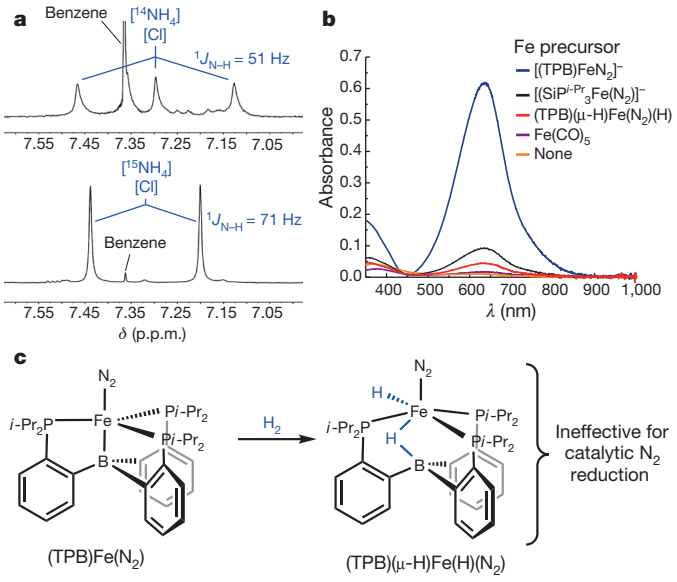
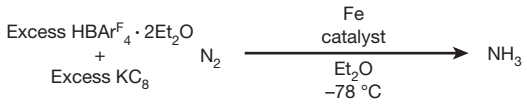


Figure 3 | Spectral data for ammonia analysis, and catalyst poisoning. **a**, ^1H NMR spectrum (300 MHz) of $[^{14}NH_4][Cl]$ in DMSO-d_6 produced from a catalytic run under $^{14}N_2$ (top) and of $[^{15}NH_4][Cl]$ in DMSO-d_6 produced from an independent catalytic run under 1 atm $^{15}N_2$. δ, chemical shift. **b**, Representative optical data for NH_3 product analysis using the indophenol method from catalytic runs using the standard conditions with the precursors indicated. **c**, H_2 binds to $(TPB)Fe(N_2)$ to generate $(TPB)(\mu\text{-}H)Fe(N_2)(H)$, which itself is ineffective for the catalytic generation of NH_3 from N_2 under the standard conditions.

Table 1 | Ammonia generation from N_2 mediated by Fe precursors



Using standard catalytic conditions as described in the text

Entry	Fe precursor	NH_3 equiv./Fe†
1	$[(TPB)Fe(N_2)][Na(12\text{-crown-}4)_2]$	7.0 ± 1‡
2	$[(TPB)Fe][BAr^F{}_4]$	6.2
3	$[(SiP^{i\text{-}Pr}{}_3)Fe(N_2)][Na(12\text{-crown-}4)_2]$	0.7
4	$(TPB)(\mu\text{-}H)Fe(N_2)(H)$	0.5
5	$FeCl_2 \cdot 1.5THF$	<0.1
6	$FeCl_3$	<0.1
7	Cp_2Fe	<0.2
8	$Fe(CO)_5$	<0.1
9	None	<0.1

Variations on standard conditions using $[(TPB)FeN_2][Na(12\text{-crown-}4)_2]$

Entry	Variation	NH_3 equiv./Fe†
10	HOTf as acid	0.4
11	[Lutidinium][$BAr^F{}_4$] as acid	<0.1
12	HCl as acid	<0.1
13	$Cp^*{}_2Co$ as reductant	0.6
14	$Cp^*{}_2Cr$ as reductant	<0.2
15	K metal as reductant	0.4

† NH_3 was collected by vacuum transfer of the reaction volatiles into HCl in diethyl ether. A dimethoxyethane solution of [Na][Ot-Bu] (20 equiv. relative to Fe) was added to the reaction vessel residue, followed by an additional vacuum transfer, to ensure complete liberation of all NH_3. The $[NH_4][Cl]$ precipitate formed in the acidic Et_2O collection vessel was reconstituted in deionized H_2O, from which an aliquot was taken for indophenol quantification. Analysis of the $[NH_4][Cl]$ formed by ^1H NMR spectroscopy in DMSO, compared with an authentic sample, provided independent confirmation of NH_3 generation. Data for individual experimental runs, and additional runs using potential precatalysts that are not presented in this table, are provided in Supplementary Information. Unless noted otherwise, all yields are reported as an average of 4 runs.

‡ Average of 16 runs.

HOTf, trifluoromethanesulphonic acid; *i*-Pr, isopropyl; lutidinium, 2,6-dimethylpyridinium; *t*-Bu, *tert*-butyl.

underscores the facility with which this Fe system mediates overall hydrogen-atom delivery to N_2.

To explore further whether a (TPB)Fe-containing precursor is needed to facilitate the overall catalysis, beyond the stoichiometric model reactions summarized above, we canvassed several Fe complexes under analogous conditions. Of most interest is the complex $[(SiP^{i\text{-}Pr}{}_3)Fe(N_2)][Na(12\text{-crown-}4)_2]$, which is isostructural to $[(TPB)Fe(N_2)][Na(12\text{-crown-}4)_2]$ but in which the B atom of TPB is replaced by a Si atom[26]. A central difference between (TPB)Fe and $(SiP^{i\text{-}Pr}{}_3)Fe$ complexes is the far greater flexibility of the Fe–B bond by comparison with the Fe–Si bond that is positioned *trans* to the apical ligand[13–15,26]. Although some NH_3 generation was observed for $[(SiP^{i\text{-}Pr}{}_3)Fe(N_2)][Na(12\text{-crown-}4)_2]$ when subjected to the standard catalytic reaction conditions described above, sub-stoichiometric yields of NH_3 relative to Fe were obtained (0.7 ± 0.5 equiv. NH_3 per Fe equiv.; Table 1, entry 3). We also conducted additional control experiments under the standard catalytic conditions with $FeCl_2 \cdot 1.5THF$, $FeCl_3$, Cp_2Fe (ref. 25) and $Fe(CO)_5$ (ref. 25) (entries 5–8) and found that only trace amounts of NH_3 (<0.2 equiv. in all cases on average; four runs) were produced by these Fe precursors[27] (Supplementary Information). The known phosphine-supported Fe(0)–N_2 complex $Fe(depe)_2(N_2)$ (ref. 28) was also subjected to the standard conditions and afforded sub-stoichiometric yields of NH_3 per Fe equiv. (Supplementary Information).

In separate work, the addition of H_2 to $(TPB)Fe(N_2)$ was shown to generate $(TPB)(\mu\text{-}H)Fe(N_2)(H)$ as a stable product[29] (Fig. 3c). We hence suspected that catalyst poisoning might occur in part through the formation of $(TPB)(\mu\text{-}H)Fe(N_2)(H)$ under the catalytic reaction conditions. In accord with this idea, when $[(TPB)Fe(N_2)][Na(12\text{-crown-}4)_2]$ was exposed to 10 equiv. $HBAr^F{}_4 \cdot 2Et_2O$ and 12 equiv. KC_8 at low temperature, infrared and ^{31}P NMR analysis of the resulting mixture showed the presence of $(TPB)(\mu\text{-}H)Fe(N_2)(H)$ through its signature spectroscopic features[29] (30% of total Fe by ^{31}P NMR integration;

Supplementary Information). (TPB)(μ-H)Fe(N_2)(H) is stable for short periods to both $HBAr^F_4 \cdot 2Et_2O$ and also KC_8 in Et_2O at room temperature, and when subjected to the standard catalytic conditions for NH_3 production liberates only 0.5 ± 0.1 equiv. NH_3 per Fe equiv. (Table 1, entry 4, and Fig. 3b).

The general absence of a functional, catalytic Fe model system over the past few decades has often led to an emphasis on Mo (ref. 30) as a plausible site of N_2 uptake and reduction at the most widely studied FeMo-cofactor. Although this may yet prove to be true, recent spectroscopic and biochemical evidence has sharpened the focus on an Fe centre as the N_2-binding site[12]. The results reported here establish that it is possible to catalyse the conversion of N_2 to NH_3 by protons and electrons using a well-defined mononuclear Fe–N_2 complex, and suggests the possibility that a single Fe-binding site of the cofactor could in principle mediate N_2 reduction catalysis[18]. To achieve this catalytic behaviour, geometric flexibility at the Fe–N_2 binding site would be beneficial as it would stabilize N_xH_y intermediates with different electronic structure requirements. Such geometric and redox flexibility, under the local three-fold symmetry presented by an Fe centre, its three neighbouring sulphides and the interstitial light atom of the FeMo-cofactor[16,17], may at least in part be achieved by attributing a hemilabile role to the interstitial C atom (Fig. 1). Such a role could serve to expose an initial Fe–N_2 binding site by Fe–C elongation. Subsequent modulation of the Fe–C interaction and, hence, the local Fe geometry as a function of the N_2 reduction state would enable the Fe centre to stabilize the various N_xH_y intermediates along a pathway to NH_3. Although it is inherently speculative, this hypothesis is rooted in the functional (TPB)Fe catalysis discussed here, along with the types of (TPB)Fe complex and stoichiometric transformation described previously for this scaffold[13–15].

Many mechanistic questions arise from the present Fe catalyst system. Although all of the model complexes relevant to the (TPB)Fe–N_xH_y system are mononuclear, we cannot exclude the possibility of bimolecular reaction intermediates. More generally, we do not know whether the N_2 reduction catalysis proceeds along a distal pathway via a terminal nitride intermediate, such as (TPB)Fe(N) or (TPB)Fe(N)$^+$, via intermediates along an alternating pathway, such as (TPB)Fe–NH–NH_2 or (TPB)Fe–NH_2–NH_2^+, or via some hybrid pathway. The tentatively assigned (TPB)Fe=N–NH_2^+ species that can be observed by EPR spectroscopy provides us with a useful starting point for such studies. In light of the identification of C as the interstitial light atom of the cofactor, it is also of interest to develop and compare synthetic model systems that feature a C atom in the ligand backbone instead of a B atom.

METHODS SUMMARY

General considerations. All complexes and reagents were prepared according to literature procedures referenced in the Supplemental Information unless otherwise noted. Manipulations were carried out under an N_2 atmosphere utilizing standard dry glove-box or Schlenk-line techniques. All solvents used were deoxygenated and dried by N_2 sparging followed by passage through an activated alumina column.

Spectroscopy. NMR measurements were made using Varian 300-MHz spectrometers. Deuterated solvents for these measurements were obtained from Cambridge Isotope Laboratories and were dried and degassed before use. All ^{1}H NMR spectra were referenced to residual solvent peaks. EPR X-band spectra were obtained on a Bruker EMX spectrometer with the aid of the Bruker WIN-EPR software suite, version 3.0. The EPR spectrometer was equipped with a rectangular cavity that operated in the TE_{102} mode. Temperature control was achieved with a liquid-N_2-filled quartz Dewar flask in which the sample was submerged during data collection. Ultraviolet–visible spectra were acquired on a Cary 50 spectrometer from 1,100 to 200 nm in the fast-scan mode. Infrared spectra were obtained using KBr pellets on a Bio-Rad Excalibur FTS 3000 spectrometer using Varian RESOLUTIONS PRO software set at a resolution of 4 cm^{-1}.

Catalysis and ammonia collection and quantification. The standard catalysis protocol involved the addition of first acid and then reductant to a suspension of the precatalyst in diethyl ether at −78 °C in a closed vessel under 1 atm N_2. Ammonia produced during each catalytic run was collected by vacuum transfer of the reaction volatiles onto anhydrous HCl in diethyl ether. The resulting slurry was dried and extracted into water and aliquots were then tested for the presence of ammonia by the indophenol method.

Full Methods and any associated references are available in the online version of the paper.

Received 8 February; accepted 4 July 2013.

1. Smil, V. *Enriching the Earth* (MIT Press, 2001).
2. Burgess, B. K. & Lowe, D. J. Mechanism of molybdenum nitrogenase. *Chem. Rev.* **96,** 2983–3012 (1996).
3. Eady, R. R. Structure-function relationships of alternative nitrogenases. *Chem. Rev.* **96,** 3013–3030 (1996).
4. Howard, J. B. & Rees, D. C. How many metals does it take to fix N_2? A mechanistic overview of biological nitrogen fixation. *Proc. Natl Acad. Sci. USA* **103,** 17088–17093 (2006).
5. Schwarz, G., Mendel, R. R. & Ribbe, M. W. Molybdenum cofactors, enzymes and pathways. *Nature* **460,** 839–847 (2009).
6. Seefeldt, L., Hoffman, B. M. & Dean, D. R. Mechanism of Mo-dependent nitrogenase. *Annu. Rev. Biochem.* **78,** 701–722 (2009).
7. Peters, J. C. & Mehn, M. P. in *Activation of Small Molecules: Organometallic and Bioinorganic Perspectives* (ed. Tolman, W. B.) 81–119 (Wiley-VCH, 2006).
8. Shilov, A. E. *et al.* New nitrogenase model for reduction of molecular nitrogen in protonic media. *Nature* **231,** 460–461 (1971).
9. Chatt, J., Dilworth, J. R. & Richards, R. L. Recent advances in chemistry of nitrogen-fixation. *Chem. Rev.* **78,** 589–625 (1978).
10. Yandulov, D. V. & Schrock, R. R. Catalytic reduction of dinitrogen to ammonia at a single molybdenum center. *Science* **301,** 76–78 (2003).
11. Arashiba, A., Miyake, Y. & Nishibayashi, Y. A molybdenum complex bearing PNP-type pincer ligands leads to the catalytic reduction of dinitrogen into ammonia. *Nature Chem.* **3,** 120–125 (2011).
12. Hoffman, B. M., Dean, D. R. & Seefeldt, L. C. Climbing nitrogenase: toward a mechanism of enzymatic nitrogen fixation. *Acc. Chem. Res.* **42,** 609–619 (2009).
13. Moret, M.-E. & Peters, J. C. N_2 functionalization at iron metallaboratranes. *J. Am. Chem. Soc.* **133,** 18118–18121 (2011).
14. Moret, M.-E. & Peters, J. C. Terminal iron dinitrogen and iron imide complexes supported by a tris(phosphino)borane ligand. *Angew. Chem. Int. Ed.* **50,** 2063–2067 (2011).
15. Anderson, J. A., Moret, M.-E. & Peters, J. C. Conversion of Fe–NH_2 to Fe–N_2 with release of NH_3. *J. Am. Chem. Soc.* **135,** 534–537 (2013).
16. Lancaster, K. M. *et al.* X-ray emission spectroscopy evidences a central carbon in the nitrogenase iron-molybdenum cofactor. *Science* **334,** 974–977 (2011).
17. Spatzal, T. M. *et al.* Evidence for interstitial carbon in nitrogenase FeMo cofactor. *Science* **334,** 940 (2011).
18. Hendrich, M. P. *et al.* On the feasibility of N_2 fixation via a single-site Fe^I/Fe^{IV} cycle: spectroscopic studies of $Fe^I(N_2)Fe^I$, $Fe^{IV}{\equiv}N$, and related species. *Proc. Natl Acad. Sci. USA* **103,** 17107–17112 (2006).
19. Pickett, C. J. & Talarmin, J. Electrosynthesis of ammonia. *Nature* **317,** 652–653 (1985).
20. Fryzuk, M. D., Love, J. B., Rettig, S. J. & Young, V. G. Transformation of coordinated dinitrogen by reaction with dihydrogen and primary silanes. *Science* **275,** 1445–1447 (1997).
21. Pool, J. A., Lobkovsky, E. & Chirik, P. J. Hydrogenation and cleavage of dinitrogen to ammonia with a zirconium complex. *Nature* **427,** 527–530 (2004).
22. Rodriguez, M. M., Bill, E., Brennessel, W. W. & Holland, P. L. N_2 reduction and hydrogenation to ammonia by a molecular iron-potassium complex. *Science* **334,** 780–783 (2011).
23. Hills, A. *et al.* Bis[1,2-bis(dimethylphosphino)ethane]dihydrogenhydridoiron(II) tetraphenylborate as a model for the function of nitrogenases. *J. Chem. Soc. Dalton Trans.* **25,** 3041–3049 (1993).
24. Crossland, J. L. & Tyler, D. R. Iron-dinitrogen coordination chemistry: dinitrogen activation and reactivity. *Coord. Chem. Rev.* **255,** 949–974 (2011).
25. Yuki, M., Tanaka, H., Miyake, Y., Yoshizawa, K. & Nishibayashi, Y. Iron-catalysed transformation of molecular dinitrogen into silylamine under ambient conditions. *Nature Commun.* **3,** 1254 (2012).
26. Lee, Y. H., Mankad, N. P. & Peters, J. C. Triggering N_2 uptake via redox-induced expulsion of coordinated NH_3 and N_2 silylation at trigonal bipyramidal iron. *Nature Chem.* **2,** 558–565 (2010).
27. Bell, L. G. & Brintzinger, H. H. Stoichiometric reduction of molecular nitrogen by iron complexes. *J. Am. Chem. Soc.* **92,** 4464–4465 (1970).
28. Hall, D. A. & Leigh, G. J. Reduction of dinitrogen bound at an iron(0) centre. *J. Chem. Soc. Dalton Trans.* 3539–3541 (1996).
29. Fong, H., Moret, M.-E., Lee, Y. H. & Peters, J. C. Heterolytic H_2 cleavage and catalytic hydrogenation by an iron metallaboratrane. *Organometallics* **32,** 3053–3062 (2013).
30. Schrock, R. R. Nitrogen reduction: molybdenum does it again. *Nature Chem.* **3,** 95–96 (2011).

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

Acknowledgements This work was supported by the NIH (GM 070757) and the Gordon and Betty Moore Foundation. A. Takaoka is thanked for developing the calibration curves used for ammonia and hydrazine quantification. D. Rees and D. Newman are acknowledged for many discussions.

Author Contributions J.S.A., J.R. and J.C.P. designed the study. J.S.A. and J.R. conducted the experiments. J.S.A., J.R. and J.C.P. interpreted the data. J.S.A., J.R. and J.C.P. 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 J.C.P. (jpeters@caltech.edu).

METHODS

General considerations. We prepared [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] (ref. 14), [(TPB)Fe][BAr^F_4] (ref. 15), (TPB)(μ-H)Fe(H)(N_2) (ref. 29), [lutidinium][BAr^F_4] (ref. 31), $HBAr^F_4 \cdot 2Et_2O$ (ref. 32), [($SiP^{i\text{-}Pr}_3$)Fe(N_2)][Na(12-crown-4)$_2$] (ref. 26), $FeCl_2 \cdot (THF)_{1.5}$ (ref. 33), KC_8 (ref. 34), [(TPB)Fe(NH_3)][BAr^F_4] (ref. 15), [(TPB)Fe(N_2H_4)][BAr^F_4] (ref. 15) and Fe(depe)$_2N_2$ (ref. 28) were prepared according to literature procedures. Labelled $^{15}N_2$ (98% purity) was obtained from Cambridge Isotope Laboratories. Solvents used for catalytic runs were additionally stirred for more than 2 h over Na/K alloy and then filtered before use, and subjected to standard sparging (Ar gas) and passage through an activated alumina column.

Ammonia quantification. A Schlenk tube was charged with HCl (3 ml of a 2.0 M solution in Et_2O, 6 mmol). Reaction mixtures were vacuum-transferred into this collection flask. Residual solid in the reaction vessel was treated with a solution of [Na][O-*t*-Bu] (40 mg, 0.4 mmol) in 1,2-dimethoxyethane (1 ml) and sealed. The resulting suspension was stirred for 10 min before all volatiles were again vacuum-transferred into the collection flask. After completion of the vacuum transfer, the flask was sealed and warmed to room temperature. Solvent was removed *in vacuo* and the remaining residue was dissolved in H_2O (1 ml). An aliquot of this solution (20 or 40 μl) was then analysed for the presence of NH_3 (trapped as [NH_4][Cl]) by the indophenol method[35]. Quantification was performed with ultraviolet–visible spectroscopy by analysing absorbance at 635 nm. The tables in Supplementary Information show the raw data for the runs. Runs with small absorbance levels ($<$0.02 absorbance units) suffer from a large degree of error due to a small signal-to-noise ratio. Catalytic runs that used a 40 μl aliquot are denoted with an asterisk, accounting for larger relative absorbances.

Standard catalytic protocol. We suspended [(TPB)Fe(N_2)][Na(12-crown-4)$_2$] (2 mg, 0.002 mmol) in Et_2O (0.5 ml) in a 20-ml scintillation vial equipped with a stir bar. This suspension was vigorously stirred and cooled to −78 °C in a cold well inside the glove box. A similarly cooled solution of $HBAr^F_4 \cdot 2Et_2O$ (93 mg, 0.092 mmol) in Et_2O (1.5 ml) was added to the suspension in one portion with rapid stirring. Any remaining acid was dissolved in cold Et_2O (0.25 ml) and added subsequently. The reaction mixture turned light yellow-orange and became homogeneous on addition of acid, and the resulting solution was stirred for 5 min before being transferred into a pre-cooled Schlenk tube equipped with a stir bar. The original reaction vial was washed with cold Et_2O (0.25 ml) and was subsequently transferred to the Schlenk tube. Solid KC_8 (15 mg, 0.100 mmol) was suspended in cold Et_2O (0.75 ml) and added drop by drop to the rapidly stirred solution in the Schlenk tube, which was then tightly sealed. The reaction was stirred for 40 min at −78 °C before being warmed to room temperature and stirred for a further 15 min.

Detailed experimental procedures and data for all experiments can be found in Supplemental Information.

31. Yandulov, D. V. & Schrock, R. R. Reduction of dinitrogen to ammonia at a well-protected reaction site in a molybdenum triamidoamine complex. *J. Am. Chem. Soc.* **124,** 6252–6253 (2002).
32. Brookhart, M., Grant, B. & Volpe, A. F., Jr. $[(3,5\text{-}(CF_3)_2C_6H_3)_4B]^-[H(OEt_2)_2]^+$: a convenient reagent for generation and stabilization of cationic, highly electrophilic organometallic complexes. *Organometallics* **11,** 3920–3922 (1992).
33. Kern, R. J. Tetrahydrofuran complexes of transition metal chlorides. *J. Inorg. Nucl. Chem.* **24,** 1105–1109 (1962).
34. Weitz, I. S. & Rabinovitz, M. The application of C_8K for organic synthesis: reduction of substituted naphthalenes. *J. Chem. Soc. Perkin Trans.* **1,** 117–120 (1993).
35. Weatherburn, M. W. Phenol-hypochlorite reaction for determination of ammonia. *Anal. Chem.* **39,** 971–974 (1967).

LETTER

doi:10.1038/nature12434

Asymmetric effects of daytime and night-time warming on Northern Hemisphere vegetation

Shushi Peng[1], Shilong Piao[1,2], Philippe Ciais[3], Ranga B. Myneni[4], Anping Chen[5], Frédéric Chevallier[3], Albertus J. Dolman[6], Ivan A. Janssens[7], Josep Peñuelas[8,9], Gengxin Zhang[2], Sara Vicca[7], Shiqiang Wan[10], Shiping Wang[2] & Hui Zeng[11]

Temperature data over the past five decades show faster warming of the global land surface during the night than during the day[1]. This asymmetric warming is expected to affect carbon assimilation and consumption in plants, because photosynthesis in most plants occurs during daytime and is more sensitive to the maximum daily temperature, T_{max}, whereas plant respiration occurs throughout the day[2] and is therefore influenced by both T_{max} and the minimum daily temperature, T_{min}. Most studies of the response of terrestrial ecosystems to climate warming, however, ignore this asymmetric forcing effect on vegetation growth and carbon dioxide (CO_2) fluxes[3–6]. Here we analyse the interannual covariations of the satellite-derived normalized difference vegetation index (NDVI, an indicator of vegetation greenness) with T_{max} and T_{min} over the Northern Hemisphere. After removing the correlation between T_{max} and T_{min}, we find that the partial correlation between T_{max} and NDVI is positive in most wet and cool ecosystems over boreal regions, but negative in dry temperate regions. In contrast, the partial correlation between T_{min} and NDVI is negative in boreal regions, and exhibits a more complex behaviour in dry temperate regions. We detect similar patterns in terrestrial net CO_2 exchange maps obtained from a global atmospheric inversion model. Additional analysis of the long-term atmospheric CO_2 concentration record of the station Point Barrow in Alaska suggests that the peak-to-peak amplitude of CO_2 increased by 23 ± 11% for a +1 °C anomaly in T_{max} from May to September over lands north of 51° N, but decreased by 28 ± 14% for a +1 °C anomaly in T_{min}. These lines of evidence suggest that asymmetric diurnal warming, a process that is currently not taken into account in many global carbon cycle models, leads to a divergent response of Northern Hemisphere vegetation growth and carbon sequestration to rising temperatures.

Both modelling and observational studies suggest that climate warming will probably enhance vegetation growth in northern terrestrial ecosystems, increasing carbon storage[3–6]. However, most of these studies do not consider the potential effects of asymmetric daytime and night-time warming. The increasing rate of global land-surface daily minimum temperature (T_{min}) over the past five decades is 1.4 times that of daily maximum temperature (T_{max})[1]. This asymmetric warming is expected to affect carbon assimilation and consumptions in plants, given that photosynthesis in most plants occurs during the daytime and thus is more sensitive to T_{max}, whereas plant respiration occurs throughout the whole day[2], and is therefore influenced by both T_{max} and T_{min}.

A few field experiments have been conducted to study the effects of asymmetric warming on ecosystems[7,8]. These and other direct field observations[9–11] reveal different effects of temperature during daytime and night time on vegetation growth and CO_2 fluxes. Yet the scarcity and short duration of field experiments makes it difficult to assess the large-scale response of vegetation to changes in T_{max} and T_{min}, which may differ between regions and ecosystems[7–10]. In this study, we use 28 years of satellite-derived NDVI data, global time-varying maps of land net CO_2 exchange (NCE) fluxes from an atmospheric inversion model, and *in situ* atmospheric CO_2 concentration measurements, together with T_{max} and T_{min} data (Methods) to study the spatial patterns of covariations between interannual changes in T_{min} and T_{max} and changes in the NDVI and NCE of northern ecosystems.

We first examine the linkage between T_{max} or T_{min} and growing-season NDVI (April to October). There is a high positive correlation between T_{max} and T_{min} (Supplementary Fig. 1), and we studied the apparent responses of NDVI to T_{max} and T_{min} with partial correlation analyses to remove the covariate effects between T_{max} and T_{min} (see Methods). When the effects of growing-season T_{min}, precipitation and solar radiation are removed in the partial correlation, the individual effect of growing-season T_{max} interannual changes on interannual NDVI shows remarkable spatial patterns (Fig. 1a, c). In most of the boreal zone (>50° N), cold mountain regions (such as the Tibetan plateau), and wet temperate regions (Japan and southern China), growing-season NDVI is positively correlated with growing-season T_{max}, with statistical significance at the 0.05 level for about 15% of the area north of 25° N (Supplementary Table 1). This relatively small percentage of area showing significant relationships can be explained by the nonlinear relationships between vegetation productivity and climate variables[12], the impact of non-growing-season factors such as frost frequency and snow depth[13,14], and the various factors controlling vegetation growth in different areas, which limit the prevalence of one dominant environmental factor such as T_{max} for vegetation growth to only a small proportion of the entire study area. For example, if we consider boreal regions where vegetation growth is limited by temperature[6,15,16], a significant positive partial correlation between growing-season NDVI and T_{max} is found in about 22% of the area (Supplementary Table 1). This positive partial covariation between growing-season T_{max} and NDVI is particularly pronounced in north-western North America and Siberia (Fig. 1a). In contrast, in drier temperate regions (such as western China, central Eurasia, central and southwestern North America), interannual T_{max} anomalies exhibit negative partial correlations with NDVI.

However, when the effects of growing-season T_{max}, precipitation, and solar radiation are removed, the partial correlation between growing-season T_{min} and NDVI is found to be negative over most boreal and wet temperate regions (Fig. 1b, d). In arid and semi-arid regions, particularly grasslands in central North America and temperate China, the growing-season NDVI is significantly positively correlated with T_{min} (Fig. 1b and Supplementary Table 2). These opposite responses of

[1]Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing 100871, China. [2]Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing 100085, China. [3]Laboratoire des Sciences du Climat et de l'Environnement (LSCE), UMR CEA-CNRS, Batîment 709, CE, L'Orme des Merisiers, F-91191 Gif-sur-Yvette, France. [4]Department of Earth and Environment, Boston University, 675 Commonwealth Avenue, Boston, Massachusetts 02215, USA. [5]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544-1003, USA. [6]Department of Earth Sciences, VU University Amsterdam, Boelelaan 1085, 1081 HV, Amsterdam, The Netherlands. [7]Department of Biology, University of Antwerp, Universiteitsplein 1, 2610 Wilrijk, Belgium. [8]Centre de Recerca Ecològica i Aplicacions Forestals (CREAF), Cerdanyola del Valles, Barcelona 08193, Catalonia, Spain. [9]Consejo Superior de Investigaciones Científicas (CSIC), Global Ecology Unit CREAF-CEAB-CSIC-UAB, Cerdanyola del Valles, Barcelona 08193, Catalonia, Spain. [10]College of Life Sciences, Henan University, Kaifeng 475001, China. [11]Peking University Shenzhen Graduate School, Shenzhen 518055, China.

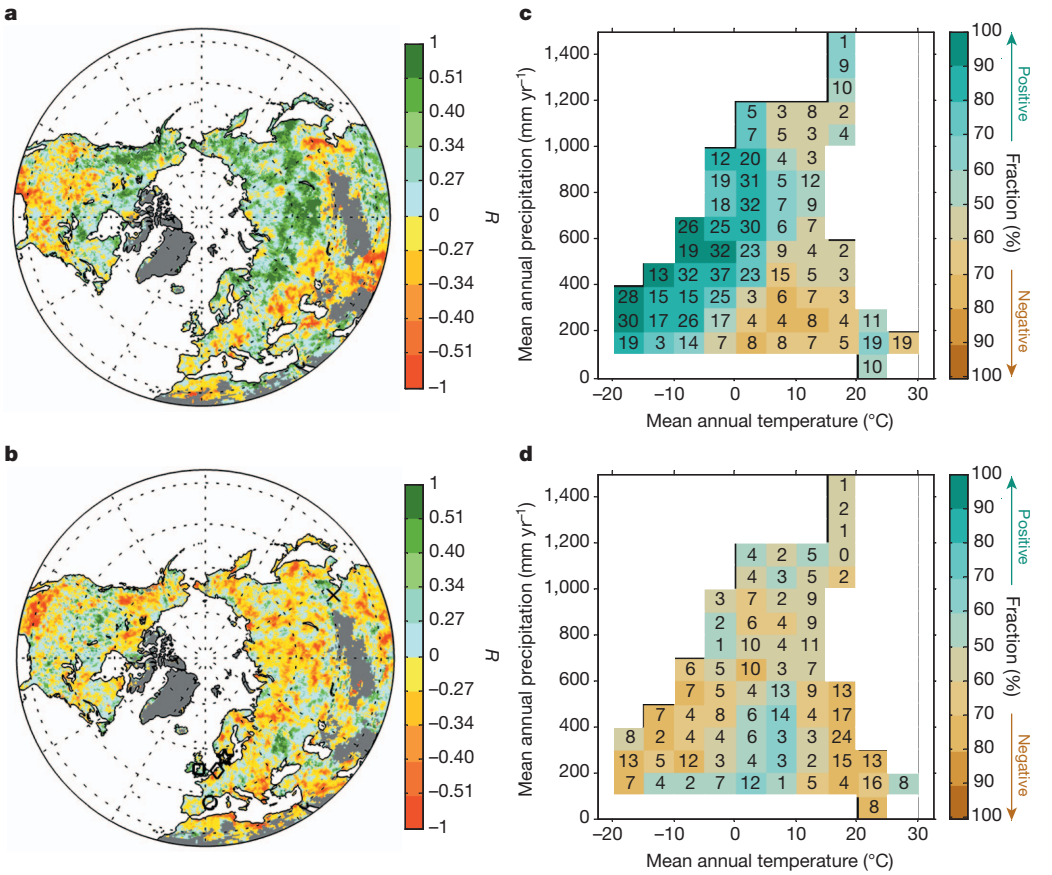


Figure 1 | The response of growing-season (from April to October) NDVI to changes in growing-season maximum temperature (T_{max}) and minimum temperature (T_{min}) in the Northern Hemisphere. **a**, Spatial distribution of the partial correlation coefficient R between growing-season NDVI and T_{max} after controlling for T_{min}, precipitation and solar radiation. **b**, Spatial distribution of R between growing-season NDVI and T_{min} after controlling for T_{max}, precipitation and solar radiation. The labels on the colour bars in **a** and **b**, $R = \pm 0.51$, $R = \pm 0.40$, $R = \pm 0.34$ and $R = \pm 0.27$ correspond to the 1%, 5%, 10% and 20% significance levels, respectively. Among the five night-time warming experiments sites marked in **b**, the results of four sites—Mols in Denmark (pentagram), Clocaenog in United Kindom (square), Garraf in Spain (circle) and Duolun in China (cross)—are consistent with the result of the partial correlation analysis between NDVI and T_{min} (Supplementary Table 2; the site of Oldebroek in the Netherlands was marked by a diamond in **b**). **c**, Percentage of pixels with dominant (positive pixels >50% or negative pixels >50%) partial correlation between NDVI and T_{max} (shown in **a**) in each 5 °C interval of mean annual temperature and 100-mm interval of mean annual precipitation climate space. **d**, Same as **c** but for the partial correlation between NDVI and T_{min} shown in **b**. The right colour bars in **c** and **d** indicate the fraction of pixels with dominant partial correlation between growing-season NDVI and T_{max}/T_{min}, and the numbers in each interval climate space of **c** and **d** indicate the percentage of pixels with significant partial correlation (for example, if the percentage of pixels with positive partial correlation exceeds 50%, the space is blue and the number is the percentage of pixels with positive and significant partial correlation, but if the percentage of pixels with positive partial correlation is less than 50%, the space is brown and the number is the percentage of pixels with negative and significant partial correlation).

NDVI to T_{min} between the wet and dry regions of the Northern Hemisphere is supported by some previous studies. For instance, an increase in T_{min} was found to reduce rice yields by 10% per °C in the Philippines[10], but to enhance growth at a temperate dry grassland site in China[8].

To test the robustness of our analysis, we also applied an independent statistical test called the ridge regression method[17]. The results show spatial patterns of the NDVI response to T_{max} and T_{min} similar to those obtained with a multiple-linear method (Supplementary Fig. 2). Nevertheless, the absolute magnitude of the sensitivities of NDVI to T_{max} and T_{min} is smaller in the ridge regression results. Patterns for the correlation between NDVI and T_{min} or T_{max} similar to those displayed in Fig. 1 were also obtained when using different growing-season definitions (May–October and May–September; Supplementary Figs 3 and 4), and using other gridded data sets of solar radiation, precipitation (or vapour pressure deficit as the moisture surrogate), instead of the climate data set from the Climate Research Unit (CRU) of the University of East Anglia, UK (Supplementary Figs 5–10). Finally, to test for possible spurious spatial correlations patterns that could be due to the spatial interpolation of meteorological station data in the CRU gridded data set, we performed the same correlation analyses using observed records from 1,736 individual meteorological stations available over the last three decades (Methods) and found the same pattern as in Fig. 1 (see Supplementary Figs 11 and 12).

Covariance between climate variables, regressed with NDVI, especially between T_{max} and T_{min} (Supplementary Fig. 1), could affect the interpretation of our partial correlation analyses. We therefore performed reduced partial correlation analyses separately using either T_{max}, precipitation and solar radiation or T_{min}, precipitation and solar radiation. The results show that the reduced partial correlation coefficients between NDVI and T_{min} (statistically controlling precipitation and solar radiation, but not T_{max}) have the same sign as between NDVI and T_{max} (statistically controlling precipitation and solar radiation, but not T_{min}), but with different strengths in most areas (Supplementary Fig. 13). In boreal ecosystems, the NDVI–T_{max} correlation is higher than the NDVI–T_{min} correlation, but in the dry temperate ecosystems it is lower. This result again confirms a differential response to warming; for example, boreal ecosystems respond more to increases in T_{max} than to increases in T_{min}.

Additionally, to minimize the effect of collinearity between T_{max} and T_{min} in a multiple linear regression statistical model, we also replaced them with their average ($T_{mean} = (T_{max} + T_{min})/2$) and difference

(DTR = $T_{max} - T_{min}$) and performed the same partial correlation analysis. The results show that boreal NDVI positively correlates with T_{mean} and DTR; whereas dry temperate NDVI exhibits negative relationships with T_{mean} and DTR (Supplementary Fig. 14). This is consistent with the results of partial correlation analysis with T_{max}, T_{min}, precipitation and solar radiation (Fig. 1). A statistical blind test also shows that the regression or partial correlation method used in our study correctly captures a relationship between NDVI and T_{max} or T_{min} if there is one (see Supplementary Information). Splitting the temperature signal into T_{max} and T_{min} thus helps the process interpretation of these results.

Second, we analysed how variations in T_{max} and T_{min} may influence interannual anomalies in the seasonal cycle of net CO_2 fluxes of northern ecosystems using the atmospheric CO_2 record of the stations of Point Barrow in Alaska (71° N) and Mauna Loa in Hawaii (19° N) (Methods). We found that the interannual variation in the amplitude of the detrended CO_2 seasonal cycle (AMP) at Point Barrow does not significantly correlate with the T_{mean} of May–September averaged over a broad boreal latitudinal zonal band (51°–90° N) ($R = -0.03$, $P = 0.897$); all variables were detrended and effects of precipitation and solar radiation were removed in the partial correlation analysis. This weak correlation of AMP with T_{mean}, however, masks its significant but opposite correlations with T_{max} and T_{min}. Indeed, multiple linear regression analyses, using AMP at Point Barrow as the dependent variable and detrended precipitation, solar radiation, T_{max} or T_{min} as independent variables, show that AMP responds positively to positive T_{max} anomalies by 3.8 ± 1.9 p.p.m. per °C (23 ± 11% of peak-to-peak AMP per °C of T_{max}) ($R = 0.38$, $P = 0.048$), but negatively to positive T_{min} anomalies by −4.8 ± 2.4 p.p.m. per °C (−28 ± 14% of peak-to-peak AMP per °C of T_{min} ($R = -0.38$, $P = 0.047$); Fig. 2 and Supplementary Table 3). At Mauna Loa, no significant partial correlation is obtained between interannual AMP variations and either the corresponding T_{max} ($R = -0.02$, $P = 0.866$) or the corresponding T_{min} ($R = 0.02$, $P = 0.893$) variations (Fig. 2), possibly owing to the mixed temperature effects across different temperate regions (Fig. 1).

Last, we investigated the covariance between the interannual anomalies of NCE (positive values indicating net CO_2 uptake by ecosystems) and those of T_{max} and T_{min} using NCE gridded estimates from a global atmospheric inversion over the last three decades[18]. For the boreal region (>50° N), the inversion-based NCE from May to September is found to respond positively to positive anomalies of T_{max}, with a sensitivity of 1.1 ± 0.6 Pg of C per °C ($R = 0.37$, $P = 0.059$), and negatively to positive anomalies of T_{min}, with a sensitivity of −1.7 ± 0.7 Pg of C per °C ($R = -0.45$, $P = 0.021$; Fig. 3 and Supplementary Table 4) (1 Pg = 10^{15} g). These relationships are qualitatively similar to the ones found between AMP versus T_{max} and T_{min} in the boreal zone. In the temperate latitude band (25°–50° N), the inversion-based NCE from May to October is not significantly correlated with either T_{max} ($R = 0.11$, $P = 0.604$) or T_{min} ($R = -0.24$, $P = 0.238$) (Fig. 3). Similar patterns were also obtained when using the inversion with flat prior information (Supplementary Fig. 15). Note that inversion NCE uncertainties are large at the grid-point scale, and that individual grid point NCE estimates are correlated with each other. The spatial distribution of the response of inversion NCE to T_{max} and T_{min} over the climatic dimensions (Supplementary Figs 16 and 17) is comparable with that derived from the (independent) NDVI (Fig. 1c and d). Interestingly, the geographical area where the growing-season NCE positively correlates with T_{max} (Supplementary Figs 16 and 17) is less extensive than the area where growing-season NDVI positively correlates with T_{max} (Fig. 1). This could be explained by a concurrent increase of soil organic matter decomposition induced by warming in boreal and temperate regions where water is readily available.

In summary, our analyses of several data streams consistently point out that interannual changes in growing-season T_{max} and T_{min} have regionally opposite effects on NDVI (greenness) and NCE. Daytime positive T_{max} anomalies correlate positively with enhanced photosynthetic activity (NDVI) and NCE over wet and cool ecosystems, but have negative effects in dry temperate regions. These regionally contrasting effects of T_{max} interannual variations on interannual vegetation greenness and growing-season CO_2 uptake are most probably associated with different ecophysiological responses. In boreal regions, photosynthetic activity and NCE are subject to temperature limitation but less subject to water limitation[6,15,16]. Rising T_{max} has been observed to enhance photosynthetic enzyme activity[19], to increase soil nitrogen mineralization and availability[20], and to extend the growing season[21]. In contrast, in dry temperate regions where soil water limitation can limit vegetation growth, daytime warming can reduce photosynthetic activity through enhanced evaporation, and reduced soil water content[8]. The hypothesis that a drop in topsoil water content (SWC) below a critical stress level in dry temperate regions is associated with warmer T_{max} during the growing season is consistent with the negative partial correlation coefficients between growing-season T_{max} and SWC retrieved from multiple microwave satellite sounders over 1988–2007 (ref. 22) (Fig. 4a, Supplementary Figs 18–21).

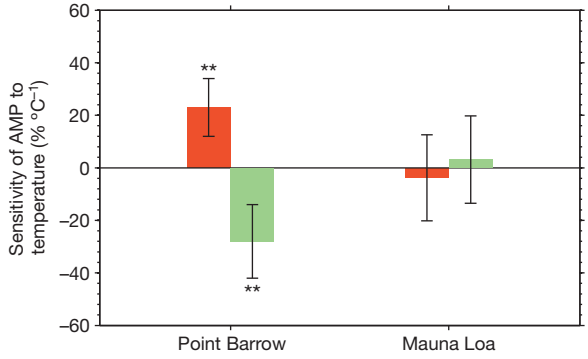


Figure 2 | T_{max} and T_{min} sensitivity of annual AMP at Point Barrow and Mauna Loa stations. The T_{max} (red) and T_{min} (green) sensitivity of AMP at each station was estimated based on a multiple linear regression analysis using detrended AMP as the dependent variable and the detrended precipitation, solar radiation, T_{max} and T_{min} during the corresponding season (May–September for Point Barrow station and May–October for Mauna Loa station) over a broad region surrounding each station by ±20 degrees of latitude (51°–90° N for Point Barrow and 1° S–39° N for Mauna Loa) as the independent variables. **Statistically significant at the 95% ($P < 0.05$) level. The error bars are standard errors of the mean of the corresponding parameters.

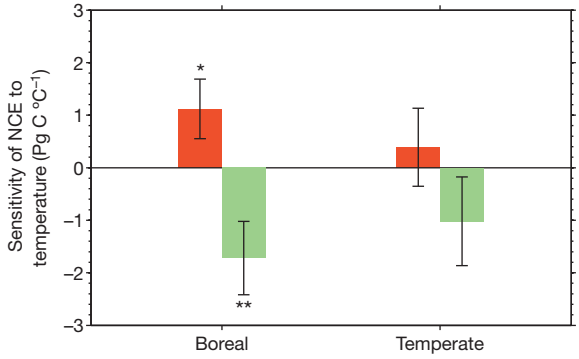


Figure 3 | T_{max} and T_{min} sensitivity of a global atmospheric inversion model estimated NCE in boreal and temperate regions. Using an approach similar to that of Fig. 2, we calculated the sensitivity of corresponding season NCE in boreal (50°–90° N) and temperate (25° –50° N) regions to changes in T_{max} and T_{min} (May–September for boreal regions and May–October for temperate regions). The T_{max} (red) and T_{min} (green) sensitivity of NCE was calculated based on a multiple linear regression analysis using detrended NCE as the dependent variable and the detrended corresponding precipitation, solar radiation, T_{max} and T_{min} as the independent variables. *Statistically significant at the 90% ($P < 0.1$) level; **Statistically significant at the 95% ($P < 0.05$) level. The error bars are standard errors of the mean of the corresponding parameters.

Night-time warming can influence vegetation productivity in two opposite ways: via enhanced autotrophic respiration, and indirectly via stimulation of plant photosynthesis during the following daytime through decreasing frost risk[13,23], and physiological regulatory mechanisms[8,11,19,24]. Leaf carbohydrates synthesized during the daytime were observed to be consumed more quickly during warmer nights because of enhanced leaf respiration[19,24], which depletes foliar carbohydrates and may produce a rebound effect of compensatory stimulated photosynthesis during the following day[8,19]. This mechanism could partly explain the positive partial correlation between NDVI and T_{min} in temperate grassland regions[8]. In contrast, the observed decrease in vegetation activity associated with warmer T_{min} in boreal and in wet temperate regions implies that the negative impact of increased autotrophic respiration is more dominant in these regions, but further studies are needed to support these inferences. In addition, changes in T_{min} could influence vegetation productivity through shifting competitive interactions among C_3 and C_4 plants in semi-arid biomes with warm summer conditions[9].

The asymmetrical response of terrestrial ecosystems to daytime versus night-time temperature anomalies found in this study suggests that most of the currently used global carbon cycle models using daily temperature forcing may neglect an essential process. If the night-time warming trend observed over the last five decades continues to be faster than that of daytime warming[1], the model using average daily temperature may overestimate the increase of boreal vegetation productivity by the end of this century[25]. Arguably, many terrestrial ecosystem models[26,27] using daily or monthly temperature data as input will not capture the response of vegetation to asymmetric diurnal temperature changes, indicating the importance of understanding the diurnal cycle of CO_2 and energy fluxes in land biosphere models. This understanding could be achieved by pattern analysis of existing eddy-covariance flux measurements[28], targeted ecosystem manipulative experiments under controlled and different T_{min} and T_{max} conditions, as well as by mesocosm studies.

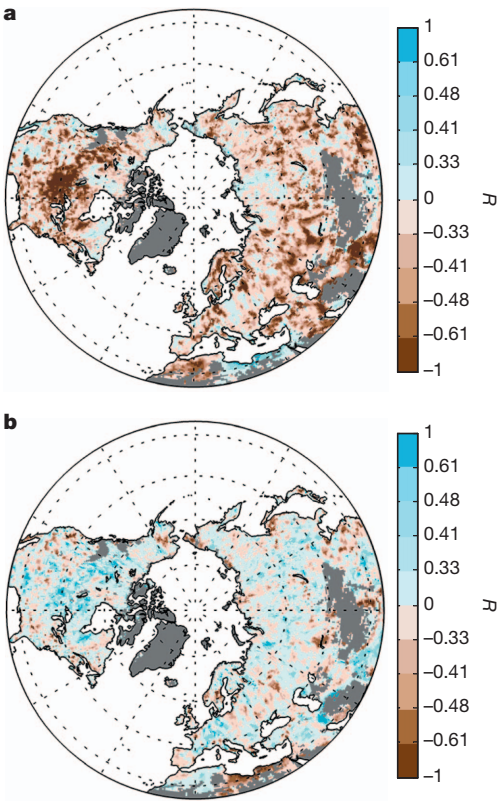


Figure 4 | The response of growing-season (April–October) SWC to changes in growing-season T_{max} and T_{min} in the Northern Hemisphere. **a**, Spatial distribution of the partial correlation coefficient R between growing-season SWC and T_{max}. **b**, Spatial distribution of the partial correlation coefficient R between growing-season SWC and T_{min}. R between growing-season SWC and T_{max} or T_{min} was calculated using the same approach for estimating R between growing-season NDVI and T_{max} or T_{min} in Fig. 1. The labels on the colour bars in **a** and **b**, $R = \pm 0.61$, $R = \pm 0.48$, $R = \pm 0.41$ and $R = \pm 0.33$ correspond to the 1%, 5%, 10% and 20% significance levels, respectively. Compared to the daytime warming effects, night-time warming exhibits a smaller negative correlation with SWC. The areas with positive correlation between SWC and T_{min} is 59% (significant over 6% of the areas). The observed positive correlation between T_{min} and SWC, particularly over the boreal and wet temperate regions, could be attributed in part to a decrease in evapotranspiration driven by a decrease in vegetation photosynthetic activity (or leaf area index), because higher T_{min} is generally accompanied by lower growing-season NDVI over these regions (Fig. 1).

METHODS SUMMARY

Satellite derived NDVI data were used to investigate the effects of daytime and night-time warming on the vegetation activity of Northern Hemispheric ecosystems. Then NCE values simulated from an inversion model, as well as from atmospheric CO_2 concentrations at Point Barrow and Mauna Loa stations, were used to investigate the response of the carbon balance to changes in T_{max} and T_{min}. Finally, the response of satellite-derived surface SWC to T_{max} and T_{min} were analysed to understand the mechanisms of asymmetric effects of daytime and night-time warming on Northern Hemisphere vegetation activity and carbon cycle. We applied a partial correlation analysis to statistically remove the effects of other factors, such as precipitation and solar radiation (all variables detrended); and the results were verified with those from another independent statistical method, ridge regression.

Full Methods and any associated references are available in the online version of the paper.

Received 15 October 2012; accepted 4 July 2013.

1. Solomon, S., *et al.* (eds) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2007).
2. Atkin, O. *et al.* Light inhibition of leaf respiration as soil fertility declines along a post-glacial chronosequence in New Zealand: an analysis using the Kok method. *Plant Soil* **367,** 163–182 (2013).
3. Keeling, R. F., Piper, S. C. & Heimann, M. Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration. *Nature* **381,** 218–221 (1996).
4. Myneni, R. B., Keeling, C. D., Tucker, C. J., Asrar, G. & Nemani, R. R. Increased plant growth in the northern high latitudes from 1981 to 1991. *Nature* **386,** 698–702 (1997).
5. Zhou, L. M. *et al.* Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *J. Geophys. Res.* **106,** 20069–20083 (2001).
6. Nemani, R. R. *et al.* Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science* **300,** 1560–1563 (2003).
7. Beier, C. *et al.* Carbon and nitrogen cycles in European ecosystems respond differently to global warming. *Sci. Total Environ.* **407,** 692–697 (2008).
8. Wan, S., Xia, J., Liu, W. & Niu, S. Photosynthetic overcompensation under nocturnal warming enhances grassland carbon sequestration. *Ecology* **90,** 2700–2710 (2009).
9. Alward, R. D., Detling, J. K. & Milchunas, D. G. Grassland vegetation changes and nocturnal global warming. *Science* **283,** 229–231 (1999).
10. Peng, S. *et al.* Rice yields decline with higher night temperature from global warming. *Proc. Natl Acad. Sci. USA* **101,** 9971–9975 (2004).
11. Prasad, P. V. V., Pisipati, S. R., Ristic, Z., Bukovnik, U. & Fritz, A. K. Impact of night-time temperature on physiology and growth of spring wheat. *Crop Sci.* **48,** 2372–2380 (2008).
12. Zhou, L., Kaufmann, R. K., Tian, Y., Myneni, R. B. & Tucker, C. J. Relation between interannual variations in satellite measures of northern forest greenness and climate between 1982 and 1999. *J. Geophys. Res.* **108,** 4004, doi:10.1029/2002JD002510 (2003).
13. Kim, Y., Kimball, J. S., Zhang, K. & McDonald, K. C. Satellite detection of increasing Northern Hemisphere non-frozen seasons from 1979 to 2008: implications for regional vegetation growth. *Remote Sens. Environ.* **121,** 472–487 (2012).
14. Kreyling, J. Winter climate change: a critical factor for temperate vegetation performance. *Ecology* **91,** 1939–1948 (2010).
15. Beer, C. *et al.* Terrestrial gross carbon dioxide uptake: global distribution and covariation with climate. *Science* **329,** 834–838 (2010).
16. Lucht, W. *et al.* Climatic control of the high-latitude vegetation greening trend and Pinatubo effect. *Science* **296,** 1687–1689 (2002).
17. Hoerl, A. E. & Kennard, R. W. Ridge regression — biased estimation for nonorthogonal problems. *Technometrics* **12,** 55–67 (1970).

18. Chevallier, F. *et al.* CO_2 surface fluxes at grid point scale estimated from a global 21 year reanalysis of atmospheric measurements. *J. Geophys. Res.* **115,** D21307 (2010).
19. Turnbull, M. H., Murthy, R. & Griffin, K. L. The relative impacts of day-time and night-time warming on photosynthetic capacity in *Populus deltoides*. *Plant Cell Environ.* **25,** 1729–1737 (2002).
20. Melillo, J. M. *et al.* Soil warming and carbon-cycle feedbacks to the climate system. *Science* **298,** 2173–2176 (2002).
21. Menzel, A. *et al.* European phenological response to climate change matches the warming pattern. *Glob. Change Biol.* **12,** 1969–1976 (2006).
22. Owe, M., de Jeu, R. & Holmes, T. Multisensor historical climatology of satellite-derived global land surface moisture. *J. Geophys. Res.* **113,** F01002, doi: 10.1029/2007JF000769 (2008).
23. Gu, L. *et al.* The 2007 eastern US spring freezes: increased cold damage in a warming world? *Bioscience* **58,** 253–262 (2008).
24. Griffin, K. L. *et al.* Leaf respiration is differentially affected by leaf vs. stand-level night-time warming. *Glob. Change Biol.* **8,** 479–485 (2002).
25. Qian, H., Joseph, R. & Zeng, N. Enhanced terrestrial carbon uptake in the Northern High Latitudes in the 21st century from the Coupled Carbon Cycle Climate Model Intercomparison Project model projections. *Glob. Change Biol.* **16,** 641–656 (2010).
26. Potter, C. S. *et al.* Terrestrial ecosystem production—a process model based on global satellite and surface data. *Glob. Biogeochem. Cycles* **7,** 811–841 (1993).
27. Sitch, S. *et al.* Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Glob. Change Biol.* **9,** 161–185 (2003).
28. Baldocchi, D. 'Breathing' of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Aust. J. Bot.* **56,** 1–26 (2008).

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

Acknowledgements This study was supported by the National Natural Science Foundation of China (grant numbers 41125004 and 31021001), the National Basic Research Program of China (grant numbers 2010CB950601 and 2013CB956303), the Foundation for Sino-EU Research Cooperation of the Ministry of Science and Technology of China (grant number 1003), and a Chinese Ministry of Environmental Protection Grant (number 201209031). We also acknowledge the GLOBALVIEW-CO_2 project based at NOAA ESRL. S.V. is a postdoctoral research associate of the Fund for Scientific Research (Flanders).

Author Contributions S. Piao, S. Peng and H.Z. designed the research. S. Peng performed analysis and calculations. S. Piao, P.C., A.C. and S. Peng drafted the paper. R.B.M. provided the remotely sensed NDVI data and contributed to the text. F.C. provided the atmospheric inverse model estimated carbon flux and contributed to the text. A.J.D. provided the remotely sensed soil moisture data and contributed to the text. I.A.J., J.P., G.Z., S.V., S. Wan, S. Wang and H.Z. contributed to the interpretation of the results and to the text.

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. Piao (slpiao@pku.edu.cn) or H.Z (zengh@pkusz.edu.cn).

METHODS

Satellite NDVI measurements. NDVI is defined as the ratio of the difference between near-infrared reflectance and red visible reflectance to their sum, and is a remote-sensed vegetation index widely used as an indicator of vegetation gross primary productivity[4,5]. The NDVI third-generation (NDVI3g) data used in this study were from the Global Inventory Monitoring and Modelling Studies (GIMMS) group derived from the NOAA/AVHRR land data set, at a spatial resolution of $8 \times 8\,km^2$ and 15-day interval, for the period January 1982 to December 2009 (ref. 29). The GIMMS-NDVI data set has been widely used for detecting vegetation growth change[5,29].

Atmospheric CO_2 measurements at two long-term monitoring stations. Weekly and monthly averaged atmospheric CO_2 concentrations at Point Barrow and Mauna Loa stations, based on continuous *in situ* observations, were obtained from the Earth System Research Laboratory of the National Oceanic and Atmospheric Administration (NOAA) (http://www.esrl.noaa.gov/gmd/ccgg/globalview). Atmospheric CO_2 concentration *in situ* station samples at Point Barrow and Mauna Loa cover the period 1979–2009 and 1959–2009, respectively.

NCE distribution from an atmospheric inversion. We use a global Bayesian inversion of NCE (a positive value indicates carbon uptake) over the past three decades[18]. Weekly CO_2 fluxes are estimated on each grid-point of a $3.75^\circ \times 2.5^\circ$ (longitude–latitude) global grid throughout the 29 years[18]. The inversion of NCE accounts for a priori time and space error correlations of NCE fluxes defined over land from the misfit between eddy-covariance observations and model simulations[18], and based on more than 128 individual atmospheric CO_2 measurements from flask and continuous-atmospheric-measurement sites (list in ref. 18). This inversion has been chosen because (1) NCE fluxes are estimated at relatively high spatial resolution for a global inversion, thus reducing the risk of spatial aggregation errors, (2) it is informed by high-temporal-resolution CO_2 measurements instead of monthly smoothed data, and (3) it covers a longer period than the TRANSCOM inversions[30], with 1981 as the beginning of the assimilation period. In addition, an inversion of NCE with a climatology of model simulations of NCE (without any interannual variations) as prior information over land was also used (Supplementary Fig. 15).

Climate data. The monthly T_{min}, T_{max} and precipitation data sets with spatial resolution of 0.5° used in this study are the CRU TS 3.1 climate data sets[31], covering NDVI, atmospheric CO_2 concentration and inversed CO_2 surface flux time series period (1959–2009). The CRU TS3.1 climate data sets are interpolated from meteorological stations based on spatial autocorrelation functions[31,32]. Because of the high spatial correlation in the interpolated CRU data sets[32], we also tested the robustness of our results with observed climate records from 1,736 individual meteorological stations (ftp://ftp.ncdc.noaa.gov/pub/data/gsod/).

Solar radiation was obtained from the CRU–NCEP data set (the period 1901–2009), which is based on the combination of the CRU TS 3.1 climate data set covering the period 1901–2009 and the NCEP reanalysis covering the period 1948–2009 (http://nacp.ornl.gov/thredds/fileServer/reccapDriver/cru_ncep/analysis/readme.htm). To assess whether the relationships between vegetation production and T_{max} and T_{min} are robust across different data sets, we also used two satellite-based solar radiation data sets—ISCCP[33] and the NASA/GEWEX Surface Radiation Budget, http://gewex-srb.larc.nasa.gov/common/php/SRB_data_products.php—covering the period 1984–2007, three precipitation data sets—GPCP[34], GPCC (http://gpcc.dwd.de) and precipitation from ref. 35—and the vapour pressure deficit calculated from CRU TS 3.1 (ref. 36).

Soil water content data. Daily SWC data ($m^3\,m^{-3}$) with spatial resolution of 0.25° over the period 1988–2007 was derived from the Special Sensor Microwave/Imager satellite using the Land Parameter Retrieval Model, which solves simultaneously for the SWC and vegetation optical depth[22]. This data set has been validated extensively over a large variety of land surfaces of sparse to moderate vegetation, showing good agreement with *in situ* observations[37].

Analyses. Several of the environmental variables tested as predictors of NDVI or NCE covary with one another. This covariance needs to be carefully dealt with when we want to detect the relationship between the dependent variable and one particular independent variable. In evaluating the relationship between NDVI or NCE and T_{max} or T_{min}, we used partial correlation analyses to exclude the confounding effects of other variables. Partial correlation analysis is a widely applied statistical tool to isolate the relationship between two variables from the confounding effects of many correlated variables[15,38,39].

To verify the asymmetric diurnal warming effect on vegetation growth derived from partial correlation analysis, we also performed a ridge regression analysis[17]. Ridge regression is a linear regularization method and constrained linear inversion method, which reduces the size of the regression coefficients by introducing a constant in the minimized residual equation[17]. The analysis was performed in the R package lm.ridge with an HKB estimate of the ridge constant (ref. 40, http://astrostatistics.psu.edu/su07/R/html/MASS/html/lm.ridge.html). We also performed the partial correlation analysis by replacing T_{max} and T_{min} with their average ($T_{mean} = (T_{max} + T_{min})/2$) and difference ($DTR = T_{max} - T_{min}$). Additionally, we conducted reduced partial correlation analyses in which the dependent variables were correlated twice: once against T_{max}, precipitation and solar radiation, and once against T_{min}, precipitation and solar radiation.

We used averages of monthly NDVI data during the growing season to study the response of vegetation growth to temperature. To match NDVI data (8 km spatial resolution) with climate data (0.5° spatial resolution), we averaged monthly NDVI data corresponding to each climate data set pixel. Monthly NDVI was obtained by choosing the monthly maximum value, which can further eliminate the disturbance from cloud, atmosphere and changes in solar altitude angle[4,5]. Following a previous study[5], growing season is defined as the period from April to October. We also used different growing-season definitions (from May to October and from May to September) in the Supplementary Information. The sensitivity of growing-season NDVI to changes in T_{max} and T_{min} was estimated based on multiple linear regression analysis using growing-season NDVI as the dependent variable and the corresponding precipitation, solar radiation, T_{max}, and T_{min} as independent variables (all variables detrended). Similarly, we also calculated the sensitivity of growing-season NCE and SWC to changes in T_{max} and T_{min}. After aggregation of daily SWC data into monthly means, the sensitivity of growing-season SWC to changes in T_{max} and T_{min} was calculated only in those pixels where SWC was recorded at least three months during the growing season. The average number of days with available daily SWC data per month during the growing season for each pixel was shown in Supplementary Fig. 22.

We used the curve-fitting procedures (CCGVU program)[41] to extract the detrended seasonal cycle from the monthly atmospheric CO_2 concentration record at Point Barrow and Mauna Loa stations. The annual maximum concentrations in the mean seasonal cycle were recorded at May for both Point Barrow and Mauna Loa, whereas the annual minimum concentrations in the mean seasonal cycle were recorded at August/September and September/October for Point Barrow and Mauna Loa, respectively. AMP is calculated as the difference between the annual maximum and minimum concentrations for each station. The temperature sensitivity of AMP at each station was estimated based on multiple linear regression analysis using detrended AMP as the dependent variable and the detrended precipitation, solar radiation, T_{max} and T_{min} during the corresponding season (from May to September for Point Barrow station and from May to October for Mauna Loa station) over a broad region surrounding each station by ± 20 degrees of latitude as independent variables. The regional climate variables were weighted by mean annual MODIS net primary production over 2000–2009 (ref. 42) to focus on the vegetated areas. Using a similar approach, we also calculated the sensitivity of corresponding season NCE in boreal (50°–90° N) and temperate (25°–50° N) regions to changes in T_{max} and T_{min} (May–September for boreal regions and May–October for temperate regions). The uncertainties of the sensitivity were estimated based on the standard errors of the corresponding parameters.

29. Tucker, C. J. *et al.* An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *Int. J. Remote Sens.* **26,** 4485–4498 (2005).
30. Baker, D. F. *et al.* TransCom 3 inversion intercomparison: impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988-2003. *Glob. Biogeochem. Cycles* **20,** GB1002 (2006).
31. Mitchell, T. D. & Jones, P. D. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.* **25,** 693–712 (2005).
32. New, M., Hulme, M. & Jones, P. Representing twentieth-century space–time climate variability. Part II: Development of 1901–96 monthly grids of terrestrial surface climate. *J. Clim.* **13,** 2217–2238 (2000).
33. Zhang, Y., Rossow, W. B. & Stackhouse, P. W. Comparison of different global information sources used in surface radiative flux calculation: radiative properties of the near-surface atmosphere. *J. Geophys. Res.* **111,** D13106 (2006).
34. Adler, R. F. *et al.* The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present). *J. Hydrometeorol.* **4,** 1147–1167 (2003).
35. Willmott, K. & Matsuura, C. J. *Terrestrial Precipitation: 1900–2008 Gridded Monthly Time Series* http://climate.geog.udel.edu/~climate/html_pages/Global2_Ts_2009/README.global_p_ts_2009.html (accessed 20 January 2013).
36. Prenger, J. J. & Ling, P. P. *Greenhouse Condensation Control: Understanding and Using Vapor Pressure Deficit (VPD)* http://ohioline.osu.edu/aex-fact/0804.html (accessed 20 January 2013).
37. de Jeu, R. A. M. *et al.* Global soil moisture patterns observed by space borne microwave radiometers and scatterometers. *Surv. Geophys.* **29,** 399–420 (2008).
38. Behera, S. K. *et al.* Paramount impact of the Indian Ocean dipole on the East African short rains: A CGCM study. *J. Clim.* **18,** 4514–4530 (2005).
39. Wang, H., Wang, B., Huang, F., Ding, Q. G. & Lee, J. Y. Interdecadal change of the boreal summer circumglobal teleconnection (1958-2010). *Geophys. Res. Lett.* **39,** L12704, doi:10.1029/2012gl052371 (2012).
40. Hoerl, A. E., Kennard, R. W. & Baldwin, K. F. Ridge regression—some simulations. *Commun. Stat. Theor. Med.* **4,** 105–123 (1975).

41. Thoning, K. W., Tans, P. P. & Komhyr, W. D. Atmospheric carbon dioxide at Mauna Loa observatory 2. analysis of the NOAA GMCC data, 1974-1985. *J. Geophys. Res.* **94,** 8549–8565 (1989).

42. Zhao, M. S. & Running, S. W. Drought-induced reduction in global terrestrial net primary production from 2000 through 2009. *Science* **329,** 940–943 (2010).

LETTER

doi:10.1038/nature12424

Evolutionary origins of the avian brain

Amy M. Balanoff[1,2]†, Gabe S. Bever[1,3], Timothy B. Rowe[4] & Mark A. Norell[1]

Features that were once considered exclusive to modern birds, such as feathers and a furcula, are now known to have first appeared in non-avian dinosaurs[1]. However, relatively little is known of the early evolutionary history of the hyperinflated brain that distinguishes birds from other living reptiles and provides the important neurological capablities required by flight[2]. Here we use high-resolution computed tomography to estimate and compare cranial volumes of extant birds, the early avialan *Archaeopteryx lithographica*, and a number of non-avian maniraptoran dinosaurs that are phylogenetically close to the origins of both Avialae and avian flight. Previous work established that avian cerebral expansion began early in theropod history and that the cranial cavity of *Archaeopteryx* was volumetrically intermediate between these early forms and modern birds[3,4]. Our new data indicate that the relative size of the cranial cavity of *Archaeopteryx* is reflective of a more generalized maniraptoran volumetric signature and in several instances is actually smaller than that of other non-avian dinosaurs. Thus, bird-like encephalization indices evolved multiple times, supporting the conclusion that if *Archaeopteryx* had the neurological capabilities required of flight, so did at least some other non-avian maniraptorans. This is congruent with recent findings that avialans were not unique among maniraptorans in their ability to fly in some form[5,6].

Birds are distinct among living reptiles in the degree to which their brains, particularly their forebrains, are expanded relative to body size.

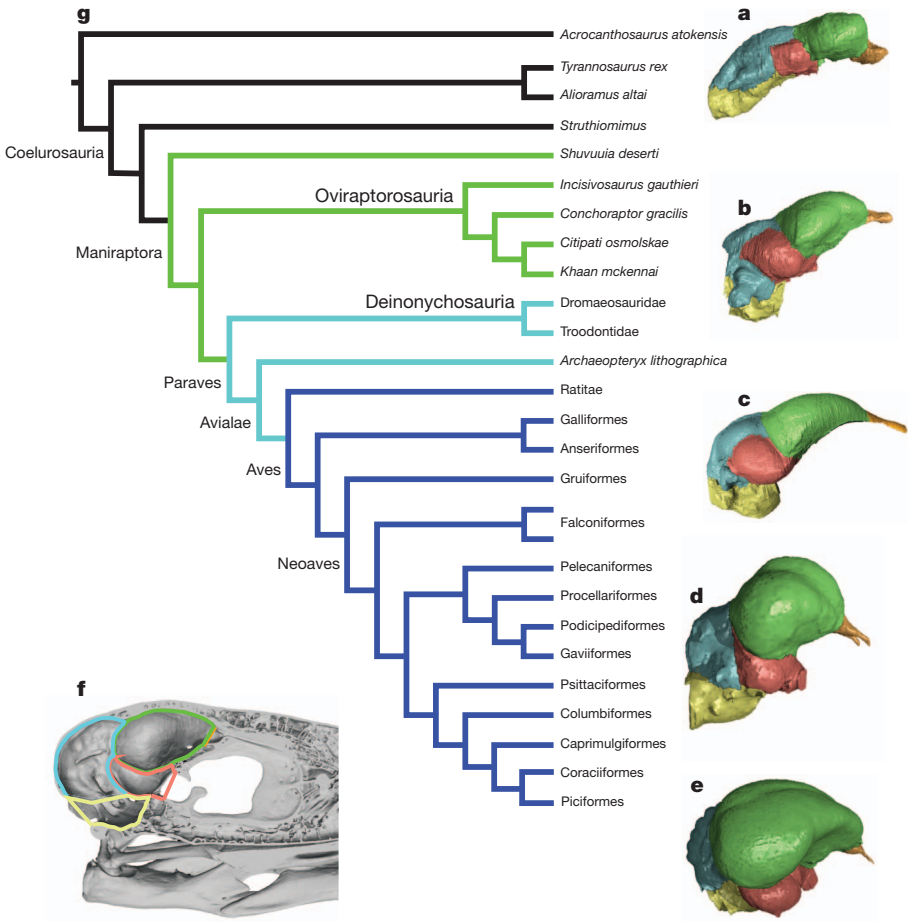


Figure 1 | Coelurosaur phylogeny and partitioned endocranial casts. **a**–**e**, Endocasts of *Citipati osmolskae* (IGM 100/978) (**a**), unnamed troodontid (IGM 100/1126) (**b**), *Archaeopteryx lithographica* (BMNH 37001) (**c**), *Struthio camelus* (ostrich) (**d**), and *Melanerpes aurifrons* (woodpecker) (**e**) divided into neuroanatomical partitions based on homologous osteological landmarks using computed tomography data. Partitions roughly correlate to the olfactory bulbs (orange), cerebrum (green), optic lobes (pink), cerebellum (blue) and brain stem (yellow). Endocasts are not scaled to size. **f**, Sagittally sectioned skull of *Phaethon rubricauda* with osteological landmarks highlighted to correspond to the regions shown in the endocasts. **g**, Phylogeny of included taxa. Proposed episodes of encephalization are indicated by changes in colour. Phylogeny adapted from ref. 30.

[1]American Museum of Natural History, Division of Paleontology, New York, New York 10024, USA. [2]Columbia University, Department of Earth and Environmental Sciences, New York, New York 10027, USA. [3]New York Institute of Technology, College of Osteopathic Medicine, Old Westbury, New York 11568, USA. [4]Jackson School of Geosciences, The University of Texas at Austin, Austin, Texas 78712, USA. †Present address: Department of Anatomical Sciences, Stony Brook University School of Medicine, Stony Brook, New York 11794, USA.

This index of encephalization ranges from six to eleven times higher in birds than other groups[7–9], and comparably large indices are known only among mammals[10–12]. The hyperinflated forebrains of birds and mammals evolved independently[11,13], possibly in response to different sensory cues; derived olfactory capabilities versus enhanced visual acuity[14,15]. Details of this neuroanatomical elaboration were recently explicated for the mammalian side of the tree based largely on fossil evidence from the latter portion of the stem and the early history of the crown[16]. No correspondingly comprehensive study exists for birds, despite broad interest in the relationship between brain size and structure, cognitive ability, and the origin of avian flight[2].

The volumetric expansion of the avian endocranium began relatively early in theropod evolution[4,17–19], and the early avialan *Archaeopteryx lithographica* is volumetrically intermediate between those of more basal theropods (for example, tyrannosaurs) and crown birds[3,4]. What remains unclear is whether the *Archaeopteryx* endocranium will continue to occupy a uniquely intermediate space between non-avialan theropods and crown birds once additional endocranial features, some with neurological implications for flight, are sampled from a wider range of bird-like, non-avian theropods. This is particularly relevant considering that recent studies argued, first, that avialans are not unique among maniraptorans in their ability to fly in some form[20], and second, that *Archaeopteryx* is more closely related to dromaeosaurs and troodontids than to modern birds[21].

We tested the relative position of *Archaeopteryx* in the evolution of avian endocranial space using comparative volumetric analyses. Volumes were obtained from digital endocasts constructed from computed tomography data sets for a diversity of crown and stem avians (Supplementary Table 1). Our study differs from previous efforts in that we sampled those theropod lineages most closely related to Avialae; Troodontidae, Dromaeosauridae, Oviraptorosauria and Alvarezsauridae (Fig. 1). In addition to considering the relationship between total endocranial volume and body size, we also divided the endocasts into volumetric partitions that estimate the major neuroanatomical regions, including the olfactory bulbs, cerebrum, optic lobes, cerebellum and brain stem (Fig. 1; Supplementary Table 2). This partitioning, accomplished using homologous osteological landmarks, enabled us to examine how the volumetric signature of different endocranial regions evolved in relation to total body size, total endocranial volume, and to one another. This approach allows detection of previously unrecognizable evolutionary complexity.

Our analysis of total endocranial volume relative to body size (estimated based on femur length[22]), recovered the apomorphically high volumetric signature for the avian crown (Fig. 2; Supplementary Tables 2 and 3)[3,4,10,23], but failed to recover *Archaeopteryx* in a uniquely transitional position between non-avialan maniraptorans and crown birds. Several oviraptorosaurs and the troodontids *Zanabazar junior* and IGM 100/1126, all have relative endocranial volumes that fall between the values of *Archaeopteryx* and crown birds (Fig. 2). The same basic pattern was recovered when cerebral volume was compared to body size. Thus, the total endocranial and cerebral volumes of *Archaeopteryx* relative to body size are not uniquely avian but reflect plesiomorphic values expected of a non-avian maniraptoran. Even the uniquely derived signature of crown birds is lost when volumes of other partitions are examined relative to body size (Supplementary Fig. 1). For each of these indices, there is direct overlap between the avian crown and one or more non-avian maniraptorans. The only partition in which *Archaeopteryx* plots within the range of crown birds is that of the olfactory bulbs, and it is not unique in this overlap. Optic lobes, cerebellum and brain-stem partitions of *Archaeopteryx* all fall outside the volumetric area defined by crown birds, *Zanabazar*, and at least some oviraptorosaurs.

Linear regressions of partitions against total endocranial volume reveal a different pattern (Fig. 3 and Supplementary Fig. 2). The strong correlation between cerebral and total volume is expected if cerebral expansion is the primary driver of general endocranial expansion along the avian stem[4,9]. The relatively high correlations between total volume and partitioned volumes (Fig. 3 and Supplementary Fig. 2) indicate that each region, with the exception of the olfactory bulbs, expanded along the avian stem. This suggests no notable structural constraint on total endocranial volume. If such a constraint were in place, then selection-driven cerebral expansion would be expected to occur at the volumetric expense of one or more of the other regions. The fact that these regions are expanding together attests to the high responsive potential of the surrounding skeleton to the tissues they envelop[24].

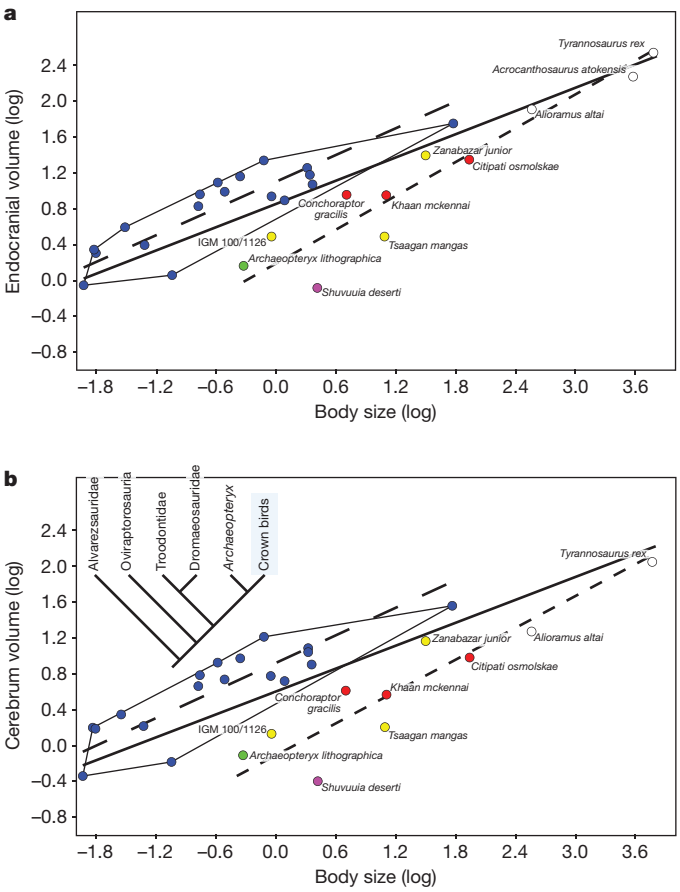


Figure 2 | Bivariate plots of log-transformed body-mass data. **a**, **b**, Body mass (kg) plotted against total endocranial volume (cm^3) (**a**) and cerebral volume (**b**). Crown birds display apomorphically high endocranial and cerebral volumes with respect to body size. Colours indicate crown birds (blue), non-maniraptoran theropods (white), *Shuvuuia deserti* (purple), oviraptorosaurs (red), deinonychosaurs (yellow), *Archaeopteryx lithographica* (green). Reduced major-axis regression line for entire sample (solid line), crown birds (large dashes), and non-avian theropods (small dashes). Regression statistics given in Supplementary Table 3.

Principal components analysis of the five partitions divided by total endocranial volume reveals complete volumetric separation between Paraves and Oviraptorosauria (Fig. 4). Principal component I accounts for approximately 82% of total sample variance (Supplementary Table 5) and is defined largely by the cerebral expansion described above. Principal component II describes approximately 11% of total sample variance and is defined largely by the inverse relationship between the relative volumes of the optic lobes and cerebellum. The relatively large cerebral volume of the troodontid, *Zanabazar*, pulls this taxon within the volumetric space defined by crown birds, so that again the crown clade is not volumetrically unique. *Archaeopteryx* plots well outside crown birds along principal component I, and even outside the alvarezsaurid *Shuvuuia deserti*, helping to define a more generalized maniraptoran space.

For volumetric indices in which *Archaeopteryx* and crown birds overlap or nearly overlap (that is, cerebellum, cerebrum and optic lobes

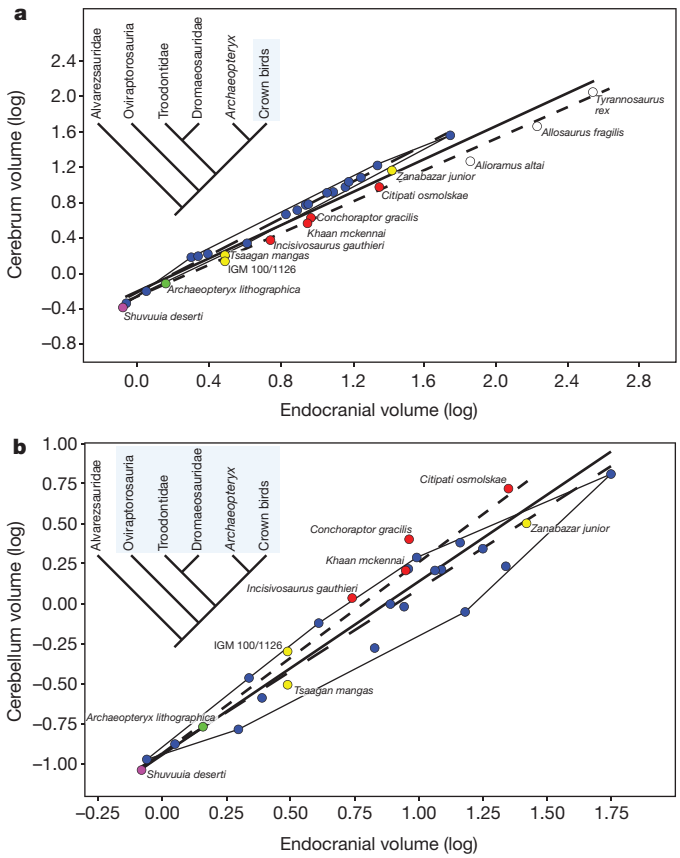


Figure 3 | Bivariate plots of log-transformed total-endocranial-volume data. **a**, **b**, Total endocranial volume (cm^3) plotted against cerebral (**a**) and cerebellar (**b**) volumes. Colours are the same as in Fig. 2. Highlighting on the tree indicates those groups that share crown-like volumes. These can be interpreted as being either homologous (synapomorphic for a more inclusive group including crown birds) or convergent (appears in crown birds and more distantly related groups). Regression statistics are given in Supplementary Table 4.

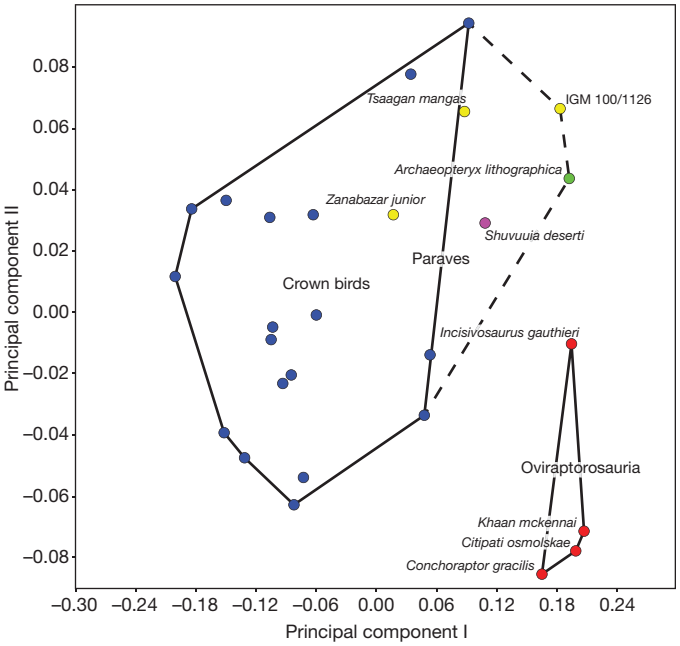


Figure 4 | Principal components analysis plot of neuroanatomical region volumes. Colours are the same as in Fig. 2. Complete volumetric isolation is indicated for Oviraptorosauria. A paravian (see Fig. 1) group is recovered in the principal components analysis (PCA) plot, but not an avialan (*Archaeopteryx* plus crown birds) volumetric grouping. PCA loadings are provided in Supplementary Table 5.

versus total endocranium; olfactory bulbs versus body size), crown-like values are also expressed in at least some non-avialan taxa. For the remaining indices, *Archaeopteryx* plots further from the crown than some other non-avialan taxa. Thus, regardless of whether *Archaeopteryx* is an early avialan or the sister to deinonychosaurs, there is no unique volumetric signature diagnosing the avialan endocranial cavity for the entirety of its evolutionary history (since Avialae diverged from Deinonychosauria).

The evolutionary history of those volumetric expansions that makes the endocranial space of modern birds so distinctive depends largely on how we interpret *Archaeopteryx*. For example, we recover at least three episodes of coelurosaur cerebral expansion relative to total endocranial volume. If the volumetric signature of the *Archaeopteryx* cerebrum reflects the ancestral avialan condition, then these three expansions optimize respectively at the base of Maniraptora, within Deinonychosauria, and somewhere within Avialae after the divergence of *Archaeopteryx* (Fig. 1). This would make the high cerebral indexes of crown birds and deinonychosaurs non-homologous, and the relatively low cerebral index of *Archaeopteryx*, a conserved condition, plesiomorphic for Paraves. Alternatively, if the relative cerebral volume of *Archaeopteryx* is autapomorphically small, then the high cerebral index expressed in deinonychosaurs and crown birds is homologous, and the three coelurosaurian expansions optimize respectively at the bases of Maniraptora, Paraves and within Avialae—again, after the divergence of *Archaeopteryx* but before the origin of the crown (Fig. 1). A cerebral expansion at Paraves corresponds roughly with an inferred paedomorphic event for the skull[25] and may be either a cause or effect of that event.

Studying cerebral expansion with regard to only body size limits the explanatory value of the results because high encephalization may be acquired through gross expansion of the neural tissue, gross reduction of body size, or a combination thereof[26]. The maintenance of relatively small body size along the lineage uniting non-avian maniraptorans with the avian crown[27] indicates that the cerebral expansions we recover cannot be explained solely as body size decreasing around a conserved endocranial volume. Gross neurological expansion was the primary driver in producing these bird-like indices.

The brain of *Archaeopteryx* is not volumetrically avian but conforms to the expectation of a generalized paravian. Thus, if *Archaeopteryx* has a 'flight-ready' brain[3], which is almost certainly the case given its postcranial morphology, then so did other paravians. The hypothesis that dromaeosaurs and troodontids had the neurological capabilities required of powered flight, gliding, or some intermediate condition is congruent with the discovery of the 'four-winged' deinonychosaurs, *Microraptor zhaoianus*[20] and *Anchiornis huxleyi*[5,6].

Endocranial volumes alone, of course, do not tell the whole story. Even volumetrically identical spaces can differ in their shape and structural arrangement with important functional and systematic implications. For example, the wulst is a neurological structure unique to crown birds used in information processing and motor control with two primary inputs: somatosensory and visual[28]. Our re-examination of *Archaeopteryx* revealed a cerebral indentation within the area occupied by the variably positioned wulst of modern birds (Supplementary Fig. 3). If homologous, the wulst appeared before the significant cerebral expansion marking the latter part of the avian stem with which the wulst is typically associated[28,29]. The behavioural implications of this structure's absence in the volumetrically expanded cerebra of *Zanabazar* and *Tsaagan* are unclear, but until evidence of the wulst is recovered from other deinonychosaurs, its homologous presence in *Archaeopteryx* would be an unambiguous synapomorphy shared with crown birds.

METHODS SUMMARY

All computed tomography scanning parameters and volumetric measurements are available in the Supplementary Information. Original computed tomography

scans performed at the University of Texas High-Resolution X-Ray Computed Tomography Facility (UTCT) and Ohio University. Endocranial cast reconstruction and segmentation were performed in VGStudioMAX and detailed in the Supplementary Information. It is important to note that these regions are casts and may contain more neurological tissues than implied by the descriptors used in the text. Bivariate and multivariate statistics performed in PAST (see Supplementary Information for full details).

Full Methods and any associated references are available in the online version of the paper.

Received 10 March; accepted 24 June 2013.
Published online 31 July 2013.

1. Norell, M. A. & Xu, X. Feathered dinosaurs. *Annu. Rev. Earth Planet. Sci.* **33,** 277–299 (2005).
2. Northcutt, R. G. Evolving large and complex brains. *Science* **332,** 926–927 (2011).
3. Dominguez Alonso, P., Milner, A. C., Ketcham, R. A., Cookson, M. J. & Rowe, T. B. The avian nature of the brain and inner ear of *Archaeopteryx*. *Nature* **430,** 666–669 (2004).
4. Larsson, H. C. E., Sereno, P. C. & Wilson, J. A. Forebrain enlargement among theropod dinosaurs. *J. Vert. Paleont.* **20,** 615–618 (2000).
5. Hu, D., Hou, L., Zhang, L. & Xu, X. A pre-*Archaeopteryx* troodontid theropod from China with long feathers on the metatarsus. *Nature* **461,** 640–643 (2009).
6. Xu, X. *et al.* A new feathered maniraptoran dinosaur fossil that fills a morphological gap in avian flight. *Chin. Sci. Bull.* **54,** 430–435 (2009).
7. Butler, A. B. & Hodos, W. *Comparative Vertebrate Neuroanatomy: Evolution and Adaptation* (Wiley, 2005).
8. Gill, F. B. *Ornithology* (W.H. Freeman, 2006).
9. Northcutt, R. G. Understanding vertebrate brain evolution. *Integr. Comp. Biol.* **42,** 743–756 (2002).
10. Jerison, H. J. *Evolution of the Brain and Intelligence* (Academic Press, 1973).
11. Northcutt, R. G. & Kaas, J. H. The emergence and evolution of mammalian neocortex. *Trends Neurosci.* **18,** 373–379 (1995).
12. Nieuwenhuys, R., Ten Donkelaar, J. H. & Nicholson, C. *The Central Nervous System of Vertebrates* (Springer, 1998).
13. Gauthier, J., Kluge, A. G. & Rowe, T. Amniote phylogeny and the importance of fossils. *Cladistics* **4,** 105–209 (1988).
14. Shimizu, T. & Karten, H. J. in *Vision, Brain and Behavior in Birds* (eds Zeigler, H. P. & Bischor H.-J.) 103–114 (MIT, 1993).
15. Rowe, T. B., Eiting, T. P., Macrini, T. E. & Ketcham, R. A. Organization of the olfactory and respiratory skeleton in the nose of the gray short-tailed opossum *Monodelphis domestica*. *J. Mamm. Evol.* **12,** 303–336 (2005).
16. Rowe, T. B., Macrini, T. E. & Luo, Z.-X. Fossil evidence on origin of the mammalian brain. *Science* **332,** 955–957 (2011).
17. Franzosa, J. W. *Evolution of the Brain in Theropoda* (*Dinosauria*). PhD thesis, Univ. Texas (2004).
18. Witmer, L. M. & Ridgely, R. C. New insights into the brain, braincase, and ear region of tyrannosaurs (Dinosauria, Theropoda), with implications for sensory organization and behavior. *Anat. Rec.* **292,** 1266–1296 (2009).
19. Bever, G. S., Brusatte, S. L., Balanoff, A. M. & Norell, M. A. Variation, variability, and the origin of the avian endocranium: insights from the anatomy of *Alioramus altai* (Theropoda: Tyrannosauroidea). *PLoS ONE* **6,** e23393 (2011).
20. Xu, X. *et al.* Four-winged dinosaurs from China. *Nature* **421,** 335–340 (2003).
21. Xu, X., You, H., Du, K. & Han, F. An *Archaeopteryx*-like theropod from China and the origin of Avialae. *Nature* **475,** 465–470 (2011).
22. Christiansen, P. & Fariña, R. A. Mass prediction in theropod dinosaurs. *Hist. Biol.* **16,** 85–92 (2004).
23. Hopson, J. A. in *Biology of the Reptilia* Vol. 9 (eds Gans, C., Northcutt, R. G. & Ulinski, P.) 39–146 (Academic Press, 1979).
24. Rowe, T. B. Coevolution of the mammalian middle ear and neocortex. *Science* **273,** 651–654 (1996).
25. Bhullar, B.-A. S. *et al.* Birds have paedomorphic dinosaur skulls. *Nature* **487,** 223–226 (2012).
26. Smaers, J. B., Dechmann, D. K. N., Goswami, A., Soligo, C. & Safi, K. Comparative analyses of evolutionary rates reveal different pathways to encephalization in bats, carnivorans, and primates. *Proc. Natl Acad. Sci. USA* **109,** 18006–18011 (2012).
27. Turner, A. H., Pol, D., Clarke, J. A., Erickson, G. M. & Norell, M. A. Basal dromaeosaurid and size evolution preceding avian flight. *Science* **317,** 1378–1381 (2007).
28. Reiner, A., Yamamoto, K. & Karten, H. J. Organization and evolution of the avian forebrain. *Anat. Rec. Pt A* **287A,** 1080–1102 (2005).
29. Milner, A. C. & Walsh, S. A. Avian brain evolution: new data from Palaeogene birds (Lower Eocene) from England. *Zool. J. Linn. Soc.* **155,** 198–219 (2009).
30. Turner, A. H., Makovicky, P. J. & Norell, M. A. A review of dromaeosaurid systematics and paravian phylogeny. *Bull. Am. Mus. Nat. Hist.* **371,** 1–206 (2012).

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

Acknowledgements Funding for this project was provided by a NSF DDIG (DEB 0909970) to A.M.B. and M.A.N., NSF IIS-0208675 and EAR-0948842 to T.B.R. and a Columbia University International Travel Fellowship to A.M.B. The University of Texas Computed Tomography Facility and the AMNH MIF helped with computed tomographic scanning and processing imagery. M. Colbert, P. Gignac, D. Ksepka, J. Flynn and J. Meng read and provided useful comments on the text.

Author Contributions A.M.B., G.S.B. and M.A.N. designed the study. A.M.B. wrote the paper, performed data entry and analytical work, and prepared figures. G.S.B. assisted in data interpretation and helped to write the paper. T.B.R. contributed computed tomography data and assisted in data interpretation. M.A.N. provided computed tomography data and assisted in writing 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 A.M.B. (abalanoff@gmail.com).

METHODS

Scanning parameters and endocast construction. Digital endocasts were extracted using computed tomography data from the cranial cavity of examined specimens (Supplementary Table 1). The endocranial casts were reconstructed using original two-dimensional imagery, either tiff or DICOM images, in the volumetric rendering program VGStudioMax 2.0.1. The reconstructions were made by adjusting grayscale contrast in the images until bone and matrix were distinguishable from one another. The cranial cavity was selected using the segmentation tools available in that program. Measurements taken from the endocast (including volume) were acquired using this same version of VGStudioMax. Endocast volume measurements were obtained by calculating the volume of negative space of the cranial cavity. For ease of description, features of the endocranial casts are referred to by the names of the soft tissues of the brain that they reflect (for example, cerebrum rather than cast of cerebrum). It is important to note that what is actually preserved is a cast of the endocranial space, which may reflect structures other than the brain, such as meninges and vascular sinuses. The endocast does help to determine relative size and shape of features of the brain as well as recognizing the branching points of the cranial nerves[23].

Osteological markers used to partition endocranium. Homologous osteological markers of the endocranium were used during the segmentation of the endocasts to distinguish five separate anatomical regions of the brain based on gross anatomical descriptions. Divisions along tissue boundaries are of course not possible when segmenting the endocranial space into separate regions. It is important to note that these regions are casts and may contain more neurological tissues than are implied by the descriptors that are used in the text. As far as preservation allowed, these regions were isolated from each endocast so that their volumes could be measured independently (Fig. 1 and Supplementary Table 2). The five segments included in this analysis comprise the olfactory bulbs and cerebrum (forebrain); the optic lobes (midbrain); and the cerebellum and brain stem (hindbrain)[7]. The pituitary body volume was also segmented but not included in the final analysis because of its sporadic preservation in the fossil taxa. The osteological landmarks used to define the regions were identified in the two-dimensional, coronal computed tomography slices (except the olfactory bulbs, which were identified in the sagittal slices) and delineated from each other using the computer program VGStudioMax 2.0.1. Standard segmentation tools from this program were used in the isolation of anatomical regions.

In this analysis, the region defined as the olfactory bulb is delineated from the rest of the forebrain anteriorly by the anterior-most constriction of the endocranial space before it opens into the nasal cavity and posteriorly by the crista that lies between the olfactory bulb fossa and the anterior cerebral cavity (Fig. 1).

The cerebral region is defined anteriorly by the same crista that delineates the posterior margin of the olfactory bulbs, dorsally by the roof of the cranial cavity formed by the frontals and parietals, posterodorsally by the crista between the cerebral and the cerebellar fossae, and ventrally by the crista between the optic lobe fossa and the cerebral fossa (Fig. 1). There is no clear posterior border of the cerebral fossa in the cranial space, so a straight line was drawn dorsally between the paired cristae that lie at the junction of the cerebral and cerebellar fossae and ventrally between the paired cristae lying at the junction of the cerebral and optic lobe fossae. These were followed posteriorly until they converged, simulating the folding of the cerebellum over the cerebrum.

The cast of the optic lobes is the only anatomical structure from the midbrain that is isolated on the endocasts. These are easily identified dorsally by drawing a line between the crista lying between the cerebral and optic lobe fossae to their opposite on the other side of the braincase. Ventrally, the border of the optic lobe cast is defined as the crista between these structures and the pituitary fossa until the fossa closes posteriorly. Posterior to the pituitary fossa a line is drawn along the crista between the fossa for the brain stem and the optic lobe fossa. The posterior border is formed as the dorsal line and ventral line converge towards the midpoint of the endocranial space (Fig. 1). The optic nerve cast is included in the midbrain and terminates anteriorly at the level of the optic foramen.

The two isolated regions of the hindbrain include the casts of the cerebellum and brain stem (Fig. 1). The cerebellar cast at its anterior margin can overlie or lie just posterior to the posterior margin of the cerebral fossa. In life, the cerebellum resides in a distinctive fossa on the roof of the cranial cavity, delineated by a surrounding crista, crista marginalis (Fig. 1). The fossa for the brain stem underlies both those of the optic lobes and the cerebellum. Again, a distinctive crista is present at the dorsal border of the brain stem fossa and traverses both of these sections (Fig. 1). The posterior margin of this region is determined by the opening of the foramen magnum (Fig. 1).

Regressions and principal components analysis. This analysis segments the endocranial cast into five separate regions that correspond closely with neuroanatomical partitions of the brain. Although previous studies of fossil taxa have partitioned the endocast (that is, cerebral cast and everything else[4]), those analyses were not able to observe changes in the various neuroanatomical divisions with respect to each other, body size, or total endocranial volume.

Several different analyses were run on the volumetric data gained from the digital endocasts of both avian and non-avian theropods. The initial analysis was a simple regression run between body mass and total endocranial volume, similar to the analyses reported by previous papers[3,4,10,17,23] but with an expanded taxonomic sampling. In addition, body mass was plotted against each of the endocranial regions. Endocast volumes were obtained from the volume of negative space as calculated by VGStudioMax 2.0.1. Endocranial volumes in crown birds were found to be an accurate measure of brain volume; and it is assumed that this also holds for non-avian theropods given that impressions of the brain on the cranial cavity indicate that the brain was sufficiently expanded so as to fill the space. Cranial nerves were cut off as close to the 'brain' as possible to minimize their influence on the total volume. Body masses for both avian and non-avian theropods were calculated using femur length (Supplementary Table 2)[22]. All volumetric and body mass data were log transformed to accommodate them onto a single chart and to facilitate ease of pattern recognition. Best-fit lines were mapped onto the data using reduced major axis regression. These lines were fit to the coelurosaurian and crown bird datapoints. A best-fit line also was drawn for the paraphyletic group 'non-avian theropods' to approximate the ancestral condition in the data (Supplementary Table 3).

The Phenotypic Diversity Analysis Program (PDAP) was used subsequently on bivariate data to test for non-independence of variables due to phylogenetic influence, and a second set of regressions was run using independent contrasts (Supplementary Table 4). As the inclusion of fossils required the use of strictly morphological trees, branch lengths within the crown were set at one. Branch lengths for fossil taxa were set according to their distribution in the fossil record. Dates were set at the midpoint of ranges.

A PCA was run on the volumes of all of the five regions of the endocranium that were isolated in this analysis (olfactory bulbs, cerebrum, optic lobes, cerebellum and brain stem). Rather than running the PCA on the absolute values of the volumes, each region was divided by the overall endocranial volume both for consistency with previous analyses performed on crown birds and mammals and to minimize errors that may be associated with shape and distortion. Minimal polygons were drawn around the taxa belonging to the clades Oviraptorosauria, Aves, and Paraves for ease of comparisons (Supplementary Table 5).

LETTER

doi:10.1038/nature12486

Video game training enhances cognitive control in older adults

J. A. Anguera[1,2,3], J. Boccanfuso[1,3], J. L. Rintoul[1,3], O. Al-Hashimi[1,2,3], F. Faraji[1,3], J. Janowich[1,3], E. Kong[1,3], Y. Larraburo[1,3], C. Rolle[1,3], E. Johnston[1] & A. Gazzaley[1,2,3,4]

Cognitive control is defined by a set of neural processes that allow us to interact with our complex environment in a goal-directed manner[1]. Humans regularly challenge these control processes when attempting to simultaneously accomplish multiple goals (multitasking), generating interference as the result of fundamental information processing limitations[2]. It is clear that multitasking behaviour has become ubiquitous in today's technologically dense world[3], and substantial evidence has accrued regarding multitasking difficulties and cognitive control deficits in our ageing population[4]. Here we show that multitasking performance, as assessed with a custom-designed three-dimensional video game (NeuroRacer), exhibits a linear age-related decline from 20 to 79 years of age. By playing an adaptive version of NeuroRacer in multitasking training mode, older adults (60 to 85 years old) reduced multitasking costs compared to both an active control group and a no-contact control group, attaining levels beyond those achieved by untrained 20-year-old participants, with gains persisting for 6 months. Furthermore, age-related deficits in neural signatures of cognitive control, as measured with electroencephalography, were remediated by multitasking training (enhanced midline frontal theta power and frontal–posterior theta coherence). Critically, this training resulted in performance benefits that extended to untrained cognitive control abilities (enhanced sustained attention and working memory), with an increase in midline frontal theta power predicting the training-induced boost in sustained attention and preservation of multitasking improvement 6 months later. These findings highlight the robust plasticity of the prefrontal cognitive control system in the ageing brain, and provide the first evidence, to our knowledge, of how a custom-designed video game can be used to assess cognitive abilities across the lifespan, evaluate underlying neural mechanisms, and serve as a powerful tool for cognitive enhancement.

In a first experiment, we evaluated multitasking performance across the adult lifespan. A total of 174 participants spanning six decades of life (ages 20–79; ~30 individuals per decade) played a diagnostic version of NeuroRacer to measure their perceptual discrimination ability ('sign task') with and without a concurrent visuomotor tracking task ('driving task'; see Supplementary Information for details of NeuroRacer). Performance was evaluated using two distinct game conditions: 'sign only' (respond as rapidly as possible to the appearance of a sign only when a green circle was present); and 'sign and drive' (simultaneously perform the sign task while maintaining a car in the centre of a winding road using a joystick (that is, 'drive'; see Fig. 1a)). Perceptual discrimination performance was evaluated using the signal detection metric of discriminability (d'). A 'cost' index was used to assess multitasking performance by calculating the percentage change in d' from 'sign only' to 'sign and drive', such that greater cost (that is, a more negative percentage cost) indicates increased interference when simultaneously engaging in the two tasks (see Methods Summary).

Prior to the assessment of multitasking costs, an adaptive staircase algorithm was used to determine the difficulty levels of the game at which each participant performed the perceptual discrimination and visuomotor tracking tasks in isolation at ~80% accuracy. These levels were then used to set the parameters of the component tasks in the multitasking condition, so that each individual played the game at a customized challenge level. This ensured that comparisons would inform differences in the ability to multitask, and not merely reflect disparities in component skills (see Methods, Supplementary Figs 1 and 2, and Supplementary Information for more details).

Multitasking performance diminished significantly across the adult lifespan in a linear fashion (that is, increasing cost, see Fig. 2a and Supplementary Table 1), with the only significant difference in cost between adjacent decades being the increase from the twenties (−26.7% cost) to the thirties (−38.6% cost). This deterioration in multitasking performance is consistent with the pattern of performance decline across the lifespan observed for fluid cognitive abilities, such as reasoning[5] and working memory[6]. Thus, using NeuroRacer as a performance assessment tool, we replicated previously evidenced age-related multitasking deficits[7,8], and revealed that multitasking performance declines linearly as we advance in age beyond our twenties.

In a second experiment, we explored whether older adults who trained by playing NeuroRacer in multitasking mode would exhibit improvements in their multitasking performance on the game[9,10] (that is, diminished NeuroRacer costs). Critically, we also assessed whether this training

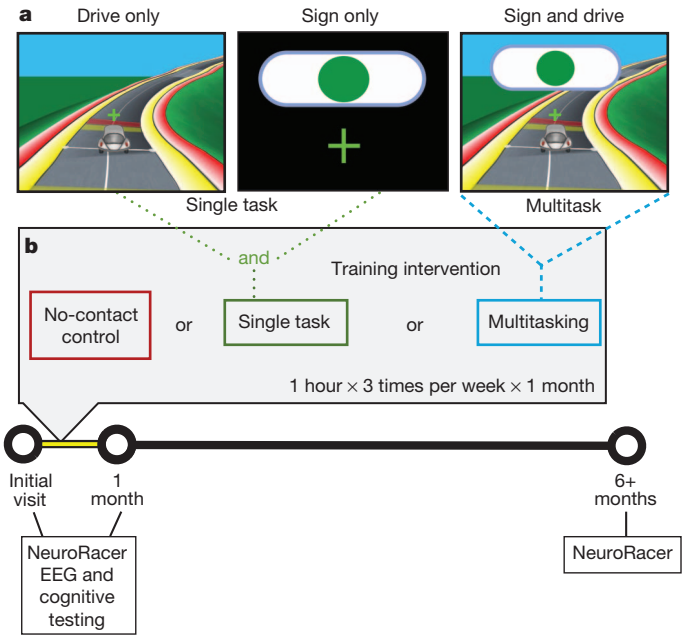


Figure 1 | NeuroRacer experimental conditions and training design. **a**, Screen shot captured during each experimental condition. **b**, Visualization of training design and measures collected at each time point.

[1]Department of Neurology, University of California, San Francisco, California 94158, USA. [2]Department of Physiology, University of California, San Francisco, California 94158, USA. [3]Center for Integrative Neuroscience, University of California, San Francisco, California 94158, USA. [4]Department of Psychiatry, University of California, San Francisco, California 94158, USA.

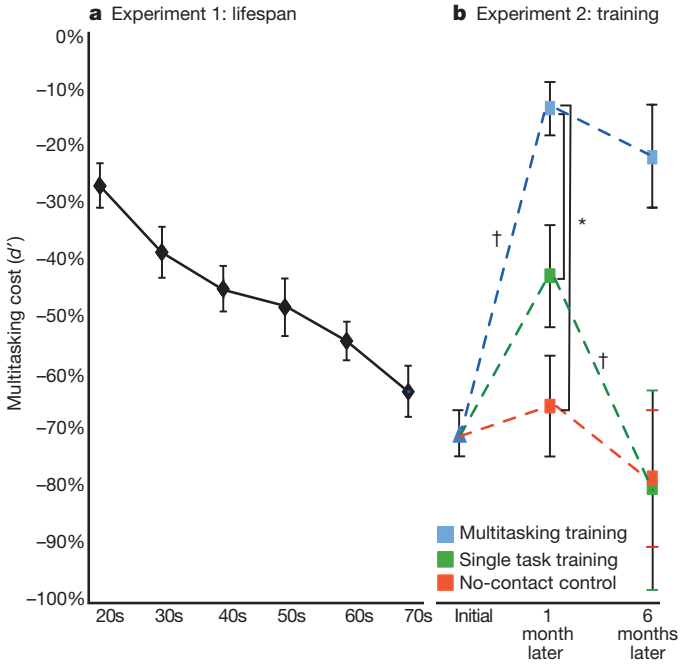


Figure 2 | NeuroRacer multitasking costs. **a**, Costs across the lifespan ($n = 174$) increased (that is, a more negative percentage) in a linear fashion when participants were grouped by decade ($F(1,5) = 135.7$, $P < 0.00001$) or analysed individually ($F(1,173) = 42.8$, $r = 0.45$, $P < 0.00001$; see Supplementary Fig. 3), with significant increases in cost observed for all age groups versus the 20-year-old group ($P < 0.05$ for each decade comparison). **b**, Costs before training, 1 month post-training, and 6 months post-training showed a session X group interaction ($F(4,72) = 7.17$, $P < 0.0001$, Cohen's $d = 1.10$), with follow-up analyses supporting a differential benefit for the MTT group (Cohen's d for MTT vs STT = 1.02; MTT vs NCC = 1.20). †$P < 0.05$ within group improvement from pre to post, *$P < 0.05$ between groups ($n = 46$). Error bars represent s.e.m.

transferred to enhancements in their cognitive control abilities[11] beyond those attained by participants who trained on the component tasks in isolation. In designing the multitasking training version of NeuroRacer, steps were taken to maintain both equivalent difficulty and engagement in the component tasks to assure a prolonged multitasking challenge throughout the training period: difficulty was maintained using an adaptive staircase algorithm to independently adjust the difficulty of the 'sign' and 'driving' tasks following each 3-min run based on task performance, and balanced task engagement was motivated by rewards given only when both component tasks improved beyond 80% on a given run.

We assessed the impact of training with NeuroRacer in a longitudinal experiment that involved randomly assigning 46 naive older adults (60–85 years old; 67.1 ± 4.2 (mean age ± s.d.)) to one of three groups: multitasking training (MTT; $n = 16$), single task training (STT; $n = 15$) as an active control, or no-contact control (NCC; $n = 15$). Training involved playing NeuroRacer on a laptop at home for 1 h a day, 3 times a week for 4 weeks (12 h of training in total), with all groups returning for a post-training assessment after 1 month, and a follow-up assessment after 6 months (Fig. 1b). The MTT group played the 'sign and drive' condition exclusively during the training period, whereas the STT group divided their time between a 'sign only' and a 'drive only' condition, and so were matched for all factors except the presence of interference. In addition to a battery of cognitive control tests used to assess the breadth of training benefits (see Supplementary Table 2), the neural basis of training effects was evaluated using electroencephalography (EEG) recorded at pre- and post-training visits while participants performed a neural assessment version of NeuroRacer.

Analysis showed that only the MTT group's multitasking performance index significantly improved from pre- (−64.2% cost) to post-training (−16.2% cost; Fig. 2b), thus supporting the role of interference during game play as a key mechanistic feature of the training approach. In addition, although cost reduction was observed only in the MTT group, equivalent improvement in component task skills was exhibited by both STT and MTT (see Supplementary Figs 4 and 5). This indicates that enhanced multitasking ability was not solely the result of enhanced component skills, but a function of learning to resolve interference generated by the two tasks when performed concurrently. Moreover, the d' cost improvement following training was not the result of a task trade-off, as driving performance costs also diminished for the MTT group from pre- to post-training (see Supplementary Information). Notably in the MTT group, the multitasking performance gains remained stable 6 months after training without booster sessions (at 6 months, −21.9% cost). Interestingly, the MTT group's post-training cost improved significantly beyond the cost level attained by a group of 20 year olds who played a single session of NeuroRacer (−36.7% cost; experiment 3; $P < 0.001$).

Next, we assessed if training with NeuroRacer led to generalized enhancements of cognitive control abilities that are known to be impaired in ageing (for example, sustained attention, divided attention, working memory; see Supplementary Table 2)[12]. We hypothesized that being immersed in a challenging, adaptive, high-interference environment for a prolonged period of time (that is, MTT) would drive enhanced cognitive performance on untrained tasks that also demanded cognitive control. Consistent with our hypothesis, significant group X session interactions and subsequent follow-up analyses evidenced pre- to post-training improvements in both working memory (delayed-recognition task with and without distraction[7]; Fig. 3a, b) and sustained attention

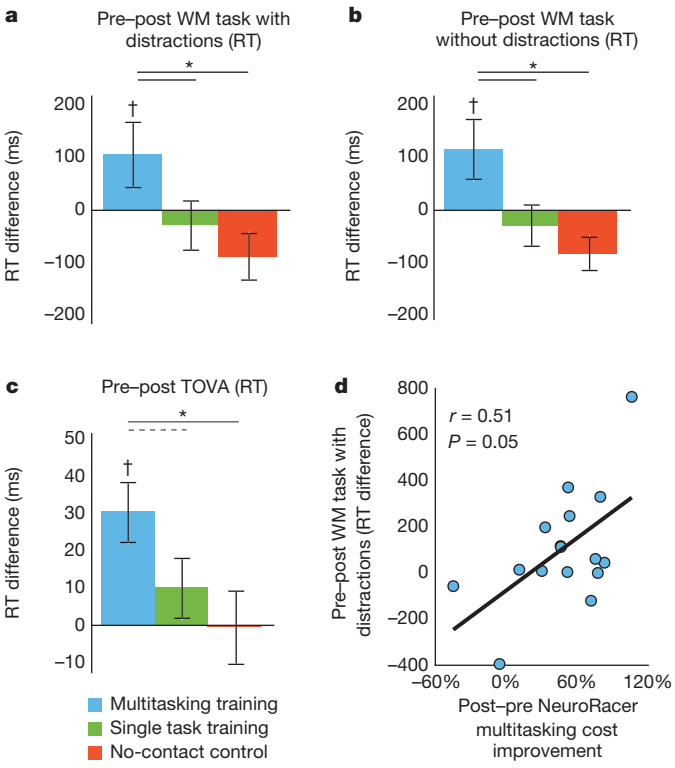


Figure 3 | Change in performance across sessions on independent tests of cognition for each experimental group. For each test, a group X session ANOVA revealed a significant interaction ($F(2,43) > 3.39$, $P < 0.04$, Cohen's $d > 0.73$), with follow-up analyses demonstrating improvement only for MTT ($n = 15$). **a**, Response time (RT) change for a delayed-recognition working memory (WM) task with the presence of distraction ($n = 46$). **b**, Response time change for a delayed-recognition WM task without distraction. **c**, Response time change for the test of variables of attention (TOVA). **d**, Correlation between data from (**a**) and NeuroRacer multitasking cost improvement 1 month after training for the MTT group ($n = 16$). †$P < 0.05$ within group improvement from pre to post, *$P < 0.05$ between groups, - - - $P = 0.08$. Error bars represent s.e.m.

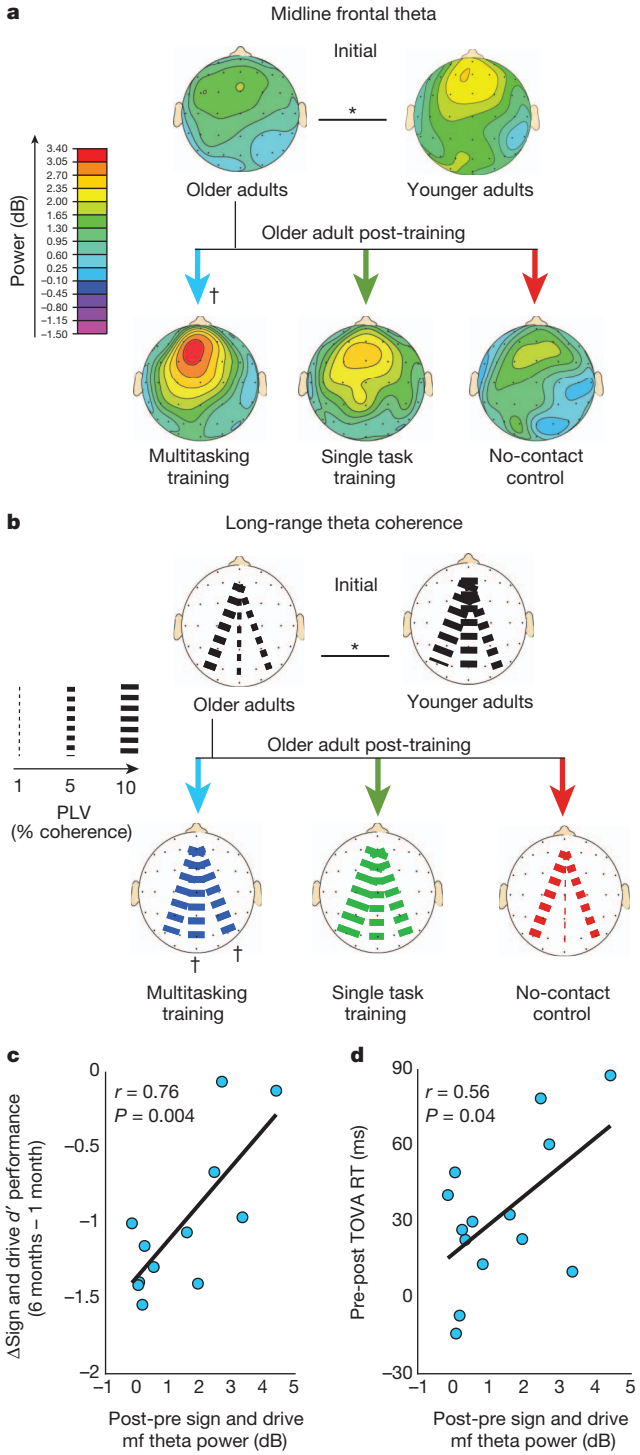


Figure 4 | 'Sign and drive' midline frontal theta activity and long-range theta coherence in younger adults and older adults pre- and post-training. **a**, **b**, For older adult training assessments, a group X session X condition ANOVA for each neural measure revealed significant interactions (in each case, $F(2,41) > 4.98$, $P < 0.01$, Cohen's $d > 0.93$; see Supplementary Fig. 6a, b), with follow-up analyses demonstrating improvement only for MTT during 'sign and drive' ($n = 15$). For younger ($n = 18$) vs older adult ($n = 44$) assessments, both neural measures revealed significant reductions in older adults (see Supplementary Fig. 8a, b). **c**, Correlation in the MTT group between the change in midline frontal theta power and multitasking behavioural gain preservation 6 months later ($n = 12$). **d**, Correlation in the MTT group between the change in midline frontal (mf) theta power and behavioural improvement on the TOVA ($n = 14$). †$P < 0.05$ within group improvement from pre- to post-training, *$P < 0.05$ between groups.

(vigilance; test of variables of attention (TOVA)[13]) only for the MTT group (Fig. 3c; see Supplementary Table 2). In addition, there were several statistical trends suggestive of improved post-training performance on other cognitive control tasks (dual-tasking, useful field of view, and change detection task; see analysis of covariance (ANCOVA) results in Supplementary Table 2). Note that although the working memory and sustained attention improvements were documented as more rapid responses to test probes, neither impulsivity (assessed with the alternative version of the TOVA) nor accuracy results showed significant group differences, revealing that training effects were not the result of a speed/accuracy trade-off. Importantly, these cognitive improvements were specific to working memory and sustained attention processes, and not the result of generalized increases in speed of processing, as no group X session interactions were found on two processing speed tasks (a stimulus detection task and the digit symbol substitution task; see Supplementary Table 2). Finally, only the MTT group exhibited a significant correlation between multitasking cost reduction (assessed with NeuroRacer) and improvements on an untrained cognitive control task (delayed-recognition with distraction) from pre- to post-training (Fig. 3d).

These important 'transfer of benefits' suggest that a common, underlying mechanism of cognitive control was challenged and enhanced by MTT with NeuroRacer. To assess this further, we examined the neural basis of training effects by quantifying event-related spectral perturbations (ERSP) and long-range phase coherence time-locked to the onset of each sign presented during NeuroRacer both pre- and post-training. We specifically assessed midline frontal theta (4–7 Hz), a well-described EEG measure of cognitive control (for example, working memory[14], sustained attention[15] and interference resolution[16]) localized to the medial prefrontal cortex. In addition, we analysed long-range theta coherence between frontal and posterior brain regions, a functional connectivity measure also associated with cognitive control (for example, working memory[14] and sustained attention[15]). Separate ANOVAs for theta power and coherence each revealed significant three-way interactions of condition ('sign and drive', 'sign only') X session (pre, post) X group (MTT, STT, NCC; see Supplementary Fig. 6). Further analysis revealed that for the 'sign and drive' condition, only the MTT group demonstrated a significant increase from pre- to post-training in both neural measures (see Fig. 4a, b). These findings are consistent with other reports of training-driven modulations in prefrontal cortical activity of older adults[9,17]. Furthermore, the coherence results demonstrate for the first time modulation of a neural network in response to cognitive training in older adults. These findings evidence a shift in the rapid engagement of prefrontal cognitive control processes less than 400 ms after a sign appears and before the motor response (see Supplementary Fig. 7 and Supplementary Table 1b), supporting training-induced neuroplasticity as the mechanistic basis of these training effects.

As described above, both MTT and STT resulted in equivalent improvements on the NeuroRacer component tasks (see Supplementary Table 2 and Supplementary Fig. 4), whereas only MTT led to broad enhancements both behaviourally (diminished multitasking costs, improved sustained attention and working memory) and neurally (enhanced midline frontal theta power and long-range coherence). This indicates that the training factor driving these effects was the interference generated when participants were motivated to engage in the two tasks simultaneously. Given that there were no clear differences in sustained attention or working memory demands between MTT and STT, transfer of benefits to these untrained tasks must have resulted from challenges to overlapping cognitive control processes. Of note, the use of a three-dimensional immersive and fun video game for training (see Methods) diverges from the sparse environments typically used in dual-task training studies[9,10], which have not documented a similar degree of far transfer[10].

Coupled with previous findings of increased midline frontal theta on a variety of cognitive control tasks[18], the current results support a common neural basis of cognitive control processes, which can be

enhanced by immersion in an adaptive, high-interference environment. This interpretation is bolstered by evidence here indicating that MTT-induced increases in midline frontal theta power during 'sign and drive' were positively correlated with both sustained multitasking performance improvements (6 months post performance, Fig. 4c) and improvements in TOVA response times (Fig. 4d). Thus, MTT-induced enhancement of midline frontal theta power was associated with the preservation of multitasking performance over time and with generalized benefits on an untrained cognitive control task, reflecting its utility as a neural signature of plastic cognitive control processes.

Finally, we questioned whether these neural measures that exhibited training effects in older adults were actually altered at baseline compared to younger adults, or if training boosted non-deficient neural processes. In a third experiment, we compared midline frontal theta power and long-range coherence from older adults before training to a naive group of younger adults who were not trained ($n = 18$; 20–29 years old (24.1 ± 2.9)). The multitasking costs for each group replicated findings of age-matched cohorts from experiment 1. Both neural measures showed a main effect of group (see Supplementary Fig. 8), indicating less theta power and coherence in older adults when processing signs in either condition ('sign and drive' depicted in Fig. 4a, b). The absence of a significant condition X age group interaction for either neural measure (see Supplementary Fig. 8) revealed that ageing was associated with a general reduction in theta power and coherence when older adults discriminate visual stimuli, regardless of whether they are multitasking or single tasking. Notably, MTT led to changes in the neural processing of signs during 'sign and drive' that reached a level comparable to neural activity patterns observed in younger adults.

The mechanism underlying these neural findings are informed by a growing literature that shows deactivation of medial prefrontal cortical activity (suppression of a node of the 'default network'[19]) during cognitively demanding tasks is associated with reduced susceptibility to internal distraction and better task performance[20]. Given that medial prefrontal activity is inversely correlated with midline frontal theta power[21], increased levels of midline frontal theta exhibited by older adults following MTT may reflect more deactivation of medial prefrontal activity. NeuroRacer training may benefit cognitive control abilities by improving the ability of older adults to suppress the default network during task engagement, a process known to be compromised in ageing[22]. Future studies using neurochemical and physiological manipulations are warranted to inform the causal nature of the relationship between medial prefrontal activity and training-induced performance effects observed here.

This study offers neural and behavioural evidence of generalized positive effects from video game training on cognitive control abilities of older adults, with enhancements comparable to those observed in younger adults who are habitual action video-game players: interference resolution[23], working memory[24] and sustained attention[25]. Although reports of transfer of benefits following cognitive training in the older population are relatively rare[11,26], the observed generalization supports the results of larger-scale training studies that demonstrate some degree of transfer to untrained cognitive tasks[27,28] and subjective measures of daily living[29]. In contrast to these studies, and most other cognitive training experiments on older adults that report small to medium effect sizes for untrained tasks, the current findings document medium to large effect sizes (all > 0.50–1.0 (using Cohen's d, see Methods)) for both cognitive control performance and neural measures versus either control group. The sustained multitasking cost reduction over time and evidence of generalizability to untrained cognitive control abilities provide optimism for the use of an adaptive, interference-rich, video game approach as a therapeutic tool for the diverse populations that suffer from cognitive control deficits (for example, attention deficit hyperactivity disorder (ADHD), depression, dementia). These findings stress the importance of a targeted training approach, as reinforced by a recent study that observed a distinct lack of transfer following non-specific online cognitive exercises[30]. In conclusion, we provide evidence of how a custom-designed video game targeting impaired neural processes in a population can be used to diagnosis deficits, assess underlying neural mechanisms and enhance cognitive abilities.

METHODS SUMMARY

All participants had normal or corrected vision, no history of neurological, psychiatric or vascular disease, and were not taking any psychotropic or hypertension medications. In addition, they were considered 'non-gamers' given that they played less than 2 h of any type of video game per month. For NeuroRacer, each participant used their left thumb for tracking and their right index finger for responding to signs on a gamepad controller (Logitech). Participants engaged in three 3-min runs of each condition in a randomized fashion. Signs were randomly presented in the same position over the fixation cross for 400 ms every 2, 2.5 or 3 s, with the speed of driving dissociated from sign presentation parameters. The multitasking cost index was calculated as follows: (('sign and drive' performance − 'sign only' performance)/'sign only' performance) × 100. EEG data for 1 MTT post-training participant and 1 STT pre-training participant were corrupted during acquisition. Two MTT participants, two STT participants and four NCC participants were unable to return to complete their six-month follow-up assessments. Critically, no between-group differences were observed for neuropsychological assessments ($P = 0.52$) or pre-training data involving: (1) NeuroRacer thresholding for both road ($P = 0.57$) and sign ($P = 0.43$); (2) NeuroRacer component task performance ($P > 0.10$ for each task); (3) NeuroRacer multitasking costs ($P = 0.63$); (4) any of the cognitive tests (all ANOVAs at pre-training: $P \geq 0.26$); (5) ERSP power for either condition ($P \geq 0.12$); and (6) coherence for either condition ($P \geq 0.54$).

Full Methods and any associated references are available in the online version of the paper.

Received 16 January; accepted 18 July 2013.

1. Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S. & Cohen, J. D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108,** 624–652 (2001).
2. Dux, P. E. *et al.* Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron* **63,** 127–138 (2009).
3. Foehr, U. G. Media multitasking among American youth: prevalence, predictors, and pairings. (Kaiser Family Foundation, 2006).
4. Gazzaley, A. in *Principles of Frontal Lobe Function* 2nd edn (eds D. T. Stuss & R. T. Knight) Top-down modulation and cognitive aging. (Oxford Univ. Press, 2013).
5. Tucker-Drob, E. M. & Salthouse, T. A. Adult age trends in the relations among cognitive abilities. *Psychol. Aging* **23,** 453–460 (2008).
6. Park, D. C. *et al.* Models of visuospatial and verbal memory across the adult life span. *Psychol. Aging* **17,** 299–320 (2002).
7. Clapp, W. C., Rubens, M. T., Sabharwal, J. & Gazzaley, A. Deficit in switching between functional brain networks underlies the impact of multitasking on working memory in older adults. *Proc. Natl Acad. Sci. USA* **108,** 7212–7217 (2011).
8. Verhaeghen, P., Steitz, D. W., Sliwinski, M. J. & Cerella, J. Aging and dual-task performance: a meta-analysis. *Psychol. Aging* **18,** 443–460 (2003).
9. Erickson, K. I. *et al.* Training-induced plasticity in older adults: effects of training on hemispheric asymmetry. *Neurobiol. Aging* **28,** 272–283 (2007).
10. Lussier, M., Gagnon, C. & Bherer, L. An investigation of response and stimulus modality transfer effects after dual-task training in younger and older. *Front. Hum. Neurosci.* **6,** 129 (2012).
11. Zelinski, E. M. Far transfer in cognitive training of older adults. *Restor. Neurol. Neurosci.* **27,** 455–471 (2009).
12. Gazzaley, A., Cooney, J. W., Rissman, J. & D'Esposito, M. Top-down suppression deficit underlies working memory impairment in normal aging. *Nature Neurosci.* **8,** 1298–1300 (2005).
13. Greenberg, L. M. *T.O.V.A. Continuous Performance Test Manual.* (The TOVA company, 1996).
14. Onton, J., Delorme, A. & Makeig, S. Frontal midline EEG dynamics during working memory. *Neuroimage* **27,** 341–356 (2005).
15. Sauseng, P., Hoppe, J., Klimesch, W., Gerloff, C. & Hummel, F. C. Dissociation of sustained attention from central executive functions: local activity and interregional connectivity in the theta range. *Eur. J. Neurosci.* **25,** 587–593 (2007).
16. Nigbur, R., Ivanova, G. & Sturmer, B. Theta power as a marker for cognitive interference. *Clin. Neurophysiol.* **122,** 2185–2194 (2011).
17. Dahlin, E., Nyberg, L., Backman, L. & Neely, A. S. Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychol. Aging* **23,** 720–730 (2008).
18. Mitchell, D. J., McNaughton, N., Flanagan, D. & Kirk, I. J. Frontal-midline theta from the perspective of hippocampal 'theta'. *Prog. Neurobiol.* **86,** 156–185 (2008).
19. Buckner, R. L., Andrews-Hanna, J. R. & Schacter, D. L. The brain's default network: anatomy, function, and relevance to disease. *Ann. NY Acad. Sci.* **1124,** 1–38 (2008).
20. Grady, C. L., Springer, M. V., Hongwanishkul, D., McIntosh, A. R. & Winocur, G. Age-related changes in brain activity across the adult lifespan. *J. Cogn. Neurosci.* **18,** 227–241 (2006).
21. Scheeringa, R. *et al.* Frontal theta EEG activity correlates negatively with the default mode network in resting state. *Int. J. Psychophysiol.* **67,** 242–251 (2008).

22. Damoiseaux, J. S. *et al.* Reduced resting-state brain activity in the 'default network' in normal aging. *Cereb. Cortex* **18,** 1856–1864 (2008).
23. Strobach, T., Frensch, P. A. & Schubert, T. Video game practice optimizes executive control skills in dual-task and task switching situations. *Acta Psychol. (Amst.)* **140,** 13–24 (2012).
24. Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M. & Gratton, G. The effects of video game playing on attention, memory, and executive control. *Acta Psychol. (Amst.)* **129,** 387–398 (2008).
25. Dye, M. W., Green, C. S. & Bavelier, D. Increasing speed of processing with action video games. *Curr. Dir. Psychol. Sci.* **18,** 321–326 (2009).
26. Berry, A. S. *et al.* The influence of perceptual training on working memory in older adults. *PLoS ONE* **5,** e11537 (2010).
27. Smith, G. E. *et al.* A cognitive training program based on principles of brain plasticity: results from the improvement in memory with plasticity-based adaptive cognitive training (IMPACT) study. *J. Am. Geriatr. Soc.* **57,** 594–603 (2009).
28. Wolinsky, F. D., Vander Weg, M. W., Howren, M. B., Jones, M. & Dotson, M. M. A randomized controlled trial of cognitive training using a visual speed of processing intervention in middle aged and older adults. *PLoS ONE* **8,** e61624 (2013).
29. Ball, K. *et al.* Effects of cognitive training interventions with older adults: a randomized controlled trial. *J. Am. Med. Assoc.* **288,** 2271–2281 (2002).
30. Owen, A. M. *et al.* Putting brain training to the test. *Nature* **465,** 775–778 (2010).

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

Acknowledgements We thank J. Avila, N. Barbahiya, M. Gugel, B. Jensen, R. Moustafa, Y. Rezaeihaghighi, P. Sztybel, C. Vong, A. Wang, B. Yang and D. Yerukhimov for their help with data collection and analyses, and B. Benson for assistance with the NeuroRacer behavioral analysis stream. Thanks to D. Ellingson, N. Falstein, and M. Omernick for insights and support of NeuroRacer development. Thanks to J. Bollinger, J. Kalkstein, J. Mishra, B. Voytek and T. Zanto for support on ERSP and coherence analyses, and Z. Chadick, W. Clapp, J. Fung, M. Hough, E. Morsella, J. Pa, M. Rubens, P. Wais, C. Walsh, and D. Ziegler for helpful discussions. Thanks to all of our participants whose time and efforts made this work possible, and Apple who generously loaned the Gazzaley laboratory all of the MacBook Pro laptops used in this study. Support for this research was provided by the Robert Wood Johnson Foundation's Pioneer Portfolio through a grant from its national program, 'Health Games Research: Advancing Effectiveness of Interactive Games for Health' (A.G.) and the National Institute of Aging (A.G.). J.A.A. was supported by a UCSF Institutional Research and Career Development Award (IRACDA).

Author Contributions J.A.A., J.B., J.L.R., O.A., E.J. and A.G. designed the experiments; J.A.A., J.L.R., O.A., E.J. and A.G. developed the NeuroRacer software; J.A.A., J.B., O.A., F.F., E.K., Y.L. and C.R. collected the data; J.A.A., J.B., O.A., J.J. and C.R. analysed the data; and J.A.A. and A.G. wrote the paper. All authors discussed the results.

Author Information 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 J.A.A. (joaquin.anguera@ucsf.edu) or A.G. (adam.gazzaley@ucsf.edu).

METHODS

Participants. All participants were recruited through online and newspaper advertisements. For experiment 1, 185 (90 male) healthy, right-handed individuals consented to participate according to procedures approved by the University of California at San Francisco. For experiments 2 and 3, 60 (33 males) older adult individuals and 18 (9 male) young adult individuals participated without having been a part of experiment 1 (see Supplementary Table 3 for demographic descriptions and Supplementary Fig. 9 for experiment 2 participant enrollment). Participants who were unable to perform the tasks, as indicated by tracking performance below 15% (6 individuals from experiment 1, 8 individuals from experiment 2), or a false positive rate greater than 70% (5 individuals from experiment 1, 6 individuals from experiment 2) during any one visit or across more than 4 individual training sessions, were excluded.

Thresholding. Prior to engaging in NeuroRacer, participants underwent an adaptive thresholding procedure for discrimination (nine 120-s runs) and tracking ability (twelve 60-s runs) to determine a 'sign' and 'drive' level that each participant would perform at ~80% accuracy (see Supplementary Figs 1 and 2). Having individuals engage each condition in their own 'space' following thresholding procedures facilitated a fairer comparison across ages and abilities. This is a frequently omitted procedure in other studies, and leads to difficulty interpreting performance differences (especially multitasking) as being the result of differences in interference processing or due to differences in component task skills.

For the perceptual discrimination thresholding, each participant's performance for a given run was determined by calculating a proportion correct score involving: (1) correctly responding to targets; (2) correctly avoiding non-targets; (3) late responses to targets; and (4) responding to non-targets. At the end of each run, if this score was greater than 82.5%, the subsequent run would be played at a higher level which had a corresponding shorter time window for responses to targets. More specifically, the adaptive algorithm would make proportional level changes depending upon participants' performance from this ~80% median, such that each 1.75% increment away from this median corresponded with a change in level (see Supplementary Fig. 1a). Thus, a 90% performance would lead to a 40-ms reduction in the time window, while a 55% (or less) performance would lead to a 100-ms lengthening of said window. Thresholding parameters for road levels followed a similar pattern with each 0.58% increment away from the same median corresponded with a change in level (see Supplementary Fig. 1b). These parameters were chosen following extensive pilot testing to: (1) minimize the number of trial runs until convergence was reached and (2) minimize convergence instability, while (3) maximizing sampling resolution of user performance.

The first three driving thresholding blocks were considered practice to familiarize participants with the driving portion of the task and were not analysed. A regression over the nine thresholding runs in each a case was computed to select the ideal time window and road speed to promote a level of ~80% accuracy on each distraction free task throughout the experiment (see Supplementary Fig. 2). All participants began the thresholding procedures at the same road (level 20) and sign levels (level 29).

Conditions. Following the driving and sign thresholding procedures, participants performed five different three minute 'missions', with each mission performed three times in a pseudo-randomized fashion. In addition to the 'sign only', 'drive only', and 'sign and drive' conditions, participants also performed a 'sign with road' condition in which the car was placed on 'auto pilot' for the duration of the run and participants responded to the signs, and a 'drive with signs' condition where participants were told to ignore the presence of signs appearing that and continue to drive as accurately as possible. Data from these two conditions are not presented here. Feedback was given at the end of each run as the proportion correct to all signs presented for the perceptual discrimination task (although we used the signal detection metric of discriminability (d')[31] to calculate our 'cost' index throughout the study), and percentage of time spent on the road (see Supplementary Fig. 10). Prior to the start of the subsequent run, participants were informed as to which condition would be engaged in next, and made aware of how many experimental runs were remaining. Including thresholding, the testing session encompassed 75 min of game play.

NeuroRacer training and testing protocol. For experiment 1, participants were seated in a quiet room in front of an Apple MacBook Pro 5.3 laptop computer at an approximate distance of 65 cm from the 15 inch screen. For experiment 2 and 3, participants were seated in a dark room with the screen ~100 cm from the participants. All training participants trained at their homes using an Apple MacBook Pro 5.3 laptop computer while sitting ~60 cm from the screen (see Supplementary Fig. 11a). For experiment 1, each perceptual discrimination-based experimental run (180 s) contained 36 relevant targets (green circles) and 36 lures (green, blue and red pentagons and squares). For experiments 2 and 3, the sign ratio was 24/48.

Prior to training, each participant was given a tutorial demonstrating how to turn on the laptop, properly set-up the joystick, navigate to the experiment, shown what the first day of training would be like in terms of the task, how to interpret what the feedback provided meant, and were encouraged to find a quiet environment in their home for their training sessions. If requested by the participant, a laboratory member would visit the participant at their home to help set up the computer and instruct training. In addition, to encourage/assess compliance and hold participants to a reasonable schedule, participants were asked to plan their training days and times with the experimenter for the entire training period and enter this information into a shared calendar. Each participant (regardless of group) was informed that their training protocol was designed to train cognitive control faculties, using the same dialogue to avoid expectancy differences between groups. There was no contact between participants of different groups, and they were encouraged to avoid discussing their training protocol with others to avoid potentially biasing participants in the other groups.

Each day of training, the participants were shown a visualization of a map that represented their 'training journey' to provide a sense of accomplishment following each training session (Supplementary Fig. 11b). They were also shown a brief video that reminded them how to hold the controller, which buttons to use, their previous level(s) reached, and what the target would be that day for the perceptual discrimination condition. In addition, the laptop's built-in video camera was also activated (indicated by a green light) for the duration of said run, providing (1) visual assessment of task engagement; (2) motivation for participants to be compliant with the training task instructions; and (3) information about any run where performance was dramatically poorer than others.

Participants were discouraged from playing 2 days in a row and they were encouraged to play at the same time of day. MTT participants were reminded that an optimal training experience was dependent upon doing well on both their sign and drive performance without sacrificing performance on one task for the other. Although the STT group were provided a 'driving' or 'sign' score following each training run, the MTT group were also provided an 'overall' score following each run as a composite of performance on both tasks (see Supplementary Figs 5 and 11). Following the completion of every fourth run, participants were rewarded with a 'fun fact' screen regarding basic human physiology (http://faculty.washington.edu/chudler/ffacts.html) before beginning their subsequent training run. To assess if training was a 'fun' experience, participants in each training group rated the training experience on their final visit to the laboratory on a scale of 1 (minimally) to 10 (maximally) (MTT: 6.5 ± 2.2; STT 6.9 ± 2.4; $t = 0.65$, $P = 0.52$). Critically, training groups did not differ on their initial thresholding values for both road ($F(2,45) = 0.58$, $P = 0.57$) and sign ($F(2,45) = 0.87$, $P = 0.43$).

Each laptop was configured to transmit NeuroRacer performance data to our secure laboratory server wirelessly using Dropbox as each run was completed. This facilitated monitoring for compliance and data integrity in a relatively real-time fashion, as participants would be contacted if (1) there was a failure to complete all 20 training runs on a scheduled training day; (2) 'sign only' and 'drive only' performance was suggestive that a problem had occurred within a given training session; or (3) a designated training day was missed. Individuals without wireless internet in their home were instructed to visit an open wireless internet location (for example, a coffee shop or public library) at least once a week to transfer data, and if this was not an option, researchers arranged for weekly home visits to acquire said data. All participants were contacted through email and/or phone calls on a weekly basis to encourage and discuss their training; similarly, in the event of any questions regarding the training procedures, participants were able to contact the research staff through phone and email.

Pre- and post-training evaluations involving cognitive testing and NeuroRacer EEG took place across 3 different days (appointment and individual test order were counterbalanced), with all sessions completed approximately within the span of a week (total number of days to complete all pre-training testing: 6.5 ± 2.2 days; post-training testing: 6.1 ± 1.5 days). Participants returned for their first post-training cognitive assessments 2.0 ± 2.2 days following their final training session. Although scheduled for 6 months after their final testing session, the follow-up visits at 6 months actually occurred on average 7.6 ± 1.1 months afterwards owing to difficulties in maintaining (and rescheduling) these distant appointments. Critically, no group differences were present regarding any of these time-of-testing measures ($F < 1.81$, $P > 0.18$ for each comparison).

Cognitive battery. The cognitive battery (see Supplementary Table 2) consisted of tasks spanning different cognitive control domains: sustained attention (TOVA; see Supplementary Fig. 12a), working memory (delayed-recognition- see Supplementary Fig. 12b), visual working memory capacity (see Supplementary Fig. 13), dual-tasking (see Supplementary Fig. 14), useful field of view (UFOV; see Supplementary Fig. 15), and two control tasks of basic motor and speed of processing (stimulus detection task, digit symbol substitution task; see Supplementary Fig. 16). Using the analysis metrics regularly reported for each measure, we performed a mixed model ANOVA of group (3: MTT, STT, NCC) X session (2: pre, post) X cognitive test (11; see Supplementary Table 2), and observed a significant three-way interaction ($F(20, 400) = 2.12$, $P = 0.004$) indicative that training had selective benefits across group

and test. To interrogate this interaction, each cognitive test was analysed separately with session X group ANOVAs to isolate those measures that changed significantly following training. We also present the *P* value associated with the ANCOVAs for each measure in Supplementary Table 2 (dependent measure = post-training performance, covariate = pre-training performance), which showed a similar pattern of effects as most of the two-way ANOVAs. The ANCOVA approach is considered to be a more suitable approach when post-test performance that is not conditional/predictable based on pre-test performance is the primary outcome of interest following treatment, as opposed to characterizing gains achieved from pre-training performance (for example, group X session interaction(s))[32]; however, both are appropriate statistical tools that have been used to assess cognitive training outcomes[27,33] (see Supplementary Fig. 17 as an example).

EEG recordings and eye movements. Neural data were recorded using an active two head cap (Cortech Solutions) with a BioSemiActiveTwo 64-channel EEG acquisition system in conjunction with BioSemiActiView software (Cortech Solutions). Signals were amplified and digitized at 1024 Hz with a 16-bit resolution. Anti-aliasing filters were used and data were band-pass filtered between 0.01–100 Hz during data acquisition.

For each EEG recording session, the NeuroRacer code was modified to flash a 1×1 inch white box for 10 ms at one of the corners on the stimulus presentation monitor upon the appearance of a sign. A photodiode (http://www.gtec.at/Products/Hardware-and-Accessories/g.TRIGbox-Specs-Features) captured this change in luminance to facilitate precise time-locking of the neural activity associated with each sign event. During the experiment, these corners were covered with tape to prevent participants from being distracted by the flashing light.

To ensure that any training effects were not due to changes in eye movement, electro-oculographic data were analysed. Using a previously validated approach[34], vertical (VEOG = FP_2-IEOG electrodes) and horizontal (HEOG = REOG-LEOG electrodes) difference waves were calculated from the raw data and baseline corrected to the mean pre-stimulus activity. The magnitude of eye movement was computed as follows: $(VEOG^2 + HEOG^2)^{0.5}$. The variance in the magnitude of eye movement was computed across trials and the mean variance was specifically examined from −200 to 1000 ms post-stimulus onset. The variance was compared: (1) between sessions for each group's performance on the 'sign and drive' and 'sign only' conditions; (2) between groups at each session for each condition; and (3) between young and older adults on each condition. We used two-tailed *t*-test that were uncorrected for multiple comparisons at every ms time point to be as conservative as possible. There was no session difference for any group on the 'sign only' condition ($P > 0.05$ for each group comparison); similarly, there were no differences for the MTT or NCC groups on the 'sign and drive' condition ($P > 0.30$ for each comparison), with the STT group showing more variance following training ($P = 0.01$). With respect to experiment 3, there were also no age differences on either condition ($P > 0.45$ for each comparison). This indicates that the training effects observed were not due to learned eye movements, and that the age effects observed were also not a function of age-related differences in eye movements as well.

EEG analysis. Preprocessing was conducted using Analyzer software (Brain Vision) then exported to EEGLAB[35] for event-related spectral perturbations (ERSP) analyses. ERSP is a powerful approach to identifying stable features in a spontaneous EEG spectrum that are induced by experimental events, and have been used to successfully isolate markers of cognitive control[36,37]. We selected this approach because we felt that a measure in the frequency domain would be more stable than other metrics given the dynamic environment of NeuroRacer. Blinks and eye-movement artefacts were removed through an independent components analysis, as were epochs with excessive peak-to-peak deflections (± 100 μV). Given the use of d', which takes into account performance on every trial, we collapsed across all trial types for all subsequent analyses. Epochs of −1000 to +1000 ms were created for ERSP total power analysis (evoked power + induced power), with theta band activity analysed by resolving 4–100 Hz activity using a complex Morlet wavelet in EEGLAB and referenced to a −900 to −700 pre-stimulus baseline (thus relative power (dB)).

Assessment of the 'sign and drive' ERSP data in 40 ms time bins collapsing across all older adult participants and experimental sessions revealed the onset of peak midline frontal activity to be between 360–400 ms post-stimulus, and so all neural findings were evaluated within this time window for the older adults (see Supplementary Fig. 7 for these topographies). For younger adults, peak theta activity occurred between 280–320 ms, and so for across-group comparisons, data from this time window was used for younger adults.

The cognitive ageing literature has demonstrated delayed neural processing in older adults using EEG[38,39]. For example, previous work has demonstrated that older adults show similar patterns of activity as younger adults in terms of selective processing, but there is a time shift to delayed processing with ageing[38]. For the data generated in this study, presented topographically in Supplementary Fig. 7, it was clear that the peak of the midline frontal theta was delayed in older versus younger adults. To fairly assess if there was a difference in power, it was necessary to select different comparison windows in an unbiased, data-driven manner for each group.

Coherence data for each channel was first filtered in multiple pass bands using a two-way, zero phase-lag, finite impulse response filter (eegfilt.m function in EEGLAB toolbox) to prevent phase distortion. We then applied a Hilbert transform to each of these time series (hilbert.m function), yielding results equivalent to sliding window FFT and wavelet approaches[40], giving a complex time series:

$$h_x[n] = a_x[n] \exp(i\phi_x[n]) \quad (1)$$

in which $a_x[n]$ and $\phi_x[n]$ are the instantaneous amplitudes and phases, respectively. The phase time series ϕ_x assumes values within ($-\pi$, π radians with a cosine phase such that π radians corresponds to the trough and 0 radians to the peak. In order to compute PLV for theta phase, for example, we extract instantaneous theta phases $\phi_\theta[n]$ by taking the angle of $h_\theta[n]$. Event-related phase time-series are then extracted and, for each time point, the mean vector length $R_\theta[n]$ is calculated across trials (circ_r.m function in CircStats toolbox)[41]. This mean vector length represents the degree of PLV in which an *R* of 1 reflects perfect phase-locking across trials and a value of 0 reflects perfectly randomly distributed phases. These PLVs were controlled for individual state differences at each session by baseline correcting each individual's PLVs using their −200 to 0 period (thus, a relative PLV score was calculated for each subject).

Statistical analyses. Mixed model ANOVAs with: (1) decade of life (experiment 1); (2) training group (experiment 2); or (3) age (experiment 3) as the between-group factor were used for all behavioural and neural comparisons, with planned follow-up *t*-tests and the Greenhouse-Geisser correction used where appropriate. One-tailed *t*-tests were used to interrogate group differences for all transfer measures given our a priori hypothesis of the direction of results following multitask training. All effect size values were calculated using Cohen's *d*[42] and corrected for small sample bias using the Hedges and Olkin[43] approach. The neural-behavioural correlations presented included only those MTT participants who demonstrated increased midline frontal theta power following training (14/15 participants). For statistical analyses, we created 1 frontal and 3 posterior composite electrodes of interest (EOI) from the average of the following electrodes: AFz, Fz, FPz, AF3, and AF4 (medial frontal), PO8, P8, and P10 (right-posterior), PO7, P7, and P9 (left-posterior); POz, Oz, O1, O2 and Iz (central-posterior), with PLVs calculated for each frontal-posterior EOI combination separately. For the coherence data, the factor of posterior EOI location (3) was modelled in the ANOVA, but did not show either a main effect or interaction with the other factors.

31. Macmillan, N. A. & Creelman, C. D. *Detection Theory: A User's Guide*. 2nd edn (Lawrence Erlbaum Associates, 2005).
32. Knapp, T. R. & Schafer, W. D. From Gain Score t to ANCOVA F (and vice versa). *Pract. Assess. Res. Eval.* **14,** 6 (2009).
33. Jaeggi, S. M., Buschkuehl, M., Jonides, J. & Perrig, W. J. Improving fluid intelligence with training on working memory. *Proc. Natl Acad. Sci. USA* **105,** 6829–6833 (2008).
34. Berry, A. S., Zanto, T. P., Rutman, A. M., Clapp, W. C. & Gazzaley, A. Practice-related improvement in working memory is modulated by changes in processing external interference. *J. Neurophysiol.* **102,** 1779–1789 (2009).
35. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* **134,** 9–21 (2004).
36. Makeig, S. Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr. Clin. Neurophysiol.* **86,** 283–293 (1993).
37. Neuper, C. & Klimesch, W. *Event-Related Dynamics of Brain Oscillations*. Vol. 159 (Elsevier Science, 2006).
38. Zanto, T. P., Toy, B. & Gazzaley, A. Delays in neural processing during working memory encoding in normal aging. *Neuropsychologia* **48,** 13–25 (2010).
39. Gazzaley, A. *et al.* Age-related top-down suppression deficit in the early stages of cortical visual memory processing. *Proc. Natl Acad. Sci. USA* **105,** 13122–13126 (2008).
40. Bruns, A. Fourier-, Hilbert- and wavelet-based signal analysis: are they really different approaches? *J. Neurosci. Methods* **137,** 321–332 (2004).
41. Berens, P. CircStat: A MATLAB Toolbox for Circular Statistics. *J. Stat. Softw.* **31,** 1–21 (2009).
42. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd edn (Lawrence Erlbaum Associates, 1988).
43. Hedges, L. V. & Olkin, I. *Statistical Methods For Meta-analysis*. (Academic Press, 1985).

LETTER

doi:10.1038/nature12402

The toxicity of antiprion antibodies is mediated by the flexible tail of the prion protein

Tiziana Sonati[1]*, Regina R. Reimann[1]*, Jeppe Falsig[1]*, Pravas Kumar Baral[2], Tracy O'Connor[1], Simone Hornemann[1], Sine Yaganoglu[1], Bei Li[1], Uli S. Herrmann[1], Barbara Wieland[2], Mridula Swayampakula[2], Muhammad Hafizur Rahman[3], Dipankar Das[3], Nat Kav[3], Roland Riek[4], Pawel P. Liberski[5], Michael N. G. James[2] & Adriano Aguzzi[1]

Prion infections cause lethal neurodegeneration. This process requires the cellular prion protein (PrP^C; ref. 1), which contains a globular domain hinged to a long amino-proximal flexible tail[2]. Here we describe rapid neurotoxicity in mice and cerebellar organotypic cultured slices exposed to ligands targeting the α1 and α3 helices of the PrP^C globular domain. Ligands included seven distinct monoclonal antibodies[3], monovalent Fab_1 fragments and recombinant single-chain variable fragment miniantibodies. Similar to prion infections[4–6], the toxicity of globular domain ligands required neuronal PrP^C, was exacerbated by PrP^C overexpression, was associated with calpain activation and was antagonized by calpain inhibitors. Neurodegeneration was accompanied by a burst of reactive oxygen species, and was suppressed by antioxidants. Furthermore, genetic ablation of the superoxide-producing enzyme NOX2 (also known as CYBB) protected mice from globular domain ligand toxicity. We also found that neurotoxicity was prevented by deletions of the octapeptide repeats within the flexible tail. These deletions did not appreciably compromise globular domain antibody binding, suggesting that the flexible tail is required to transmit toxic signals that originate from the globular domain and trigger oxidative stress and calpain activation. Supporting this view, various octapeptide ligands were not only innocuous to both cerebellar organotypic cultured slices and mice, but also prevented the toxicity of globular domain ligands while not interfering with their binding. We conclude that PrP^C consists of two functionally distinct modules, with the globular domain and the flexible tail exerting regulatory and executive functions, respectively. Octapeptide ligands also prolonged the life of mice expressing the toxic PrP^C mutant[7], PrP(Δ94–134), indicating that the flexible tail mediates toxicity in two distinct PrP^C-related conditions. Flexible tail-mediated toxicity may conceivably play a role in further prion pathologies, such as familial Creutzfeldt-Jakob disease in humans bearing supernumerary octapeptides.

The infectious prion PrP^{Sc}, which causes transmissible spongiform encephalopathies, is a pathologically misfolded conformational variant of PrP^C, a glycosylphosphatidyl-inositol (GPI) linked cell-surface protein encoded by the *Prnp* locus[8]. Only PrP^C-expressing cells are damaged by prion infections[1], suggesting that neurotoxicity is triggered by PrP^C–PrP^{Sc} interactions on cell membranes. In an effort to model such interactions, we exposed cerebellar organotypic cultured slices (COCS)[4,9] to a panel of anti-PrP antibodies directed against well-defined PrP^C epitopes[3]. Supplementary Fig. 1 summarizes our main findings and Supplementary Table 1 lists all abbreviations.

PrP^C-overexpressing COCS (derived from tga20 mice that are hypersensitive to prion diseases[5]) were treated with 5 antibodies against the flexible tail (FT) of PrP^C (residues 23–123) (ref. 10) and with 12 antibodies against the globular domain (GD) of PrP^C (residues 124–230; Supplementary Table 2; Fig. 1a). Seven of the antibodies targeting the GD induced dramatic loss of cerebellar granule cells (CGC) at 67 nM and one was toxic at 333 nM, whereas four were innocuous (Fig. 1b). None of three high-affinity antibodies to the octapeptide repeats (OR, residues 50–90 embedded within the FT) were neurotoxic (Fig. 1b). Antibodies POM3 and D13, which bind the 'charged cluster-2' (ref. 11) (CC2, residues 95–110 of the FT), were innocuous at 67 nM but neurotoxic at 200 nM (Fig. 1b). None of the tested antibodies were toxic to $Prnp^{o/o}$ COCS (Supplementary Fig. 2a). The identity of the targeted epitopes appeared to be a better predictor of PrP^C antibody binding induced toxicity than their affinity to PrP^C, suggesting that neurotoxicity resulted from the interaction of antibodies with specific PrP^C domains (Supplementary Table 2).

The mechanisms of neurotoxicity were further explored using POM1, a highly toxic antibody targeting the GD. Wild-type and tga20 COCS lost most granule cells (CGC) within 28 and 14 days post-exposure (dpe) to POM1, respectively (Fig. 2a–c). Controls included POM1-treated $Prnp^{o/o}$ COCS[12], tga20 COCS treated with pooled mouse immunoglobulins (IgG) and tga20 COCS treated with POM1 preincubated with full-length recombinant mouse PrP (PrP(23–230)). All of these controls were unaffected (Fig. 2d), indicating that POM1 toxicity was PrP^C-dependent. Propidium iodide (PI) staining, a marker of cell death, revealed that CGC membrane integrity was compromised before overt neuronal loss in POM1-treated tga20 COCS. The prevalence of PI^+ cells peaked at 3–5 days and decreased thereafter, suggesting removal of dead cells (Fig. 2e, Supplementary Fig. 3). Therefore POM1-mediated neurotoxicity was dependent on *Prnp* gene dosage as well as duration and dosage of the antibody treatment. Similarly to neurotoxicity in prion infections[6], neuronal PrP^C expression was necessary and sufficient for POM1 toxicity (Supplementary Fig. 4a–c), suggesting that its effects were cell-autonomous to neurons. Finally, treatment of COCS with an anti-NOGO (also known as RTN4) receptor antibody was innocuous, indicating that toxicity was not a generic consequence of binding to GPI-linked receptors (Supplementary Fig. 4d).

It was previously reported that antibody-mediated crosslinking of PrP^C at the cell surface leads to neurotoxicity[11]. However, both monovalent and divalent POM1, POM17 and POM19 derivatives (Fab_1 fragments, Fab_2 fragments and recombinant single-chain variable fragment (scFv) miniantibodies) induced neurotoxicity in tga20 COCS but not in $Prnp^{o/o}$ COCS (Fig. 2f, g and Supplementary Figs 2b and 5). Hence neurotoxicity was rapidly induced by structurally diverse GD ligands, and did not require antibody effector functions mediated by Fc receptors (Fig. 2f, g) or complement (Supplementary Fig. 6).

Neurotoxicity may result from conformational changes imposed onto PrP^C by POM1. We therefore solved the crystal structure of the protein complex containing POM1 Fab_1 bound to recombinant mouse PrP residues 120–230, (PrP(120–230)) (Supplementary Fig. 7; Supplementary Tables 2 and 3). The POM1 epitope on PrP(120–230) was discontinuous

[1]Institute of Neuropathology, University Hospital Zurich, Schmelzbergstrasse 12, CH-8091 Zurich, Switzerland. [2]Department of Biochemistry, Faculty of Medicine & Dentistry, University of Alberta 474 Medical Sciences Building, Edmonton, Alberta T6G 2H7, Canada. [3]Department of Agricultural, Food and Nutritional Science, University of Alberta, 410 Agriculture/Forestry Centre, Edmonton, Alberta T6G 2P5, Canada. [4]ETH Zurich, Physical Chemistry, ETH Hönggerberg, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland. [5]Laboratory of Electron Microscopy and Neuropathology, Department of Molecular Pathology and Neuropathology, Medical University of Lodz, Czechoslowacka 8/10, 92-216 Lodz, Poland.
*These authors contributed equally to this work.

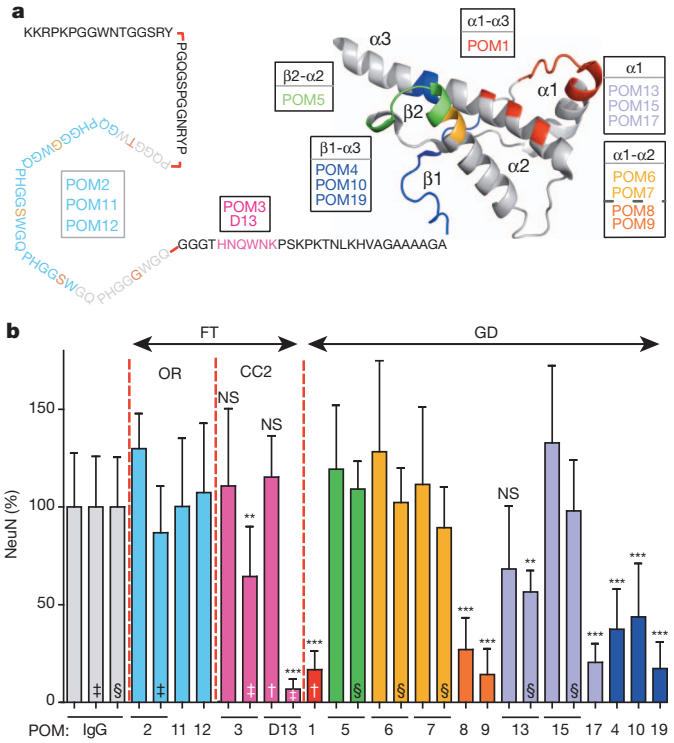


Figure 1 | Epitope specificity of the neurotoxicity mediated by antiprion antibodies. **a**, Antibody binding sites on PrP(23–230). POM6–9, 13, 15 and 17 bind helix α1 and compete with POM1 (ref. 3). POM6 and POM7 (light orange) show additional binding sites at helix α2 (dark orange), whereas POM5 recognizes a unique epitope at the β2–α2 transition (green, Supplementary Table 2). The GD image was rendered using the PDB structure 2L1H. **b**, NeuN morphometry of tga20 COCS treated with various antibodies to FT and GD epitopes (same colour codes as in panel **a**). Treatment modalities: 67 nM, 10 days (†); 200 nM, 21 days (‡); 333 nM, 14 days (§); or 67 nM, 21 days (unlabelled columns). CGC loss was induced by antibodies to helices α1 (light violet, orange, red) and α3 (blue). Antibodies binding the CC2 (magenta) were toxic only at higher concentrations and after protracted incubation. Antibodies against the OR (turquoise) were consistently innocuous. Graph shows percentage of $NeuN^+$ pixels within the CGC layer, normalized against treatment with a random pool of IgG antibodies. One-way ANOVA with Dunnett's post-hoc test, $^{***}P < 0.001$, $^{**}P < 0.01$, error bars represent mean ± s.d., $n = 9$ biological replicates. NS, not significant.

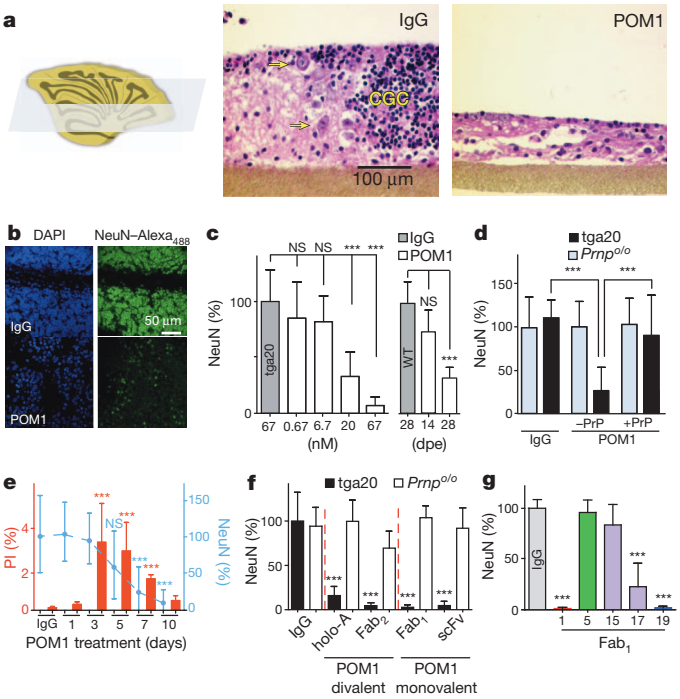


Figure 2 | Antibody-mediated toxicity in cerebellar slices does not require crosslinking or antibody effector functions. **a**, tga20 COCS were treated with POM1 (67 nM) or with pooled IgG (67 nM, used here and in all following figures as the negative control) for 14 dpe, cut transversally as indicated in the left panel, and stained with haematoxylin and eosin. POM1 treatment caused COCS thinning and CGC ablation. Arrows indicate Purkinje cells. **b**, tga20 COCS were exposed to POM1 or to control pooled IgG (67 nM, 10 dpe), immunostained for the neuronal marker NeuN and imaged by confocal laser scanning microscopy. POM1 (lower panels) induced loss of NeuN immunoreactivity. **c**, Concentration and time-dependent CGC loss in POM1-exposed tga20 (left panel; 14 dpe) and wild-type (WT) COCS (right panel; antibody concentration: 270 nM). One-way ANOVA with Dunnett's post-hoc test. **d**, Preincubation of POM1 with PrP(23–230) (molar ratio 1:60) suppressed neurodegeneration of tga20 COCS (antibody treatment: 67 nM, 14 dpe). One-way ANOVA with Tukey's post-hoc test. **e**, COCS were treated with POM1 (67 nM) and scored for $NeuN^+$ (blue) and PI^+ (red) area coverage within the cerebellar granular layer. PI incorporation peaked at 3 dpe, coinciding with the maximum slope of NeuN decay. **f**, $Prnp^{o/o}$ and tga20 COCS were incubated with POM1 holoantibody, divalent Fab_2, monovalent Fab_1, or scFv derivatives of POM1. NeuN morphometry identified severe neurodegeneration after exposure (10 dpe) to each ligand irrespective of its valence. Red dashed lines (here and elsewhere) delineate independent experiments normalized against their respective IgG control assays. One-way ANOVA with Dunnett's post-hoc test. **g**, Treatment of tga20 COCS for 21 days with Fab_1 fragments of POM17 and POM19 (134 nM) resulted in similar toxicity as with holoantibodies, whereas Fab_1 POM5 and Fab_1 POM15 were innocuous. One-way ANOVA with Tukey's post-hoc test. For all experiments, data in bar charts are presented as mean ± s.d., $n = 9$ biological replicates; $^{***}P < 0.001$. NS, not significant.

and included the C-terminal part of the β1–α1 loop, the N-terminal part of helix α1, as well as a segment of helix α3. The surface area of the interface was around 580 $Å^2$ and involved primarily the POM1 heavy chain (Supplementary Video 1 and Supplementary Fig. 7). Comparisons with published structures[2,10,13,14] indicate that Fab_1 POM1 binding did not cause conformational distortions in the GD. We also studied the binding interface of $scFv^{POM1}$ complexed with PrP(23–230) by solution nuclear magnetic resonance (NMR) spectroscopy. Overlays of the NMR spectra of PrP(23–230) with those of the $scFv^{POM1}$–PrP(23–230) complex did not reveal appreciable perturbations of the FT (Supplementary Fig. 8), whereas PrP^C residues in contact with $scFv^{POM1}$ in the GD displayed chemical shift changes as expected from the crystal structure. Moreover, we found subtle perturbations in the chemical shifts in the β1/β2 regions that were not present in the crystal structure (Supplementary Fig. 8).

To investigate whether the FT contributes to the neurotoxicity of GD ligands, we administered POM1 to COCS expressing either PrP(Δ32–93) (lacking most of the FT) or PrP(Δ94–110) (lacking the CC2)[15,16] (Supplementary 9a). Although PrP(Δ94–110) COCS experienced POM1-dependent neurodegeneration, COCS overexpressing PrP(Δ32–93) were resistant to all tested neurotoxic antibodies, POM1, POM4, POM8 and POM19 (Fig. 3a, Supplementary Figs 4b and 9b, c). The affinities of POM1 Fab_1 fragments for N-terminally truncated prion protein PrP(90–230) and full-length recombinant PrP(23–230) were similar (2.9 and 2.5 nM, respectively; Supplementary Table 2), confirming that the resistance of PrP(Δ32–93) COCS to POM1 effect was not due to any steric changes compromising the POM1 binding site.

These results suggested that the FT is required for GD ligand-mediated toxicity. If so, engaging the FT with antibodies might prevent the toxicity of GD antibodies. Indeed, preincubation of tga20 COCS with OR antibodies (POM2 or POM11), or with scFv derivatives, blocked POM1 toxicity (Fig. 3b, c). Interestingly, low concentrations of POM3 and D13 also blocked POM1 toxicity, although these antibodies were toxic at higher concentrations (Fig. 3b). Hence antibodies binding

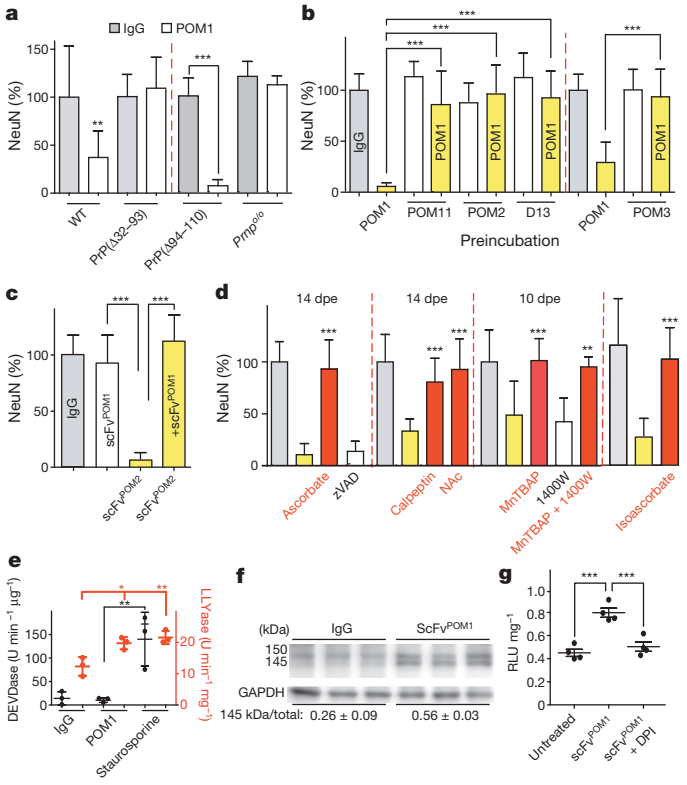


Figure 3 | Neuroprotection against globular domain ligand toxicity. **a**, NeuN morphometry of COCS treated with POM1 antibody (167 nM, 28 dpe). Wild-type (WT) and PrP(Δ94–110) COCS (line L46), but neither PrP (Δ32–93) nor *Prnp*$^{o/o}$ COCS, experienced POM1 toxicity. One-way ANOVA with Tukey's post-hoc test. **b**, tga20 COCS were pre-treated with various antibodies directed against the FT (37°C, 30 min) and exposed to POM1 (yellow bars: 67 nM, 14 dpe). POM2, POM11, POM3 and D13 (67 nM) prevented POM1 toxicity. White bars: omission of POM1. One-way ANOVA with Dunnett's post-hoc test. **c**, Pre-treatment with scFv^{POM2} (264 nM) protected tga20 COCS from scFv^{POM1} (264 nM; 10 dpe). One-way ANOVA with Dunnett's post-hoc test. **d**, Effect of various compounds (listed in Supplementary Table 4) on POM1 toxicity (67 nM). Compound concentrations are reported in the Supplementary Information. Yellow: no pre-treatment. Red: protective compounds. White: non-protective compounds. Grey: IgG control-treated slices. POM1 neurotoxicity was blocked by antioxidants and calpain inhibitors, but not by caspase or iNOS inhibitors. Isoascorbate, a cell impermeable analogue of ascorbate, was also protective. One-way ANOVA with Dunnett's post-hoc test. **e**, Cultures (3 dpe) were assayed for DEVDase (black) and LLYase activity (red), indicative of caspase and calpain activity, respectively. Although staurosporine (24 h) increased both activities, POM1 treatment increased LLYase but not DEVDase activity. Protease activity was normalized to protein amount ± s.d. One-way-ANOVA with Dunnett's post-test, mean ± s.d., $n = 3$ pools of 18 slices. **f**, Immunodetection of α-fodrin in tga20 COCS (3 dpe). ScFv^{POM1} treatment increased total α-fodrin cleavage and the calpain-specific fragment of 145 kDa ($n = 3$ pools of 6 slices; mean ± s.d. underneath blot; two-tailed Student's t-test, $P = 0.0054$). **g**, ScFv^{POM1}-treated COCS (400 nM, 24h) were collected and analysed for ROS production by lucigenin assay in the presence or absence of DPI. One-way ANOVA with Dunnett's post-hoc test, mean ± s.d., $n = 4$, each data point represents a pool of 9 slices. For all experiments, data in bar charts are presented as mean ± s.d., $n = 9$ biological replicates; $^{***}P < 0.001$, $^{**}P < 0.01$.

the CC2 acted similarly to pharmacological 'partial antagonists', triggering opposite effects depending on the presence or absence of additional ligands. Crucially, neither POM2 nor POM3 sterically hinders POM1 binding[3,4] and the non-toxic GD antibody POM5, which did not compete for the POM1 epitope[3] (Supplementary Fig. 9d) did not rescue neurotoxicity (Supplementary Fig. 9e). We conclude that the domain executing toxicity resides within the FT.

Although we suspected that POM1 treatment would result in apoptosis, electron micrographs of POM1-treated COCS revealed cell death with morphology distinct from typical apoptosis (Supplementary Fig. 10). Furthermore, POM1 did not induce caspase activity (Fig. 3e), and the caspase antagonist zVAD conferred no neuroprotection against POM1 (Fig. 3d). Instead, proteolysis of the calpain substrates, α-fodrin[17] and Ac-LLY-AFC, was significantly raised after scFv^{POM1} and POM1 treatment (Fig. 3e, f), and calpain inhibition prevented POM1-mediated toxicity (Fig. 3d). None of the compounds used in the above experiments affected the viability of IgG-treated COCS (Supplementary Fig. 11a). Hence POM1 neurotoxicity results in calpain-dependent cell death, similar to cell death in prion-infected COCS[4].

Oxidative stress is a common promoter of neurodegeneration[18], raising the question of its possible involvement in the toxicity of GD ligands. Indeed, the reactive oxygen species (ROS) scavengers, ascorbate (0.1–2.5 mM) and *N*-acetyl cysteine (NAc), conferred neuroprotection against POM1-mediated toxicity (Fig. 3d, Supplementary Fig. 11a, b). Long-term incubation of POM1 with ascorbate did not reduce the binding of POM1 to PrP(23–230), ruling out the possibility that antioxidants affected antibody stability (Supplementary Fig. 11c). The cell-impermeable molecule isoascorbate was also protective, suggesting that the relevant ROS were extracellular (Fig. 3d, Supplementary 11a). Moreover, inhibition of nitric oxide production did not show a protective effect, whereas the superoxide dismutase mimetic MnTBAP abolished POM1 neurotoxicity (Fig. 3d and Supplementary Fig. 11a), pointing to superoxide rather than nitric oxide as the critical ROS species. NADPH oxidases (NOX) are an important source of extracellular superoxide[19], and indeed scFv^{POM1} treatment induced a superoxide spike in COCS that was suppressed by the NOX inhibitor, diphenyleneiodonium chloride (DPI; Fig. 3g). Hence GD ligands initiate a cascade of events which ultimately triggers the activation of NADPH oxidases.

We then investigated the effects of GD ligands in intact mice. POM1 or IgG were stereotactically injected into the cerebella of tga20 and *Prnp*$^{o/o}$ mice (Fig. 4a, Supplementary Fig. 12a). Sequential manganese-enhanced magnetic resonance imaging (MEMRI)[20] revealed hypointense (dark) lesions indicative of oedema formation arising at 4–72 h post-injection. At their peak, such lesions occupied up to 2.5% and 20% of the total cerebellar volume of wild-type and tga20 mice, respectively (Fig. 4a, b and Supplementary Fig. 12b). Histology revealed CGC chromatin condensation beginning at 5 h post-injection (Supplementary Fig. 13a), conspicuous tissue oedema by 24 h (Supplementary Fig. 13b), and pyknotic CGC death with limited perifocal haemorrhages by 72 h (Fig. 4a). After one week, we found substantial neuronal loss and glial scars (Supplementary Fig. 13c). No such changes were observed in IgG-treated tga20 or POM1-treated *Prnp*$^{o/o}$ mice (Fig. 4a, b, and Supplementary Fig. 13). Hypointense lesions also arose after stereotactic injection of POM1 or D13 into tga20 hippocampi, histologically identified as sites of extensive degeneration of CA3 and dentate gyrus neurons (Supplementary Fig. 14). Intracerebellar injection of scFv^{POM1} also induced extensive lesions in tga20 cerebella and smaller lesions in wild-type cerebella (Fig. 4c). Lesions did not arise when scFv^{POM1} was pre-incubated with PrP(23–230), and when scFv^{POM1} was administered to *Prnp*$^{o/o}$ mice (Fig. 4c). ScFv^{POM2} was innocuous to tga20 mice (Fig. 4c). Hence these *in vivo* results confirm that toxicity was dependent on GD binding and on PrP^{C} expression levels, but neither on antigen cross-linking nor on any antibody effector functions. As with COCS, neuron-restricted PrP^{C} expression sufficed to confer scFv^{POM1} susceptibility *in vivo* (Fig. 4d), whereas mice expressing PrP(Δ32–93)[15] were resistant to scFv^{POM1} toxicity despite higher PrP(Δ32–93) expression, PrP(Δ94–110) mice[16] lacking only the CC2 domain of PrP^{C} experienced toxicity (Fig. 4e and Supplementary Figs 4b and 9a and c).

Mice expressing PrP(Δ94–134), which lacks the central domain that connects the FT and the GD of PrP^{C}, develop spontaneous neurodegeneration whose pathogenesis is poorly understood[7] (Supplementary Fig. 9a). If this pathology were mediated by the FT, it might be alleviated by anti-FT antibodies. Indeed, a single intraventricular injection

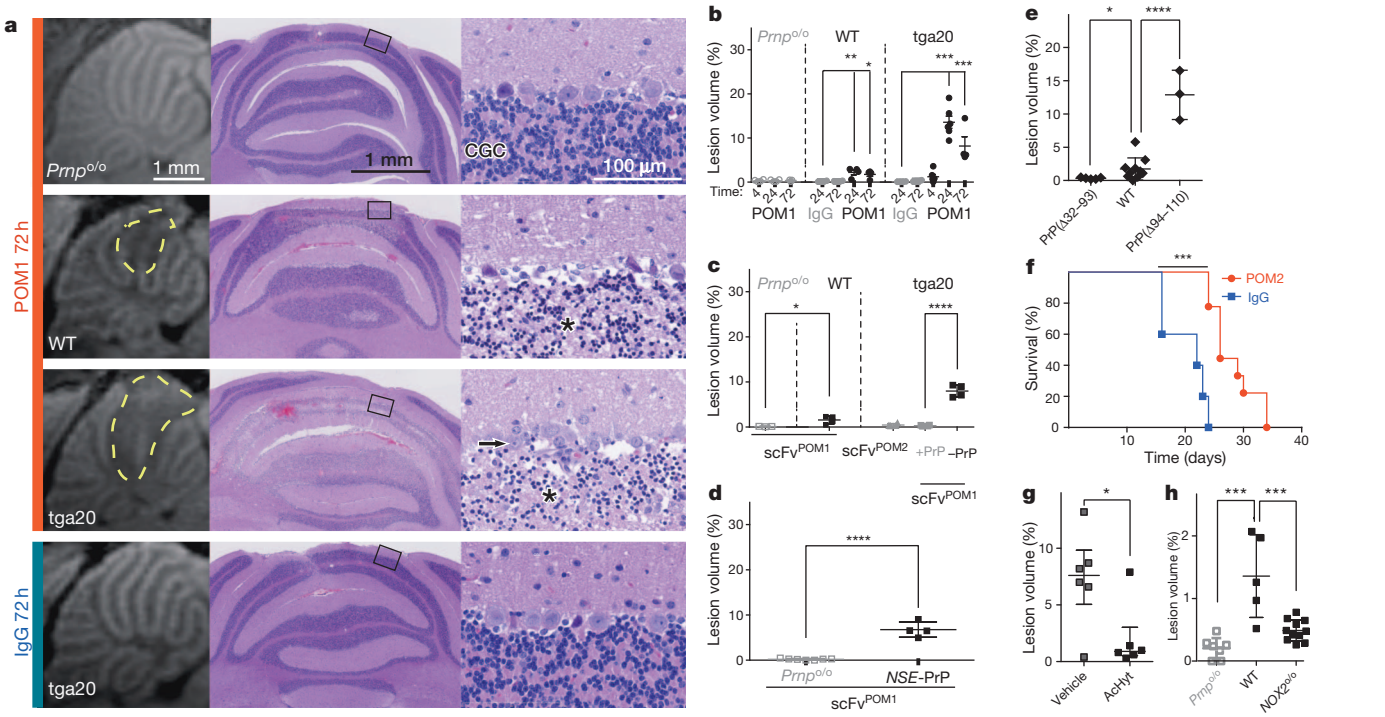


Figure 4 | *In vivo* toxicity of antibodies binding to PrP. **a**, Manganese-enhanced magnetic resonance imaging (MEMRI) scans showed hypointense lesions in POM1-treated (6 µg in 2 µl) wild-type (WT) and tga20 mice (left panels) 72 h post intracerebellar injection. Yellow dashed lines delineate the lesions. Histology (middle and right panels) revealed CGC destruction (asterisks) and oedema. Inset rectangles in the middle images indicate the regions magnified in the corresponding panels on the right. tga20 Purkinje cells (arrow), which do not express PrP^C, were largely preserved. **b**, Lesion volumes at 4, 24 and 72 h post intracerebellar injection as percentages of total cerebellar volumes in wild type, tga20 and $Prnp^{o/o}$. Grey: control injections (6 µg, mean ± s.d., $n = 4$ except for $Prnp^{o/o}$ at 72 h $n = 3$, tga20 at 4 h and 24 h $n = 6$, one-way ANOVA with Dunnett's post-hoc test). **c**, Lesions occurred in tga20 ($n = 4$) and wild-type ($n = 4$) mice after injection of $scFv^{POM1}$ (2 µg) but not after $scFv^{POM2}$ ($n = 4$), or $scFv^{POM1}$ (2 µg) preincubated with recombinant PrP(23–230) (molar ratio 1:5, 24 h, $n = 4$). −PrP indicates that no recombinant PrP was administered. $ScFv^{POM1}$-treatment was innocuous to $Prnp^{o/o}$ ($n = 3$) mice (mean ± s.d., two-tailed Student's *t*-test). **d**, MEMRI volumetry (24 h post intracerebellar injection) of neuron-specific enolase (NSE, also known as Eno2) *NSE*-PrP (*NSE* promoter driving specifically neuronal PrP expression) ($n = 4$) and $Prnp^{o/o}$ ($n = 7$) mice injected with $scFv^{POM1}$ (3 µg; mean ± s.d., two-tailed Student's *t*-test). **e**, MEMRI volumetry 24 h post intracerebellar injection of $scFv^{POM1}$ injected into PrP(Δ94–110) (L52, $n = 3$), PrP(Δ32–93) ($n = 5$) and wild type ($n = 8$) (mean ± s.d.; two-tailed Student's *t*-test). **f**, Survival of mice expressing a PrP(Δ94–134) transgene was prolonged by intraventricular POM2 ($n = 9$) treatment versus IgG control ($n = 5$) (4 µg, Mantel-Cox test). **g**, tga20 mice 24 h post intracerebellar injection of $scFv^{POM1}$ (2 µg) and pre-treated with AcHyt (mean ± s.d., $n = 6$, one-tailed student *t*-test. $P = 0.013$). **h**, MEMRI volumetry at 24 h post intracerebellar injection of $scFv^{POM1}$ (3 µg) injected into $Prnp^{o/o}$ ($n = 7$), wild-type ($n = 5$) and NOX2-deficient ($n = 11$) mice (mean ± s.d., one-tailed Student's *t*-test).

of POM2 to newborn (P0) PrP(Δ94–134) mice resulted in significantly increased median survival (26 vs. 22 days, $n = 9$; Fig. 4f), suggesting that the FT is also involved in the neurotoxicity of these PrP^C mutants.

To determine whether ROS are involved in POM1 neurotoxicity *in vivo*, we administered the enterically activated antioxidant, acetylated hydroxytyrosol[21] (AcHyt) to tga20 mice. Five of six untreated, but only one of six AcHyt-treated tga20 mice developed cerebellar lesions upon $scFv^{POM1}$ injection (Fig. 4g; Supplementary Fig. 15). POM1-treated NADPH oxidase 2 (NOX2) deficient mice[22] displayed much smaller lesions than POM1-treated wild-type controls (Fig. 4h), implicating NOX2 as the crucial source of ROS in this model.

The data reported above uncover a critical role for the FT as the effector domain of PrP^C-mediated neuronal death. FT-mediated toxicity appears to arise from inappropriate gain of deleterious functions, rather than loss of homeostatic interactions, because shortening or steric masking of the FT conferred protection from neurotoxicity (Supplementary Fig. 1a, b), and because most *Prnp* transgenes bearing interstitial FT deletions[5,15,23] are innocuous. Similar mechanisms may be operative in mice[24] and humans[25–27] expressing PrP^C variants with supernumerary OR (Supplementary Fig. 1c). These syndromes are surprising because the OR lie outside the PrP^{Sc} amyloid core and are dispensable for prion generation[15,23]. If the disease caused by supernumerary OR is mechanistically equivalent to that elicited by GD ligands, agents effective against GD ligand toxicity may also be beneficial to patients with OR insertions.

Might GD ligands be toxic by emulating the docking of PrP^{Sc} to PrP^C? This conjecture is supported by the role of activated calpains, which are involved in mediating the neurodegeneration occurring in prion infections. Perhaps PrP^{Sc} oligomers engage PrP^C through its α1/α3 helices in a similar way to the GD ligands (Supplementary Fig. 1c). In this scenario soluble PrP^C may reduce prion toxicity[28] by intercepting PrP^{Sc} oligomers and preventing them from interacting with membrane-bound PrP^C. Furthermore, the dramatic effects of GD ligands *in vivo* suggest that anti-GD autoimmunity may cause neurological conditions, and that it may be worthwhile screening patients with idiopathic neurodegeneration for such autoantibodies.

METHODS SUMMARY

Slices were prepared using a vibratome from 10–12-day-old mice according to a previously published protocol[9]. Viability was assessed by quantifying propidium iodide incorporation or NeuN immunoreactive area using analysis software analySIS v.5.0 (Olympus BioSystem). Fab_1 and Fab_2 fragments were generated by ficin digestion, isolated after protein A elution and further purified by size exclusion chromatography. Single chain fragments of POM1 and POM2 were expressed in *Escherichia coli* (periplasm) and purified by Ni-NTA agarose columns. Recombinant mouse prion proteins (23–230 and 90–230) were generated in bacteria and purified as reported[29,30]. The POM1 epitope was determined by X-ray and NMR analyses as described in the Methods section. Antibodies and respective fragments were stereotactically injected into mouse brains; lesions were visualized by MEMRI and quantified with ParaVision software (v.5 Opl3, Bruker). Surface

plasmon resonance experiments were performed with a BIAcore X100 instrument and interpreted with BIAcore T100 Evaluation Software, v.2.0.3.

Full Methods and any associated references are available in the online version of the paper.

Received 27 November 2012; accepted 21 June 2013.
Published online 31 July; corrected online 4 September 2013 (see full-text HTML version for details).

1. Brandner, S. *et al.* Normal host prion protein necessary for scrapie-induced neurotoxicity. *Nature* **379,** 339–343 (1996).
2. Riek, R. *et al.* NMR structure of the mouse prion protein domain PrP(121–231). *Nature* **382,** 180–182 (1996).
3. Polymenidou, M. *et al.* The POM monoclonals: a comprehensive set of antibodies to non-overlapping prion protein epitopes. *PLoS ONE* **3,** e3872 (2008).
4. Falsig, J. *et al.* Prion pathogenesis is faithfully reproduced in cerebellar organotypic slice cultures. *PLoS Pathog.* **8,** e1002985 (2012).
5. Fischer, M. *et al.* Prion protein (PrP) with amino-proximal deletions restoring susceptibility of PrP knockout mice to scrapie. *EMBO J.* **15,** 1255–1264 (1996).
6. Mallucci, G. *et al.* Depleting neuronal PrP in prion infection prevents disease and reverses spongiosis. *Science* **302,** 871–874 (2003).
7. Baumann, F. *et al.* Lethal recessive myelin toxicity of prion protein lacking its central domain. *EMBO J.* **26,** 538–547 (2007).
8. Aguzzi, A. & Calella, A. M. Prions: protein aggregation and infectious diseases. *Physiol. Rev.* **89,** 1105–1152 (2009).
9. Falsig, J. & Aguzzi, A. The prion organotypic slice culture assay—POSCA. *Nature Protocols* **3,** 555–562 (2008).
10. Riek, R., Hornemann, S., Wider, G., Glockshuber, R. & Wüthrich, K. NMR characterization of the full-length recombinant murine prion protein, mPrP(23–231). *FEBS Lett.* **413,** 282–288 (1997).
11. Solforosi, L. *et al.* Cross-linking cellular prion protein triggers neuronal apoptosis *in vivo*. *Science* **303,** 1514–1516 (2004).
12. Büeler, H. *et al.* Normal development and behaviour of mice lacking the neuronal cell-surface PrP protein. *Nature* **356,** 577–582 (1992).
13. Zahn, R. *et al.* NMR solution structure of the human prion protein. *Proc. Natl Acad. Sci. USA* **97,** 145–150 (2000).
14. Wüthrich, K. & Riek, R. Three-dimensional structures of prion proteins. *Adv. Protein Chem.* **57,** 55–82 (2001).
15. Flechsig, E. *et al.* Prion protein devoid of the octapeptide repeat region restores susceptibility to scrapie in PrP knockout mice. *Neuron* **27,** 399–408 (2000).
16. Bremer, J. *et al.* Axonal prion protein is required for peripheral myelin maintenance. *Nature Neurosci.* **13,** 310–318 (2010).
17. Wang, K. K. Calpain and caspase: can you tell the difference? *Trends Neurosci.* **23,** 20–26 (2000).
18. Fatokun, A. A., Stone, T. W. & Smith, R. A. Oxidative stress in neurodegeneration and available means of protection. *Front. Biosci.* **13,** 3288–3311 (2008).
19. Sorce, S. & Krause, K. H. NOX enzymes in the central nervous system: from signaling to disease. *Antioxid. Redox Signal.* **11,** 2481–2504 (2009).
20. Silva, A. C., Lee, J. H., Aoki, I. & Koretsky, A. P. Manganese-enhanced magnetic resonance imaging (MEMRI): methodological and practical considerations. *NMR Biomed.* **17,** 532–543 (2004).
21. Granados-Principal, S., Quiles, J. L., Ramirez-Tortosa, C. L., Sanchez-Rovira, P. & Ramirez-Tortosa, M. C. Hydroxytyrosol: from laboratory investigations to future clinical trials. *Nutr. Rev.* **68,** 191–206 (2010).
22. Pollock, J. D. *et al.* Mouse model of X-linked chronic granulomatous disease, an inherited defect in phagocyte superoxide production. *Nature Genet.* **9,** 202–209 (1995).
23. Yamaguchi, Y. *et al.* Biological and biochemical characterization of mice expressing prion protein devoid of the octapeptide repeat region after infection with prions. *PLoS ONE* **7,** e43540 (2012).
24. Chiesa, R., Piccardo, P., Ghetti, B. & Harris, D. A. Neurological illness in transgenic mice expressing a prion protein with an insertional mutation. *Neuron* **21,** 1339–1351 (1998).
25. Krasemann, S. *et al.* Prion disease associated with a novel nine octapeptide repeat insertion in the PRNP gene. *Brain Res. Mol. Brain Res.* **34,** 173–176 (1995).
26. Mead, S. *et al.* Inherited prion disease with six octapeptide repeat insertional mutation–molecular analysis of phenotypic heterogeneity. *Brain* **129,** 2297–2317 (2006).
27. Vital, C., Gray, F., Vital, A., Ferrer, X. & Julien, J. Prion disease with octapeptide repeat insertion. *Clin. Exp. Pathol.* **47,** 153–159 (1999).
28. Meier, P. *et al.* Soluble dimeric prion protein binds PrP^{Sc} *in vivo* and antagonizes prion disease. *Cell* **113,** 49–60 (2003).
29. Zahn, R., von Schroetter, C. & Wüthrich, K. Human prion proteins expressed in *Escherichia coli* and purified by high- affinity column refolding. *FEBS Lett.* **417,** 400–404 (1997).
30. Lysek, D. A. & Wuthrich, K. Prion protein interaction with the C-terminal SH3 domain of Grb2 studied using NMR and optical spectroscopy. *Biochemistry* **43,** 10393–10399 (2004).

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

Acknowledgements We thank S. Izui, H. Monyer, D. Burton and G. Mallucci for reagents and mice, S. Schauer and the Functional Genomics Center Zurich for advice and help with affinity determinations, A. Steingötter, U. Ungethüm, M. Polymenidou, A. Lau and A. Keller for input, R. Moos, B. Sikorska, C. Tiberi, P. Schwarz, A. Varol, K. Arroyo and M. Delic for technical help. A.A. is the recipient of an Advanced Grant of the European Research Council and is supported by grants from the European Union (PRIORITY, LUPAS and NEURINOX), the Swiss National Foundation, the Foundation Alliance BioSecure, the Clinical Research Priority Program (KFSP) of the University of Zurich, and the Novartis Research Foundation. J.F. is supported by a career development award of the University of Zurich. Research support from PrioNet Canada and Alberta Prion Research Institute (APRI) for the work conducted in the Canadian laboratories is gratefully acknowledged. P.P.L. and A.A. are supported by Polish Swiss Research grant nr PSPB-062/2010. This paper is dedicated to the memory of Dr Marek Fischer, who created the tga20 mouse line.

Author Contributions T.S., J.F. and A.A. conceived the study. Planning and execution were performed by T.S. with significant contributions from R.R.R., J.F. and additional contribution from T.O'C., S.H., S.Y., B.L. and U.S.H. MEMRI was established and performed by R.R.R.; electron microscopy was performed by P.P.L.; scFv were cloned and produced by B.W., M.S., M.H.R., D.D. and N.K.; X-ray crystallography was performed by P.K.B. and M.N.G.J.; and NMR experiments were performed by S.H. with additional contributions from R.R. T.S., J.F. and A.A. wrote the manuscript.

Author Information X-ray crystallographic coordinates and structure factor files have been deposited in the RSCB Protein Data Bank (PDB) database under the accession code number 4H88. 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.A. (Adriano.Aguzzi@usz.ch).

METHODS

Chemicals and mice. All compounds were purchased from Sigma-Aldrich unless otherwise stated. GABA-Aα6-Cre mice were generated on a C57BL/6 × CBA background and backcrossed to a *Prnp*$^{o/o}$ background[12,31]. Tg37 mice allowing for conditional *PrP* deletion were generated on a *Prnp*$^{o/o}$ FVB background[32]. GABA-$A_{\alpha 6}$-Cre^{-}; *loxPrP*-tg37 littermates were used as negative controls (PrP^{CGC+}). B6.129S6-*Cybb*^{tm1Din}/J (NOX2)-deficient mice were generated on a mixed C57BL/6xSv129 background and backcrossed to C57BL/6[22]. *Prnp*$^{o/o}$, *Prnp*$^{o/o}$; tga20$^{+/+}$ (tga20), *Prnp*$^{o/o}$; C4/C4 (PrP(Δ32–93)), *Prnp*$^{o/o}$;1046 (PrP(Δ94–134))[7], *Prnp*$^{o/o}$;L46 (PrP(Δ94–110))[16], *Prnp*$^{o/o}$;L52 (PrP(Δ94–110)), *Prnp*$^{o/o}$;*NSE-PrP*$^{+/+}$ (*NSE-PrP*) and *Prnp*$^{o/o}$; *MBP-PrP*$^{+/+}$ (*MBP-PrP*) (MBP, myelin basic protein) mice were on a mixed 129Sv/BL6 background and wild-type mice on a C57BL/6 background[5,12,15,16,33,34]. *Prnp*$^{o/o}$;L52 (PrP(Δ94–110)) was previously reported[16]. CD11b-HSVTK mice[35] on a C57BL/6 background were crossed to *Prnp*$^{o/o}$; tga20$^{+/+}$. All offspring were *Prnp*$^{+/o}$; tga20^{+} and CD11b-HSVTK positive offspring were referred to as tga20^{TK+}. All mouse experiments were performed on 6–9-month-old males and females, in accordance with Swiss law, and were approved by the Animal Experimentation Committee of the Canton of Zurich (permits 200/2007 and 130/2008).

Animal surgery. Mice were anaesthetized with isoflurane and placed in a motorized stereotaxic frame controlled by software with a three-dimensional brain map, allowing for real-time monitoring of needle placement (Neurostar). The skull was exposed by cutting along the midline and a small hole was drilled using a surgical drill and the needle mounted in an electronic micro-injector unit was placed for cerebellar injection at the following lambda coordinates: AP −2.3 mm, ML 0 mm, DV 2 mm and for hippocampal injection at the following bregma coordinates: AP −2 mm, ML 1.7 mm, DV 2.2 mm, angle in ML/DV plane 15°. Antibodies (2 µl) were injected at a flow-rate of 0.5 µl min^{-1} and the needle was left in place for 1 min. Mice were sutured and received an injection of buprenorphinum (0.1 µg per g of body weight).

Hydroxytyrosol treatment. Mice were treated with acetylated hydroxytyrosol starting 7 days before injection with 2 g l^{-1} in drinking water.

Newborn injections. Newborn mice (P0) were injected with a thin Hamilton syringe into the lateral ventricle (2 µl of 2 mg ml^{-1} IgG or POM2 solution into each hemisphere) according to previously reported protocols[36].

Manganese-enhanced magnetic resonance imaging (MEMRI). Mice received five intraperitoneal injections of $MnCl_2$ (40 mg kg^{-1}, 20 mM in H_2O and bicine, pH 7.4) at 12 h intervals[37]. The last injection was administered immediately after the stereotaxic injection. Mice were imaged under isoflurane anaesthesia at 4 h, 24 h and 72 h post-surgery in a 4.7-Tesla small-animal MRI system (Bruker Pharmascan). Mice were placed in a bed equipped with a mouse whole-body radio frequency transmitter coil and a mouse head surface-coil receiver (Bruker Pharmascan). Body temperature was maintained with a warming blanket. T-1 weighted brain images were obtained using a 3D gradient-echo sequence (TR: 15 ms, TE: 2.5 ms, FA: 20 deg, Average:10, Matrix: 265/265/126, Field of View: 2 × 2.56 × 2 cm^3, acquisition time: 1 h, voxel size: 78 × 100 × 156 µm^3)[38].

Quantification was performed with ParaVision software (Version 5, Opl3, Bruker). Cerebellar lesions were quantified by two region of interests (ROIs) corresponding to the lesioned and total cerebellar area were set for each optical slice of the 3D data set. For hippocampal lesion quantification, the volume of non-affected CA3 was measured. ROIs were set on the ipsilateral and contralateral side of injection. Volumes for each ROI were calculated by multiplying the sum of the ROI area times the voxel height. Data are presented as the lesion volume divided by the total cerebellar volume, that is, lesion volume (%) or as CA3 volume (mm^3), separated by ipsilateral versus contralateral side.

Organotypic brain culture preparation. Organotypic cerebellar slice cultures, 350-µm thick, were prepared from 10–11-day-old mice according to a previously published protocol[9]. Free-floating sections were cut in Gey's balanced salt solution (GBSS) (NaCl 8 g l^{-1}, KCl 0.37 g l^{-1}, Na_2HPO_4 0.12 g l^{-1}, $CaCl_2$ $2H_2O$ 0.22 g l^{-1}, KH_2PO_4 0.09 g l^{-1}, $MgSO_4$ $7H_2O$ 0.07 g l^{-1}, $MgCl_2$ $6H_2O$ 0.210 g l^{-1}, $NaHCO_3$ 0.227 g l^{-1}) supplemented with the glutamate receptor antagonist kynurenic acid (1 mM) (GBSSK) at 4 °C. Slices were placed on a 6-well Millicell-CM Biopore PTFE membrane insert (Millipore). Residual buffer was removed and the inserts were transferred to a cell culture plate and cultured in 'slice-culture medium' (50% vol/vol MEM, 25% vol/vol basal medium Eagle and 25% vol/vol horse serum supplemented with 0.65% glucose (w/vol), penicillin/streptomycin and glutamax (Invitrogen)). Cultures were kept in a standard cell incubator (37 °C, 5% CO_2, 95% humidity) and the culture medium was exchanged three times per week.

Antibody treatment was randomly assigned to individual wells and initiated after a 10–14-day recovery period allowing the initial gliosis induced by tissue preparation to subside. Slices were collected for biochemical analyses or fixed for immunocytochemical analysis at various time points. For microglia depletion experiments, tga20^{TK+} slices were treated with ganciclovir (GCV, 5 µg ml^{-1}) for 14 days before antibody treatment. At this time point, less than 1% of microglia were left in the tissue[39].

Pharmacological treatment of COCS. Drug treatment was initiated at the time of antibody addition and drugs were added again at every medium change. Appropriate drug concentrations were determined by literature search, assuming that slice culture uptake of compounds were similar to other cell culture systems. The toxicity of each compound was tested in parallel on IgG and POM1 slices; if toxicity occurred, drugs were retested at a lower concentration. Drug and concentration used were ascorbate (1.5 mM), isoascorbate (1.5 mM), MnTBAP (100 µM), benzyloxycarbonyl-Val-Ala-Asp (OMe) fluoromethylketone (zVAD-fmk, 40 µM), diphenyleneiodonium chloride (DPI, 5 µM), N-([3-(Aminomethyl)phenyl]methyl)-ethanimidamide dihydrochloride (1400W, 20 µM), N-benzyloxycarbonyl-L-leucylnorleucinal (calpeptin, 20 µM), *N*-acetylcysteine (NaC, 1 mM).

Protein analysis. COCS were washed twice in PBS. Cerebellar tissue was then scraped off the membrane using 10 µl per slice of PBS with 0.5% DOC, 0.5% NP-40 supplemented with PMSF (1 µM) and complete mini protease inhibitor cocktail (Roche) and homogenized by trituration using a 30G syringe. Protein concentration was determined using the bicinchoninic acid assay (Pierce). Samples were prepared in loading buffer (NuPAGE, Invitrogen) and boiled at 95 °C for 5 min. Proteins (10 µg per lane) were separated on a 12% Bis-Tris polyacrylamide gel or for higher molecular weight proteins on a 4–12% gradient gel (NuPAGE, Invitrogen) and blotted onto a nitrocellulose membrane. Membranes were blocked with 5% w/vol Topblock (Fluka) in Tris-buffered saline supplemented with Tween (150 mM NaCl, 10 mM Tris HCl, 0.05% Tween 20 (vol/vol)) and incubated with primary antibodies in 1% Topblock. Primary mouse monoclonal antibodies used were: POM1, POM2, mouse IgG_1 antibody raised against PrP^{C} (anti-PrP^{C}; 200 ng ml^{-1}), mouse anti-α-fodrin (AA6, 100 ng ml^{-1}, Millipore) and anti-GAPDH (200 ng ml^{-1}, Millipore). Secondary antibodies were horseradish peroxidase (HRP)-conjugated rabbit anti–mouse IgG_1 (1:10,000, Zymed), goat anti–rabbit IgG_1 (1:10,000, Zymed) and rabbit anti–goat IgG_1 (1:10,000, Zymed). Blots were developed using SuperSignal West Pico chemiluminescent substrate (Pierce) and visualized using the VersaDoc system (model 3000, Bio-Rad). PNGase treatment was performed using a commercially available kit, according to the manufacturer's protocol (New England Biolabs). In brief, 10 µg protein was treated with 2 µl denaturation buffer in a 20 µl reaction and incubated for 15 min at 95 °C. A reaction mixture of 2.6 µl G7, 2.6 µl NP-40 (10%), as well as 0.5 µl PNGase was added and samples were incubated for 2 h at 37 °C. Samples were then mixed with loading dye, cooked and analysed by western blotting. Cerebella from 10-day-old mice were homogenized (10% w/v) in 0.32 M sucrose in 1× PBS supplemented with 4-(2-Aminoethyl)-Benzene sulfonyl fluoride HCL (AEBSF, 1 mM) and complete mini protease inhibitor cocktail (Roche) using a tissue Ribolyser (Qiagen). PrP^{C} expression in the brains of PrP deletion mutants was analysed by PrP-specific sandwich ELISA[3]. Briefly, samples were incubated in POM1-coated ELISA plates (overnight at 4 °C). Plates were then washed and incubated with biotinylated POM19 (1.6 ng ml^{-1}) for 1 h at 37 °C. Avidin-HRP (1:1,000) (BD-Pharmigen) was used as detection antibody (1 h, 37 °C). After 3 washes, stabilized chromogen (Invitrogen) was added, incubated for 30 min at 37°C and absorbance was read at 450 nM. All samples were analysed at dilutions falling within the logarithmic range of the calibration curve (PrP(23–230): 0.62–80 ng ml^{-1}).

Generation of POM monoclonals, Fab_1 POM1 and Fab_2 POM1 fragments. POM monoclonal antibodies were generated as published[3]. Fab_1 POM1 and Fab_2 POM1 fragments were generated by ficin digestion and purified on a protein A column according to manufacturer's protocol (Pierce). Protein A flow-through containing Fab_1 POM1 was collected, concentrated by Amicon ultra-15 centrifugal units (Millipore). Fab_1 POM1 was further purified by size exclusion chromatography (Superdex75 10/300 GL, GE Healthcare) using PBS and a flow rate of 0.5 ml min^{-1}. Pure fractions were pooled, concentrated and stored at −20 °C. The concentrations of the IgG and Fab_1 POM fragments were determined by measuring their absorbance at 280 nM in a UV-VIS photometer using an extinction coefficient (E) 0.1% at 280 nM of 1.35. Aldolase (158 kDa), conalbumin (75 kDa) and ovalbumin (43 kDa) were used as control molecular weight markers. The purity of the fragments was checked by Coomassie-stained SDS–PAGE.

For K_d determination, Fab_1 POM fragments were generated from the POM antibodies using immobilized papain (10–20 mg of antibody ml^{-1} of immobilized papain as packed resin) activated in digestion buffer (20 mM sodium phosphate, 10 mM EDTA, 20 mM cysteine hydrochloride, pH 7.0) according to the manufacturer's instructions (Pierce). The reaction mixture was incubated overnight with rigorous shaking at 37 °C. The digest was separated from the immobilized papain with a resin separator. The immobilized papain was then washed with binding buffer (Protein A IgG binding buffer, Pierce) which was added to the digest.

To remove undigested IgG and Fc fragments from the Fab_1 fragments, the digest was loaded onto a 2 ml Protein A plus agarose column (Pierce), equilibrated in binding buffer and washed with 6 ml binding buffer. The Fab_1-containing flow-through

and wash fractions were pooled and further purified by size exclusion chromatography as described above.

Periplasmic expression and purification of scFvPOM1, scFvPOM2, PrP(23–230) and PrP(90–230). For periplasmic expression and purification of $scFv^{POM2}$, the pET-22b(+) vector (Novagen) was used, as it has both periplasmic localization and C-term 6×His tag sequences. The scFvPOM2 gene, constructed previously using the phage display method[3], was inserted into the pET-22b(+) vector between the restriction sites EcoRI and XhoI. The resulting plasmid was transformed into competent Rosetta (DE3) pLysS cells by heat-shock.

A fresh, single colony of transformed Rosetta cells was inoculated in 2×YT media with 100 μg ml^{-1} ampicillin and 34 μg ml^{-1} chloramphenicol (2×YT-AC media) and grown overnight at 37 °C at 250 r.p.m. The overnight culture was further propagated (1:1,000) in fresh 2×YT-AC media until the $D_{600\,nm}$ reached ~0.4–0.5. The culture was then cooled down to 25 °C in a cold water bath and induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. The induction was allowed to continue overnight at 25 °C and 250 r.p.m. The cells were collected and resuspended in 50 ml l^{-1} of the original culture volume of freshly prepared, pre-chilled periplasmic extraction buffer (200 mM Tris-HCl, 20% (w/v) sucrose, 1 mM EDTA pH 8.0) and incubated on ice for 1 h with occasional stirring. The suspension was spun down at 15,000*g* for 30 min at 4 °C and the supernatant was collected as the soluble protein extract. The extract was thoroughly dialysed against 50 mM Tris-HCl, 100 mM NaCl, pH 7.0 at 4 °C and further clarified by centrifugation at 15,000*g* for 30 min at 4 °C (ref. 40).

The clarified soluble protein extract was loaded onto a Ni-NTA agarose (Qiagen) column equilibrated with 50 mM Tris and 100 mM NaCl, pH 7.0. The unbound protein and impurities were washed off with 50 mM Tris, 100 mM NaCl and 50 mM imidazole pH 7.0 and the elution was performed with 50 mM Tris, 100 mM NaCl and 500 mM Imidazole pH 7.0. The purity of the eluates was checked using gel electrophoresis. Pure $scFv^{POM2}$ samples were dialysed against 50 mM Tris, 100 mM NaCl and 50 mM Imidazole pH 7.0 and concentrated to ~3 mg ml^{-1}. The same procedures applied to expression and purification of periplasmic $scFv^{POM1}$ except the induction was at 17 °C. Recombinant mouse PrP was generated in bacteria and purified as reported elsewhere[29,30].

SPR measurements and binding affinity determination. Amino-terminal truncated PrP(90–230) and full-length PrP(23–230) were immobilized on a carboxymethylated-dextran sensor chip (CM5, Biacore) by standard EDC/NHS chemistry. The amount of protein immobilized on the surface corresponded to 1,000 resonance units (RU) of PrP(23–230) and 4600 RU of PrP(90–230) for comparing Fab_1 POM1 binding in the absence of the OR. For comparing different Fab_1 affinities, a density of 100 RU of PrP(23–230) was used (Supplementary Table 1). Various concentrations of Fab_1 measurements were performed on a Biacore T100 at 37 °C. Fab_1 fragments, diluted in running buffer (HBS−EP+) and at a concentration range of 0.01–5.12 nM were injected for 350 s at a flow rate of 30 μl min^{-1}. The sensor surface was regenerated between each measurement with 20 mM NaOH. Binding affinity was determined by using the instrument software (Biacore T100 Evaluation software, v.2.0.3).

Immunocytochemistry. For immunocytochemistry, organotypic slices were washed twice in PBS and fixed in 4% formalin overnight at 4 °C. Membrane inserts were washed and incubated for 1 h in blocking buffer (0.05% vol/vol Triton X-100 and 3% vol/vol goat serum dissolved in PBS) and incubated with conjugated mouse anti-NeuN-$Alexa^{488}$ (0.5 μg ml^{-1}, Millipore) antibody diluted in blocking buffer at 4 °C for 3 d. Inserts were washed four times with PBS and counterstained with 4,6-diamidino-2-phenylindole (DAPI) (1 μg ml^{-1}). For NeuN morphometry images were recorded at ×4 magnification on a fluorescence microscope (BX-61, Olympus) equipped with a cooled black/white charge coupled device (CCD) camera at identical exposure times. The area of immunoreactivity was determined by morphometry with image analysis software analySIS v.5.0 using identical grey-scale threshold settings for identifying positive pixels.

Mice were euthanized after the last scan (either 5, 24 or 72 h post injection) and brains fixed in 4% formalin. Cerebella were paraffin embedded and 2-μm coronal sections were cut and stained. Hematoxylin and eosin staining and NeuN immunohistochemistry were performed according to standard protocols.

Proteolytic assays. Slices were collected in pools of 18 slices in PBS, 0.5% DOC, 0.5% NP-40 with 2% β-mercaptoethanol and homogenized by trituration. Homogenates were analysed immediately for DEVDase activity using caspase 3 fluorometric detection kit (Caspase-3 activity assay kit, Enzo Life Sciences) and normalized to protein concentration. Slices (pools of 18) were homogenized and processed according to the manufacturer's instructions (Calpain Activity Assay kit, Biovision). Calpain activity (Ac-LLY-ase) was measured by detecting cleavage of a specific fluorogenic calpain substrate, Ac-LLY-AFC on a fluorescence plate reader (360-nm excitation filter, 440-nm emission filter) and normalizing to protein amount.

Electron microscopy. Slices were washed in Na-phosphate buffer, fixed in freshly prepared 2% PFA plus 2.5% glutaraldehydein 0.1 M Na-phosphate buffer 0.1 M pH 7.4, postfixed in 1% osmium tetroxide for 1–2 h, dehydrated through a series of graded ethanols and propylene oxide, embedded in Epon and processed for electron microscopy using standard procedures. Grids were examined and photographed in JEOL JEM 100 CX and JEOL JEM 1011 transmission electron microscopes at 80 kV. Nuclear morphology (as presented in Supplementary Fig. 10) was assigned by counting 11 fields for IgG-treated COCS at 3, 7 and 10 dpe. For POM1-treated COCS, 10 fields at 3 dpe, 12 fields at 5 and 10 dpe, were analysed. When defining necrotic, apoptotic and normal POM1-treated slices at 7 dpe, respectively 15, 16 and 20 fields were counted.

Viability and ROS assays. For propidium iodide (PI) incorporation, slices were incubated for 30 min with PI (5 μg ml^{-1}) and images were recorded in living tissue using a fluorescent microscope (Axiovert 200) equipped with a cooled CCD camera using a ×5 objective and analysed by morphometry. Lucigenin conversion assay was performed at room temperature as follows. Inserts containing 5–10 slices each were washed in PBS and collected in Krebs-Ringer solution supplemented with complete mini protease inhibitor cocktail (Roche). Samples were triturated with a 30G syringe and 50 μl of each sample was mixed with 175 μl assay buffer and 0.25 μl lucigenin (10 mM). Background activity was measured using a chemiluminescence reader. Subsequently, 50 μl NADPH (1 mM) was added to each well and the plate was read again. Background activity was subtracted from the NADPH dependent signal and data (each bar: average of 4 inserts ± s.d.) are presented as relative light unit per mg of total protein.

Erythrocyte complement lysis assay. One millilitre of tail vein blood was collected from $Prnp^{o/o}$ and $Prnp^{o/o}$;tga20$^{+/+}$ mice in 50 μl, 0.5 M EDTA on ice. Red blood cells (RBCs) were pelleted (371*g* for 10 min at room temperature) and washed in ice-cold PBS three times. RBCs were resuspended cells in GVB^{2+} buffer (0.15 mM $CaCl_2$, 141 mM NaCl, 0.5 mM $MgCl_2$, 0.1% gelatin, 1.8 mM sodium barbital and 3.1 mM barbituric acid, pH 7.3 and aseptically filtered) and sensitized to complement lysis by incubation with complement activating antibody (34–3C (ref. 41), 100 μg ml^{-1} final concentration) or different POM antibodies (100 μg ml^{-1} final concentration) for 1 h at room temperature. RBCs were washed in ice-cold PBS three times, resuspended in GVB^{2+} buffer and 5×10^7 cells were incubated with different concentrations of serum diluted in GVB^{2+} buffer in a total volume of 125 μl. RBCs were incubated for 30 min at 37 °C, pelleted by centrifugation and 100 μl supernatant was transferred to a transparent 96-well plate and the absorbance at 414 nM was determined. Relative lytic activity was defined as the serum dilution factor leading to 30% erythrocyte lysis, calculated from the fitted serum dilution curves.

Protein expression and isotopic labelling. For the expression of uniform [^{2}H,^{15}N]-labelled PrP(23–230), M9 medium was produced from D_2O. M9 medium was supplemented with magnesium chloride, vitamins, ampicillin (final concentration: 100 μg ml^{-1}) and ammonium chloride (2g l^{-1}). The main culture was started by transferring cells from a 2 l preculture of LB medium grown to $D_{600\,nm}$ of 1.6 at 37 °C into 2 l of M9 D_2O medium. Cells were further incubated for 1 h at 37 °C and induced by addition of 1 mM IPTG. After 6 h cells were collected and PrP(23–230) was purified according to our standard protocol[29,42].

NMR experiments. The [^{15}N,^{2}H]- transverse relaxation optimized spectroscopy (TROSY) NMR[43] experiments of the free ^{2}H,^{15}N-labelled PrP(23–230) and in complex with $scFv^{POM1}$ were performed on a Bruker Avance 700 MHz spectrometer equipped with a 5 mM triple-resonance cryoprobe and a single pulsed field gradient in 10 mM sodium phosphate, pH 7.3 and 5% (v/v) D_2O at 20 °C. The data were processed by the software XwinNMR, version 3.5 (Bruker, Germany) and further analysed by the program CARA (http://www.nmr.ch)[44]. For the assignment of residues 89–230 the assignments deposited in the BioMagResBank (accession numbers: PrP(90–231), 16071; numeration corresponding human PrP) were used, whereas for residues 23–88 a tentative assignment was performed based on the available assignments at pH 4.5 (ref. 45).

Statistical analysis of COCS. One-way ANOVA with Tukey's post-hoc test for multi-column comparison, or Dunnett's post-hoc test for comparison of all columns to a control column, were used for statistical analysis of experiments involving the comparison of three or more samples. A paired Student's *t*-test was used for comparing the two samples. Results are displayed as the average of replicas ± s.d.

Statistical analysis of MEMRI data sets. For statistical analysis of experiments involving the comparison of three or more data sets, we used one-way ANOVA with Dunnett's post-hoc test for comparison of all data sets to a control data set. One or two-tailed unpaired Student's *t*-test was used for comparing two data sets. Results are displayed as mean ± s.d. $*P < 0.05$; $**P < 0.01$; $***P < 0.001$, $****P < 0.0001$.

Data collection and structure determination. Diffraction quality POM1 Fab_1–PrP(120–230) protein complex crystals were grown by the vapour diffusion method at room temperature[46]. X-ray diffraction data were measured at the Stanford Synchrotron Radiation Laboratory (SSRL), beamline 9–2 (refs 47–49). The data were processed in space group *C*2 to a resolution of 1.9 Å using the program HKL2000 (ref. 50). The data collection details as well as the refinement statistics are presented

in Supplementary Table 3. The structure of POM1 Fab_1–PrP(120–230) protein complex was solved by the molecular replacement method using the program MOLREP[51] of the CCP4 package[52]. The coordinates of POM1 Fab_1–huPrP(120–230) protein complex, (PDB accession code: 4DGI), were used as the template. Solution from the molecular replacement was then refined by restrained refinement, implemented in the refinement program of the PHENIX package[53]. The progress of the refinement process was monitored by a reduction in both the R_{work} and R_{free} factors to final values of 19.9% and 23.4%, respectively. Solvent molecules are added to the model by an automated PHENIX program and those water molecules were accepted only when well defined positive peaks were present in both the $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ electron density maps and there was a satisfactory hydrogen-bonding network with either protein atoms or other water molecules. Model building was performed with the program COOT[54]. The final structural coordinates for the POM1 Fab_1–PrP(120–230) complex were validated with MOLPROBITY[55] and deposited in the RSCB PDB (accession code 4H88) along with the structure factors. MOLPROBITY showed that 96.8% of the amino acid residues were in the most favoured region of the Ramachandran plot. All illustrations were created with the program PyMOL (http://www.pymol.org).

31. Aller, M. I. *et al.* Cerebellar granule cell Cre recombinase expression. *Genesis* **36,** 97–103 (2003).
32. Mallucci, G. R. *et al.* Post-natal knockout of prion protein alters hippocampal CA1 properties, but does not result in neurodegeneration. *EMBO J.* **21,** 202–210 (2002).
33. Radovanovic, I. *et al.* Truncated prion protein and Doppel are myelinotoxic in the absence of oligodendrocytic PrP^C. *J. Neurosci.* **25,** 4879–4888 (2005).
34. Prinz, M. *et al.* Intrinsic resistance of oligodendrocytes to prion infection. *J. Neurosci.* **24,** 5974–5981 (2004).
35. Heppner, F. L. *et al.* Experimental autoimmune encephalomyelitis repressed by microglial paralysis. *Nature Med.* **11,** 146–152 (2005).
36. Levites, Y. *et al.* Intracranial adeno-associated virus-mediated delivery of anti-pan amyloid β, amyloid β40, and amyloid β42 single-chain variable fragments attenuates plaque pathology in amyloid precursor protein mice. *J. Neurosci.* **26,** 11923–11928 (2006).
37. Grünecker, B. *et al.* Fractionated manganese injections: effects on MRI contrast enhancement and physiological measures in C57BL/6 mice. *NMR Biomed.* **23,** 913–921 (2010).
38. Faas, H. *et al.* Context-dependent perturbation of neural systems in transgenic mice expressing a cytosolic prion protein. *Neuroimage* **49,** 2607–2617 (2010).
39. Falsig, J. *et al.* A versatile prion replication assay in organotypic brain slices. *Nature Neurosci.* **11,** 109–117 (2008).
40. Das, D., Allen, T. M. & Suresh, M. R. Comparative evaluation of two purification methods of anti-CD19-c-myc-His_6-Cys scFv. *Protein Expr. Purif.* **39,** 199–208 (2005).
41. Baudino, L. *et al.* IgM and IgA anti-erythrocyte autoantibodies induce anemia in a mouse model through multivalency-dependent hemagglutination but not through complement activation. *Blood* **109,** 5355–5362 (2007).
42. Hornemann, S., Christen, B., von Schroetter, C., Perez, D. R. & Wuthrich, K. Prion protein library of recombinant constructs for structural biology. *FEBS J.* **276,** 2359–2367 (2009).
43. Pervushin, K., Riek, R., Wider, G. & Wuthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA* **94,** 12366–12371 (1997).
44. Keller, R. L. J. *The Computer-aided Resonance Assignment Tutorial CARA* (Cantina Verlag 2004) (2004).
45. Riek, R. *NMR of the mouse prion protein*, PhD thesis, ETH Zürich (1998).
46. Baral, P. K. *et al.* Crystallization and preliminary X-ray diffraction analysis of prion protein bound to the Fab fragment of the POM1 antibody. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* **67,** 1211–1213 (2011).
47. Cohen, A. E., Ellis, P. J., Miller, M. D., Deacon, A. M. & Phizackerley, R. P. An automated system to mount cryo-cooled protein crystals on a synchrotron beamline, using compact sample cassettes and a small-scale robot. *J. Appl. Crystallogr.* **35,** 720–726 (2002).
48. González, A. *et al. Web-Ice*: integrated data collection and analysis for macromolecular crystallography. *J. Appl. Crystallogr.* **41,** 176–184 (2008).
49. McPhillips, T. M. *et al. Blu-Ice* and the *Distributed Control System*: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* **9,** 401–406 (2002).
50. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
51. Vagin, A. & Teplyakov, A. *MOLREP*: an automated program for molecular replacement. *J. Appl. Crystallogr.* **30,** 1022–1025 (1997).
52. Winn, M. D. *et al.* Overview of the *CCP*4 suite and current developments. *Acta Crystallogr. D* **67,** 235–242 (2011).
53. Echols, N. *et al.* Graphical tools for macromolecular crystallography in *PHENIX*. *J. Appl. Crystallogr.* **45,** 581–586 (2012).
54. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
55. Chen, V. B. *et al. MolProbity*: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66,** 12–21 (2010).

LETTER

doi:10.1038/nature12416

Induction of intestinal stem cells by R-spondin 1 and Slit2 augments chemoradioprotection

Wei-Jie Zhou[1], Zhen H. Geng[1], Jason R. Spence[2] & Jian-Guo Geng[1]

Cancer research has been rightly and successfully focused on prevention, early detection, and identification of specific molecular targets that distinguish the malignant cells from the neighbouring benign cells[1]. However, reducing lethal tissue injury caused by intensive chemoradiotherapy during treatment of late-stage metastatic cancers remains a key clinical challenge. Here we tested whether the induction of adult stem cells could repair chemoradiation-induced tissue injury and prolong overall survival in mice. We found that intestinal stem cells (ISCs)[2] expressed Slit2 and its single-span transmembrane cell-surface receptor roundabout 1 (Robo1)[3,4]. Partial genetic deletion of *Robo1* decreased ISC numbers and caused villus hypotrophy, whereas a *Slit2* transgene increased ISC numbers and triggered villus hypertrophy. During lethal dosages of chemoradiation, administering a short pulse of R-spondin 1 (Rspo1; a Wnt agonist)[5–14] plus Slit2 reduced ISC loss, mitigated gut impairment and protected animals from death, without concomitantly decreasing tumour sensitivity to chemotherapy. Therefore Rspo1 and Slit2 may act as therapeutic adjuvants to enhance host tolerance to aggressive chemoradiotherapy for eradicating metastatic cancers.

With the use of two-colour fluorescent *in situ* hybridization (FISH), we found that *Slit2* and *Robo1* messenger RNAs not only co-localized, but also were expressed in markedly higher amounts in the crypts when compared to the villi in the small intestines of adult C57BL/6 (wild-type) mice (Fig. 1a and Supplementary Fig. 1a). Ki67^{+} transient amplifying cells and Lgr5^{+} ISCs expressed *Slit2* and *Robo1* mRNAs (Fig. 1b, c). By contrast, neither *Slit2* nor *Robo1* mRNA was visible in lysozyme^{+} Paneth cells. Using the intestinal specimens collected from Lgr5-enhanced green fluorescent protein (eGFP)-internal ribosome entry site (IRES)-CreERT2 (Lgr5–GFP) mice[15], we confirmed that *Slit2* and *Robo1* mRNAs were expressed by GFP^{high} ISCs at the bottom of the crypt (Fig. 1d). The distribution of Slit2 and Robo1 is apparently higher at the crypts than in the villi of the small intestine, which mirrors the well-characterized gradient of active β-catenin in the crypt–villus axis[16].

Complete or partial genetic deletion of *Robo1* is embryonic lethal, whereas mice with partial genetic deletion of both *Robo1* and *Robo2* are viable[17,18]. Thus, we examined the intestinal morphology in *Robo1*$^{+/-}$ *Robo2*$^{+/-}$ double heterozygotes (referred to here as *Robo1/2* mutants). *Robo1/2* mutants displayed reduced expression of Robo1 protein, but not Slit2 or α-tubulin (Supplementary Fig. 2a). Notably, *Robo2* mRNA was not detectable in the wild-type small intestine, even though it clearly was present in the cerebellum (Supplementary Fig. 1b). Compared to wild-type littermates, *Robo1/2* mutants had noticeably sparser, shorter and floppier villi throughout the entire small intestine (Fig. 2a and Supplementary Fig. 3a), displaying markedly fewer Ki67^{+} transient amplifying cells (Fig. 2b), Lgr5^{+} ISCs (Fig. 2c), villin (Vil1)$^{+}$ enterocytes (Fig. 2d), and a reduced number of proliferating intestinal cells visualized by 5-bromodeoxyuridine (BrdU) staining (Supplementary Fig. 4a). The mutant animals also had a higher level of endotoxin (Supplementary Fig. 5), a functional indicator of intestinal impairment[13]. Importantly, *Robo1* single heterozygotes (*Robo1*$^{+/-}$ or *Robo1* mutants) phenocopied the *Robo1/2* double heterozygotes (Supplementary Fig. 6a–d), whereas no such alterations were detected in the intestines of *Robo2* single heterozygotes (*Robo2*$^{+/-}$ or *Robo2* mutants; Supplementary Fig. 7a–d). We also isolated intestinal crypts from *Robo1/2* mutants and found that *Robo1*$^{+/-}$ *Robo2*$^{+/-}$ crypts failed to form intestinal organoids, also

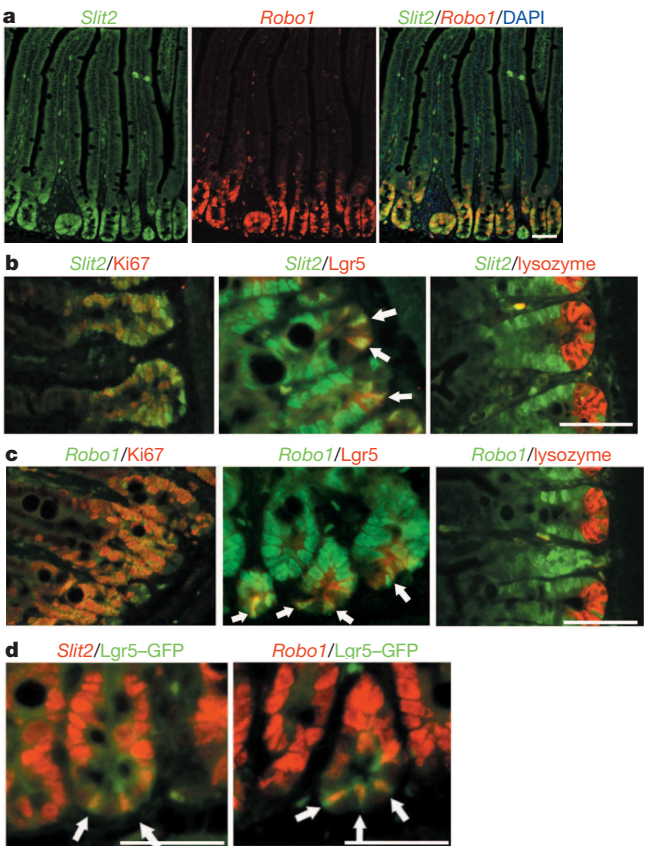


Figure 1 | Expression of Slit2 and Robo1 in the mouse small intestine. **a**, Expression and co-localization of *Slit2* and *Robo1* mRNAs in the crypts of small intestine. *Slit2* and *Robo1* mRNAs in the wild-type small intestines were detected by using the digoxigenin (DIG)- or biotin-conjugated antisense *Slit2* and *Robo1* mRNA probes. Slides were counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Immunofluorescent images were observed under a laser scanning confocal microscope, and the recorded fluorescent images were then merged. **b**–**d**, Cellular distribution of mRNAs for *Slit2* (**b** and **d**) and *Robo1* (**c** and **d**). *Slit2*, *Robo1*, Ki67 (a marker for proliferating transient amplifying cells), Lgr5 (a marker for ISCs) and lysozyme (a marker for Paneth cells) were found at the crypt of small intestines using FISH (*Slit2* and *Robo1*) and immunofluorescent staining with their respective antibodies (Ki67, Lgr5 and lysozyme). Alternatively, the intestinal tissues isolated from Lgr5–GFP mice were stained by the anti-GFP antibody to detect GFP^{high} cells (**d**). White arrows indicate Lgr5^{+} cells (**b**, **c**) and GFP^{high} cells (**d**) co-localized with *Slit2* or *Robo1* mRNA. Results represent at least three separate experiments (five mice per group; 8 weeks old). All scale bars, 50 μm.

[1]Department of Biologic and Materials Sciences, University of Michigan School of Dentistry, Ann Arbor, Michigan 48109, USA. [2]Departments of Cell and Developmental Biology and Internal Medicine, Division of Gastroenterology, University of Michigan School of Medicine, Ann Arbor, Michigan 48109, USA.

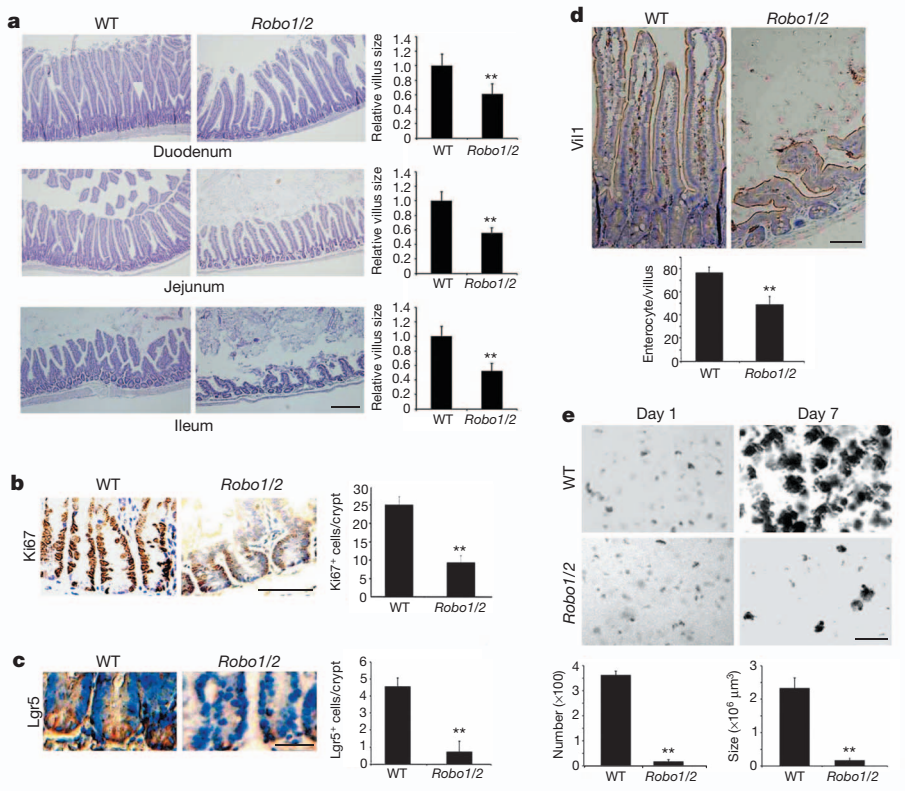


Figure 2 | Phenotypic aberration in *Robo1/2* small intestine. **a**, Morphology of small intestines. Paraffin-embedded intestinal sections from wild-type (WT) littermates and *Robo1/2* mutants were stained with haematoxylin & eosin (H&E). Relative villus sizes were measured and statistically analysed. **b–d**, Effects of partial *Robo1/2* deficiency on intestinal cells. Intestinal sections from wild-type littermates and *Robo1/2* mutants were immunohistochemically stained for Ki67$^+$ transient amplifying cells (**b**), Lgr5$^+$ ISCs (**c**) and Vil1$^+$ enterocytes (**d**). The numbers of positive cells were counted in each crypt (**b–d**). **e**, Formation of intestinal organoids. The intestinal crypts isolated from wild-type littermates and *Robo1/2* mutants were cultured and the numbers and sizes of intestinal organoids were measured at day 7. Results represent 50 tissue specimens in each group ($n = 5$) and the mean ± s.d. values. Scale bars, 200 μm (**a**, **e**) and 50 μm (**b–d**). $^{**}P < 0.01$.

called mini-guts or enteroids[5–8], as compared to their wild-type littermates (Fig. 2e, and Supplementary Fig. 8a for enlarged images). The expression of *Lgr5*, *CD133* (also known as *Prom1*), *Sox9*, *Bmi1* and *mTERT* (also known as *Tert*^{tm1Fish}) mRNAs as ISC markers[19] was significantly reduced in the *Robo1/2* mutant intestines relative to their wild-type counterparts (Supplementary Fig. 9a). As Slit–Robo signalling re-models neovasculature[3] and induces epithelial–mesenchymal transition[4], we examined and found no aberrant localization and distribution of CD31 (also known as Pecam1)$^+$ vascular endothelial cells in the vasculatures, including those distributing alongside the transient amplifying cell compartment and the capillary bed in the villi (Supplementary Fig. 10a, d), or of α-smooth muscle actin$^+$ intestinal smooth muscle cells (Supplementary Fig. 11a) in the intestines of *Robo1/2* mutants as compared to their wild-type littermates. These data collectively indicate that the Slit2–Robo1 interaction induces ISCs during physiological maintenance of intestinal homeostasis.

To complement our genetic findings, we used R5, a monoclonal antibody that binds to Robo1, but not Robo2, Robo3 or Robo4, and neutralizes Slit2 binding to Robo1 (refs 3, 4). Wild-type mice were treated daily for 6 days with an intraperitoneal injection of isotype-matched irrelevant mouse IgG or R5. As expected, R5, but not IgG, reduced intestinal villus size (Supplementary Fig. 12a) and number (Supplementary Fig. 3b), decreased the numbers of Ki67$^+$ transient amplifying cells (Supplementary Fig. 12b), Lgr5$^+$ ISCs (Supplementary Fig. 12c), Vil1$^+$ enterocytes (Supplementary Fig. 12d) and BrdU$^+$ intestinal cells (Supplementary Fig. 4b), without concomitantly affecting the distributions of vascular endothelial cells (Supplementary Fig. 10b, e) and smooth muscle cells (Supplementary Fig. 11b). Compared to IgG, R5 potently inhibited the *in vitro* formation of intestinal organoids[5–8] isolated from wild-type mice (Supplementary Figs 12e, and 8b for enlarged images), and the expression of *Lgr5*, *CD133*, *Sox9*, *Bmi1* and *mTERT* mRNAs[19] (Supplementary Fig. 9b).

We next asked whether ectopically expressed *Slit2* (Robo1 ligand) might augment ISCs and their daughter cells, leading to villus hypertrophy. We therefore examined the intestinal phenotypic changes in *Slit2* transgenic (*Slit2*-Tg) mice driven by the pCMV promoter for efficient but non-selective expression of human *Slit2* transgene[20]. Compared to wild-type mice, *Slit2*-Tg mice displayed an increased mosaic expression of Slit2 protein, but not Robo1 or α–tubulin, in the small intestines (Supplementary Fig. 2b). When compared to wild-type mice, *Slit2*-Tg mice had noticeably thicker, longer, enlarged and outnumbered villi (Fig. 3a and Supplementary Fig. 3c), with increased numbers of Ki67$^+$ transient amplifying cells (Fig. 3b), Lgr5$^+$ ISCs (Fig. 3c), BrdU$^+$ intestinal cells (Supplementary Fig. 4c) and Vil1$^+$ enterocytes (Fig. 3d), without concomitantly affecting the distribution of CD31$^+$ vascular endothelial cells (Supplementary Fig. 10c, f) or α-smooth muscle actin$^+$ smooth muscle cells (Supplementary Fig. 11c). The compartment of Ki67$^+$ transient amplifying cells appeared to be enlarged, which could be caused by accelerated proliferation (Supplementary Fig. 4c) and/or aberrant directional migration of transient amplifying cells triggered by Slit–Robo signalling[3,4]. Notably, the intestinal crypts isolated from *Slit2*–Tg mice formed more and larger intestinal organoids[5–8] relative to their wild-type counterparts (Fig. 3e, and Supplementary Fig. 8c for enlarged images). Compared to wild-type mice, the expression of *Lgr5*, *CD133*, *Sox9*, *Bmi1* and *mTERT* mRNAs[19] (Supplementary Fig. 9c) was significantly augmented in the small intestine of *Slit2*-Tg mice. Taken together, our results indicate the functional significance of Slit–Robo signalling in intestinal regeneration through modulation of ISCs.

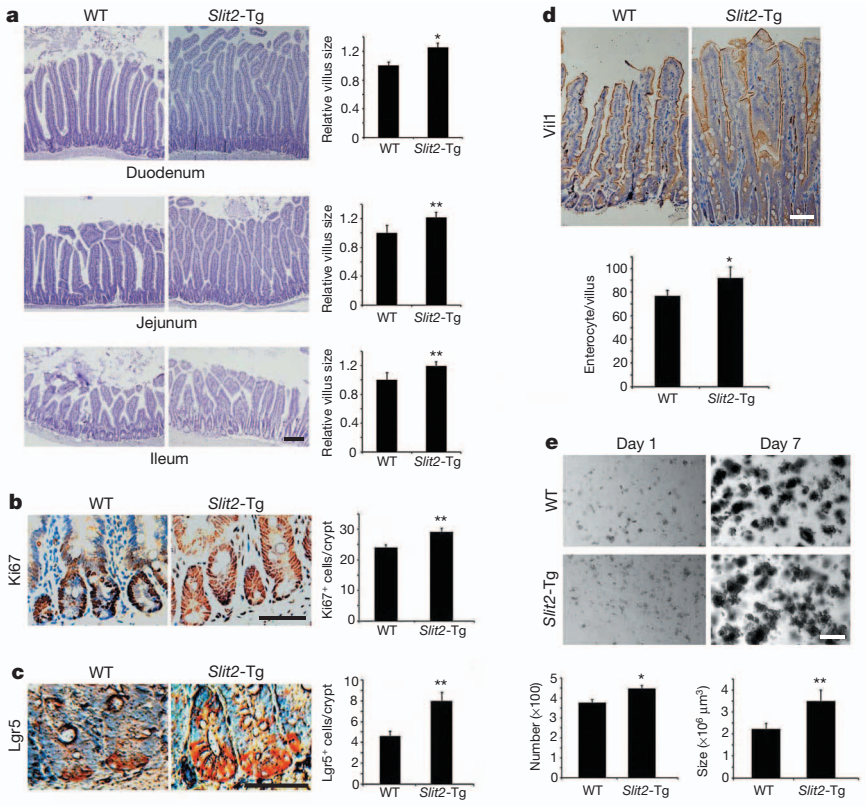


Figure 3 | *Slit2* transgene enhances small intestinal regeneration. **a**, Effects of *Slit2* overexpression on intestinal morphology. Small intestines from wild-type and *Slit2*-Tg mice were stained with H&E, and the relative villus sizes were measured and statistically analysed. **b**–**d**, Ectopic Slit2 expression affects the number and distribution of intestinal cells. Small intestines from wild-type and *Slit2*-Tg mice were immunohistochemically stained for Ki67$^+$ transient amplifying cells (**b**), Lgr5$^+$ ISCs (**c**) and Vil1$^+$ enterocytes (**d**). The numbers of positive cells were counted in each crypt. **e**, The intestinal crypts isolated from wild-type and *Slit2*-Tg mice were *in vitro* cultured and the numbers and sizes of intestinal organoids were measured at day 7. Results represent 50 tissue specimens in each group (five mice per group) and the mean ± s.d. of ten tissue sections per mouse. Scale bars, 100 μm (**a**, **e**) and 50 μm (**b**–**d**). $^*P < 0.05$; $^{**}P < 0.01$.

We further tested whether recombinant Slit2 (rSlit2 (ref. 4)) could substitute and potentiate recombinant Rspo1 (rRspo1; Supplementary Fig. 13a) for inducing intestinal organoids *in vitro*. As predicted, rSlit2 acted synergistically with rRspo1 to promote *in vitro* formation and growth of intestinal organoids in terms of their number and size (Supplementary Figs 14a, and 8d for enlarged images). In the absence of rRspo1 (rRspo1−), rSlit2 alone—at 0.5 or 1 μg ml^{-1}—was capable of inducing intestinal organoids. The intestinal crypts isolated from *Slit2*-Tg mice also formed intestinal organoids without added rRspo1 or rSlit2. Using single cell sorting gated for GFP[5], we isolated GFP^{high} ISCs from Lgr5–GFP mice[15] and found that rSlit2 by itself induced intestinal organoids (Supplementary Figs 14b, and 8e for enlarged images). Importantly, rSlit2 acted cooperatively with rRspo1 in the formation of intestinal organoids *in vitro*. To further demonstrate whether Slit2 could induce the self-renewal and pluripotential capacities of ISCs, we performed serial passages (up to four passages; every 2 weeks for each passage) of cultured intestinal organoids and found that rSlit2 not only substituted for rRspo1, but also functioned synergistically with rRspo1, using both the isolated wild-type intestinal crypts (Supplementary Fig. 15a) and the sorted GFP^{high} ISCs (Supplementary Fig. 15b).

Considering that Lgr5 is a specific ISC marker[15] and a direct targeting gene of canonical Wnt signalling[2,5–8], we treated Lgr5–GFP mice with mIgG and R5 and found that, compared to untreated (normal) or IgG-treated Lgr5–GFP mice, R5 inhibited the number of GFP^{high} ISCs (~50% inhibition; Supplementary Fig. 16a). Compared to untreated Lgr5–GFP mice, treatment with rRspo1 plus rSlit2 increased the number of GFP^{high} ISCs (~30% increase; Supplementary Fig. 16b). Treatment with rRspo1 or rSlit2 alone also slightly augmented the number of GFP^{high} ISCs. Our results thus provide *in vivo* evidence for the biological importance of our newly discovered Slit2 and Rspo1 cooperation in the induction of ISCs and the activation of Wnt/β-catenin signalling.

We also tested the functional importance of Robo1 in Rspo1-induced intestinal repair *in vivo*. Consistent with previous reports[9–14], intravenous administration of rRspo1 potently promoted villus growth in the wild-type jejunum (Supplementary Fig. 17a). Surprisingly, rRspo1 (0.1 mg per mouse per day for 5 days) failed to accelerate growth of intestinal epithelial cells in *Robo1/2* mutants. Compared to their rRspo1-treated wild-type counterparts, rRspo1 also failed to significantly augment the numbers of Lgr5$^+$ ISCs (Supplementary Fig. 17b) and Ki67$^+$ transient amplifying cells (Supplementary Fig. 17c) in the jejunum crypts of *Robo1/2* mutants. A prolonged pulse of rRspo1 (0.1 mg per mouse per day for 10 days) robustly enhanced crypt size (Supplementary Fig. 18a) and augmented the numbers of Ki67$^+$ transient amplifying cells (Supplementary Fig. 18b), BrdU$^+$ intestinal cells (Supplementary Fig. 18c) and Lgr5$^+$ ISCs (Supplementary Fig. 18d) in wild-type littermates, but not in *Robo1/2* mutants. As predicted, treatment of wild-type littermates with rRspo1 markedly induced the cytoplasmic and nuclear translocation of β-catenin (Supplementary Fig. 19a) and increased the expression of c-Myc (Supplementary Fig. 19b) compared to *Robo1/2* mutants.

Stimulation of Wnt/β-catenin signalling with Rspo1 can ameliorate 5-fluorouracil (5-FU) and radiation-induced gut damage, including radiation-induced gastrointestinal syndrome (RIGS)[9–13]. Relatedly, we found that the therapeutic dosage of 5-FU[9], a well-characterized chemotherapy medicine, markedly shortened the villus length (Supplementary Fig. 20a) and reduced the numbers of Lgr5$^+$ ISCs (Supplementary Fig. 20b) and Ki67$^+$ transient amplifying cells (Supplementary Fig. 20c) in the wild-type jejunum, but not in the *Slit2*-Tg jejunum. These

findings demonstrate the functional significance of Slit2 for promoting chemotherapy-induced intestinal repair.

Considering our findings that Slit2 potentiated Rspo1-induced ISCs and synergized with Rspo1 for reducing chemotherapy-induced gut injury, we asked whether Slit2, in combination with Rspo1, could induce ISCs and consequently prolong overall survival in mice receiving lethal challenges of chemoradiation. Surprisingly, in mice with the *Slit2* transgene, there was a 70% survival rate in mice receiving a lethal dose of 5-FU[21], whereas this same dosage caused the death of all wild-type mice within 2 weeks (Fig. 4a). The lethal dose of 5-FU abolished >90% of GFP^{high} ISCs (Supplementary Fig. 21a). However, a 3-day treatment of rSlit2 or rRspo1 alone protected ~40% of GFP^{high} ISCs. A notable finding was that the same regimen of rRspo1 plus rSlit2 preserved ~80% of GFP^{high} ISCs. Our findings thus indicate that Slit2 acts synergistically with Rspo1 for cooperative induction of ISCs, leading to prolongation of overall survival in response to the lethal dose of chemotherapy.

Because aberrant Wnt signalling causally contributes to colorectal carcinogenesis[2], because Slit2 and Robo1 expression inversely correlates with the overall survival rate in colorectal cancer[4], and because recurrent gene fusion of Rspo2 and Rspo3 contributes to colorectal cancer[22], we explored whether a short 3-day pulse of rSlit2 and/or rRspo1 could accelerate the development of intestinal carcinogenesis and/or decrease its sensitivity to chemoradiotherapy. To test this, $Apc^{Min/+}$ mice with spontaneous intestinal adenomas were treated with dextran sulphate sodium (DSS) to induce inflammation-related intestinal carcinogenesis, a murine model that is thought to closely mimic multifactorial human colorectal cancer[23]. Treatment of DSS-treated $Apc^{Min/+}$ mice with rSlit2 or rRspo1 alone led to a 20–30% survival rate upon the lethal dosage of 5-FU (Fig. 4b). Importantly, a combination of rSlit2 and rRspo1 led to a 60% survival rate ($P = 0.003$ between the rRspo1 plus rSlit2 group compared to either group alone), demonstrating the functional cooperation between Slit2 and Rspo1 leading to increased host tolerance to the lethal dose of chemotherapy in this murine model of carcinogenesis.

Consistent with previous reports using recombinant flagellin in xenografted mouse sarcoma and malignant melanoma[24] and adenoviral Rspo1 in xenografted human colorectal carcinoma[12], neither acceleration of intestinal cancer development nor desensitization to 5-FU chemotherapy was detected with our rRspo1 and rSlit2 combination treatment (Fig. 4c). By contrast, upon administration of rRspo1 and rSlit2, a single 'lethal' dose of 5-FU markedly eliminated the number of intestinal tumours, suggesting that Rspo1 plus Slit2 treatment may act as adjuvants before, during or after intensive chemotherapy for cancer eradication.

Furthermore, the combination of rSlit2 plus rRspo1 led to a 30% survival rate in mice receiving a lethal dose of whole body irradiation (WBI; Fig. 4d) and a 50% survival rate in mice receiving a lethal dose of abdominal irradiation (AIR; Fig. 4h). We also observed concomitant prolongations of the villus length (Fig. 4e, i), augmentations of $Lgr5^+$ ISCs (Fig. 4f, j) and $Ki67^+$ transient amplifying cells (Fig. 4g, k) in the jejunum, which are in analogous to previous finding that $Lgr5^+$ ISCs in the small intestine are resistant to radiation[25]. Mechanistically, although the lethal dose of WBI abolished >80% of GFP^{high} ISCs (Supplementary Fig. 21b), a 3-day treatment of rRspo1 or rSlit2 alone protected ~40%

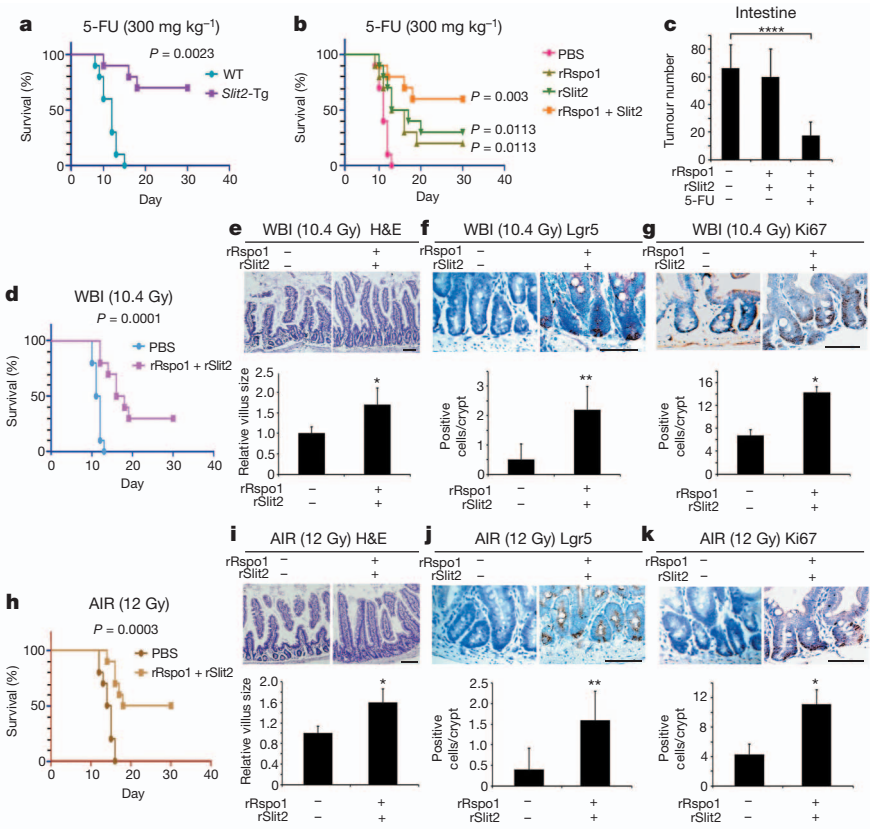


Figure 4 | Slit2 and Rspo1 cooperatively induce ISCs, reduce intestinal damage and prolong overall survival. **a**, *Slit2* transgene increases resistance to 5-FU. A single lethal dose of 5-FU was given to wild-type and *Slit2*-Tg mice and the death rates were recorded. **b**, **c**, Slit2 plus Rspo1 prolongs overall survival in tumour-bearing mice. rSlit2 or rRspo1 alone or in combination were given intravenously to DSS-treated $Apc^{Min/+}$ mice for 3 days. On day 2, a single lethal dose of 5-FU was given and the death rates were recorded (**b**). The tumour numbers in the small intestines of surviving mice were also counted at day 30 (**c**). **d**–**k**, Slit2 plus Rspo1 decreases radiation-induced death. Wild-type mice were treated with rSlit2 plus rRspo1. On day 2, they were irradiated and the death rates were recorded (**d**, **h**). The villus sizes (**e**, **i**) and the numbers of $Lgr5^+$ ISCs (**f**, **j**) and $Ki67^+$ transient amplifying (**g**, **k**) cells in the jejunum were also determined. Results are derived from 10 or 12 mice per group (**a**–**d**, **h**) or the mean ± s.d. of 10 tissue sections/mouse (**e**–**g**, **i**–**k**). Kaplan–Meier survival curves were constructed and analysed by a log-rank test (**a**, **b**, **d**, **h**). All scale bars, 50 μm. $*P < 0.05$; $**P < 0.01$; $****P < 0.0001$.

of GFP^{high}-ISCs, whereas rRspo1 plus rSlit2 preserved >80% of GFP^{high} ISCs in Lgr5–GFP mice.

In summary, our study indicates that Slit2 and Rspo1 cooperatively induce ISCs for intestinal homeostasis and repair and significantly prolong overall survival following lethal doses of chemoradiotherapy. To the best of our knowledge, the combined Slit2/Rspo1 treatment, used as an adjuvant approach, is the first example to demonstrate the feasibility of inducing endogenous adult tissue-specific stem cells for organ and tissue repair with far-reaching medical impact.

METHODS SUMMARY

Linearized DIG or biotin-labelled antisense or sense RNA probes for murine *Slit2*, *Robo1* and *Robo2* were employed for FISH. Various strains of genetic engineering mice were either untreated or treated with mouse IgG versus R5, DSS, 5-FU versus irradiation, rSlit2 and/or rRspo1, or BrdU. The tissue specimens of mouse small intestines were collected and examined by various histological, immunohistological and immunofluorescent staining methods. The isolated crypts of mouse small intestines were cultured *in vitro*. The baculovirus expression system was used to express rSlit2, rRspo1 and rNog in Sf9 insect cells, which were affinity purified to homogeneity from their corresponding serum-free media.

Full Methods and any associated references are available in the online version of the paper.

Received 13 February; accepted 24 June 2013.
Published online 31 July 2013.

1. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144,** 646–674 (2011).
2. Clevers, H. & Nusse, R. Wnt/β-catenin signaling and disease. *Cell* **149,** 1192–1205 (2012).
3. Wang, B. *et al.* Induction of tumor angiogenesis by Slit-Robo signaling and inhibition of cancer growth by blocking Robo activity. *Cancer Cell* **4,** 19–29 (2003).
4. Zhou, W. J. *et al.* Slit-Robo signaling induces malignant transformation through Hakai-mediated E-cadherin degradation during colorectal epithelial cell carcinogenesis. *Cell Res.* **21,** 609–626 (2011).
5. Sato, T. *et al.* Single Lgr5 stem cells build crypt–villus structures *in vitro* without a mesenchymal niche. *Nature* **459,** 262–265 (2009).
6. Ootani, A. *et al.* Sustained *in vitro* intestinal epithelial culture within a Wnt-dependent stem cell niche. *Nature Med.* **15,** 701–706 (2009).
7. Jung, P. *et al.* Isolation and *in vitro* expansion of human colonic stem cells. *Nature Med.* **17,** 1225–1227 (2011).
8. Spence, J. R. *et al.* Directed differentiation of human pluripotent stem cells into intestinal tissue *in vitro*. *Nature* **470,** 105–109 (2011).
9. Kim, K. A. *et al.* Mitogenic influence of human R-spondin1 on the intestinal epithelium. *Science* **309,** 1256–1259 (2005).
10. Zhao, J. *et al.* R-Spondin1, a novel intestinotrophic mitogen, ameliorates experimental colitis in mice. *Gastroenterology* **132,** 1331–1343 (2007).
11. Zhao, J. *et al.* R-Spondin1 protects mice from chemotherapy or radiation-induced oral mucositis through the canonical Wnt/β-catenin pathway. *Proc. Natl Acad. Sci. USA* **106,** 2331–2336 (2009).
12. Bhanja, P. *et al.* Protective role of R-spondin1, an intestinal stem cell growth factor, against radiation-induced gastrointestinal syndrome in mice. *PLoS ONE* **4,** e8014 (2009).
13. Takashima, S. *et al.* The Wnt agonist R-spondin1 regulates systemic graft-versus-host disease by protecting intestinal stem cells. *J. Exp. Med.* **208,** 285–294 (2011).
14. de Lau, W. *et al.* Lgr5 homologues associate with Wnt receptors and mediate R-spondin signalling. *Nature* **476,** 293–297 (2011).
15. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449,** 1003–1007 (2007).
16. Solanas, G. & Batlle, E. Control of cell adhesion and compartmentalization in the intestinal epithelium. *Exp. Cell Res.* **317,** 2695–2701 (2011).
17. Grieshammer, U. *et al.* SLIT2-mediated ROBO2 signaling restricts kidney induction to a single site. *Dev. Cell* **6,** 709–717 (2004).
18. Long, H. *et al.* Conserved roles for Slit and Robo proteins in midline commissural axon guidance. *Neuron* **42,** 213–223 (2004).
19. Yan, K. S. *et al.* The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations. *Proc. Natl Acad. Sci. USA* **109,** 466–471 (2012).
20. Ye, B. Q., Geng, Z. H., Ma, L. & Geng, J.-G. Slit2 regulates attractive eosinophil and repulsive neutrophil chemotaxis through differential srGAP1 expression during lung inflammation. *J. Immunol.* **185,** 6294–6305 (2010).
21. Martin, D. S. *et al.* High-dose 5-fluorouracil with delayed uridine ''rescue'' in mice. *Cancer Res.* **42,** 3964–3970 (1982).
22. Seshagiri, S. *et al.* Recurrent R-spondin fusions in colon cancer. *Nature* **488,** 660–664 (2012).
23. Phutthaphadoong, S. *et al.* Chemopreventive effect of fermented brown rice and rice bran (FBRA) on the inflammation-related colorectal carcinogenesis in $Apc^{Min/+}$ mice. *Oncol. Rep.* **23,** 53–59 (2010).
24. Burdelya, L. G. *et al.* An agonist of toll-like receptor 5 has radioprotective activity in mouse and primate models. *Science* **320,** 226–230 (2008).
25. Hua, G. *et al.* Crypt base columnar stem cells in small intestines of mice are radioresistant. *Gastroenterology* **143,** 1266–1276 (2012).

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

Acknowledgements We thank L. Ma for generating recombinant proteins, M. H. Geng for technical assistance, X. Yang for helping with the *in situ* hybridization and J. Connett for critical reading of the manuscript. This work was supported by a grant from the National Institutes of Health (CA126897 to J.-G.G.).

Author Contributions W.-J.Z. and Z.H.G. carried out all experiments, collected and analysed data; J.R.S. contributed *in vitro* intestinal crypt culture; and J.-G.G. proposed the hypothesis, designed experiments and wrote the manuscript.

Author Information 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 J.-G.G. (jgeng@umich.edu).

METHODS

Mouse experiments and histology. C57BL6/J (wild type; stock no. 005304), Lgr5-eGFP-IRES-CreERT2 (Lgr5–GFP; stock no. 008875)[15] and $Apc^{Min/+}$ (stock no. 002020) mice were purchased from Jackson Laboratory. $Robo1^{+/-}$ $Robo2^{+/-}$ mice[17,18] were purchased from MMRRC/University of Missouri and their wild-type littermates were used as controls. During the process of expanding our cohort of $Robo1^{+/-}$ $Robo2^{+/-}$ mice, we unexpectedly obtained several adult single *Robo1* ($Robo1^{+/-}$; less than 3% of total mouse population) or *Robo2* ($Robo2^{+/-}$) heterozygotes that were not embryonically lethal. Notably, we have never obtained any viable $Robo1^{-/-}$ homozygotes after multiple trials. *Slit2*-Tg mice were generated and characterized as described[20]. Male and female mice were equally divided without randomization. They were used at 8 weeks old unless specifically indicated. For measurement of cell proliferation, mice were injected with BrdU (1 mg per 100 g of body weight) for 2 h before euthanization. For antibody treatment, wild-type mice were injected intraperitoneally with isotype-matched irrelevant mouse IgG or R5 monoclonal antobody[3,4] (1 mg per mouse per day for 6 consecutive days). For induction of inflammation-related intestinal carcinogenesis, $Apc^{Min/+}$ mice were treated with 2% DSS in drinking water for 1 week and were used experimentally 2 weeks later[23]. For 'chemoradiation' experiments, wild-type and DSS-treated $Apc^{Min/+}$ mice (10 weeks old) were injected intravenously with Slit2 or Rspo1 alone or in combination (0.1 mg per mouse per day for three consecutive days). On day 2, they were given 5-FU intraperitoneally (30 mg per kg per mouse per day for 5 days for the therapeutic dose[9] or 300 mg kg^{-1} once for the lethal high dose[21]) or were irradiated (10.4 Gy per mouse once for WBI or 12 Gy per mouse once for AIR[12]). Alternatively, *Slit2*-Tg mice[20] (10 weeks old) were given the lethal high dose of 5-FU intraperitoneally (300 mg kg^{-1} once[21]). Tissue sections of small intestines were stained with H&E and the relative villus sizes and numbers at the duodenum, jejunum and ileum were double-blindly measured. The length ratios of 1:3:2 of the entire small intestine were defined as the duodenum, jejunum and ileum[26], respectively. Serum lipopolysaccharide (LPS) levels were measured using the limulus amebocyte lysate assay (Lonza). Mouse experiments were approved by the University Committee on Use and Care of Animals (UCUCA) of the University of Michigan.

Fluorescent *in situ* hybridization. cDNA segments of mouse *Slit2* (base pairs 3,960–4,566), mouse *Robo1* (base pairs 3,443–4,956) and mouse *Robo2* (base pairs 2,328–4,527) were reverse transcribed and labelled with DIG or biotin according to the manufacturer's instructions (Roche). Intestinal tissues were fixed with 4% paraformaldehyde and tissue sections (5-μm thick) were incubated with 1 μg ml^{-1} linearized DIG- or biotin-labelled antisense or sense RNA probe[27]. For immunofluorescent staining, Western Blocking Reagent (11921673001, Roche) was diluted 1 to 10 with 20 mM Tris-HCl, pH 7.4, containing 0.1% Tween-20 (TBST) for blocking non-specific staining and for diluting all antibodies, including 1:1,000 dilution of sheep anti-DIG antibody (11333089001, Roche) or mouse anti-biotin monoclonal (200-002-211, Jackson ImmunoResearch), and 1:200 dilution of horseradish peroxidase (HRP)-conjugated rabbit anti-sheep antibody (313-035-047, Jackson ImmunoResearch). Tyramide Signal Amplification Kits (Alexa Fluor 488 and 555; Invitrogen) were used according to the manufacturer's protocol.

Immunofluorescent and immunohistochemical staining. Intestinal tissues were fixed with 4% paraformaldehyde, sectioned (5-μm thick) and permeabilized with 0.05% Triton X-100 in PBS, pH 7.4. Samples were blocked with 1% BSA (Sigma) and incubated with primary antibody at 37 °C for 1 h, including antibodies against Ki67 (ab15580, Abcam), Lgr5 (ab75850, Abcam), lysozyme (ab36362, Abcam), CD31 (ab28364, Abcam), α-smooth muscle actin (ab5694, Abcam), Vil1 (610358, BD Biosciences), c-Myc (sc-40, Santa Cruz Biotechnology), GFP (NB600-303, Novus Biologicals) and BrdU (B2531, Sigma Aldrich). After washing extensively, samples were incubated with appropriate fluorescent dye- or HRP-conjugated secondary antibody at 37 °C for 1 h. Sections were counterstained with DAPI or H&E. Slides were then washed and mounted for observation under a scanning confocal microscope (Leica TCS SP2) or a fluorescence stereomicroscope (Leica M205 FA). The immunohistochemical staining data were quantified double blindly using ImageTool Software.

***In vitro* culture of intestinal crypts.** The intestinal crypts of mouse small intestines were isolated[28] and cultured *in vitro*[5–8]. Specifically, mouse ileums (~6 cm) were dissected out of the animals, flushed with ice-cold sterile PBS, cut open lengthwise and into 1-cm pieces, and transferred into the Petri dish with ice-cold sterile PBS supplemented with 1% penicillin/streptomycin (Invitrogen). This step was repeated twice. They were then transferred to a 50-ml conical tube containing 30 ml of ice-cold sterile shaking buffer (PBS supplemented with 15 mM EDTA and 1% penicillin/streptomycin) and rocked in the cold room for 30 min, then vortexed at maximal speed for a total of 2 min in 30-s intervals. The shaking buffers containing separated crypts were filtered through a 70-μm filter into a new 50-ml conical tube, counted, transferred to a round-bottom tube (500 crypts per tube), and centrifuged at 200*g* at 4 °C for 10 min. After gently removing the supernatant, the crypt pellets were re-suspended in 20 μl of the gut media (Advanced DMEM/F12, 1% L-glutamine, 1% penicillin/streptomycin, 10 μM HEPES (all from Invitrogen), 1% N2 supplement (R&D Systems), 2% B27 supplement (Invitrogen), 0.5 μg ml^{-1} rRspo1 (Supplementary Fig. 13a), 0.1 μg ml^{-1} rNog (Supplementary Fig. 13b), 0.1 μg ml^{-1} EGF (R&D Systems), 0.1 μg ml^{-1} rWnt3a (R&D Systems) or 0.5 or 1 μg ml^{-1} rSlit2 (ref. 4)). These crypts were mixed with ice-cold 50 μl Matrigel (BD Biosciences) and plated onto one well of the 24-well cell culture plate, using a pre-chilled pipette tip. The plate was placed in the 37 °C incubator for 30 min and an aliquot of 0.5 ml gut media was overlaid. The media was replaced every 2–4 days.

Isolation and cell sorting of GFP-positive ISCs expressing Lgr5–GFP was performed according to the previous publication[10]. Specifically, the small intestinal crypts of Lgr5–GFP mice were isolated and dissociated with TrypLE express (Invitrogen) and 2,000 U ml^{-1} DNase (Sigma Aldrich) at 37 °C for 30 min. Dissociated cells were passed through a 20-μm cell strainer and washed with sterile PBS. GFP^{high} ISCs were isolated using Moflo cell sorter. An aliquot of 10,000 sorted GFP^{high} ISCs were suspended in 20 μl gut media, mixed well with 50 μl Matrigel, and transferred to one well of 24-well plate. The plate was placed in the 37 °C incubator for 30 min and an aliquot of 0.5 ml gut media supplemented with Y-27632 (10 μM, to prevent anoikis; Sigma Aldrich) was overlaid. The efficiency of organoids generation from isolated GFP^{high} ISCs was ~3%, which was similar to the previous reports[5,29]. Indicated growth factors were added every other day and the entire medium was changed every 4 days. For serial passage, the cultured intestinal organoids were removed from the Matrigel culture and mechanically dissociated into the single crypt domains. After mixing with gut media and Matrigel, they were transferred to the new well and an aliquot of 0.5 ml gut media was overlaid. This procedure was repeated during a 2-week period, without loss of replicating efficiency.

Isolation and quantification of GFP^+ ISCs. The small intestinal crypts of Lgr5–GFP mice were isolated and dissociated with TrypLE express (Invitrogen) and 2,000 U ml^{-1} DNase (Sigma Aldrich) at 37 °C for 30 min. Dissociated cells were passed through a 20-μm cell strainer and washed with sterile PBS. GFP^{high} ISCs were stained with the anti-GFP antibody (NB600-303, Novus Biologicals), followed by an appropriate FITC-conjugated secondary antibody for quantification by flow cytometry.

Quantitative (q)RT–PCR. All mouse qRT–PCR primer pairs were purchased from SA Biosciences. The crypts of mouse small intestine were isolated[28] for qRT–PCR as previously described[4]. Notably, we routinely isolated smaller amounts of total mRNAs from the atrophied *Robo1/2* and R5-treated intestines and higher amounts of total mRNAs from the hypertrophied *Slit2*-Tg intestines than the wild-type intestines (data not shown). However, the adjusted amounts of total isolated mRNAs were used as the starting materials in the identical quantity for the purpose of comparison.

Immunoblotting. The isolated intestinal crypts were washed with ice-cold PBS, pH 7.4, and lysed with ice-cold radioimmunoprecipitation assay lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1 mM sodium orthovanadate, 10 mM sodium fluoride, 1 mM phenylmethylsulfonyl fluoride, 2 μg ml^{-1} aprotinin, 2 μg ml^{-1} leupeptin, 1 μg ml^{-1} pepstatin A, 15 μg ml^{-1} benzamidine, 0.5% Nonidet P-40, 0.15% BSA and 10% glycerol) at 4 °C for 1 h. Samples were centrifuged at 12,000*g* for 15 min at 4 °C. Samples were subjected to SDS–PAGE, transferred to polyvinylidene difluoride (PVDF) membranes (EMD Millipore) and detected with appropriate primary antibodies followed by HRP-conjugated goat anti-mouse or rabbit IgG. The blotting signals were detected using SuperSignal West Dura Extended Duration Substrate (Pierce). Quantitative analyses of immunoblotting signals on Fuji Films were obtained by densitometry analysis using LAS4000 Image Software. Notably, we routinely isolated smaller amounts of total proteins from the atrophied *Robo1/2* and R5-treated intestines and higher amounts of total proteins from the hypertrophied *Slit2*-Tg intestines than the wild-type intestines (data not shown). However, the adjusted amounts of total isolated proteins were used as the starting materials in the identical quantity for the purpose of comparison. The primary antibodies against pan-Slit and Robo1 (S1 and R4 (refs 3, 4)), α-tubulin (T6074, Sigma Aldrich) and β-actin (A1978, BD Biosciences) were used.

Baculovirus expression of Rspo1 and Nog. The cDNAs of human *Rspo1* and *Nog* (Open Biosystems) were amplified for construction of 6-His fusion proteins, using the forward primer 5′-TTGCGGCCGCATGCGGCTTGGGCTGTG-3′ and the reverse primer 5′-GGGAATTCGGCAGGCCCTGCAGATGTGAGTGGCC-3′ for *Rspo1*, and the forward primer 5′-TAGCGGCCGCATGGAGCGCTGCCCC-3′ and the reverse primer 5′-GGGAATTCGCACGAGCACTTGCACTCGGAATG ATGG-3′ for *Nog*. The inserts of *Rspo1* and *Nog* were digested with NotI/EcoRI. They were ligated into the pVL1392 vector (BD Pharmingen).

Recombinant *Rspo1* and *Nog* were expressed in Sf9 insect cells using the baculovirus expression system (BaculoGold; BD Pharmingen) and purified to homogeneity from the serum-free supernatant of Sf9 cells infected with their respective viral stocks (multiplicity of infection ~2×10^8 ml^{-1}) by Talon metal affinity chromatography (BD Clontech)[30]. Endotoxin levels of these isolated recombinant proteins

were <0.1 U mg^{-1} of proteins measured by limulus amoebocyte lysate (LAL) from Cape Cod.

Statistical analysis. The experimental data were statistically analysed by Student's *t*-test or Mann–Whitney test. Kaplan–Meier survival curves were constructed and analysed by a log-rank test. A *P* value of less than 0.05 or 0.01 was considered statistically significant or very significant, respectively.

26. Duan, L. P., Wang, H. H. & Wang, D. Q. Cholesterol absorption is mainly regulated by the jejunal and ileal ATP-binding cassette sterol efflux transporters Abcg5 and Abcg8 in mice. *J. Lipid Res.* **45,** 1312–1323 (2004).
27. Yang, X., Chrisman, H. & Weijer, C. J. PDGF signalling controls the migration of mesoderm cells during chick gastrulation by regulating N-cadherin expression. *Development* **135,** 3521–3530 (2008).
28. Booth, C., O'Shea, J. A. & Potten, C. S. Maintenance of functional stem cells in isolated and cultured adult intestinal epithelium. *Exp. Cell Res.* **249,** 359–366 (1999).
29. Merlos-Suárez, A. *et al.* The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* **8,** 511–524 (2011).
30. Wang, H.-B. *et al.* P-selectin primes leukocyte integrin activation during inflammation. *Nature Immunol.* **8,** 882–892 (2007).

LETTER

doi:10.1038/nature12496

Microbial colonization influences early B-lineage development in the gut lamina propria

Duane R. Wesemann[1,2,3,4], Andrew J. Portuguese[1,2,3], Robin M. Meyers[1,2,3], Michael P. Gallagher[1,2,3], Kendra Cluff-Jones[1,2,3], Jennifer M. Magee[1,2,3], Rohit A. Panchakshari[1,2,3], Scott J. Rodig[5], Thomas B. Kepler[6] & Frederick W. Alt[1,2,3]

The RAG1/RAG2 endonuclease (RAG) initiates the V(D)J recombination reaction that assembles immunoglobulin heavy (*IgH*) and light (*IgL*) chain variable region exons from germline gene segments to generate primary antibody repertoires[1]. *IgH* V(D)J assembly occurs in progenitor (pro-) B cells followed by that of *IgL* in precursor (pre-) B cells. Expression of IgH μ and IgL (Igκ or Igλ) chains generates IgM, which is expressed on immature B cells as the B-cell antigen-binding receptor (BCR). *Rag* expression can continue in immature B cells[2], allowing continued *Igκ* V(D)J recombination that replaces the initial VκJκ exon with one that generates a new specificity[3–5]. This 'receptor editing' process, which can also lead to *Igλ* V(D)J recombination and expression[3,6,7], provides a mechanism whereby antigen encounter at the *Rag*-expressing immature B-cell stage helps shape pre-immune BCR repertoires. As the major site of postnatal B-cell development, the bone marrow is the principal location of primary immunoglobulin repertoire diversification in mice. Here we report that early B-cell development also occurs within the mouse intestinal lamina propria (LP), where the associated V(D)J recombination/receptor editing processes modulate primary LP immunoglobulin repertoires. At weanling age in normally housed mice, the LP contains a population of *Rag*-expressing B-lineage cells that harbour intermediates indicative of ongoing V(D)J recombination and which contain cells with pro-B, pre-B and editing phenotypes. Consistent with LP-specific receptor editing, *Rag*-expressing LP B-lineage cells have similar V_H repertoires, but significantly different *Vκ* repertoires, compared to those of *Rag2*-expressing bone marrow counterparts. Moreover, colonization of germ-free mice leads to an increased ratio of *Igλ*-expressing versus *Igκ*-expressing B cells specifically in the LP. We conclude that B-cell development occurs in the intestinal mucosa, where it is regulated by extracellular signals from commensal microbes that influence gut immunoglobulin repertoires.

Pre-immune immunoglobulin diversification occurs within the gut or gut-associated structures in several vertebrate species, including sheep, rabbits, cattle, pigs and chicken[8–10]. In these species, cells harbouring a limited RAG-mediated V(D)J repertoire migrate to gut-associated structures to undergo further diversification through somatic mutation and/or gene conversion to generate a full pre-immune immunoglobulin repertoire[11]. These examples raise the notion that the gut environment may provide some benefit to the process of primary immunoglobulin diversification. In this regard, commensal bacteria are required for primary antibody repertoire diversification in pigs and rabbits[8,12], and may have an important role in stimulating this process in cattle, sheep and chickens shortly after birth[11]. In contrast to the above species, RAG-mediated V(D)J recombination is the major driver of pre-immune diversification in mice and humans. In this context, stimulated by our previous finding of mouse B-lineage tumours that arise in mesenteric lymph nodes from apparent receptor-editing B cells[13,14], we proposed that B-cell development might occur in the mouse gut and, thereby, allow *Rag* expression to diversify gut B cell pre-immune repertoires.

To test for RAG expression, we used our *Rag2* reporter mice, which contain a functional fusion with the green fluorescent protein gene (*Rag2-Gfp*) within the endogenous *Rag2* locus that provides *Rag2* expression functionally equivalent to that of the endogenous R*ag2* gene[15,16]. We used flow cytometry to test for RAG2–GFP in lymphocytes from mesenteric lymph nodes, small intestinal LP and intraepithelial lymphocytes of 3-week-old mice. Cells were gated on the CD19 pan-B-lineage marker, and GFP was plotted against the B220 pan-B-lineage marker. Staining with dual B-cell markers was done to optimize true GFP signal over background auto-fluorescence, which in wild-type controls was approximately 0.1% (Fig. 1a). With this method, we found essentially no *Rag2*-expressing B-lineage cells in the intraepithelial lymphocytes and mesenteric lymph nodes (Fig. 1a). However, we did find a population of *Rag2*-expressing, $CD19^+$ $B220^{low}$ cells within the LP that comprised approximately 3% of total $CD19^+$ cells (Fig. 1a). Quantitative PCR (qPCR) revealed *Rag1* and *Rag2* expression in wild-type small intestinal LP at a level of about 1–10% that of total bone marrow (BM), but little or no *Rag1* or *Rag2* expression in mesenteric lymph node or intraepithelial lymphocyte cells (Supplementary Fig. 1), confirming the flow cytometry results found with the *Rag2-Gfp* reporter mice. Large intestinal LP contained GFP^+ B-lineage cells as well, but at a lower level compared to that in the small intestinal LP (Supplementary Fig. 2). We did not find RAG2–GFP in Peyer's patch B cells (Supplementary Fig. 3) or mucosal T cells (Supplementary Fig. 4).

We examined various stages of early postnatal development to determine if levels of $RAG2^+$ B-lineage cells in the gut LP change over time. The proportion of LP RAG2–GFP^+ cells among total $CD19^+$ cells was low (<0.5%) in the first week of life; however, after that it gradually increased with levels peaking at approximately 4% at age 18–23 days before decreasing to undetectable levels by postnatal day 35 (Fig. 1b). In contrast, the $CD19^+$ B-cell population of peripheral blood contained 20–40% RAG2–GFP^+ cells during the first week of life, which then decreased over time to undetectable levels over the next 4 weeks (Fig. 1c). Similarly, RAG2–GFP^+ cell levels in the spleen appeared highest (10–15%) in the first week of life before decreasing to undetectable levels (Fig. 1d). The finding of low (<0.5%) levels of RAG2–GFP^+ cells in the gut LP in the first week of life, despite the presence of substantial proportions of RAG2–GFP^+ cells in the peripheral blood and spleen, suggests that the mechanism responsible for the later emergence of $RAG2^+$ cells in the gut may not be due to nonspecific dissemination driven by high levels of these cells in the blood. As $RAG2^+$ LP B-lineage cells do not express proteins known to promote gut lymphocyte tropism such as the α4β7 integrin or the CCR9 chemokine receptor (Supplementary Fig. 5), mechanisms underlying their appearance in the gut remain to be determined.

Sixteen days after intraperitoneal alum injection, RAG2–GFP^+ cells accumulate in the peripheral blood and spleens of adult mice owing to

[1]Program in Cellular and Molecular Medicine and Department of Medicine, Children's Hospital Boston, Boston, Massachusetts 02115, USA. [2]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Howard Hughes Medical Institute, Boston, Massachusetts 02115, USA. [4]Division of Rheumatology, Immunology and Allergy, Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [5]Department of Pathology, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA. [6]Department of Microbiology, Boston University School of Medicine, Boston, Massachusetts 02215, USA.

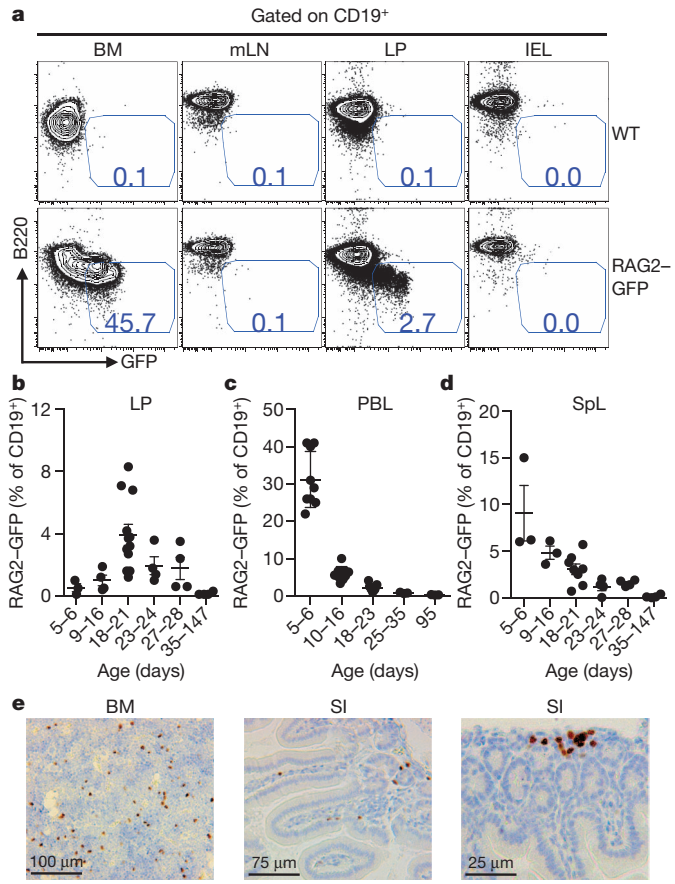


Figure 1 | Gut LP RAG2$^+$ B-lineage cells in weanling age mice. **a**, FACS plots of CD19$^+$ cells from the indicated tissues taken from wild-type (WT) (top) or homozygous *Rag2-Gfp* knock-in (bottom) mice. B220 expression is plotted against GFP fluorescence. Numbers denote percentage of CD19$^+$ B220^{low} RAG2–GFP$^+$ cells. IEL, intraepithelial lymphocyte; mLN, mesenteric lymph node. **b**–**d**, Dot plots showing percentage of RAG2–GFP$^+$ cells in LP (**b**), peripheral blood lymphocytes (PBL) (**c**) and spleen (SpL) (**d**) from indicated post-natal ages. Each point represents one mouse. Horizontal bars indicate mean values ± s.e.m. **e**, Immunohistochemistry of paraffin-embedded sections from bone marrow (BM) and small intestine (SI) stained with an anti-TdT antibody. Dark brown indicates TdT reactivity.

increased bone marrow output after initial alum-mediated bone marrow suppression[17,18]. To determine if the gut LP in adult mice maintains ability to support RAG2$^+$ B-lineage cells, we injected 4–6-month-old *Rag2-Gfp* mice with intraperitoneal alum and examined gut tissues on day 16. Following alum injection, low levels of RAG2–GFP$^+$ B-lineage cells appeared in intraepithelial lymphocytes, mesenteric lymph nodes and Peyer's patches; however, the most striking accumulation was in the LP, where RAG2–GFP$^+$ cells made up about 2.5% of total CD19$^+$ cells (Supplementary Fig. 6a, b). Appearance of RAG$^+$ B-lineage cells in the spleen and blood following alum injection is mediated by tumour necrosis factor alpha (TNF-α)[19]. To determine whether appearance of RAG$^+$ B-lineage cells in the small intestinal LP at weaning age is also TNF-α-dependent, we measured LP *Rag1* and *Rag2* expression by qPCR in 3-week-old *Tnfα* (also known as *Tnf*) knockout and wild-type control mice and found no differences (Supplementary Fig. 6c, d). Thus, the mechanism of gut LP RAG$^+$ B-lineage cell accumulation at weaning age seems distinct from that of peripheral RAG$^+$ cell accumulation that occurs after alum immunization.

RAG-expressing B-lineage cells in the BM comprise a heterogeneous population of early developmental subsets including pro-B, pre-B and immature B cells undergoing receptor editing[20]. The expression patterns of Igμ and Igκ (which accounts for ~95% of mouse IgL) can be used to distinguish these groups[5,21]. In this context, productive assembly of *Igh* V(D)J exons in pro-B cells leads to the cytoplasmic Igμ$^+$ pre-B cell stage[21]. Assembly of *Igκ* VJ exons in pre-B cells leads to formation of IgM and differentiation to the surface IgM$^+$ immature B cell stage. *Rag*-expressing cells with cytoplasmic Igκ and low or absent surface IgM have been defined as immature B cells undergoing receptor editing[2,5,17]. Staining of fixed/permeabilized CD19$^+$ B220^{low} RAG2–GFP$^+$ LP B cells for cytoplasmic Igμ and Igκ revealed similar relative levels of pro-B cells (Igμ$^-$, Igκ$^-$), pre-B cells (Igμ$^+$, Igκ$^-$) and editing B cells (Igμ$^+$, Igκ$^+$, RAG2$^+$), respectively, to those of the BM (Fig. 2a and Supplementary Fig. 7). Live CD19$^+$ B220^{low} RAG2–GFP$^+$ LP B cells that were surface IgM positive also had similarly low IgM levels to those of this putative editing B-lineage subset in BM (Fig. 2a and Supplementary Fig. 7). In addition, ligation-mediated qPCR further showed that sorted RAG2–GFP$^+$ LP B cells had a similar level of RAG-dependent DNA double-strand breaks at *Jκ* as RAG2–GFP$^+$ BM B-lineage cells (Fig. 2b), demonstrating that *Igκ* V(D)J recombination takes place in the RAG2$^+$ LP B cells at levels similar to those in RAG2$^+$ BM B-lineage cells. Finally, microarray analysis of RAG2$^+$ B-lineage cells in weanling age gut LP revealed no significant differences in general transcript expression profiles in RAG2$^+$ LP and RAG2$^+$ BM B-lineage cells (Supplementary Fig. 8), demonstrating a strong similarity between RAG2$^+$ cells in these two sites. Overall, these data indicate that the LP *Rag2*-expressing B-lineage cells contain early B-lineage developmental subsets representative of those found in the BM, supporting the occurrence of B-cell development in the LP.

We performed immunohistochemistry to confirm the LP localization of early B-lineage cells in mice at postnatal day 18–23. The terminal deoxynucleotidyl transferase (TdT) enzyme is present in pro-B cells and mediates addition of random nucleotides to *IgH* V(D)J DNA ends during V(D)J recombination to increase *IgH* V(D)J junctional diversity[22]. Staining for TdT, which is expressed similarly in BM and LP RAG2$^+$ B-lineage cells (Supplementary Fig. 8), revealed a dense nuclear expression pattern (brown) in small intestinal LP cells similar to TdT$^+$ cells in the BM (Fig. 1e). Small intestine sections were also subjected to dual staining for B220 plus TdT, which showed that the TdT$^+$ cells were also faintly B220-positive, analogous to TdT$^+$ cells in the BM (Supplementary Fig. 9). There were also cells in the small intestinal LP that stained strongly for B220 but were TdT$^-$, representing more mature B-cell populations (Supplementary Fig. 9, black arrows). TdT$^+$ cells were distributed throughout the LP, including within villi; however, they generally appeared to be more proximal to bases of villi, closer to the serosal

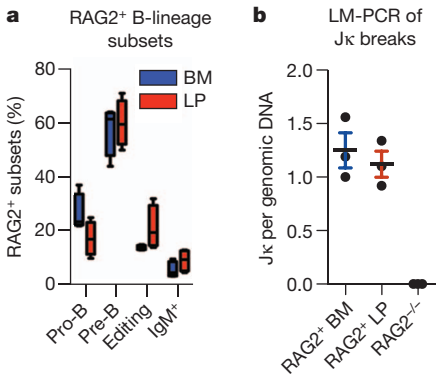


Figure 2 | RAG2–GFP$^+$ LP B-lineage developmental subsets. **a**, Plots show the relative percentage of *Rag2*-expressing pro-B cells, pre-B cells, editing B cells and surface IgM$^+$ B cells (see text for definition of each) in the bone marrow (blue bars) and lamina propria (red bars). Plotted are mean values ± s.e.m. and each are derived from experiments of 4 independent mice at postnatal day 17–24 (see Supplementary Fig. 7 for more details). **b**, Plots show quantitative ligation-mediated PCR (LM-PCR) of RAG2$^+$ BM B cells and RAG2$^+$ LP B cells normalized to genomic DNA. BM cells from *Rag2*$^{-/-}$ mice were a negative control. Values on the *y* axis are units relative to the signal obtained from RAG2$^+$ BM B-cell samples.

(antiluminal) intestinal surface compared to B220^{high} TdT^{-} B cells (Supplementary Fig. 10). These data indicate that gut-resident early B-lineage cells inhabit a generally distinct location within the LP compared to more mature B cells. Immunohistochemistry studies of human fetal intestine also identified serosally positioned pre-B cells in the intestinal LP, suggesting that similar early B-cell development may occur in the human gut[23].

Given our finding of primary B-cell development in the mouse intestinal LP, we asked whether this process contributes to differential diversification of pre-immune repertoires in developing RAG2^{+} LP B-lineage cells versus those from BM. To test this, we isolated RNA from sorted RAG2–GFP^{+} cells from the BM and LP (Supplementary Fig. 11) of 3-week-old *Rag2-Gfp* mice, and assessed productive V_H and $V\kappa$ use through 5′ rapid amplification of complementary DNA ends (5′RACE) generated from mature immunoglobulin gene transcripts, followed by 454 sequencing. To best visualize potential *V* segment usage differences between RAG2–GFP^{+} cells from LP versus BM, we plotted *V* segment use from the two sources against individual in-frame *V* segments in the order of highest to lowest usage in RAG2–GFP^{+} BM cells (Fig. 3 and Supplementary Fig. 12). These studies showed that V_H usage was very similar in RAG2–GFP^{+} LP cells and RAG2–GFP^{+} BM cells (χ^2 test, $P = 0.235$) (Fig. 3a and Supplementary Fig. 12a). As a positive control for the method, comparison of V_H usage between RAG2^{+} BM and total non-sorted splenic B cells showed an expected highly significant difference (χ^2 test, $P = 2.2 \times 10^{-16}$) (Supplementary Fig. 13a). The similar V_H repertoires of RAG2^{+} BM and LP B-lineage cells suggests that V_H use during primary IgH V(D)J recombination occurs similarly in the two locations. In contrast, we observed prominent and highly significant differences in $V\kappa$ usage in RAG2–GFP^{+} cells from LP versus BM (χ^2 test, $P = 0.00084$) (Fig. 3b and Supplementary Fig. 12b); as a negative control, we observed no significant overall differences in $V\kappa$ usage between RAG2^{+} BM samples compared with other RAG2^{+} BM samples; or RAG2^{+} LP samples compared with other RAG2^{+} LP samples from separate pools of mice (χ^2 test, $P = 0.560$ and 0.545, respectively) (Supplementary Fig. 13b, c). The finding of different $V\kappa$ repertoires within RAG2^{+} LP and BM B-lineage cells indicates that the LP versus BM location of B-cell development may influence $V\kappa$ usage in developing B cells. In this regard, the very similar V_H repertoires of LP and BM RAG2^{+} B-lineage cell indicates that a probable explanation for the marked difference in the LP $V\kappa$ repertoires from those of BM is that they are generated in the RAG2^{+} receptor-editing LP B-cell population.

Intestinal microflora have been shown to influence immune cell development in terms of lymphoid organization and T-cell subset accumulation and activity, both locally in the gut as well as systemically[24]. To assess the influence of microflora on B-cell development, we co-housed 3-week-old wild-type Swiss-Webster germ-free mice with regular specific pathogen free (SPF) mice for 7 days. Gram staining of small intestine contents was performed to confirm bacterial colonization of co-housed mice (Supplementary Fig. 14). We used qPCR to assay colonized mice and germ-free littermates for *Rag1* and *Rag2* expression, which was normalized to *Cd19* expression. In accord with ability of microflora to induce systemic effects on immune cell development[24], we observed colonization-dependent increases in *Rag1* and *Rag2* expression in the BM and spleen, as well as the gut LP (Supplementary Fig. 15). We also found pro-B cells (identified as CD19^{+} B220^{low} CD43^{+}) to represent an increased percentage of total CD19^{+} B cells in colonized mouse BM and LP, and potentially in the colonized spleen (Fig. 4a and Supplementary Fig. 16), correspondingly, the increased *Rag* expression in these tissues probably derives from the increased percentage of pro-B cells. We conclude that gut microflora induce increased levels of pro-B cells systemically, including in the gut LP, in previously un-colonized weaning-age mice.

Given our finding of different $V\kappa$ repertoires in LP versus BM RAG2^{+} B-lineage populations that might be generated via receptor editing, we examined the ratio of Igλ^{+} to Igκ^{+} B cells in the LP of colonized mice versus their germ-free mice littermates. In this regard, increased Igλ usage in the B-cell repertoire is another marker of receptor editing[3,6,7]. Notably, colonization led to a significant and reproducible increase of the ratio of Igλ^{+} to Igκ^{+} B cells in the LP but not the BM or spleen (Fig. 4b and Supplementary Fig. 17), consistent with a commensal-dependent process leading to increased editing specifically within LP B cells. However, as these analyses were not performed in *Rag2-Gfp* mice, only total B-cell populations could be analysed. Therefore, a non-mutually exclusive possibility would be selection for Igλ^{+} B cells in the gut after colonization with commensal microbes, a phenomenon not previously described. In this regard, we do find significant differences in both V_H and $V\kappa$ segment usage in the LP IgM^{+} B-cell population of colonized mice compared to germ-free littermates (Supplementary Fig. 18), indicating presence of commensals influences both IgH and IgL mature B-cell repertoires.

Consistent with growing evidence demonstrating that the microflora act as a regulators of T-lymphocyte subsets[24], we find that weanlings harbour an intestinal LP B-cell developmental process that is influenced in germ-free mice by microbial colonization. As weaning is concurrent with microbial expansion[25], occurrence of *Rag*-expressing B-lineage cell accumulation in weanlings may have evolved to allow B-cell primary repertoires to be modulated in response to colonization. In this regard, our findings also suggest that this LP B-cell developmental process includes BCR editing, which may contribute to significant differences in the primary $V\kappa$ repertoire of LP versus BM B-lineage populations. Past studies have implicated BCR editing in the BM as a negative selection process, largely based on studies of B cells engineered to make specific self-reactive, high-affinity BCRs. Given the natural repertoire in our studies, the degree to which the observed BCR-editing process in the gut represents a tolerance mechanism is unclear. However, given the potential special role of the gut in pre-immune diversification

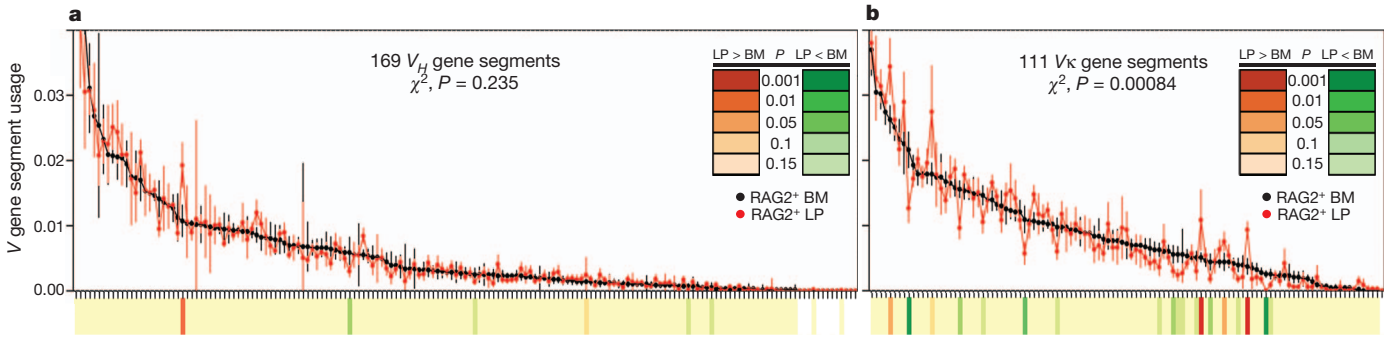


Figure 3 | Distinct $V\kappa$ segment usage in RAG2^{+} cells from BM versus LP. **a**, **b**, Dot plots show contributions (in order of highest to lowest BM use) of different V_Hs (**a**) and $V\kappa$s (**b**) to in-frame rearrangements in RAG2–GFP^{+} BM (black dots) and LP (red dots) cells. The two most highly used V_Hs are omitted to increase plot resolution. Each point shows mean ± s.e.m. of at least 4 experiments. The χ^2 calculated *P* values for overall differences between BM and LP are indicated. Significant *V* segment usage differences between BM and LP are indicated on heat map (*P* values scale indicated in inset). Full data set at increased resolution is in Supplementary Fig. 12.

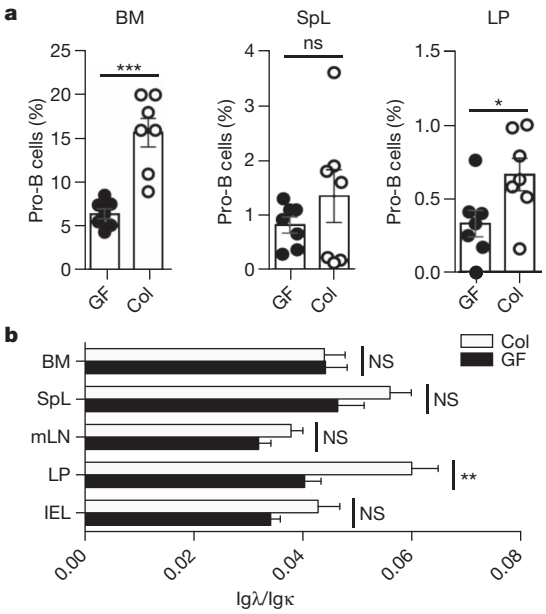


Figure 4 | Effects of gut colonization on development of LP B-lineage cells. **a**, Plots of percentage of pro-B cells versus total $CD19^+$ B-lineage cells from bone marrow (BM), spleen (SpL) and lamina propria (LP) of 4-week-old germ-free (GF) mice and littermates colonized (Col) by co-housing with specific pathogen-free mice for 7 days. **b**, Bar graphs show ratios of $Ig\lambda^+$ versus $Ig\kappa^+$ B cells within mesenteric lymph nodes (mLN), inter-epithelial lymphocyes (IEL) and tissues indicated in **a** from germ-free mice and littermates colonized by co-housing with specific pathogen-free mice. Mean values and s.e.m. are shown. $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$. NS, not significant. (Details in Supplementary Figs 16 and 17.)

in other vertebrates, RAG-dependent editing in the LP may also contribute diversification-related roles. In addition, immature gut-derived B cells also may have specialized roles, such as those suggested for immature splenic B cells[26]. Finally, primary B-cell development in the intestine, including mucosal B-cell receptor editing, might allow both luminal antigens and peripheral host mucosal components opportunities to shape the pre-immune repertoire. In this regard, the transient nature of LP B-cell development implies that there may be windows of opportunity for this influence to occur.

METHODS SUMMARY

All experiments with mice followed the protocols approved by the Animal Resources at Children's Hospital (ARCH) review committee and performed in accordance with NIH guidelines. Mice harbouring the *Rag2-Gfp* knock-in fusion gene at the endogenous locus were described previously[15] and were maintained on a 129/SvJ background. Wild-type BALB/c mice were purchased from Jackson Laboratories. Swiss-Webster germ-free mice were purchased from Taconic Farms. Peyer's patches, intraepithelial lymphocytes and LP lymphocytes were isolated essentially as described[27].

Full Methods and any associated references are available in the online version of the paper.

Received 22 February; accepted 23 July 2013.
Published online 21 August 2013.

1. Jung, D., Giallourakis, C., Mostoslavsky, R. & Alt, F. W. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* **24,** 541–570 (2006).
2. Yu, W. *et al.* Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature* **400,** 682–687 (1999).
3. Tiegs, S. L., Russell, D. M. & Nemazee, D. Receptor editing in self-reactive bone marrow B cells. *J. Exp. Med.* **177,** 1009–1020 (1993).
4. Gay, D., Saunders, T., Camper, S. & Weigert, M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *J. Exp. Med.* **177,** 999–1008 (1993).
5. Pelanda, R. *et al.* Receptor editing in a transgenic mouse model: site, efficiency, and role in B cell tolerance and antibody diversification. *Immunity* **7,** 765–775 (1997).
6. Retter, M. W. & Nemazee, D. Receptor editing occurs frequently during normal B cell development. *J. Exp. Med.* **188,** 1231–1238 (1998).
7. Hertz, M. & Nemazee, D. BCR ligation induces receptor editing in IgM^+IgD^- bone marrow B cells *in vitro*. *Immunity* **6,** 429–436 (1997).
8. Lanning, D., Zhu, X., Zhai, S. K. & Knight, K. L. Development of the antibody repertoire in rabbit: gut-associated lymphoid tissue, microbes, and selection. *Immunol. Rev.* **175,** 214–228 (2000).
9. Jenne, C. N., Kennedy, L. J. & Reynolds, J. D. Antibody repertoire development in the sheep. *Dev. Comp. Immunol.* **30,** 165–174 (2006).
10. Ratcliffe, M. J. Antibodies, immunoglobulin genes and the bursa of Fabricius in chicken B cell development. *Dev. Comp. Immunol.* **30,** 101–118 (2006).
11. Lanning, D. K., Rhee, K. J. & Knight, K. L. Intestinal bacteria and development of the B-lymphocyte repertoire. *Trends Immunol.* **26,** 419–425 (2005).
12. Butler, J. E. *et al.* Antibody repertoire development in fetal and neonatal piglets. VIII. Colonization is required for newborn piglets to make serum antibodies to T-dependent and type 2 T-independent antigens. *J. Immunol.* **169,** 6822–6830 (2002).
13. Wang, J. H. *et al.* Mechanisms promoting translocations in editing and switching peripheral B cells. *Nature* **460,** 231–236 (2009).
14. Wang, J. H. *et al.* Oncogenic transformation in the absence of Xrcc4 targets peripheral B cells that have undergone editing and switching. *J. Exp. Med.* **205,** 3079–3090 (2008).
15. Monroe, R. J. *et al.* RAG2:GFP knockin mice reveal novel aspects of RAG2 expression in primary and peripheral lymphoid tissues. *Immunity* **11,** 201–212 (1999).
16. Nagaoka, H., Yu, W. & Nussenzweig, M. C. Regulation of RAG expression in developing lymphocytes. *Curr. Opin. Immunol.* **12,** 187–190 (2000).
17. Nagaoka, H., Gonzalez-Aseguinolaza, G., Tsuji, M. & Nussenzweig, M. C. Immunization and infection change the number of recombination activating gene (RAG)-expressing B cells in the periphery by altering immature lymphocyte production. *J. Exp. Med.* **191,** 2113–2120 (2000).
18. Gärtner, F., Alt, F. W., Monroe, R. J. & Seidl, K. J. Antigen-independent appearance of recombination activating gene (Rag)-positive bone marrow B cells in the spleens of immunized mice. *J. Exp. Med.* **192,** 1745–1754 (2000).
19. Ueda, Y., Yang, K., Foster, S. J., Kondo, M. & Kelsoe, G. Inflammation controls B lymphopoiesis by regulating chemokine CXCL12 expression. *J. Exp. Med.* **199,** 47–58 (2004).
20. Jankovic, M., Casellas, R., Yannoutsos, N., Wardemann, H. & Nussenzweig, M. C. RAGs and regulation of autoantibodies. *Annu. Rev. Immunol.* **22,** 485–501 (2004).
21. Raff, M. C., Megson, M., Owen, J. J. & Cooper, M. D. Early production of intracellular IgM by B-lymphocyte precursors in mouse. *Nature* **259,** 224–226 (1976).
22. Desiderio, S. V. *et al.* Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. *Nature* **311,** 752–755 (1984).
23. Golby, S. *et al.* B cell development and proliferation of mature B cells in human fetal intestine. *J. Leukoc. Biol.* **72,** 279–284 (2002).
24. Hooper, L. V., Littman, D. R. & Macpherson, A. J. Interactions between the microbiota and the immune system. *Science* **336,** 1268–1273 (2012).
25. Mackie, R. I., Sghir, A. & Gaskins, H. R. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am. J. Clin. Nutr.* **69,** 1035S–1045S (1999).
26. Ueda, Y., Liao, D., Yang, K., Patel, A. & Kelsoe, G. T-independent activation-induced cytidine deaminase expression, class-switch recombination, and antibody production by immature/transitional 1 B cells. *J. Immunol.* **178,** 3593–3601 (2007).
27. Lefrancois, L. & Lycke, N. Isolation of mouse small intestinal intraepithelial lymphocytes, Peyer's patch, and lamina propria cells. *Curr. Protoc. Immunol.* Ch. 3, Unit 3.19 (2001).

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

Acknowledgements This work was supported by National Institutes of Health grants AI020047 (to F.W.A.) and AI89972 (to D.R.W.), Lymphoma and Leukemia SCOR 7009-12 (to F.W.A.), and National Institutes of Health research contract HHSN272201000053C (to T.B.K.). D.R.W. was also supported by an award from the American Academy of Allergy Asthma and Immunology and CSL-Behring and holds a Career Award for Medical Scientists from the Burroughs Wellcome Fund. F.W.A. is an Investigator of the Howard Hughes Medical Institute.

Author Contributions D.R.W. and F.W.A. designed the study; D.R.W., A.J.P., M.P.G., K.C.-J., J.M.M. and R.A.P. performed experiments; R.M.M. and T.B.K. performed computational analysis of sequencing data; S.J.R. performed immunohistochemistry experiments; D.R.W. and F.W.A. wrote the paper.

Author Information Microarray data have been deposited in MIAME format into the Gene Expression Omnibus (GEO) database under accession number GSE48870, and repertoire sequencing data have been deposited into the GEO database under accession number GSE48805. 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 F.W.A. (alt@enders.tch.harvard.edu) or D.R.W. (dwesemann@research.bwh.harvard.edu).

METHODS

Mice, immunizations and colonization. Mice harbouring the *Rag2-Gfp* knock-in fusion gene at the endogenous locus were described previously[15] and were maintained on a 129/SvJ background. Wild-type BALB/c mice were purchased from Jackson Laboratories. Swiss-Webster germ-free mice were purchased from Taconic Farms. For each germ-free/colonization experiment, littermate germ-free mice were used as controls. Germ-free status and colonization status were confirmed by Gram staining of caecal and small intestinal contents as well as microbial culture. Immunization experiments were performed with alum as described[18]. All experiments with mice followed the protocols approved by the Boston Animal Care Facility of the Children's Hospital.

Cell isolation and flow cytometry. Peyer's patches, intraepithelial lymphocytes and LP lymphocytes were isolated essentially as described[27]. Peyer's patches were excised from the small intestine, and the remaining tissue was incubated with 1× Hank's balanced salt solution with 1 mM EDTA 10% FBS PBS for 30 min and room temperature three times for intraepithelial lymphocyte extraction. Residual intestinal tissue was digested in 20% FBS RPMI with 0.05% collagenase from *Clostridium histolyticum* (Sigma) for 1 h at 37 °C three times. Intraepithelial lymphocytes and LP cells were centrifuged over Lympholyte (Cedar Lane) per manufacturer's recommendations to minimize mucus contamination. Single-cell suspensions of mesenteric lymph nodes, Peyer's patches and spleen were prepared by mashing through a cell strainer (70 μm). Cells were stained with fluorophore-conjugated mouse antibodies, and flow cytometry was performed.

qPCR with reverse transcription and microarray analysis. For qPCR, total RNA was extracted using the TRIzol method (Invitrogen) and reverse transcribed into cDNA using qScript (Quanta Biosciences). *Rag1* and *Rag2* transcripts were then quantified using TaqMan qPCR assays Mm01270936_m1 and Mm00501300_m1, respectively (Applied Biosystems). The comparative C_t method was used to quantify transcripts that were normalized with respect to *Cd19* expression (Taqman assay Mm00515420_m1, Applied Biosystems). For comparative transcriptome analysis, B cells were isolated from BM and small intestinal LP of 3-week-old mice. Cells were sorted (BD FACSAria) into TRIzol on the basis of the following cell surface markers: $CD19^+$, $B220^{low}$, GFP^+. RNA was extracted and then amplified, labelled, and hybridized to Affymetrix GeneChip Mouse Gene 1.0 ST arrays (Expression Analysis). Raw data were normalized with the RMA algorithm implemented in the Expression File Creator module from the GenePattern suite[28]. Data were visualized with the Multiplot and Hierarchical Clustering Viewer modules. All cell populations analysed were generated in triplicate from independent experiments consisting of a pool of at least 8 mice for each experiment.

Immunohistochemistry. Immunohistochemistry was performed using 4-mm-thick formalin-fixed paraffin-embedded (FFPE) tissue sections. Slides were soaked in xylene, passed through graded alcohols, and put in distilled water. Slides were pretreated with EDTA (pH 8.0) retrieval solution (Zymed) in a steam pressure cooker (Biocare Decloaking Chamber CD2008US, Biocare Biomedical) at manufacturer's recommended settings. All further steps are performed at room temperature in a hydrate chamber. The slides were blocked for endogenous peroxidase activity with peroxidase block (DAKO), washed 5 min in buffer, and followed by 20 min incubation with serum free protein block (DAKO). For TdT single staining, a polyclonal rabbit antibody (DAKO catalogue no. A3524) was applied at 1:100 dilution for 1 h at room temperature followed by washing. The detection of antibody used DAKO Rabbit Envision and DAB according to the manufacturer's directions. For TdT/B220 double staining, rabbit anti-TDT was followed with Mach-2 Rabbit AP polymer (Biocare) and developed with Vulcan Fast Red (Biocare). Subsequently, rat anti-B220 (BD Pharmingen, catalogue no. 550286), was applied for 1 h at 1:200 dilution followed by Goat anti-Rat-HRP (Millipore) and developed with DAB. All slides were counterstained with Harris haematoxylin. Stained slides were scanned at ×200 magnification using an Aperio ScanScope XT workstation (Aperio Technology). Images were visualized, annotated and microscopic distances quantified using ImageScope software (version 10.0.35.1800, Aperio Technology).

Ligation-mediated PCR. Sorted $CD19^+$ $B220^{low}$ RAG2–GFP^+ B-lineage cells from BM and LP were lysed in SDS lysis buffer (5 mM EDTA, 200 mM NaCl, 100 mM Tris-HCl pH 8.0, 0.2% SDS) with proteinase K (200 mg ml^{-1}) overnight at room temperature followed by incubation at 37 °C for 1 h. Following DNA isolation by phenol:chloroform separation and isopropanol precipitation, blunt-end ligation reactions were performed using an oligonucleotide duplex linker consisting of BW-1 (5′-GCGGTGACCCGGGAGATCTGAATTC-3′) and BW-2 (5′-GAATTCAGATC-3′). DNA was ligated overnight at 16 °C in ligation buffer (50 mM Tris pH 7.5, 10 mM $MgCl_2$, 10 mM DTT, 1 mM ATP) and T4 DNA ligase (Promega). Ligase was inactivated by incubation for 10 min at 70 °C. Ligation reaction was diluted 1:3 in H_2O before being used for PCR. Nested PCR was used to detect ligation products resulting from both *Jκ1* and *Jκ2* double-stranded DNA breaks. The first round of amplification was performed using Qiagen Hot Star Taq (1.25 U per reaction) and primers Ko3 (5′-AGTGCCACTAACTGCTGAGAAACCT-3′) and BW-1H (5′-CCGGGAGATCTGAATTCCAC-3′). The PCR reaction was performed as follows: 95 °C for 15 min, followed by 26 cycles of 94 °C for 45 s, 57 °C for 45 s and 72 °C for 50 s, followed by 72 °C for 5 min. The second round of amplification was performed as qPCR using primers Ko (5′-CCACGCATGCTTGGAGAGGGGGTT-3′) and BW-1H primers with an internal Jκ probe (5′-56-FAM/ZEN-3-Iowa Black-TGAGGAGGGTTTTTGTACAGCCAGA-3′). Signals were normalized to actin amplified from genomic DNA (Mm00607939_s1). Each sample was calculated as a per cent of a BM standard run on each PCR plate to determine variability within BM and LP samples from three biologic replicates isolated from independent pools of 4–8 mice per experiment.

Repertoire sequencing. Total RNA was obtained from purified B cells using TRIzol reagents (Invitrogen). κ chain and μ chain cDNAs from each sample were synthesized using a SMARTer-RACE cDNA amplification kit (Clontech), according to the manufacturer's protocol. *Cμ*-specific (5′-CAGGTGAAGGAAATGGTGCT-3′) and *Cκ*-specific (5′-TTAACTGCTCACTGGATGGTG-3′) primers were used in lieu of oligo dT primers for cDNA synthesis. A total of 0.05 to 0.2 μg total RNA per sample was used. PCR was performed using Phusion DNA Polymerase (Thermo Scientific) and 12.5 μl of first-strand reactions with long and short universal primers (5′-CTAATACGACTCACTATAGGGCAAGCAGTGGTAACAACGCAGAGT-3′ and 5′-CTAATACGACTCACTATAGGGC-3′) together with either a biotinylated round-1 *Cμ* primer (5′-BIO-CTTATCAGACAGGGGGCTCTC-3′) or round-1 *Cκ* primer (5′-BIO-TCACTGGATGGTGGGAAGAT-3′) specific primers. First round PCR reaction conditions were then followed as described elsewhere[29]. PCR round-1 product sizes of 500–700 base pairs (bp) were extracted from agarose gels using a QIAquick gel extraction kit (Qiagen), enriched on streptavidin-coupled Dynabeads (Invitrogen), and purified with a QIAquick PCR purification columns (Qiagen) per manufacturer's instructions. Purified round-1 products were then subjected to a second round of PCR with nested primers containing 'A' and 'B' adaptor sequences as well as distinct 10-bp barcode sequences (to distinguish source material) using the nested universal primer (5′-CGTATCGCCTCCCTCGCGCCATCAG[unique 10-bp barcode sequence]ACGACTCACTATAGGGCAAGCAG-3′) together with either nested *Cμ* primer (5′-CTATGCGCCTTGCCAGCCCGCTCAG[unique 10-bp barcode sequence]GGGAAGACATTTGGGAAGGA-3′) or *Cκ* primer (5′-CTATGCGCCTTGCCAGCCCGCTCAG[unique 10-bp barcode sequence]TGGATGGTGGGAAGATGGAT-3′). PCR products (size 500–700 bp) were extracted from agarose gel, and 100 ng of each amplicon library was combined and used for 454 sequencing analysis. GS FLX Titanium sequencing kit XLR70 (Roche) was used for sample preparation. Data were collected at the sequencing core and at the University of Illinois. Data were analysed using the empirical Bayes procedure as described[30]. Clonotypes and clone assignments were determined using a recursive set of hypothesis tests on the equality of the V-gene segment mutation rate and that of CDR3. To control for false detection rate, comparisons were made between the same tissues of repeat experiments of BM and LP *Vκ* samples (Supplementary Fig. 13). V_H segment usage from $RAG2^+$ BM was also compared to total splenic B cell V_H segment usage to ensure that our analysis could identify the expected differences that occur due to selection between these populations.

Statistics. If not otherwise stated, data were expressed as arithmetic means ± s.e.m., and statistical analyses were made by unpaired *t*-test, exact test, or χ^2 test where appropriate. $P < 0.05$ was considered statistically significant.

28. Reich, M. *et al.* GenePattern 2.0. *Nature Genet.* **38,** 500–501 (2006).
29. Warren, R. L. *et al.* Exhaustive T-cell repertoire sequencing of human peripheral blood samples reveals signatures of antigen selection and a directly measured repertoire size of at least 1 million clonotypes. *Genome Res.* **21,** 790–797 (2011).
30. Kepler, T. B. Reconstructing a B-cell clonal lineage. I. Statistical inference of unobserved ancestors [v1; ref status: indexed, http://f1000r.es/z6], *F1000 Res.* **2,** 103 (2013).

LETTER

doi:10.1038/nature12423

Vesicular and non-vesicular transport feed distinct glycosylation pathways in the Golgi

Giovanni D'Angelo[1,2], Takefumi Uemura[3], Chia-Chen Chuang[4], Elena Polishchuk[1], Michele Santoro[1], Henna Ohvo-Rekilä[5], Takashi Sato[3], Giuseppe Di Tullio[6], Antonio Varriale[2], Sabato D'Auria[2], Tiziana Daniele[6], Fabrizio Capuani[1], Ludger Johannes[7,8], Peter Mattjus[5], Maria Monti[9], Piero Pucci[9], Roger L. Williams[10], John E. Burke[10], Frances M. Platt[4], Akihiro Harada[3,11] & Maria Antonietta De Matteis[1]

Newly synthesized proteins and lipids are transported across the Golgi complex via different mechanisms whose respective roles are not completely clear. We previously identified a non-vesicular intra-Golgi transport pathway for glucosylceramide (GlcCer)—the common precursor of the different series of glycosphingolipids—that is operated by the cytosolic GlcCer-transfer protein FAPP2 (also known as PLEKHA8) (ref. 1). However, the molecular determinants of the FAPP2-mediated transfer of GlcCer from the *cis*-Golgi to the *trans*-Golgi network, as well as the physiological relevance of maintaining two parallel transport pathways of GlcCer—vesicular and non-vesicular—through the Golgi, remain poorly defined. Here, using mouse and cell models, we clarify the molecular mechanisms underlying the intra-Golgi vectorial transfer of GlcCer by FAPP2 and show that GlcCer is channelled by vesicular and non-vesicular transport to two topologically distinct glycosylation tracks in the Golgi cisternae and the *trans*-Golgi network, respectively. Our results indicate that the transport modality across the Golgi complex is a key determinant for the glycosylation pattern of a cargo and establish a new paradigm for the branching of the glycosphingolipid synthetic pathway.

Complex glycosphingolipids (GSLs), which have key roles in cell signalling, adhesion, proliferation and differentiation[2], are synthesized in the Golgi complex from GlcCer, which is synthesized from ceramide at the cytosolic leaflet of early Golgi membranes[3,4]. Upon translocation to the luminal leaflet, GlcCer is galactosylated to lactosylceramide (LacCer), which can then be converted into complex GSLs in later Golgi compartments (Fig. 1a)[5]. GlcCer can be transported through the Golgi complex via membrane trafficking and via non-vesicular transfer owing to the action of the cytosolic GlcCer-transfer protein FAPP2, which fosters GSL synthesis[1,6]. However, the respective roles of the vesicular and non-vesicular transport of GlcCer remain to be defined[7].

We have addressed this question by assessing the consequences of *FAPP2* gene ablation in mice (Fig. 1b–d and Supplementary Fig. 1a). *FAPP2*$^{-/-}$ mice showed no overt phenotype. However, measurement of GSL levels in the kidneys, where FAPP2 is highly expressed (Fig. 1b and Supplementary Fig. 1b), highlighted a specific decrease in the globoside series of GSL, in particular in globotriaosylceramide (Gb3), in *FAPP2*$^{-/-}$ mice (Fig. 1e). Visualization of GSLs using the Shiga toxin B fragment (ShTxB) that binds Gb3 (ref. 8), the cholera toxin B fragment (ChTxB) that binds monosialotetrahexosylganglioside (GM1) (ref. 9), and anti-monosialodihexosylganglioside (GM3) antibodies confirmed the previously reported distribution of Gb3 in the mouse kidney[10,11] and showed a selective reduction of Gb3 staining in the *FAPP2*$^{-/-}$ kidneys (Fig. 1f, g) and in kidney tubular cells isolated from *FAPP2*$^{-/-}$ mice (Supplementary Fig. 2).

Thus, FAPP2, in line with its rather recent evolutionary appearance coincident with the divergence of multiple GSL branches[12], selectively controls one GSL branch *in vivo*. We then performed the analysis of GSL in cells knocked down for *FAPP2* by short interfering RNA (siRNA) treatment (Supplementary Fig. 3). As reported previously[1], *FAPP2* knockdown induced a 40% decrease in the levels of total GSLs (taken as the sum of LacCer, GM3 and Gb3) but, considering the individual GSL species, it lowered the levels of LacCer and Gb3 but not GM3 (Fig. 2a). In agreement with these biochemical measurements, a selective decrease in ShTxB staining was also observed in *FAPP2* knockdown cells (Supplementary Fig. 4a).

When we analysed the impact of *FAPP2* knockdown on newly synthesized GSLs in ^{3}H-sphingosine-labelled HeLa cells, we found that FAPP2 depletion inhibited the synthesis of ^{3}H-LacCer and ^{3}H-Gb3 at all time points (Fig. 2b and Supplementary Fig. 4b). *FAPP2* knockdown, however, also inhibited the synthesis of ^{3}H-GlcCer at early time points, in agreement with previous results[1], and as a consequence lowered the levels of ^{3}H-GM3 (Fig. 2b and Supplementary Fig. 4b). To circumvent changes in complex GSLs that might be secondary to the inhibition of GlcCer synthesis, we bypassed GlcCer synthesis by labelling the cells with fluorescent GlcCer (C12 BODIPY-GlcCer). Under these conditions, FAPP2 depletion selectively inhibited the synthesis of C12-BODIPY-Gb3 but not of C12-BODIPY-GM3 (Fig. 2c), indicating that the decrease in Gb3 but not in GM3 synthesis was a direct consequence of FAPP2 depletion. Systematic silencing of enzymes involved in GSL biosynthesis (Figs 1a, 2d and Supplementary Fig. 4c, d) highlighted that the GSL profile induced by FAPP2 silencing was similar to that induced by LacCer synthase (LCS) silencing in terms of a decrease in LacCer, but they differed in their effects on downstream GSL species: LCS knockdown induced a uniform decrease in Gb3 and GM3 whereas *FAPP2* knockdown selectively decreased Gb3. These results indicate that the GlcCer transported via FAPP2 feeds a pool of LacCer specifically destined to globoside (that is, Gb3) synthesis. Dynamic assessment of GSL metabolic fluxes followed by mathematical modelling corroborated this conclusion (Supplementary Fig. 5).

To search for the mechanisms responsible for the different sensitivities of Gb3 and GM3 synthesis to FAPP2 depletion, we studied the intra-Golgi distribution of Gb3 synthase (Gb3S) and of GM3 synthase (GM3S) by two independent approaches[7]. First, we measured the synthesis of Gb3 and GM3 in cells treated with brefeldin A (BFA), a toxin that redistributes the Golgi cisternae (but not the *trans*-Golgi network (TGN)) into the endoplasmic reticulum (ER) (generating an ER–Golgi intermixed compartment), interrupts vesicular trafficking from this intermixed compartment to the TGN[13], and releases FAPP2 from Golgi membranes[14]. BFA treatment decreased the synthesis of Gb3

[1]Telethon Institute of Genetics and Medicine, Via Pietro Castellino 111, 80131 Naples, Italy. [2]Institute of Protein Biochemistry National Research Council, Via Pietro Castellino 111, 80131 Naples, Italy. [3]Laboratory of Molecular Traffic, Department of Molecular and Cellular Biology, Institute for Molecular and Cellular Regulation, Gunma University, Gunma 371-8512, Japan. [4]Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK. [5]Department of Biosciences, Biochemistry, Åbo Akademi University, Artillerigatan 6 A III, BioCity, FI-20520 Turku, Finland. [6]Department of Cell Biology and Oncology, Consorzio Mario Negri Sud, Via Nazionale 8/A, 66030 Santa Maria Imbaro, Chieti, Italy. [7]Institut Curie Centre de Recherche, 26 rue d'Ulm, 75248 Paris Cedex 05, France. [8]CNRS, UMR144, F-75248 Paris, France. [9]Dipartimento di Scienze Chimiche and CEINGE Biotecnologie Avanzate, Università di Napoli Federicoll, ViaGaetano Salvatore 482, 80145 Napoli, Italy. [10]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 0QH, UK. [11]Department of Cell Biology, Osaka University, Osaka 565-0871, Japan.

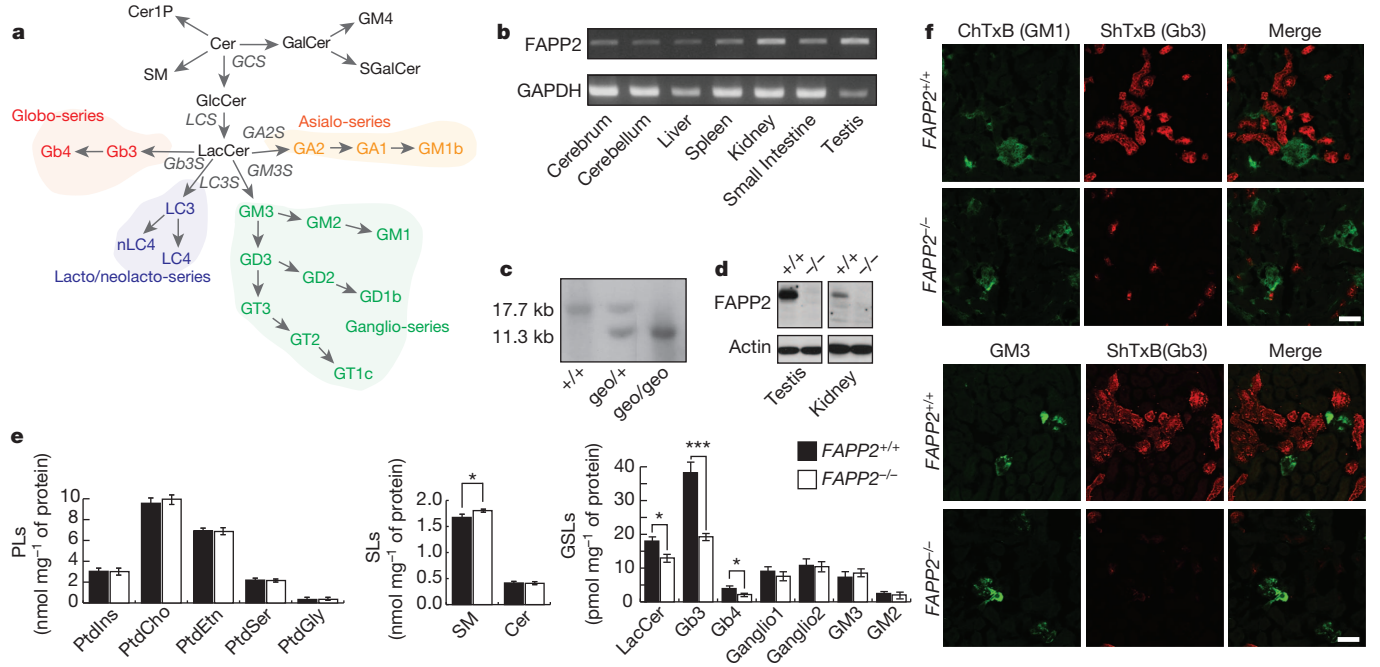


Figure 1 | FAPP2 selectively controls the levels of globosides *in vivo*. **a**, Simplified representation of the GSL synthetic pathway in vertebrates. GA2S, GA2 synthase; GCS, GlcCer synthase; LC3S, LC3 synthase. **b**, Expression of FAPP2 in mouse tissues. **c**, Southern blot of wild-type ($FAPP2^{+/+}$) and recombinant ($FAPP2^{geo/+}$ and $FAPP2^{geo/geo}$) mice tail DNA. See Supplementary Fig. 1b. **d**, FAPP2 levels in $FAPP2^{+/+}$ and $FAPP2^{-/-}$ testes and kidneys. **e**, Phospholipid (PL) and sphingolipid (SL) mass-spectrometry analysis, and GSL measurements of kidneys from $FAPP2^{+/+}$ ($n = 5$) and $FAPP2^{-/-}$ ($n = 10$) mice. Ganglio1, 2, unidentified gangliosides; Gb4, globoside; GM2, monosyalotrihexosylceramide; PtdCho, phosphatidylcholine; PtdIns, phosphatidylinositol; PtdEtn, phosphatidylethanolamine; PtdGly, phosphatidylglycine; PtdSer, phosphatidylserine; SM, sphingomyelin. Means ± s.e.m. **f**, ChTxB, ShTxB anti-GM3 staining of $FAPP2^{+/+}$ and $FAPP2^{-/-}$ kidney cortex sections. Representative pictures from at least five $FAPP2^{+/+}$ and five $FAPP2^{-/-}$ mice. Scale bars, 50 μm. $*P < 0.05$; $***P < 0.005$.

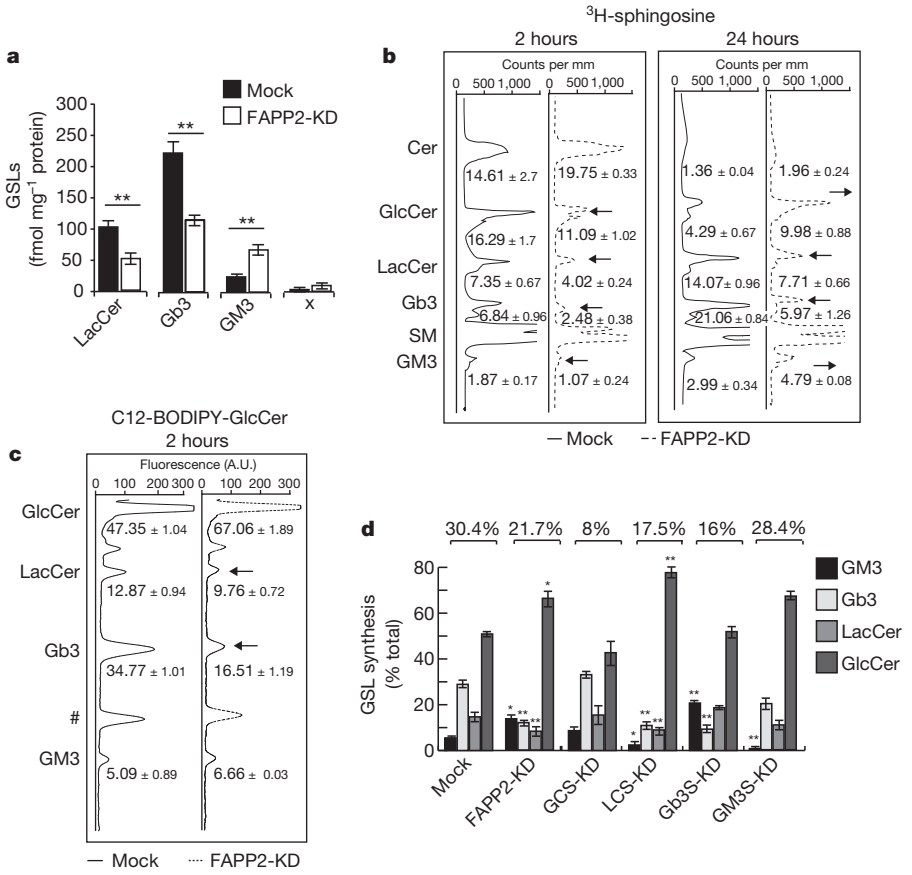


Figure 2 | FAPP2 is selectively required for Gb3 synthesis. **a**, GSL levels in HeLa cells. KD, knockdown; X, unassigned anionic GSL. **b**, High-performance thin-layer chromatography (HPTLC) profile of ^{3}H-sphingosine-labelled HeLa cells. Arrows denote changes induced by *FAPP2* knockdown. Numbers represent the percentage of each GSL species on total sphingolipids. **c**, HPTLC profile of C12-BODIPY-GlcCer-labelled HeLa cells. Arrows denote GSLs reduced by *FAPP2* knockdown; # represents unassigned peak and numbers represent the percentage of each GSL species on total GSL. A.U., arbitrary units. **d**, Effect of silencing of FAPP2 and of GSL synthetic enzymes on GSL species (expressed as percentage of total GSLs). Numbers represent the percentage of total GSLs on total sphingolipids. Means ± s.d. of at least three independent experiments. $*P < 0.05$; $**P < 0.01$.

but not that of GM3, indicating that the main fraction of endogenous Gb3S (but not of GM3S) resides in the TGN and thus remains segregated from its substrates that are synthesized in the BFA-induced intermixed ER–Golgi compartment (Fig. 3a). Second, we analysed the distribution of GM3S and Gb3S (Fig. 3b, c). In agreement with previous reports[6,15], we found that Gb3S is enriched in the TGN whereas GM3S is enriched in the Golgi cisternae. Moreover, consistent with its effect on GSL synthesis (Fig. 3a), BFA redistributed GM3S, but not Gb3S, to the ER (Supplementary Fig. 6).

The analysis of the role of FAPP2 was extended to other cell lines that also synthesize more complex gangliosides that, like GM3, were insensitive to FAPP2 depletion (Supplementary Fig. 7).

Thus, the synthesis of globosides at the TGN relies on the non-vesicular transport of GlcCer operated by FAPP2, whereas the synthesis of GM3 in the Golgi cisternae does not, eliciting the question as to whether GM3 synthesis depends instead on the vesicular transport of GlcCer. To address this question we inhibited intra-Golgi membrane trafficking[1] by treating cells with dicoumarol[1], by depleting the TRAPP component BET3 (also known as TRAPPC3) (ref. 1), or by depleting cytosolic phospholipase A_2 (cPLA$_2$) (ref. 16), and followed the transport of the temperature-sensitive mutant of vesicular stomatitis virus G protein (ts045-VSVG)[14]. As reported previously[1,16], these treatments suppressed the intra-Golgi progression of VSVG and strongly inhibited GM3 synthesis, but not Gb3 synthesis (Fig. 3d, e).

These results led us to propose that the vesicular transport of GlcCer feeds a pool of LacCer that is made in the Golgi cisternae and used for GM3 biosynthesis, whereas the non-vesicular transport of GlcCer via FAPP2 feeds a pool of LacCer that is made in the TGN and used in this compartment for globoside synthesis. This hypothesis generated some key predictions: (1) LCS should be present not only in the Golgi cisternae but also in the TGN; (2) other LacCer derivatives that, similarly to Gb3, are made at the TGN should depend on FAPP2; and (3) shifting the localization of GM3S from the Golgi cisternae to the TGN should make GM3 synthesis sensitive to FAPP2 depletion. We verified all of these predictions. First, LCS was found to localize both to the Golgi cisternae and to the TGN (Supplementary Fig. 8); second, *FAPP2* knockdown in SK-N-MC human neuronal cells selectively lowered the synthesis of gangliotriosylceramide (GA2), which is made in the TGN from LacCer (Fig. 1a and Supplementary Fig. 9a); third, GM3 synthesis, which is normally insensitive, becomes sensitive to FAPP2 depletion when a substantial fraction of GM3S is forced to localize at the TGN by expressing the enzyme at high levels (Supplementary Fig. 9b–d).

The selective requirement of FAPP2 for GSLs synthesized at the TGN, together with our previous observation that cells depleted of FAPP2 fail to concentrate GlcCer at the TGN[1], indicate that FAPP2 drives the transfer of GlcCer from the *cis*-Golgi—where GlcCer is synthesized—to the TGN, and raises the question of how this vectorial transport is sustained. We reasoned that in order to mediate the *cis*-Golgi-to-TGN transfer, apo-FAPP2 should be targeted to early Golgi membranes whereas GlcCer-bound FAPP2 should be targeted to the TGN. Thus, we compared the distribution of wild-type FAPP2 with that of a FAPP2 mutant, which is unable to bind GlcCer and thus is permanently apo (FAPP2(W407A))[1]. Although the main fraction of wild-type FAPP2 localizes at the TGN, the main fraction of FAPP2(W407A) localizes to the Golgi cisternae (Fig. 4a, b). Moreover, wild-type FAPP2 failed to localize at the TGN and was present mainly in the Golgi cisternae

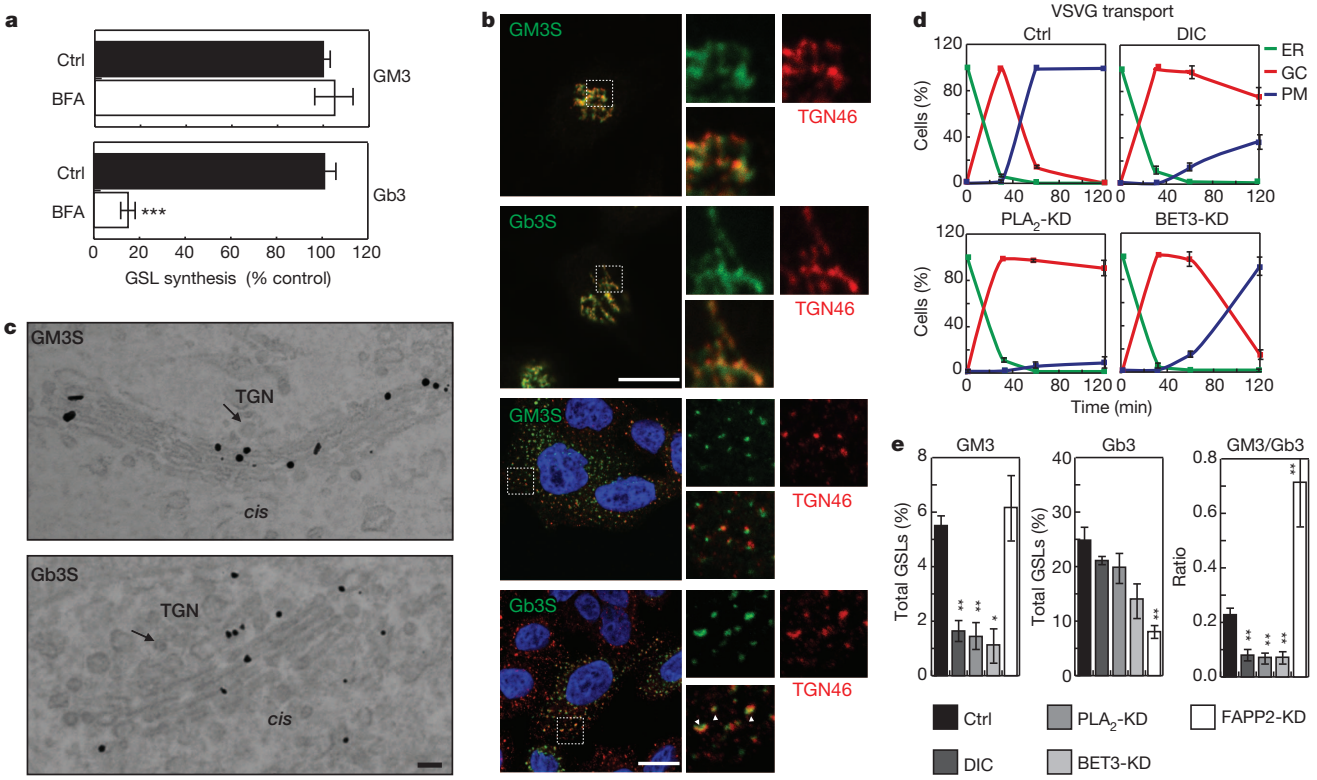


Figure 3 | Vesicular GlcCer transport feeds GM3 synthesis in the Golgi cisternae, whereas non-vesicular GlcCer transport feeds Gb3 synthesis in the TGN. **a**–**c**, Effect of BFA (5 μg ml^{-1}) on Gb3 and GM3 synthesis (**a**). Distribution of haemagglutinin (HA)–Gb3S and HA–GM3S evaluated by immunofluorescence (**b**) and by immunoelectron microscopy (**c**). In **b**, top panels represent untreated cells and bottom panels represent nocodazole-treated cells (3 h, 33 μM). Insets: enlargement of the boxed areas. The co-localization of HA–Gb3S and HA–GM3S with TGN46 was 50% and 14%, respectively. Data are representative of at least 30 cells per condition. Scale bar, 10 μm. **c**, Arrows represent clathrin-coated profiles at the TGN. Data are representative of at least 30 stacks. Scale bar, 100 nm. **d**, **e**, Effect of intra-Golgi trafficking blockage on VSVG transport (means ± s.d. in three independent experiments for at least 100 cells per time point) (**d**) and on GM3 and Gb3 synthesis (3-h ^{3}H-sphingosine pulse) (**e**). Means ± s.d. of three independent experiments. DIC, dicoumarol (200 μM); GC, Golgi complex; PM, plasma membrane. $*P < 0.05$; $**P < 0.01$; $***P < 0.005$.

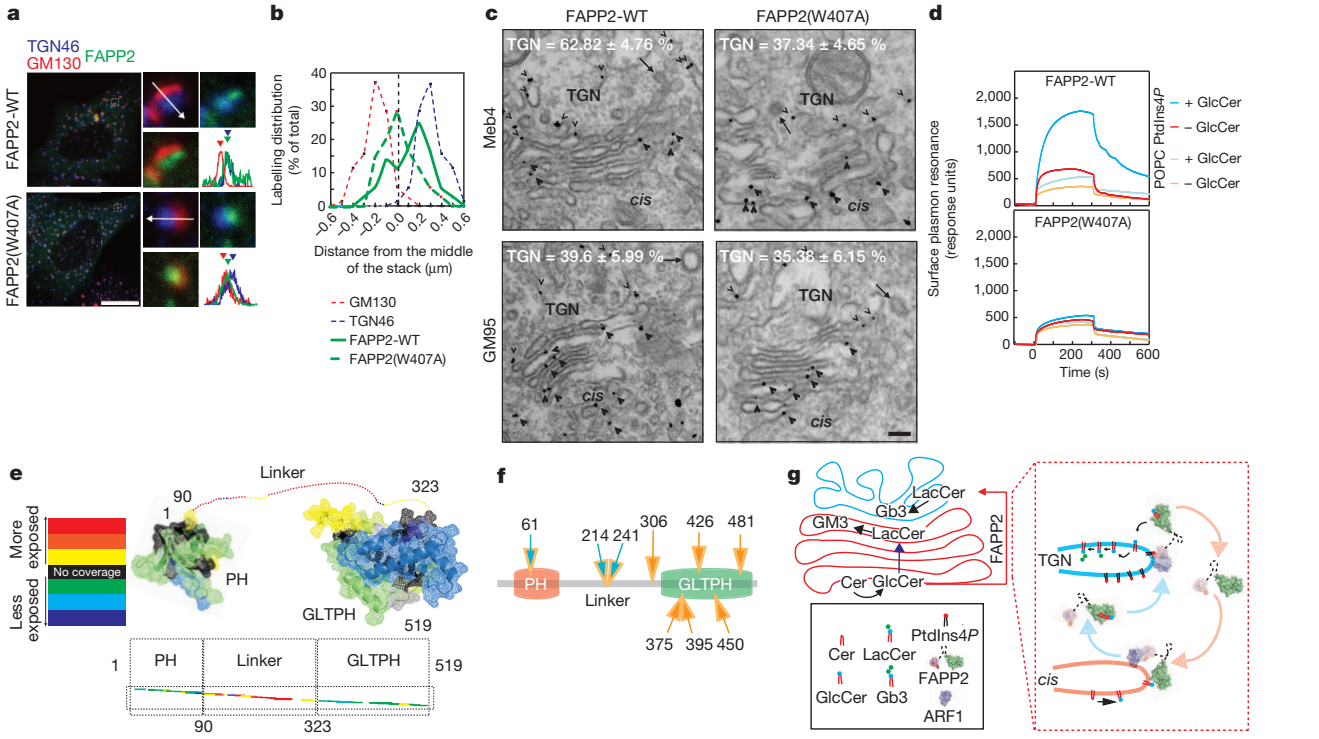


Figure 4 | GlcCer binding enhances the binding of FAPP2 to PtdIns4*P* and targets it to the TGN. **a**, Intra-Golgi distribution of wild-type FAPP2 (FAPP2-WT) and FAPP2(W407A) in nocodazole-treated cells (3 h, 33 μM). Right, enlargement of boxed areas and distribution of the maximal fluorescence intensity of FAPP2 along the *cis*-Golgi-to-TGN axis (white arrows). Scale bar, 10 μm. **b**, Quantification of the maximal labelling distribution of wild-type FAPP2 and FAPP2(W407A). The middle of the stack (0, black dashed line) is taken as a plane equidistant from GM130 and TGN46 fluorescence intensity peaks in 50 stacks per condition. **c**, Intra-Golgi distribution of wild-type FAPP2 and FAPP2(W407A) in Meb4 and GlcCer-deficient GM95 mouse cells. The percentage of TGN labelling is indicated. Means ± s.e.m. of at least 30 stacks per condition. Arrowheads represent Golgi cisternae staining, wedges represent TGN staining and arrows represent clathrin-coated profiles. Scale bar, 100 nm. **d**, Surface plasmon resonance analysis of the effect of GlcCer on FAPP2 binding to *1*-palmitoyl, *2*-oleoyl phosphatidylcholine (POPC)-PtdIns4*P* (98:2 mol:mol) liposomes. Results are representative of at least three independent experiments. PtdIns4*P* K_d value: 24 ± 0.9 μM for apo-FAPP2 and 6 ± 0.4 μM for GlcCer-bound FAPP2. **e**, Regions of FAPP2 showing different accessibility in HDX-MS were mapped on FAPP2. FAPP2 domains were modelled on the FAPP1 PH domain[18] and GLTPH domain[1,20]. **f**, Schematic representation of conformational changes induced by GlcCer loading on FAPP2 as assessed by HDX-MS (orange arrows) and controlled proteolysis-MS (blue arrows). Detailed results are presented in Supplementary Figs 12 and 13. **g**, Schematic representation of intra-Golgi non-vesicular (red arrow) and vesicular (blue arrow) transport of GlcCer. Inset: mechanism of FAPP2-mediated GlcCer-transfer directionality (cyan profiles, TGN; red profiles, Golgi cisternae; red arrows, targeting of apo-FAPP2 to the Golgi cisternae; cyan arrows, targeting of GlcCer-bound FAPP2 to the TGN.).

in cells that do not synthesize GlcCer[17], where FAPP2 is always apo (Fig. 4c). These results suggested that GlcCer binding positively regulates the targeting of FAPP2 to the TGN.

FAPP2 localization at the TGN is determined by its pleckstrin homology (PH) domain that coincidentally and independently[18] binds the small GTPase ARF1 and phosphatidylinositol 4-phosphate (PtdIns4*P*)[14], a phosphoinositide enriched at the TGN[19]. Single point mutations either in the PtdIns4*P*[14] or in the ARF-binding site[18] (Supplementary Fig. 10) abolish the recruitment of the PH domain to the Golgi complex (ref. 14 and data not shown), indicating a requirement for both binding sites. Interestingly, however, when these mutations are introduced into tandem forms of the FAPP PH-domain (di-PH), such that the di-PH has two binding sites either for ARF (di-PH-R18L) or for Ptdins4*P* (di-PH-E50A), the chimaeric proteins are able to localize to the Golgi complex, although with significantly different intra-Golgi distributions. In particular, a mutant FAPP-PH domain with a lower affinity for PtdIns4*P* and a higher affinity for ARF1 (di-PH-R18L) distributes throughout the Golgi stacks[14], whereas a mutant FAPP-PH domain with a lower affinity for ARF1 and a higher affinity for PtdIns4*P* (di-PH-E50A)[18] preferentially localizes to the TGN (Supplementary Fig. 10b), indicating that PtdIns4*P* dictates the TGN targeting of FAPP2.

We therefore assessed whether the preferential TGN association of GlcCer-bound FAPP2 could be due to the fact that it has a higher affinity for PtdIns4*P*. We found indeed that GlcCer loading (Supplementary information and Supplementary Fig. 11) increased FAPP2 binding to PtdIns4*P* (Fig. 4d), whereas it did not significantly affect the ability to bind ARF1 *in vitro* (not shown).

As FAPP2 binds GlcCer through its carboxy-terminal GLTP homology (GLTPH) domain[1,20] and PtdIns4*P* through its amino-terminal PH domain[14], the increase in PtdIns4*P* affinity induced by GlcCer binding suggested that GlcCer binding triggers a conformational change that is not limited to the GLTPH domain but is transmitted to more N-terminal regions. To gain insight into these conformational changes, we probed FAPP2 by hydrogen-deuterium exchange mass spectrometry (HDX-MS)[21]. The general profile of HDX-MS of FAPP2 indicated the presence of scarcely solvent-accessible N-terminal and C-terminal regions, corresponding to the PH-domain and the GLTPH domain, and the presence of a highly accessible and flexible intervening linker region (Fig. 4e). We then analysed the effects of GlcCer binding on the HDX-MS profile of FAPP2 and integrated this analysis with that of controlled proteolysis of FAPP2. The results of these analyses (Fig. 4f and Supplementary Figs 12 and 13) showed that the binding of GlcCer not only induced a stabilization of the GLTPH domain of FAPP2 but also affected the linker region interposed between the GLTPH domain and the PH domain, and on the PH domain itself, thus possibly having an impact on its PtdIns4*P* binding.

At this point a 'FAPP2 cycle' can be delineated (Fig. 4g): apo-FAPP2 associates with the *cis*-Golgi where it acquires GlcCer, resulting in a higher affinity of FAPP2 for PtdIns4*P*. FAPP2 then relocates to the PtdIns4*P*-enriched TGN where it delivers GlcCer.

Our findings establish a new paradigm for GSL biosynthesis whereby two branches receive their common precursor, GlcCer, from two parallel transport routes (Fig. 4g): the vesicular route feeds the LacCer pool used to make the ganglio-series in the Golgi cisternae, whereas the non-vesicular route mediated by FAPP2, which bypasses the intervening cisternae and delivers GlcCer to the TGN, feeds a TGN pool of LacCer converted *in loco* into GSLs of the globo- or asialo-series (Figs 1a and 4g). In a wider context, our results show how different modes of transporting a cargo through the Golgi complex channel the cargo itself towards distinct and otherwise potentially competing glycosylation pathways.

METHODS SUMMARY

FAPP2$^{-/-}$ mice were obtained following the procedure described in Supplementary Data and in ref. 22. High-performance liquid chromatography-based GSL measurements, metabolic labelling with ^{3}H-sphingosine, GSL extraction and HPTLC and analysis, as well as immunofluorescence and immunoelectron microscopy studies for subcellular protein localization assessments, were performed as described in ref. 1. Transport of ts045-VSVG was assessed as described previously[23]. Protein purification, fluorescence, circular dichroism and surface plasmon resonance studies were performed as described in refs 24, 25. HDX-MS was performed as in ref. 26. For statistical analysis, two-tailed Student *t*-tests were applied to the data. $*P < 0.05$; $**P < 0.01$; $***P < 0.005$.

Full Methods and any associated references are available in the online version of the paper.

Received 21 June 2012; accepted 25 June 2013.
Published online 4 August 2013.

1. D'Angelo, G. *et al.* Glycosphingolipid synthesis requires FAPP2 transfer of glucosylceramide. *Nature* **449,** 62–67 (2007).
2. Hakomori, S. I. Structure and function of glycosphingolipids and sphingolipids: recollections and future trends. *Biochim. Biophys. Acta* **1780,** 325–346 (2008).
3. Jeckel, D., Karrenbauer, A., Burger, K. N., van Meer, G. & Wieland, F. Glucosylceramide is synthesized at the cytosolic surface of various Golgi subfractions. *J. Cell Biol.* **117,** 259–267 (1992).
4. Futerman, A. H. & Pagano, R. E. Determination of the intracellular sites and topology of glucosylceramide synthesis in rat liver. *Biochem. J.* **280,** 295–302 (1991).
5. Hannun, Y. A. & Obeid, L. M. Principles of bioactive lipid signalling: lessons from sphingolipids. *Nature Rev. Mol. Cell Biol.* **9,** 139–150 (2008).
6. Halter, D. *et al.* Pre- and post-Golgi translocation of glucosylceramide in glycosphingolipid synthesis. *J. Cell Biol.* **179,** 101–115 (2007).
7. Maccioni, H. J., Quiroga, R. & Ferrari, M. L. Cellular and molecular biology of glycosphingolipid glycosylation. *J. Neurochem.* **117,** 589–602 (2011).
8. Jacewicz, M., Clausen, H., Nudelman, E., Donohue-Rolfe, A. & Keusch, G. T. Pathogenesis of shigella diarrhea. XI. Isolation of a shigella toxin-binding glycolipid from rabbit jejunum and HeLa cells and its identification as globotriaosylceramide. *J. Exp. Med.* **163,** 1391–1404 (1986).
9. van Heyningen, S. V. Cholera toxin: interaction of subunits with ganglioside G_{M1}. *Science* **183,** 656–657 (1974).
10. Psotka, M. A. *et al.* Shiga toxin 2 targets the murine renal collecting duct epithelium. *Infect. Immun.* **77,** 959–969 (2009).
11. Okuda, T. *et al.* Targeted disruption of Gb3/CD77 synthase gene resulted in the complete deletion of globo-series glycosphingolipids and loss of sensitivity to verotoxins. *J. Biol. Chem.* **281,** 10230–10235 (2006).
12. D'Angelo, G., Vicinanza, M. & De Matteis, M. A. Lipid-transfer proteins in biosynthetic pathways. *Curr. Opin. Cell Biol.* **20,** 360–370 (2008).
13. Chege, N. W. & Pfeffer, S. R. Compartmentation of the Golgi complex: brefeldin-A distinguishes trans-Golgi cisternae from the trans-Golgi network. *J. Cell Biol.* **111,** 893–899 (1990).
14. Godi, A. *et al.* FAPPs control Golgi-to-cell-surface membrane traffic by binding to ARF and PtdIns(4)P. *Nature Cell Biol.* **6,** 393–404 (2004).
15. Yamaji, T., Nishikawa, K. & Hanada, K. Transmembrane BAX inhibitor motif containing (TMBIM) family proteins perturbs a *trans*-Golgi network enzyme, Gb3 synthase, and reduces Gb3 biosynthesis. *J. Biol. Chem.* **285,** 35505–35518 (2010).
16. San Pietro, E. *et al.* Group IV phospholipase $A_2\alpha$ controls the formation of inter-cisternal continuities involved in intra-Golgi transport. *PLoS Biol.* **7,** e1000194 (2009).
17. Ichikawa, S., Nakajo, N., Sakiyama, H. & Hirabayashi, Y. A mouse B16 melanoma mutant deficient in glycolipids. *Proc. Natl Acad. Sci. USA* **91,** 2703–2707 (1994).
18. He, J. *et al.* Molecular basis of phosphatidylinositol 4-phosphate and ARF1 GTPase recognition by the FAPP1 pleckstrin homology (PH) domain. *J. Biol. Chem.* **286,** 18650–18657 (2011).
19. D'Angelo, G., Vicinanza, M., Di Campli, A. & De Matteis, M. A. The multiple roles of PtdIns(4)*P*–not just the precursor of PtdIns(4,5)*P*2. *J. Cell Sci.* **121,** 1955–1963 (2008).
20. Kamlekar, R. K. *et al.* The glycolipid transfer protein (GLTP) domain of phosphoinositol 4-phosphate adaptor protein-2 (FAPP2): structure drives preference for simple neutral glycosphingolipids. *Biochim. Biophys. Acta* **1831,** 417–427 (2013).
21. Yan, X., Watson, J., Ho, P. S. & Deinzer, M. L. Mass spectrometric approaches using electrospray ionization charge states and hydrogen-deuterium exchange for determining protein structures and their conformational changes. *Mol. Cell. Proteomics* **3,** 10–23 (2004).
22. Sato, T. *et al.* The Rab8 GTPase regulates apical protein localization in intestinal cells. *Nature* **448,** 366–369 (2007).
23. D'Angelo, G. *et al.* GRASP65 and GRASP55 sequentially promote the transport of C-terminal valine-bearing cargos to and through the Golgi complex. *J. Biol. Chem.* **284,** 34849–34860 (2009).
24. Zhai, X. *et al.* Glycolipid acquisition by human glycolipid transfer protein dramatically alters intrinsic tryptophan fluorescence: insights into glycolipid binding affinity. *J. Biol. Chem.* **284,** 13620–13628 (2009).
25. Ohvo-Rekilä, H. & Mattjus, P. Monitoring glycolipid transfer protein activity and membrane interaction with the surface plasmon resonance technique. *Biochim. Biophys. Acta* **1808,** 47–54 (2011).
26. Burke, J. E., Perisic, O., Masson, G. R., Vadas, O. & Williams, R. L. Oncogenic mutations mimic and enhance dynamic events in the natural activation of phosphoinositide 3-kinase p110alpha (*PIK3CA*). *Proc. Natl Acad. Sci. USA* **109,** 15259–15264 (2012).

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

Acknowledgements We thank A. Luini, C. Wilson and D. Priestman for discussions, A. Egorova for help with electron microscopy, G. Liebisch, A. Sigruener and G. Schmitz for lipidomic analysis. M.A.D.M. acknowledges the support of Telethon (GSP08002 and GGP06166), Associazione Italiana per la Ricerca sul Cancro (AIRC) (IG 8623), and the EU (FP7 Lipidomicnet). G.D.'A. acknowledges the support of AIRC (MFAG 10585). P.M. acknowledges the support of Academy of Finland and Sigrid Jusélius Foundation. C.-C.C. was funded by a Study Abroad Scholarship from the Taiwan Ministry of Education.

Author Contributions M.A.D.M. supervised the entire project; M.A.D.M. and G.D.'A. wrote the manuscript with comments from all co-authors; G.D.'A., with the help of M.S., designed and conducted at TIGEM the experiments of sphingolipid labelling, membrane trafficking, immuno-localization and controlled proteolysis. M.S. designed the strategy and produced plasmid vectors. M.S. and G.D.T. prepared recombinant proteins, anti-FAPP2 and anti-BET3 antibodies. T.U. and T.S. generated and characterized *FAPP2*$^{geo/geo}$ and *FAPP2*$^{-/-}$ mice under the supervision of A.H. C.-C.C. conducted the HPLC measurements of GSLs under the supervision of F.M.P. L.J. provided the Cy3-ShTxB. E.P. and T.D. conducted the electron microscopy experiments. H.O.-R. conducted the surface plasmon resonance experiments under the supervision of P.M. A.V. conducted the tryptophan fluorescence and circular dichroism experiments under the supervision of S.D'.A. F.C. and G.D'.A. produced the mathematical model for GSL metabolism. M.M. and P.P. performed and interpreted the MS analysis; R.L.W. and J.E.B. performed and interpreted the HDX analysis.

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 M.A.D.M. (dematteis@tigem.it).

METHODS

Reagents and antibodies. All chemical reagents were of analytical grade or higher and purchased from Sigma unless otherwise specified. Cell culture media were from Invitrogen. Polyclonal antibodies against human FAPP2, BET3, cPLA$_2$IVa and GM130 were raised in rabbits using glutathione *S*-transferase fusion proteins as immunogens. All were affinity purified on their corresponding immunogens. The anti-VSVG clone P5D4, anti-Flag M2 and anti-HA monoclonal antibodies, and the anti-rabbit and anti-mouse IgG Cy3-conjugated antibodies were from Sigma. The Alexa 488-conjugated ChTxB fragment was from Invitrogen. The Cy3-conjugated ShTxB fragment was prepared as described[27]. The mouse monoclonal antibody against GM3 (clone 2590) was from Cosmo Bio Co. Sheep polyclonal antibodies against TGN46 were from AbD Serotech. The Alexa 488 goat anti-mouse and anti-rabbit IgG antibodies were from Molecular Probes. All unlabelled purified lipids were from Avanti Polar Lipids. ^{3}H-sphingosine was from PerkinElmer. Stock solutions of GSLs were prepared in chloroform/methanol (2:1 by volume) and of other lipids in hexane/2-propanol (3:2 by volume). Lipid solutions were stored in the dark at −20 °C and warmed to room temperature before use.

FAPP2 Knockout mice. Mice were from the C57BL/6 strain. All animal procedures were performed in accordance with the guidelines of the Animal Care and Experimentation Committee of Gunma University, and all animals were bred in the Institute of Animal Experience Research of Gunma University. *FAPP2* knockout mice were generated following the procedure described in Supplementary information. Histological, immunofluorescence microscopy and X-gal staining were performed as previously described[22,28,29].

Cell culture. HeLa, Meb4, GM95, HepG2, HK2, COS7 and MDCK cells were grown and transiently transfected by TransIT-LT1 (Mirus Bio) as described in ref. 1. Stably-expressing HeLa-GM3S cells were obtained after transfection of the 3×HA–GM3S coding plasmid and selection in the presence of G418 (Invitrogen) and screening of monoclonal colonies by indirect immunofluorescence.

Plasmids and constructs. Most of the constructs used in this study were obtained from HeLa RNA by RT–PCR and cloning into appropriate vectors. In brief, total RNA was isolated from HeLa cells; RT–PCR was performed using a poly dT oligo as a primer. The complementary DNA obtained was used as a template for PCR, using the following primers: for Gb3S: forward 5′-GTTGAATTCGATCTGG GGATACCATGTCC-3′, reverse 5′-CACCTCGAGCAAGTACATTTTCATGG CCTC-3′; for GM3S: forward 5′-CAGGAATTCAGAATGAGAAGGCCCA GCTTGTTA-3′, reverse 5′-AACGCGGCCGCTGAAATTCACGATCAATGC CTCCA-3′; for LCS: forward 5′-ATAGAATTCTGGCTGCAGCATGCGCGC-3′, reverse 5′-CGCGATATCAAGTACTCGTTCACCTGAGCCA-3′. The PCR products were cloned into a linearized pCR2.1 vector, and processed for automatic sequencing. All of the cloned sequences matched the sequence reported in databases for human Gb3S (AF513325), GM3S (AY152815.2) and LCS (B4GALT5; (NM_004776)). The DNAs corresponding to the various coding sequences were then subcloned into EcoRI/XhoI (Gb3S), EcoRI/NotI (GM3S), EcoRI/EcoRV sites of pCDNA3-3XHA at C-terminus or p3XFLAG-CMV-14.

Green fluorescent protein (GFP)–FAPP2 wild-type and W407A constructs were obtained as described in ref. 1; recombinant GST–FAPP2 wild-type, GST–FAPP2(W407A), His-FAPP2 wild type and His-FAPP2(W407A) were produced in *Escherichia coli* as described in ref. 1.

GFP–diFAPP2PH-wild type and E50A were obtained as follows: GFP–FAPP2 wild-type DNA was used as a template for two distinct PCR reactions using as primers: PCR (a): 5′-TCTGAATTCATGGAGGGGGTGCTGTACA-3, 5′-TAT GGTACCGAGCAAGCAAGCCTTGGCTGATCCC-3; PCR (b): 5′-TATGGT ACCTTGCTGGAGGGGGTGCTGTACAAGTG-3, 5′-TCACTCGAGTTAGCA AGCCTTGGCTGATCC-3. Products from PCR (a) were subcloned into EcoRI/KpnI sites of vector pEGFP-C1 to obtain construct pEGFP-FAPP2PH. Subsequently, products from PCR (b) were subcloned into KpnI/XhoI sites of the pEGFP-FAPP2PH construct to obtain GFP–diFAPP2PH-wild type. A similar procedure was applied to obtain the GFP–diFAPP2PH(E50A) mutant using as a template GFP–FAPP2(E50A) DNA. GFP–FAPP2(E50A) was obtained from wild-type GFP–FAPP2 by site-directed mutagenesis using the primers 5′-GAGCATACAAATGGCAGTCTGTGCAATTCAAGTTCATTCTGTAG-3′ and 5′-CTACAGAATGAACTTGAATTGCACAGACTGCCATTTGTATGCTC-3′.

siRNA treatments. The siRNAs for human FAPP2 (NM_001197026), GCS (NM_003358), BET3 (NM_014408), B4GALT5 (NM_004776), B4GALT6 (NM_004775), SIAT9/GM3S (AY152815.2), A4GALT/Gb3S (NM_017436), PLA$_2$ (NM_001199562) comprised mixtures of at least three siRNA duplexes (Supplementary Table 1) and were obtained from Dharmacon. HeLa, HK2, HepG2 and MDCK cells were plated at 30% confluence in 12-well plates and transfected with 120–150 pmol of siRNAs with Oligofectamine (Invitrogen) or Dharmafect4 (Dharmacon), in accordance with the manufacturer's protocol. At 72 h after the initial treatment with siRNA, the cells were processed directly. Silencing efficiency was evaluated either by western blot (Supplementary Fig. 2) or by quantitative PCR (Supplementary Fig. 4b) using specific primers (Supplementary Table 1).

Measurement of GSLs. Metabolic labelling with ^{3}H-sphingosine or ^{14}C-galactose, GSL extraction and HPTLC and analysis were performed as described in refs 1, 6.

Immunofluorescence and morphometric analysis. All immunofluorescence experiments were performed as described previously[1,14]. Images are confocal optical slices obtained using an LSM 710 (Zeiss) confocal microscope. Co-localization analysis was performed as described in ref. 14 or by using an object-based co-localization method included in the JACoP v2.0 application for ImageJ[30]. In brief, individual mini-stacks in nocodazole-treated cells were considered as objects whose mass-centre position was calculated after segmentation. The perfect coincidence of mass-centre positions for two distinct labellings (that is, Gb3S/TGN46 or GM3S/TGN46) in a single mini-stack was considered as a positive co-localization event.

Immunoelectron microscopy. Immunoelectron microscopy was performed in transfected HeLa, Meb4 and GM95 cells as described previously[1].

VSVG intracellular transport assay. Transport of ts045-VSVG was assessed as described previously[23].

Statistical analysis. For statistical analysis, two-tailed Student *t*-tests were applied to the data. $*P < 0.05$; $**P < 0.01$; $***P < 0.005$.

27. Mallard, F. & Johannes, L. Shiga toxin B-subunit as a tool to study retrograde transport. *Methods Mol. Med.* **73,** 209–220 (2003).
28. Muramatsu, K. *et al.* Neuron-specific recombination by Cre recombinase inserted into the murine *tau* locus. *Biochem. Biophys. Res. Commun.* **370,** 419–423 (2008).
29. Hashimoto, Y. *et al.* Neuron-specific and inducible recombination by Cre recombinase in the mouse. *Neuroreport* **19,** 621–624 (2008).
30. Bolte, S. & Cordelieres, F. P. A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224,** 213–232 (2006).

LETTER

doi:10.1038/nature12395

Recovery from slow inactivation in K^+ channels is controlled by water molecules

Jared Ostmeyer[1], Sudha Chakrapani[2], Albert C. Pan[1]†, Eduardo Perozo[1] & Benoît Roux[1]

Application of a specific stimulus opens the intracellular gate of a K^+ channel (activation), yielding a transient period of ion conduction until the selectivity filter spontaneously undergoes a conformational change towards a non-conductive state (inactivation). Removal of the stimulus closes the gate and allows the selectivity filter to interconvert back to its conductive conformation (recovery). Given that the structural differences between the conductive and inactivated filter are very small, it is unclear why the recovery process can take up to several seconds. The bacterial K^+ channel KcsA from *Streptomyces lividans* can be used to help elucidate questions about channel inactivation and recovery at the atomic level. Although KcsA contains only a pore domain, without voltage-sensing machinery, it has the structural elements necessary for ion conduction, activation and inactivation[1–7]. Here we reveal, by means of a series of long molecular dynamics simulations, how the selectivity filter is sterically locked in the inactive conformation by buried water molecules bound behind the selectivity filter. Potential of mean force calculations show how the recovery process is affected by the buried water molecules and the rebinding of an external K^+ ion. A kinetic model deduced from the simulations shows how releasing the buried water molecules can stretch the timescale of recovery to seconds. This leads to the prediction that reducing the occupancy of the buried water molecules by imposing a high osmotic stress should accelerate the rate of recovery, which was verified experimentally by measuring the recovery rate in the presence of a 2-molar sucrose concentration.

Available X-ray structures of the KcsA channel provide an atomic view of the four most important functional states in which the intracellular gate is either closed or open, and the selectivity filter is either conductive or inactivated[8–10]. These are (in terms of their Protein Data Bank entries): the closed-conductive (1K4C), open-conductive (3FB7), open-inactivated (3F5W), and the closed-inactivated (1K4D) states (Fig. 1a). When the filter is in the conductive state (1K4C and 3FB7), it is occupied by several K^+ ions arranged in single file over five binding sites (S0 to S4); ion occupancy is considerably reduced when the filter is in the non-conductive inactivated state (1K4D and 3F5W) because it is 'pinched' at the central glycine residue of the signature sequence TTVGYGD. The conductive conformation of the filter is virtually unchanged when the gate is closed (1K4C) or open (3FB7). The pinched non-conductive conformation of the filter is similar in the open-inactivated conformation (3F5W) and in the closed structure (1K4D) crystallized at low K^+ concentration, indicating that the latter provides a realistic representation of the KcsA channel in the closed-inactivated state[10].

Although available X-ray crystal structures reveal the conformations of the functional states, they do not explain the experimentally observed timescales. Recovery from inactivation is an extremely slow, ion-dependent process[11]. Macroscopic current measurements show that the KcsA channel inactivates within 1–3 s and that recovery takes place in 5–20 s (Fig. 1b, c), depending on the external K^+ concentration[2,3]. However, the structural difference between the conductive and non-conductive filter is less than ~1 Å root mean square deviation (r.m.s.d.), barely larger than the thermal fluctuations of proteins under ambient conditions (Fig. 1d). According to a naive Eyring rate theory argument, the long timescale corresponds to an activation free energy barrier of about 15–20 kcal mol^{-1}, a value that is difficult to reconcile with the very slight structural difference between the conductive and inactive filter (Supplementary Fig. 1). Explaining the molecular origin of the extraordinarily long timescale of the recovery process is an unresolved issue.

To identify the rate limiting step opposing recovery, a molecular dynamics simulation of KcsA with the non-conductive filter and the closed gate was carried out in the presence of a high K^+ concentration. Although these conditions were designed to favour a spontaneous transition of the filter towards the conductive state, the filter remained

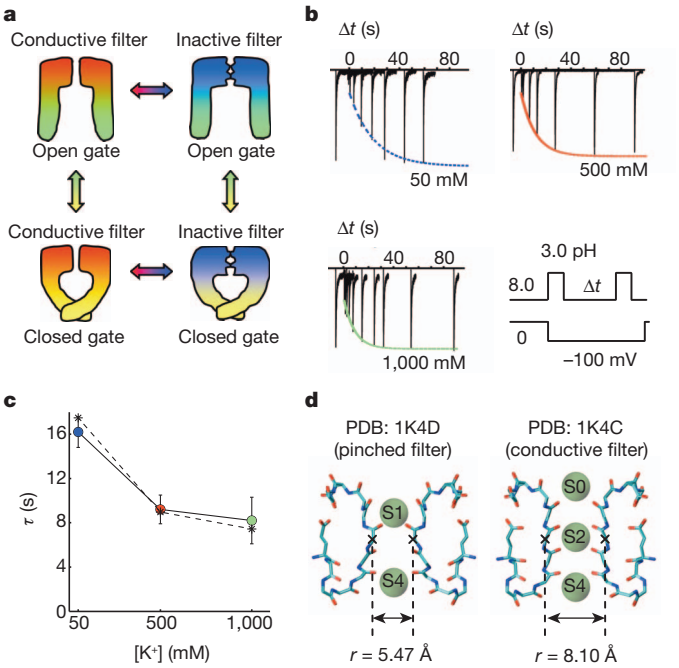


Figure 1 | Gating, inactivation and recovery in KcsA channels. **a**, Schematic depiction of the four dominant functional states. **b**, Recovery of the selectivity filter from slow-inactivation measured from inward K^+ current during a double pulse protocol that cycles the intracellular gate from open (inactivating the filter) to closed (promoting filter recovery) to open (for measuring the extent of recovery). Currents were measured at indicated symmetrical K^+ concentration. The fraction of recovered channels increases as a function the inter-pulse duration (Δt) in each of the three patches, and may be fit to a single exponential function. **c**, The fit time constants for recovery for various symmetric K^+ concentrations demonstrate that K^+ accelerates recovery. Results based on a kinetic model are shown as a dashed line with asterisks (error bars denote standard deviation with $n = 3$). **d**, The conductive state of the selectivity filter compared to the pinched state is characterized by a relatively small increase in the minimum inter-subunit distance between backbone atoms of the filter and an increase in ion occupancy.

[1]Department of Biochemistry and Molecular Biology, The University of Chicago, 929 E57th Street, Chicago, Illinois 60637, USA. [2]Department of Physiology and Biophysics, School of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA. †Present address: D. E. Shaw Research, New York, New York 10036, USA.

in the non-conductive state during the 17-μs-long simulation (Fig. 2a). Throughout the simulation, the selectivity filter stayed near the crystallographic conformation 1K4D and no ion translocation event was observed (Supplementary Fig. 2). The stability of the non-conductive filter over this long timescale indicates that the simulation captured the rate-limiting step opposing a rapid recovery of the filter.

Examination of the 17 μs trajectory led us to focus on three buried water molecules in an ~8-Å-long cavity located behind the selectivity filter of each subunit. Whereas water molecules bound near the protein surface are expected to undergo rapid dynamical fluctuations and exchange easily with the bulk[12], the buried waters maintained distinct positions and orientations throughout the entire simulation. The buried waters are stabilized by a network of water–water and water–protein hydrogen bonds (Fig. 2b), displaying long residence times (Supplementary Table 1). The buried water molecules thereby become an integral part of the protein structure when the selectivity filter is in the pinched conformation. On average, a given subunit cavity is occupied by three buried water molecules ~99% of the time as each cavity releases its waters into the bulk no faster than 10–15 μs and refills in 80–100 ns (Supplementary Fig. 5). The Tyr 82 side chain acts as a lid, controlling

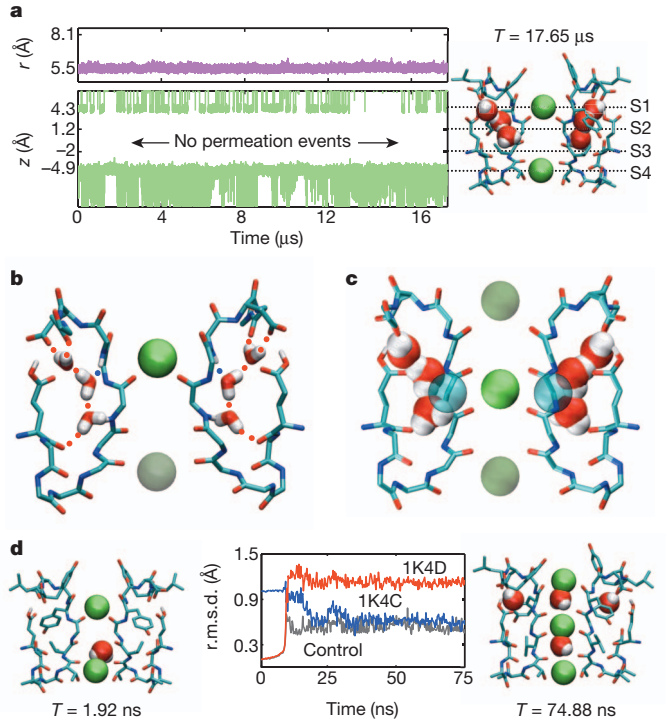


Figure 2 | Molecular dynamics simulations reveal mechanism of recovery from inactivation. **a**, Results from a simulation of the pinched filter show that the width of the filter remained near $r = 5.5$Å where r is defined as the cross-subunit pinching distance between the Cα atoms of Gly 77 (magenta line, top plot). Plot of the height of K^+ ions, z, above the centre of mass of the selectivity filter (green traces, bottom plot). Projection of the aforementioned plot onto a snapshot of the selectivity filter taken at time $T = 17.65$ μs at the end of the simulation (right). **b**, Network of hydrogen bonds matches each water with two donors and an acceptor, keeping water molecules trapped behind the selectivity filter throughout the simulation. **c**, Water positions behind the pinched filter of 1K4D sterically clash with the conductive filter 1K4C. Rendering both water positions from 1K4D and the Cα glycine atoms from 1K4C as van der Waals spheres reveals that unfavourable steric clashes of ~1 Å in magnitude would exist between the protein and water molecules. **d**, In simulations without water behind the filter, the r.m.s.d. distance of the pinched filter relative to its original crystallographic coordinates (red) increases as the r.m.s.d. distance to the conductive filter in 1K4C (blue) decreases. The r.m.s.d. distance versus 1K4C falls to ~0.6 Å, comparable with a control simulation (dark grey) starting from the conductive filter 1K4C. Snapshots taken at time $T = 1.92$ and 74.88 μs are shown.

access to the cavity. In the cavity, the buried water molecules remain near the location observed in the structure with the pinched non-conductive filter (1K4D), which is also consistent with nuclear magnetic resonance (NMR) data[5,13]. The outermost water molecule is observed in the X-ray structure of the conductive filter (1K4C), whereas the two others are absent. Structural modelling shows that the latter would clash with the backbone Cα of the central glycine residue (Gly 77) if the filter adopted a conductive conformation (Fig. 2c). These results indicate that the presence of the buried water molecules bound behind the filter locks its conformation into the pinched non-conductive state and prevents a spontaneous transition towards the conductive state.

To demonstrate that these 'inactivating water molecules' prevent recovery, additional molecular dynamics simulations were carried out in which they were removed (Fig. 2d). Within a few nanoseconds, the filter made a transition towards the conductive conformation (1K4C). The recovery transition was accompanied by K^+ ions translocating inward towards the filter from the extracellular side and the binding of water molecules behind the filter at the positions observed in 1K4C. While one K^+ ion moved to the site S2 of the filter, another entered the site S0 from the extracellular bulk. These two K^+ ions, combined with a third K^+ ion in S4 that remained present at its initial location during the entire simulation, resulted in a 'S0–S2–S4' ion configuration in the filter. The involvement of K^+ ions is expected, as the timescale of recovery increases at low extracellular K^+ concentration. To clarify the role of the ions further, molecular dynamics simulations were carried out with no free K^+ ions in solution. In all cases, the filter remained in the pinched non-conductive conformation, consistent with the experimental observation that low extracellular K^+ inhibits recovery (Fig. 1c). In contrast, additional simulations in 1 M KCl showed that whether or not a K^+ ion was initially bound in the site S4 or in the intracellular cavity of the channel, a K^+ ion entered the filter from the extracellular side to bind to the site S2. Additional molecular dynamics experiments were carried out to determine the outcome from different intermediate starting conditions (Supplementary Table 2), indicating that the most likely pathway leading to recovery requires the absence of all three inactivating water from all four subunits to permit the binding of an external K^+ ion to the filter.

To characterize the free energy landscape controlling the recovery process, we calculated the two dimensional potential of mean force (PMF) as a function of the pinched glycine Cα–Cα distance of the filter r and the position of the external K^+ ion Z (Supplementary Fig. 6), with and without the inactivating water molecules (Fig. 3). These two states illustrate the coupling between the inactivating waters and the external K^+ ion on the recovery process. When the inactivating water molecules are present (Fig. 3a), no significant free energy barrier is encountered to move a K^+ ion from the bulk to the binding site S1 of the pinched filter pinched filter (snapshot *i*), but the subsequent transition of this external K^+ to the binding sites S2 or S3 is energetically forbidden; the free energy landscape rises sharply, up to 30 kcal mol^{-1} relative to the local minimum, as the K^+ is brought deeper into the filter towards the site S2 along the Z coordinate. There is a large energy barrier because the non-conductive filter is too narrow for a K^+ ion to fit through. The energy barrier to gradually open the filter to a conductive conformation is also highly unfavourable, rising by ~25 kcal mol^{-1} when an ion is in the site S1 (snapshot *j*). Opening the pinched filter by moving along the r coordinate is opposed by a large free energy barrier because the backbone of the filter clashes with the buried inactivating waters. The PMF with the inactivating waters removed is markedly different (Fig. 3b). Once a K^+ ion reaches the site S1 (snapshot *k*), the recovery process becomes energetically downhill. A free energy basin guides the recovery process as the free energy decreases (snapshot *l*). The filter first makes a transition from the pinched to the conductive conformation, the ion in S1 then permeates down to the site S2 (snapshot *m*), ending in a conformation that is highly similar to the 1K4C structure. In the absence of inactivating waters, the channel will spontaneously and rapidly recover to a conductive conformation upon the

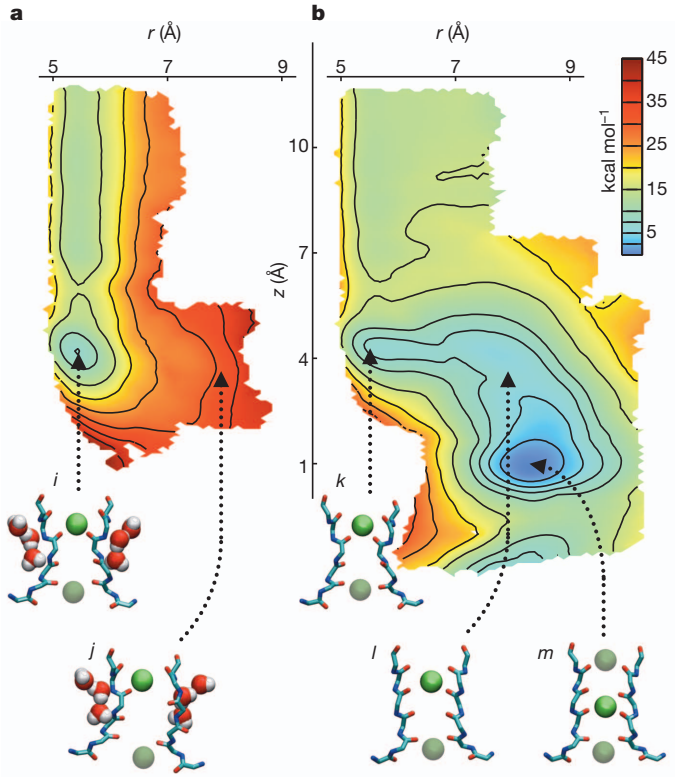


Figure 3 | Two-dimensional free energy landscape of the recovery process. The horizontal reaction coordinate r describes the width of the selectivity filter and is defined as the cross-subunit pinching distance between the Cα atoms of Gly 77. The vertical reaction coordinate z is the height of a K^+ ion relative to the centre of mass of the selectivity filter. **a**, PMF calculated with inactivating water molecules present behind the selectivity filter. The pinched filter rests in a free energy minimum with a K^+ in position S1 (snapshot *i*). The transition from a pinched to a conductive conformation (snapshot *j*) of the selectivity filter is impeded by a ~25 kcal mol^{-1} free energy barrier relative to the local minimum, resulting in an unstable conformation of the conductive filter. **b**, PMF calculated with the inactivating water molecules absent. The pinched filter with a K^+ ion in position S1 (snapshot *k*) recovers spontaneously, following the downhill slope of the free energy landscape. The filter recovers to a conductive conformation by moving first to an open conformation (snapshot *l*) before ions in the filter adopt a conductive configuration (snapshot *m*).

binding of a K^+ ion from the extracellular solution. But the filter remains locked in the non-conductive state, even with an incoming external K^+ ion, as long as the inactivating waters are present.

These computational experiments suggest that two factors control the conformational transition of the selectivity filter from the non-conductive pinched state to the conductive state: the release of the inactivating water molecules and the binding of an external K^+ ion. The buried inactivating waters act as gatekeepers, locking the filter in the pinched non-conductive conformation; the binding of an external K^+ ion cannot lead to recovery if they are present. The external K^+ ion binding to the filter acts as a catalyst for the final step; releasing the inactivating water molecules does not lead to recovery without the binding of an external K^+ ion. The inactivating waters must first be released to unlock the pinched state, and an external K^+ ion must then bind to the filter to catalyse the final step to recovery.

Although these simulations suggest a plausible mechanistic scenario, a complete and spontaneous recovery from an inactivation event was not observed during these experiments. The recovery process, which occurs over a timescale of ~10 s, is longer than what can be achieved with state-of-the-art computing technologies. Despite these limitations, it is important to bridge the results from molecular dynamics experiments to the extremely long timescales of the recovery process observed in macroscopic current measurements. For this purpose, we rely on a kinetic model (Fig. 4a) based on the assumption that occupancy of the cavity of each of the four subunits by inactivating waters is independent and uncorrelated, and that recovery can proceed to its final stage through the binding of an external K^+ ion only after all of the inactivating water molecules have been released. Although it is possible that recovery might occur very infrequently under different circumstances (Supplementary Table 2), these constraints capture the dominant mechanism of recovery. Occupancy changes take place with forward and backward rates of k_f and k_b, followed by a final concerted step corresponding to the binding of an external K^+ ion to the selectivity filter. By virtue of the stoichiometric constraint posed by the four subunits, the timescale of recovery predicted by this model is tens of seconds (Fig. 1c, dashed line) using microscopic rates that are consistent with the current molecular dynamics simulations (Fig. 4). Moreover, the model displays the correct sensitivity to the external concentration of K^+, becoming slower at low concentration and reaching a plateau at high concentration.

This mechanism predicts that lowering the occupancy of the subunit cavities by inactivating water molecules should accelerate the rate of recovery from inactivation. This is consistent with the experimentally observed hydrostatic pressure dependence of recovery for the Shaker K^+ channel[14], which indicates that ~8 water molecules must be released from the inactivated state to return to the conductive filter. An alternative approach to vary the occupancy of the cavities is to apply a high osmotic stress to the external solution[15]. Increasing the osmotic stress from the external solution should accelerate the process of recovery; the stability of the inactivating waters in 2 M sucrose would decrease due to the high osmotic stress, which should lead to a more rapid time of recovery (Supplementary Fig. 7). To test this prediction, we measured the rate of recovery of wild-type KcsA channel in the presence of 5 mM of K^+ and 2 M sucrose in the external solution (Fig. 4b). In the

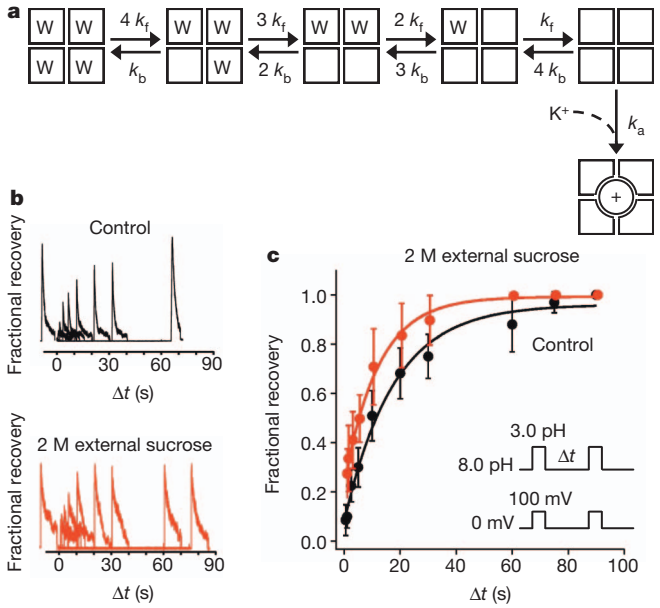


Figure 4 | Impact of water molecules on recovery process. a, Kinetic scheme incorporating the main findings of the molecular dynamics simulations. The model was used to produce the simulated recovery times as a function of $[K^+]$ shown in Fig. 1c (dashed line) using the rate constants $1/k_f = 11$ μs, $1/k_b = 79$ ns, $1/k_a = 8.3$ ns × (150 mM/$[K^+]$). When starting the kinetic model in the state with no inactivating water molecules (W), the channel reaches the active conductive state rather than the fully inactive state with a probability of ~0.5 at 50 mM of $[K^+]$. This probability rises to ~0.7 at 150 mM of $[K^+]$. **b**, **c**, Effect of external sucrose on the time course of recovery. **b**, Outward currents recorded with external 5K^+/145 mM NMG and internal 150 mM K^+ in the absence and presence of 2 M external sucrose. **c**, Fractional recovery averaged from eight patches plotted as a function of the inter-pulse interval. Error bars denote standard deviation.

absence of sucrose, the average recovery time is 17.06 ± 4.75 s, and in the presence of sucrose it decreases to 12.21 ± 3.73 s, consistent with our prediction (mean $\pm$ s.d.; Supplementary Fig. 7). Although the overall effect is small (the waters bound to one subunit are destabilized by only $\sim 0.7\ k_B T$, where k_B is Boltzman's constant and T is the temperature), protein dynamics and the diffusion and binding of an external K^+ ion to the selectivity filter that is critical to catalyse recovery are expected to considerably slow down in a viscous environment[16]. The fact that the recovery process is accelerated in 2 M sucrose—despite a viscosity that is 30 times that of pure water—is strong evidence that the applied osmotic stress decreased the occupancy of the gatekeeper inactivating water molecules locking in the non-conductive state of the selectivity filter.

Our molecular dynamics simulations and macroscopic current measurements suggest that the selectivity filter of K^+ channels functions as a ligand-gated pore through a built-in osmometer, where water molecules from the external solution are the 'ligands' responsible for the gating stimulus. It is unlikely that such a mechanism could have an important physiological role in higher organisms, as there are negligible variations in internal osmotic strength; however, it is possible that such a mechanism could regulate bacterial K^+ channels, as microorganisms are often exposed to widely varying external conditions. Examination of homology models of various K^+ channels indicates that the subunit cavities are conserved structurally and in terms of the hydrogen bonding interactions that they could provide through the peptide backbone (Supplementary Fig. 3), indicating that the role of inactivating waters acting as gatekeepers of the pinched non-conductive filter is likely to be conserved in the K^+ channel superfamily. The present analysis may provide a basis to understand the wide range of inactivation and recovery phenotypes arising from point mutations in the neighbourhood of the selectivity filter[1,4,7,9,17].

METHODS SUMMARY

Initial coordinates for the molecular dynamic simulations were taken from the crystal structures 1K4C and 1K4D. Crystallographic water molecules around the single subunit in each of the two crystal structures were placed around all four subunits in the tetramer. Residues were assigned their standard protonation state at pH 7.0, except for residue Glu 71, which was protonated. The channel was embedded in a bilayer of POPC lipids and solvated in 1 M KCl using the web service CHARMM-GUI (total number of atoms, 45,898). All-atom simulations were run using the CHARMM PARAM27 force field[18] under constant NVT conditions at a temperature of 310 K. Molecular dynamic simulations were carried out on the special purpose computer Anton[20] on loan to the Pittsburgh Supercomputer Center (PSC). A total of 114 molecular dynamic simulations were carried out starting from different initial conditions with respect to the inactivating water molecules, for an aggregate total time of 29.5 μs of molecular dynamics.

Full Methods and any associated references are available in the online version of the paper.

Received 26 October 2012; accepted 18 June 2013.
Published online 28 July 2013.

1. Cordero-Morales, J. F. *et al.* Molecular determinants of gating at the potassium-channel selectivity filter. *Nature Struct. Mol. Biol.* **13,** 311–318 (2006).
2. Chakrapani, S., Cordero-Morales, J. F. & Perozo, E. A quantitative description of KcsA gating II: single-channel currents. *J. Gen. Physiol.* **130,** 479–496 (2007).
3. Chakrapani, S., Cordero-Morales, J. F. & Perozo, E. A quantitative description of KcsA gating I: macroscopic currents. *J. Gen. Physiol.* **130,** 465–478 (2007).
4. Cordero-Morales, J. F. *et al.* Molecular driving forces determining potassium channel slow inactivation. *Nature Struct. Mol. Biol.* **14,** 1062–1069 (2007).
5. Imai, S., Osawa, M., Takeuchi, K. & Shimada, I. Structural basis underlying the dual gate properties of KcsA. *Proc. Natl Acad. Sci. USA* **107,** 6216–6221 (2010).
6. Chakrapani, S. *et al.* On the structural basis of modal gating behavior in K^+ channels. *Nature Struct. Mol. Biol.* **18,** 67–74 (2011).
7. Cordero-Morales, J. F., Jogini, V., Chakrapani, S. & Perozo, E. A multipoint hydrogen-bond network underlying KcsA C-type inactivation. *Biophys. J.* **100,** 2387–2393 (2011).
8. Zhou, Y., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0 Å resolution. *Nature* **414,** 43–48 (2001).
9. Cuello, L. G. *et al.* Structural basis for the coupling between activation and inactivation gates in K^+ channels. *Nature* **466,** 272–275 (2010).
10. Cuello, L. G., Jogini, V., Cortes, D. M. & Perozo, E. Structural mechanism of C-type inactivation in K^+ channels. *Nature* **466,** 203–208 (2010).
11. Levy, D. I. & Deutsch, C. Recovery from C-type inactivation is modulated by extracellular potassium. *Biophys. J.* **70,** 798–805 (1996).
12. Roux, B. & Karplus, M. Ion-transport in a gramicidin-like channel—dynamics and mobility. *J. Phys. Chem.* **95,** 4856–4868 (1991).
13. Bhate, M. P., Wylie, B. J., Tian, L. & McDermott, A. E. Conformational dynamics in the selectivity filter of KcsA in response to potassium ion concentration. *J. Mol. Biol.* **401,** 155–166 (2010).
14. Meyer, R. & Heinemann, S. H. Temperature and pressure dependence of Shaker K^+ channel N- and C-type inactivation. *Eur. Biophys. J.* **26,** 433–445 (1997).
15. Parsegian, V. A., Rand, R. P. & Rau, D. C. Osmotic stress for the direct measurement of intermolecular forces. *Methods Enzymol.* **127,** 400–416 (1986).
16. Rector, K., Jiang, J., Berg, M. & Fayer, M. Effects of solvent viscosity on protein dynamics: infrared vibrational echo experiments and theory. *J. Phys. Chem. B* **105,** 1081–1092 (2001).
17. Pan, A. C., Cuello, L. G., Perozo, E. & Roux, B. Thermodynamic coupling between activation and inactivation gating in potassium channels revealed by free energy molecular dynamics simulations. *J. Gen. Physiol.* **138,** 571–580 (2011).
18. MacKerell, A. D. *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102,** 3586–3616 (1998).
19. Dror, R. O., Jensen, M. O., Borhani, D. W. & Shaw, D. E. Exploring atomic resolution physiology on a femtosecond to millisecond timescale using molecular dynamics simulations. *J. Gen. Physiol.* **135,** 555–562 (2010).
20. Shaw, D. E. *Millisecond-Scale Molecular Dynamics Simulations on Anton* Vol. SC09 (ACM Press, 2009).

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

Acknowledgements This work was supported by the National Institute of Health through grant R01-GM062342 (J.O. and B.R) and R01-GM57846 (S.C. and E.P.). R. Hulse and C. Palka provided purified KcsA. J.O. is a student in the Graduate Program in Computational Neuroscience at the University of Chicago. This research used resources of the Oak Ridge Leadership Computing Facility located in the Oak Ridge National Laboratory, which is supported by the Office of Science of the Department of Energy under Contract DE-AC05-00OR22725. Anton computer time was provided by the National Resource for Biomedical Supercomputing and the Pittsburgh Supercomputing Center (PSC) through Grant RC2GM093307 from the National Institutes of Health and from a generous loan from D. E. Shaw. We are most grateful for the opportunity to use Anton[20] and for the support from R. Roskies and M. Dittrich at the PSC.

Author Contributions J.O. carried out all the final molecular dynamics simulations and 2D-PMF calculations and the computational analysis; A.C.P. initiated the molecular dynamics simulations and the 2D-PMF calculations; B.R. designed and simulated the kinetic models; S.C. and E.P. planned, carried out and analysed the experiments; B.R. conceived and supervised the entire project. All authors contributed to writing 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 B.R. (roux@uchicago.edu).

METHODS

Initial coordinates for the molecular dynamics simulations were taken from the crystal structures 1K4C and 1K4D. Crystallographic waters around the single subunit in each of the two crystal structures were placed around all four subunits in the tetramer. Residues were assigned their standard protonation state at pH 7.0, except for residue Glu 71, which was protonated. The channel was embedded in a bilayer of POPC lipids and solvated in 1 M KCl using the web service CHARMM-GUI (total number of atoms, 45,898). All-atom simulations were run using the CHARMM PARAM27 force field[18] under constant NVT conditions at a temperature of 310 K. Molecular dynamics simulations were carried out on the special purpose computer Anton[20] on loan to the Pittsburgh Supercomputer Center (PSC). A total of 114 molecular dynamics simulations were carried out starting from different initial conditions with respect to the inactivating water molecules, for an aggregate total time of 29.5 μs of molecular dynamics (Supplementary Table 2).

The two-dimensional PMFs (Fig. 3) with respect to the two coordinates r and z (Supplementary Fig. 6) were calculated using NAMD 2.9 (ref. 21), which were carried out on the petaflop supercomputer Jaguar, located at the National Center for Computational Sciences (NCCS). The region of interest in the (r, z) space was covered by a grid of equally spaced umbrella sampling (US) windows. To improve the statistical sampling, the US calculations were performed using Hamiltonian Replica Exchange MD (US/H-REMD) simulations (also called US window swapping)[22]. Initial coordinates for the US windows were taken from the unbiased trajectories carried out on Anton. Initial coordinates for the missing windows were obtained by gradually dragging neighbouring windows along the reaction coordinates to the centre of each of the missing windows. Initial coordinates for the first PMF came from the 17-μs-long unbiased molecular dynamics simulation of the stable, pinched filter (Fig. 2a). Windows taken from this trajectory already contained three inactivating water molecules lodged behind the selectivity filter in the cavities of each of the four subunits. A series of flat-bottom harmonic distance restraints were added between channel atoms and the water oxygen of the inactivating waters to ensure that they remained present behind the filter. The flat portion of the restraints spanned 8 Å above the pore helix of the channel—providing the water molecules freedom to move around a region of space purposefully designed to be larger than the overall size of the cavity. The initial coordinates for the second PMF came from the molecular dynamics trajectory of the recovery process (Fig. 2d), providing windows spanning the conformational transition from the pinched to the conductive filter. Windows taken from this trajectory lacked all inactivating water molecules behind the filter. To maintain the system in this occupancy state during the US/H-REMD simulations, dummy atoms (no non-bonded interactions with the channel) positioned above the pore helix were introduced to repel any water molecule trying to enter into the cavities behind the filter from the bulk region. The first PMF comprised 137 windows whereas the second PMF consisted of 226 windows spanning the complete recovery process. All windows were equilibrated $\geq$0.5 ns before starting REMD. Exchange attempts were made every 1,000 steps (or 1 ps of simulation of time), and neighbouring windows were swapped if the Metropolis Monte Carlo exchange probability was satisfied. US/H-REMD simulations for the two PMFs were run for more than 1 ns. The total aggregate simulation time used to produce the two US/H-REMD calculations is more than 800 ns. Windows were unbiased using the Weighted Histogram Analysis Method (WHAM)[23,24], which only required that the umbrella sampling windows were generated according to Boltzmann statistics.

Electrophysiological measurements were made by patch-clamp recordings in channel-reconstituted liposomes prepared as described earlier[25,26]. Purified protein was reconstituted in asolectin vesicles (in 1:100 (mass:mass) protein to lipid ratio) by dilution with 200 mM KCl and 10 mM MOPS buffer at pH 7.0. Residual detergent was further removed by incubation with biobeads (Bio-Rad Laboratories). Channel-incorporated liposome suspension was then centrifuged for 2 h at 100,000g and the pellet was re-suspended in 60 μl of KCl/MOPS buffer. A drop of the proteoliposome was placed on a glass slide and dried overnight in a desiccator at 4 °C. The sample was then rehydrated with 20 μl of buffer, which yielded giant liposomes. This preparation was suitable for patch-clamp recordings after ~2 h. Currents were recorded in 10 mM MOPS buffer with indicated salt concentration. For measurements under KCl gradient, *N*-methyl-D-glucamine was used to replace KCl in the pipette. Recording pipettes were pulled from thin-walled borosilicate glass and heat polished such that they had a bath resistance of 1–2 MΩ when filled with 200 mM KCl, 10 mM MOPS solution. All measurements in this study were conducted in the inside-out configuration of the patch clamp technique. Experiments were performed at room temperature (20–22 °C). Currents were elicited in response to pH jumps from 8.0 to 3.0 using an RCS-160 fast solution exchanger (Biologic) fed by gravity. Macroscopic currents were sampled at 5 kHz using Axon 200-B patch-clamp amplifier. For each experiment in the absence or presence of sucrose, the fractional recovery was measured at different inter-pulse intervals. There are eight sets of experiment without sucrose, and eight with sucrose. For the two conditions, the average recovery curve and its standard deviation was determined from the eight different sets, and the time constant was determined from a single-exponential least-square fit. To determine if the recovery time constant in the absence and presence of sucrose is significantly different, we fitted each of the eight individual data sets (for the two conditions) by a single exponential and determined the recovery time constant for each experiment. The two sets of eight values were subjected to an unpaired Student's t-test (degrees of freedom = 4), which yielded a P value of 0.03969.

21. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–1802 (2005).
22. Jiang, W., Luo, Y., Maragliano, L. & Roux, B. Calculation of free energy landscape in multi-dimensions with Hamiltonian-exchange umbrella sampling on petascale supercomputer. *J. Chem. Theory Comput.* **8,** 4672–4680 (2012).
23. Kumar, S., Bouzida, D., Swendsen, R. H., Kollman, P. A. & Rosenberg, J. M. The weighted histogram analysis method for the free-energy calculations on biomolecules. *J. Comput. Chem.* **13,** 1011–1021 (1992).
24. Roux, B. The calculation of the potential of mean force using computer simulations. *Comput. Phys. Commun.* **91,** 275–282 (1995).
25. Delcour, A. H., Martinac, B., Adler, J. & Kung, C. Modified reconstitution method used in patch-clamp studies of *Escherichia coli* ion channels. *Biophys. J.* **56,** 631–636 (1989).
26. Cortes, D. M., Cuello, L. G. & Perozo, E. Molecular architecture of full-length KcsA: role of cytoplasmic domains in ion permeation and activation gating. *J. Gen. Physiol.* **117,** 165–180 (2001).

CAREERS

COLUMN Romance between scientists is not just about the two-body problem **p.127**

NATUREJOBS FACEBOOK Science-careers advice and information **go.nature.com/4lzxra**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

OCEAN/CORBIS

CHEMISTRY

Greener pastures

For chemists with an interest in the environment, making a move into green chemistry can be fruitful.

BY NEIL SAVAGE

One night in spring 2005, Laura Muollo stayed up late trying to choose which chemistry graduate programme to attend. Although she had been accepted by big-name universities including Yale, Johns Hopkins and Northwestern, she settled on the lesser-known University of Massachusetts Lowell — even though she would receive a lower stipend. Lowell offered her something that the other institutions could not at the time: the opportunity to focus heavily on green chemistry. "I decided that the name of the school couldn't make me happy, but doing what I wanted would," she says.

By 2009, Muollo had gained a PhD in organic chemistry, with a green-chemistry option that catered to her interest in the environment. Her classes included toxicology, environmental law, and energy and the environment. And even before she graduated, Muollo had found a job at the Warner Babcock Institute for Green Chemistry in Wilmington, Massachusetts, which was founded and is helmed by John Warner, one of the pioneers of green chemistry. The institute designs and develops materials and processes that reduce the use and generation of hazardous substances in industry.

Now one of the directors of research programmes at Warner Babcock, Muollo is working on, among other things, a chemical method for recovering metals from used electronics that is less energy-intensive than conventional means.

Green chemists not only create new molecules, but also take particular account of how those molecules will behave in the environment — whether they will be toxic or otherwise undesirable. When designing a new product, green chemists consider all the stages of its life cycle, from the raw materials to its production and use and what happens when the discarded product breaks down in the environment.

Green chemistry is not so much a separate job track as an addition to general chemistry training. "You can't make a career being a green chemist," says Warner. "You do green chemistry in a career." Aspiring chemists hoping to pursue a career with a green emphasis need to acquire some training in areas beyond conventional chemistry, but relatively few formal master's or PhD programmes in the subject exist. The American Chemical Society (ACS) in Washington DC currently lists 26 institutions around the world that have some sort of green focus (see go.nature.com/keja5h).

CLEAN CHEMISTRY

One aspect of green chemistry involves designing chemical production processes that are more environmentally friendly than those currently available. This does not necessarily mean achieving complete non-toxicity. The aim may be to use less-hazardous solvents or safer methods of synthesis, to produce less waste, to use renewable feedstocks or to develop a more energy-efficient process. And to be acceptable to industry, the greener version has to work as well as — or better than — any existing version, and be no more expensive.

Training in green chemistry is rarely a formal requirement for an industry post. But industrial interest in the field is growing. Chemical manufacturers are likely to look favourably on applicants who are trained to consider the health and environmental impacts of chemical processes and products. Experts in the field say that this view is partly attributable to an increase in government rules, such as the European Union's Registration, Evaluation, Authorisation and Restriction of Chemical Substances (REACH) regulation. REACH, which came into effect in 2007, requires industry to evaluate the risks posed by the chemicals that it uses, and to seek safer alternatives.

Green chemistry has many potential benefits for industry. In principle, safer processes can mean reduced expense for the disposal of ▶

hazardous waste and lower liability costs, as workers and consumers get less exposure to toxic substances. Many consumers are becoming increasingly concerned about the environmental effects of the products they buy, so companies also see green as a means of increasing their share of the market. Employing scientists who can recognize, and correct, problems of toxicity or environmental hazard early in the cycle of research and development can save time and money (see 'Green chemistry in action'). "Having people work for us who have a background understanding and knowledge of green-chemistry process cycles gives us an advantage," says Karen Koster, executive vice-president for environmental health and safety at Momentive, a chemical manufacturer based in Columbus, Ohio.

Koster says that Momentive looks for people who have green-chemistry training as well as some experience in a market that the company deals with, so that they understand the products that the chemicals will go into. Much of the company's work is for the automotive industry — creating speciality coatings for composite materials that serve as lightweight alternatives to metal parts, and helping car makers to meet new emissions standards. Other clients include the construction industry and suppliers of parts for turbines. Someone in a chemical-engineering graduate programme who has had an internship in one of these industries would be an attractive hire, says Koster. "There's more of a focus on applications and markets and less of a focus on 'OK, I need an expert in polymer chemistry and they're going to come in and design molecules,'" she says.

Some companies offer training on the job. At the multinational company 3M, based in St Paul, Minnesota, which has about 80,000 employees worldwide, developmental chemists and engineers can take six two-hour training sessions on toxicology. It is important for the company to train employees who went through graduate education before green chemistry made it into academic programmes, says Robert Skoglund, senior laboratory manager in 3M's medical department. "We're trying to give them an understanding of the hazards of chemicals."

Laura Muollo studied green chemistry in her PhD.

LAUREN PLAVISCH/WARNER BABCOCK INST.

GLOBAL SHIFTS

One study suggests that green chemistry could be a means for US chemical manufacturers to strengthen their global foothold (see go.nature.com/7o8yqy). Plastics production and non-pharmaceutical chemical manufacturing in the United States have seen a drop from about 807,000 workers in each area in 1992 to 626,000 and 504,000 in 2010, respectively. The study's authors, political economists James Heintz and Robert Pollin at the University of Massachusetts, Amherst, argue that focusing on green chemistry could increase jobs across the board, in part by making US products more attractive to environmentally conscious consumers.

Chemical production is growing in China, Singapore and India, partly because of the shift of heavy industry away from the West, says James Clark, head of the Green Chemistry Centre of Excellence at the University of York, UK. China is by far the largest manufacturer of chemicals worldwide, according to the European Chemical Industry Council in Brussels, with US$982 billion in sales in 2011, nearly double those of the United States in second place. Clark has had many students from China who return home after their course to work in the chemical industry. "They find it relatively easy to walk into jobs," he says.

Clark adds that large manufacturers in Europe have not entirely shifted focus to green chemistry, but there are still plenty of jobs for his graduates. Many small companies are developing products that use chemical ingredients derived from biomass rather than from scarcer, carbon-generating petroleum. These firms are mostly start-ups, he says, although some large firms are moving in the same direction — one is DuPont, which is based in Wilmington, Delaware, but has a big European presence.

The Delft University of Technology in the Netherlands provides all its chemistry students with green training to make them more employable, says Isabel Arends, head of the biotechnology department. "All of our students find a job

SUCCESS STORIES

Green chemistry in action

Green chemistry encompasses anything that renders a process or product more environmentally friendly, whether by reducing waste, replacing toxic solvents with less toxic ones, using renewable feedstocks such as biomass, or making more efficient use of inputs such as energy or water. Here are a few examples of green-chemistry successes, some taken from past winners of the US Environmental Protection Agency's Presidential Green Chemistry Challenge.

- Momentive in Columbus, Ohio, has produced a coupling agent for silica-containing tyres that eliminates most of the ethanol released during manufacture. It also improves the tyre's resistance and traction, thus reducing fuel consumption.

- Nike in Beaverton, Oregon, has changed the rubber formulation it uses for making its shoes, reducing, for example, the total zinc content by 80% and replacing solvent-based processes with water-based ones. Nike says it has decreased the amount of hazardous waste generated in making a pair of shoes by more than one-third since 2005.

- BASF in Ludwigshafen, Germany, and Dow Chemical in Midland, Michigan, have developed a process to convert hydrogen peroxide into propylene oxide, a key component in products ranging from domestic appliances and furniture to paint and pharmaceuticals. The companies say that their process reduces waste water by 70–80% and energy use by 35%.

- Codexis in Redwood City, California, has engineered a more efficient enzyme and come up with a low-cost feedstock for making simvastatin, a leading cholesterol-lowering drug. The process cut the number of steps involved in synthesizing the drug and reduced the use of hazardous solvents.

- Geoffrey Coates of Cornell University, Ithaca, New York, developed a family of catalysts to convert carbon monoxide and carbon dioxide efficiently into polymers used in commercial plastics, some of which are biodegradable. Start-up company Novomer in Waltham, Massachusetts, is using the catalysts to produce a polycarbonate coating for use by electronics manufacturers. The production process is expected to use only half the amount of energy typically required to make similar products.

- Genomatica in San Diego, California, has developed a microbe that ferments sugar to synthesize 1,4-butanediol (BDO), a chemical building block for plastics such as spandex. The process is estimated to use 60% less energy and produce 70% less carbon emissions than BDO synthesis from natural gas. N.S.

within two months after they graduate," she says. Some have studied industrial ecology and get policy-oriented jobs, whereas others take manufacturing jobs in pharmaceuticals, polymers or food engineering.

PAESMEM PHOTOGRAPHY

"You can't make a career being a green chemist. You do green chemistry in a career."
John Warner

Students who cannot get onto a formal master's or PhD course that includes green chemistry should "take it upon yourself to get that information anyway", says Warner. One way to do that is through programmes aimed at recent graduates, including a week-long summer school run by the ACS in conjunction with its annual conference. Alternatively, meetings with a green-chemistry focus enable networking and an overview of developments in the field. A Gordon Research Conference in Green Chemistry and an associated Gordon Research Seminar will be held in Hong Kong next July, and the University of Nottingham, UK, last month hosted the 6th International Conference on Green and Sustainable Chemistry. Warner Babcock's non-profit education arm, Beyond Benign, also offers resources (www.beyondbenign.org).

Fortunately, green-chemistry courses are becoming more common. Yale University in New Haven, Connecticut, which Muollo turned down in 2005, now has a Center for Green Chemistry and Green Engineering. Founded in 2007 by current director Paul Anastas, who co-wrote the textbook *Green Chemistry* (Oxford University Press, 2000) with Warner, the centre offers courses on green engineering and sustainable design, product life-cycle assessment, water-resources management and how businesses can become greener.

The field, says Muollo, has grown since 2005. Eventually, she says, green will be integral to chemical manufacturing. "I think green chemistry is going to become chemistry." ■

Neil Savage *is a freelance writer based in Lowell, Massachusetts.*

CORRECTION
The Turning Point about Jason Weber (*Nature* **500,** 493; 2013) wrongly identified Senator Dick Durbin as a Republican. He is a Democrat.

COLUMN
Two-body blessing

Romantic partnerships between scientists offer plenty of benefits, argues **J. T. Neal**.

EMPRIZE/SHUTTERSTOCK

I am frequently amazed by how many of my fellow scientists are unwilling to pursue romantic relationships with other scientists. No matter how many times I tell them how great it is that my wife is also in science, too many have heard horror stories from a mentor or colleague about spouses forced to find jobs with carnivals, impact-factor-related divorces or children lost to the liberal arts.

"Don't you just talk about work all the time?" they ask. What? No! Sure, our dinner conversations sometimes cover the advantages of stimulated emission depletion over confocal microscopy, the finer points of TAL-effector-nuclease design, or whether Reviewer 3 would be more appropriately classified as an Old or a New World monkey (I say New World; you would need a prehensile tail to grasp at that many straws!). But we are just as likely to talk about, you know, the weather. Or that movie — the one we are totally going to see next week. Probably. Or that local sports team.

It is true that science can be hard on a relationship. Late-night data-gathering, grant-application deadlines and manuscript revisions can play havoc with domestic bliss. But who is going to understand these challenges better than another researcher? A scientist-partner will accept you stumbling, bleary eyed, into the house in the middle of a week night without accusing you of carousing or philandering (I spilled 70% ethanol on myself, I swear!). A scientist-partner will sympathize when your undergraduate research assistant kinks the light guide on your tissue-culture scope the day before the lab meeting (I'm looking at you, Ben).

A scientist-partner really understands the myriad daily successes and failures that are unique to science, and either too obscure or too unimaginably mundane for the rest of the world to care about. My wife and I drank champagne when her paper was accepted and ate tiny, expensive French food when I got my fellowship. We spent a week in sweatpants when my paper was rejected for the 16th time, and ate fast food and ice cream when she got scooped. I pack leftovers for her to eat when she gets back from her 1 a.m. experiment. She massages my pipetting arm after 12 straight hours in the tissue-culture hood. It's awesome.

And that is not all. Your scientist-partner can act as a 24-hour sounding board for your next great scientific idea. "Hey honey, what do you think about a forward genetic screen in African elephants? No? I guess you're right. That'll never fly in this funding climate."

"But what about the two-body problem?" you ask. "It's hard enough to get one faculty job as it is." Ah, the dreaded two-body problem. It sounds like some kind of development patterning defect. Sure, getting two science jobs in the same time zone can be tricky, but what about all those Egyptologist couples? The two-body problem is far from unique to academia, or to science in particular, because, no matter how you slice it, balancing two careers is hard (see *Nature* **466,** 1144–1145; 2010). We scientists like to imagine that we are exceptional in all things (and we mostly are), but all relationships, scientific or otherwise, require hard work and patience to succeed. Thankfully, scientists have these in spades. (Wine also helps.)

"A scientist-partner really understands the myriad daily successes and failures that are unique to science."

As for my wife and me, I am not sure how it will all shake out when we finish our postdocs and start looking for positions in academia and/or industry. But I know one thing: I wouldn't trade our situation for anything. Not even a full professorship. ■

J. T. Neal *is a postdoctoral fellow in cancer research at the Stanford University School of Medicine in California.*

THE SCENT OF THINGS TO COME

Breathing is believing.

BY J. R. JOHNSON

JACEY

The big male roared as the door behind me opened. Wade walked onto the observation deck just as a thick, wet splat hit the far side of our one-way mirror. I tried to hide a snigger and failed.

"I don't care what BlackSky's techs say, Elsa, those bastards can sense us through the divider," Wade said, eying the dripping pile of excrement and wrinkling his nose. "This place reeks. He does that every time I come in."

I frowned, but not in disgust. "Really?" I brought up old readouts on my lab tab by flicking a finger. "This is the first time I've seen Phil do that."

Wade snorted. "Phil? You're naming them now?"

Despite the three PhDs that made him my boss, Wade was an idiot. I was an evolutionary biologist tasked with understanding why this dying colony had regressed into something less — or other — than human. Of course I'd named them.

"You'd be pissed too, if you were one of the last of your kind."

A man of limited imagination, Wade just shrugged. Still, he wasn't wrong about the smell.

A mix of proteins, pheromones, gaseous chemicals and some unknown component gave the air a pungent aroma that often defeated even the lab's filtration system. I hadn't noticed it for months.

"Got anything for me?"

Ah. Time for our weekly '(why aren't you making more) progress' meeting, followed by what Wade called research and I called a travesty.

"I'm studying the natives' use of chemosignalling."

Wade sighed but let me continue.

"Their olfactory cortex is huge. It's active in social identification and behaviour, but the data also indicate that odour performs physiological functions. Functions that are in the process of transforming."

His eyes glazed over but I was desperate to discuss the puzzle. I pulled up my research results. "See these curves? Phil is pumping out high levels of hormone elevators. Could this be a reaction to their declining birthrate?"

Wade frowned. "I don't care, and neither should you. We're here to analyse the cave's chemical make-up and exploit it. That's it."

Pompous ass, so what if his last post was lecturer on MIT's flagship vessel? Everyone knew he didn't get tenure.

"You do want to keep your job, right, Elsa?"

I knew what he was about to say and didn't want to hear it.

"If you're going to dig yourself out of this rock and back to civilization, you need a project with legs. Something sexy."

Unlike Wade, I no longer thought about returning to Earth. What I wanted was my reputation back, and work worthy enough to make up for past misdeeds.

I'd find a way to get the birthrate back on track with or without Wade.

When company surveyors reported that the locale's chemistry promised an extensive range of new materials and that the inhabitants were on their last legs, BlackSky couldn't set up fast enough.

Hollowed from the mountain, BlackSky's lab was supposed to let us collect data on Phil and his depleted clan undetected. The plan was a bust. An hour after completion, one of Phil's relatives began knocking on the mirror, which to him should have looked like any other rock face.

The one-way pane overlaid the dim cave view with readouts, but I watched the occupants. Ostensibly my job was to keep the poor creatures from gasping their last — at least until they signed resource-extraction contracts in Galactaspeak. Not entirely ethical, but I needed the job. My gut twisted at the thought of screwing up again.

Phil was cute, in an off-kilter way. I could ignore the extra phalanges and the sensory glands like open wounds at the sides of his neck. Our research showed that they were still genetically quite similar to humans. Close enough, Wade had said, eyeing a well-endowed female.

This was my chance, and after that containment incident with Omega Corp's asteroid lab, well … Trying to save the crew was all that kept me credentialled.

"We're done here," Wade said. "Want to get a drink?"

I didn't bother looking at him, knowing that he would be ogling me with what we in the trade call 'space goggles'.

"I should keep working. Or else, right?" The painful frisson of fear tightening my throat was better motivation than Wade's clumsy threats.

As the door whispered shut, Phil turned and looked straight at me. The nodes on either side of his neck flared a deep red, and a scent more pungent than any yet recorded flooded the sensors. Strong, yes, but it no longer disgusted me. It made me feel … hungry.

I checked the readout and frowned back down at my tablet. The numbers looked off.

The odour wasn't all external.

Hunched over in fear and anticipation, I inhaled the new smell coming from my body, then winced as a rough patch on my neck rubbed against my jumpsuit's collar. Fresh skin, transforming.

That wasn't the only change. I could sense Phil like a beacon through the lab wall, his scent a melody I almost recognized. He stood in silence, then extended his hand.

The self-doubt that had haunted me since Omega melted away; I knew what I needed to do. What I wanted to do.

My tablet chimed with a new message from Wade: "Remember, Elsa, something new, something hot!"

No one saw me cross the hall to the cave's access hatch, or heard the squeal as I jerked it open and slipped inside — to Phil, and to the way I could fix the birthrate problem and stop BlackSky at the same time.

With quarantine alarms ringing, I sent my final message as a BlackSky employee: "Wade. I've found my 'sexy idea'. We'll be in touch." ■

NATURE.COM
Follow Futures:
@NatureFutures
go.nature.com/mtoodm

J. R. Johnson *has a PhD from the Massachusetts Institute of Technology, researches questionable business practices, and fuels her writing with cookies. For more on her latest projects visit jrjohnson.me.*