THIS WEEK

EDITORIALS

WORLD VIEW A call to protect research serving public needs **p.7**

FUKUSHIMA Cleaning up the damaged reactors could take 40 years **p.8**

TRENDS New class of drugs to treat depression takes a second hit **p.8**

Reach out about climate

Where political leadership on climate change is lacking, scientists must be prepared to stick their heads above the parapet.

Consider the following as a statement of national ambition: "The Federal Climate Change Action Plan presents a strategy for launching a transformation in public attitudes and behavior towards climate-change risk. Key state, industry and nonprofit sector allies stand ready to build on the federal strategy to create and sustain a national climate-change risk reduction campaign. The national campaign will increase the public understanding of the risk; advance effective national, state and local climate-change policy; and deliver financing and other incentives to help citizens mitigate climate change. This national climate-change effort — led jointly by the federal government and key national partners — will fundamentally change citizens' expectations and behavior."

That is the wording of a US federal action plan produced last summer, with just one change introduced by *Nature*: the original was not about climate change but referred to indoor radon — a naturally occurring radioactive gas that contributes to lung cancer. Sadly, the altered statement is politically impossible in today's United States. Yet it would be an entirely sensible response to the vastly greater global and local risks posed by climate change as described in the international scientific literature and in national impact analyses conducted by the US government itself. Indeed, it would be a welcome response by any government.

With US politics in gridlock, Europe in financial turmoil and minimal progress at the climate conference in South Africa in December, 2011 was a bad year for political progress in tackling climate change. In addition, surveys of public opinion show a declining belief that climate change is an urgent problem. Clearly, the need to make the public aware of the threat has never been greater. In the face of climate-change contrarians and denialists, some of them with political clout and voices amplified by the media, climate scientists must be even more energetic in taking their message to citizens.

COMMUNICATING RISK

The radon-awareness campaign offers lessons to climate-change communicators. The health risk of radon is unlike the risks of climate change, being uncontroversial, local and directly identifiable. But, like climate change, the risks are not immediately apparent and they are easily ignored. Whether to invest in mitigating measures is the individual's decision, but in the case of radon the US government — like many others — has decided that it has a duty to advise and encourage homeowners to make the changes.

Such campaigns need a strategy for communicating risk that will persuade citizens to spend their own money. Those already involved in risk communication will be familiar with the strategies recommended by the World Health Organization to deal with the dangers of indoor radon: identify core messages, understand and engage with your target audiences — both direct (householders) and indirect (such as teachers and bankers) — through surveys and in-depth discussions, develop information sheets and websites, use trusted networks and ensure that your message is coherently delivered across multiple channels.

So what can climate scientists learn from such strategies? What should their core messages be? Should they relate current trends in local weather to the predicted trends, or show what the 'four-degree-warmer' world — which on current emissions trends lies ahead of us — actually looks like? Either way, there are freely available online resources to call on. Some countries have produced national climate-change impact studies. For example, a 2009 US government report examines both regional and economic sector impacts under high- and low-emission scenarios (see go.nature.com/9fnsk1) in measured tones — here the numbers tell the story.

Those wishing to draw attention to disastrous but entirely possible futures can use reports of an international meeting in 2009 that put together multidisciplinary studies of a world that warms by 4° C or more this century (see go.nature.com/mj8c8f), and on a summary produced by the UK government (see go.nature.com/zu2frk).

"2011 was a bad year for political progress in tackling climate change."

As many scientists as possible should convey these messages through outreach to local or national organizations, the media, in blogs and in policy discussions. Even better if one can be extra-creative and provide people with interactive tools to explore the possible scenarios, such as the energy-pathway calculator launched last month by David MacKay, chief scientific adviser to the UK Department of Energy and Climate Change (see go.nature.com/1wfvnx).

A more taxing and wearying task is to actively counter misrepresentation — whether in the form of crass errors made by politicians and public figures, or more subtle assertions that require detailed examination. The latter can be scientifically revealing, as discussed by climatologist Ben Santer at last month's meeting of the American Geophysical Union (see video at go.nature.com/mwwleu).

Two challenges face those who communicate the science of climate change to the public. The first is to make the messages from models and observations as vivid as possible while maintaining scientific probity — avoiding the blurring of dispassionate discussions of the science and the equally important individual right of advocacy. The second is to find the right ways of conveying uncertainties without losing grip on the central, generally agreed, conclusions. Training in communication is advisable (see, for example, climatecommunication.org). Those who engage with the media could do worse than take on board the maxims of the late Stephen Schneider's 'mediarology' website: know thy audience, know thyself, and know thy stuff (see go.nature.com/dehvsf).

Even if governments find it difficult to achieve the same clarity of national action on climate change as they can for radon, scientists and their organizations need to do more to help citizens engage with the issues and not be misled by travesties of the evidence. Let that be a resolution for 2012. ■

Blue-sky bias should be brought down to Earth

High-prestige research hogs the money, while the needs — and value — of the US science agencies closest to the public are ignored, says **Daniel Sarewitz**.

The Republican majority swept into Congress promising to reduce government spending. One of the first targets was the US Geological Survey (USGS), an agency within the Department of the Interior. The survey had long prided itself on the excellence of its science, but its relatively small budget and low profile made it a good candidate for political sacrifice. The year was 1995.

As 2012 begins, we are entering the most important and decisive period for US science and technology policy since the late 1940s. After 60 years dominated by growing federal expenditure, US science now faces a long period of budgetary stasis, or even contraction. From today's vantage point, we can see the 1995 assault on the USGS as a harbinger of this new era.

The USGS did survive. An important factor in this was the 1,400 state and local organizations that collaborated with the agency to monitor and manage water resources. When these groups let their elected representatives know about the survey's importance for the well-being of the nation, Congress took the USGS off the chopping block. But over the past 15 years, mission agencies such as the USGS that seek principally to serve public goals rather than to advance science have experienced minimal budgetary growth, in some cases not even keeping up with inflation. Since 1996, research funds at the USGS have risen by a mere 16%; at the National Oceanic and Atmospheric Administration (NOAA), 11%; the Environmental Protection Agency, 33%; the National Institute of Standards and Technology (NIST), 38%; and the Centers for Disease Control and Prevention, 45%. Even Department of Defense research has grown relatively modestly, by 60% in 15 years.

Yet, over this same period, government funding for research doubled. Most of the increase went to the National Institutes of Health (NIH) and the National Science Foundation (NSF). The NIH's budget has tripled; the NSF's more than doubled. Together, they captured three-quarters of all the spending increases for federal science. (Although the NIH is in some respects a mission agency, its priorities, its work force and the image it has cultivated focus on fundamental science, a reality acknowledged in director Francis Collins's efforts to create an institute to translate research into useful technology).

Why is this a problem? As the USGS story shows, mission-agency research is directed at, and by, users' needs. Mission agencies provide knowledge — fundamental and applied — to address social problems as diverse as preventing and preparing for natural and technological disasters, providing the scientific basis for environmental monitoring and protection, dealing with public-health threats such as obesity and emerging infectious diseases, and developing the standards and measures that facilitate technological innovation. Indeed, just over a year ago, NOAA director Jane Lubchenco earned a place on the cover of *Nature* for guiding her agency's response to the Deepwater Horizon oil spill.

AS THE CURRENT BUDGET CRISIS UNFOLDS, THE EROSION OF MISSION-ORIENTED RESEARCH IS LIKELY TO ACCELERATE.

But as the current budget crisis unfolds, the erosion of mission-oriented research is likely to accelerate. For example, the spending bill passed in late November increased the NSF's budget by 2.5%, flat-funded NIST and cut NOAA's by 4.3%.

It wasn't supposed to be this way. America's pragmatic culture has long been assumed to favour applied investigation over fundamental science, a notion that goes back at least to Alexis de Tocqueville's nineteenth-century classic *Democracy in America.* And the foundational text of modern US science policy, Vannevar Bush's 1945 report *Science the Endless Frontier*, builds its case on the claim that the government will naturally support applied research, but must be compelled to support basic work.

Why, then, the neglect of the mission agencies? One important reason may be that the leading public voices speaking on behalf of research funding come mostly from the high-prestige frontiers of science, and from the institutions associated with such research — universities, the National Academies, the professional scientific societies, and so on.

Last November, for example, the head of the American Association for the Advancement of Science called for "rethinking the science system" to make the funding of university researchers more efficient (A. I. Leshner *Science* **334,** 738; 2011). This is a worthy goal, but nowhere in his editorial, or in the many similar examples of hand-wringing, is it acknowledged that the main goal of rethinking science should be to ensure that the scientific enterprise continues to meet existing and future challenges to public well-being, not simply to protect science for its own sake.

Defending science for its own sake disproportionately benefits the fundamental-science agencies, which can claim to be doing the most prestigious and therefore the most apparently worthwhile science. In the face of the new budgetary reality, advocacy for science must take a new, strategic approach — one that insists on balance between the fundamental-science agencies and the mission agencies that link science to the public good. Otherwise, the value of the public investment in science will decline right along with the budget. ■

➲ NATURE.COM
Discuss this article online at:
go.nature.com/4xg9w9

Daniel Sarewitz *is co-director of the Consortium for Science, Policy and Outcomes at Arizona State University, and is based in Washington DC.*
e-mail: daniel.sarewitz@gmail.com

SEVEN DAYS *The news in brief*

PEOPLE

Stem-cell arrest

A US university scientist was arrested on 27 December, accused of supplying stem cells for use in unapproved therapies. The US Department of Justice says that Vincent Dammai, a researcher at the Medical University of South Carolina in Charleston, supplied the stem cells to Global Laboratories in Scottsdale, Arizona, without the approval of his university or of the US Food and Drug Administration. Three other men are accused in the case, which may be the first in which a researcher has been charged with supplying stem-cell clinics. See go.nature.com/vvxznh for more.

Lab death charges

Three years after a researcher died following a lab fire at the University of California, Los Angeles (UCLA), her supervisor and the university are facing criminal charges. On 27 December, the Los Angeles district attorney charged organic chemist Patrick Harran and the regents of the University of California

NAVEEN SANGJI

with wilful violation of health and safety standards. UCLA had already paid fines relating to the death of 23-year-old Sheharbano Sangji (**pictured**); the latest charges are thought to be the first instance of criminal prosecution over an accident in a US academic lab. See go.nature.com/cvxyii for more.

RESEARCH

XMRV retractions

Two controversial papers reporting links between chronic fatigue syndrome and viruses have been retracted. Partially retracted last October, the first study to make the connection to the virus XMRV (V. C. Lombardi *et al. Science* **326,** 585–589; 2009) was fully withdrawn on 22 December — unusually, without the full agreement of the authors. This was followed, five days later, by the retraction of a paper that found links to related viruses (S.-C. Lo *et al. Proc. Natl Acad. Sci. USA* **107,** 15874–15879; 2010). See go.nature.com/ixulai for more.

Neutrino test OK

Japan's T2K neutrino experiment is back online, nine months after the Tohoku quake forced it to close down. On 24 December, neutrinos were again generated by a proton accelerator in Tokai; the experiment sends the particles 295 kilometres to a detector in Kamioka. Researchers hope to confirm tentative observations of muon neutrinos changing into electron neutrinos (see go.nature.com/ty8t3i), and are discussing an upgrade that would allow T2K to test the claim that neutrinos can travel faster than light.

POLICY

Fukushima clean-up

It could take 40 years to clean up and decommission the three reactors that went into meltdown at the Fukushima Daiichi nuclear plant, Japan's government said on 21 December. Under a newly approved clean-up plan, plant operators the Tokyo Electric Power Company will remove spent fuel rods in two to three years, try to repair building leaks over six years, and start removing melted fuel in ten years. The plan says that research is needed on what to do with the radioactive waste.

Curbs on mercury

The US Environmental Protection Agency announced on 21 December that it had passed air-pollution regulations that will sharply curb emissions of mercury and other toxic chemicals from some 600 power plants across the country. The rules set limits on mercury emissions for the first time, requiring a 90% reduction from present levels at an estimated cost of US$9.6 billion. See go.nature.com/iasvgw for more.

COMING UP

5–8 JANUARY
To celebrate the 70th birthday of Stephen Hawking on 8 January, theoretical physicists meet in Cambridge, UK, for a conference on 'The State of the Universe'.
go.nature.com/igptyq

10 JANUARY
At the American Astronomical Society's meeting in Austin, Texas, the US National Science Foundation announces a new name for the Very Large Array radio telescopes in New Mexico.
go.nature.com/cobpfu

TREND WATCH

The drug TC-5214 once seemed set to become the first of a new class of therapies that aim to treat depression by targeting neuronal nicotinic receptors. But it has now failed two phase III clinical trials, with the latest negative result announced on 20 December by co-developers Targacept, based in Winston-Salem, North Carolina, and AstraZeneca in London. Two additional clinical trials of TC-5214 against depression are ongoing, but the drug has lost the support of investors (see chart).

DEPRESSION DRUG DECLINE

Targacept's share price fell again in late 2011 after its promising antidepressant TC-5214 failed a second late-phase clinical trial.

SOURCE: NASDAQ

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

NEWS IN FOCUS

KIN CHEUNG/AP PHOTO

More than 17,000 chickens were culled in Hong Kong last month after one was found to have H5N1.

BIOSECURITY

Call to censor flu studies draws fire

Critics say controversy over publication of mutant-H5N1 research highlights biosecurity weaknesses.

BY HEIDI LEDFORD

"I don't like to scare people," says microbiologist Paul Keim. "But the worst-case scenarios here are just enormous."

Keim, who chairs the US National Science Advisory Board for Biosecurity (NSABB), is reflecting on its unprecedented recommendation to censor two scientific papers describing how to make a more transmissible form of the H5N1 avian flu virus. On 20 December, the board said that although the general conclusions could be published, the papers (currently under review at *Nature* and *Science*) should not include "the methodological and other details that could enable replication of the experiments by those who would seek to do harm".

That advice has brought the NSABB flak from all sides. Although the board has said that selected researchers should have access to the full details of the work, some scientists have accused it of trying to curtail academic communication, and suggest that partial publication could hamper research intended to help to defend humans against a flu pandemic. Others complain, however, that the NSABB has done too little, too late, to protect the world from a potentially dangerous man-made pathogen.

Perhaps most worryingly, some argue that the NSABB — a purely advisory body — is not equipped to oversee research that could pose a biosecurity threat. "The NSABB was set up not to do anything," says John Steinbruner, a security expert at the University of Maryland in College Park. "It is just a way of pretending there is some kind of oversight when there isn't."

The roots of the NSABB go back to the 2001 anthrax attacks in the United States, when letters laced with the pathogen killed five people and infected more than a dozen others. The attacks prompted calls to limit the publication of scientific research that terrorists could use to design biological weapons. Scientists in turn feared that the government would step in to regulate sensitive work. "There was a sense, whether right or wrong, that if the community did not act to protect the integrity of science, government would overreach and there would be censorship," recalls Ronald Atlas, a microbiologist at the University of Louisville in Kentucky, who at the time was president-elect of the American Society for Microbiology. That society was one of many that pushed for a system in which researchers would regulate themselves, with the help of an advisory board made up of scientists. The US government agreed, and the NSABB was born in 2004.

"We're being accused of being the bad guys," says Keim, based at Northern Arizona University in Flagstaff. "But most of what we've done is to push back against harsher regulations." Since its inception, Keim says that the NSABB has only been asked to review six papers, including two in 2005 that described the reconstruction of the 1918 influenza virus that is thought to have killed more than 20 million people. In that case, the board recommended that the papers simply be amended to spell out the public-health benefits of the research.

But the stakes are higher for the H5N1 work, because the altered viruses readily spread between laboratory ferrets breathing the same air. If the same were true in humans, the new strains could combine H5N1's high death rate — much higher than the 1918 flu — with seasonal flu's rapid transmission (see *Nature* **480,** 421–422; 2011). Add in uncertainties about the efficacy and availability of vaccines and drugs to combat the virus, and the risk of misuse becomes more frightening than any other case that the board has considered, says NSABB member Kenneth Berns, a microbiologist at the University of Florida in Gainesville.

The editors-in-chief of *Nature* and *Science* have both acknowledged the NSABB's ▶

concerns, but say that they are reserving judgement about whether to censor the papers until the US government provides details of how it will allow genuine researchers to obtain redacted information.

Steinbruner is frustrated that no such mechanism exists. About five years ago, he and his collaborators studied a scenario similar to the H5N1 transmissibility studies, and realized that the government would need to construct a system to disseminate such information to a limited, vetted set of users (J. Steinbruner *et al. Controlling Dangerous Pathogens: A Prototype Protective Oversight System;* Center for International and Security Studies at Maryland, Univ. Maryland, 2007). He says they notified the NSABB of their findings, to no effect.

Keim says that working out how to distribute sensitive information is "not our job", adding that it may fall to the US government's Department of Health and Human Services (DHHS), possibly with guidance from the White House Office of Science and Technology Policy.

Other biosecurity experts say that the NSABB should have been involved in reviewing the H5N1 work earlier. "The time when action is needed is at the grant-application level," says Laura Kahn, a security expert at Princeton University in New Jersey. But the NSABB can only evaluate a project when the DHHS explicitly asks it to do so. Keim and several other board members hope that in the future they will be asked to weigh in earlier.

"All along, we hoped to assist the US government in putting in place a much more far-reaching, proactive oversight system," says David Relman, a microbiologist at Stanford University in Palo Alto, California, and a member of the board since it began. But "everything of this sort comes at some cost, not just in terms of work and burden and time, but also dollars and bodies".

For security experts such as Steinbruner, the best solution is to replace the board, or to supplement its role, with a mandatory biosecurity oversight system. Even some scientists who have argued strongly in favour of self-regulation say that they have been disappointed by the NSABB's performance. "I wanted to see them do more," says Atlas.

He may get his wish. In the course of its deliberations over the H5N1 papers, the NSABB became aware of additional work on H5N1 transmissibility that was nearing publication. Keim says the board is now considering whether to recommend a voluntary moratorium on the publication of such work until the community can discuss further precautions to prevent misuse. He expects the board to vote on this in the next few weeks, and adds: "It is time for us to have a broad and global discussion." ■

D. LONG, SDSS-III

Holes in aluminium plates channel light from individual galaxies to calculate how fast each is receding.

COSMOLOGY

Survey tunes in to dark energy

Sonic yardstick offers a measure of Universe's expansion.

BY ERIC HAND

David Schlegel's tool for exploring dark energy, one of nature's biggest mysteries, is deceptively simple. It is an aluminium plate the size of a manhole cover — or rather, 2,200 of them, each with a specific pattern of holes drilled to match the arrangement of galaxies in a particular section of the sky. Every plate is used once, for an hour, at the prime focus of the 2.5-metre telescope at Apache Point Observatory in New Mexico. When the telescope is pointing at the correct spot, light from each galaxy streams through its corresponding hole. The light is then broken up into its constituent wavelengths and used to clock how fast each galaxy is being carried away from us by the relentless expansion of space.

The study, which began in 2009, will ultimately gather data from 1.5 million galaxies. Its goal is to measure dark energy — a phenomenon thought to be driving the Universe to expand at an ever-increasing rate — and discern whether its influence has remained constant or has varied slightly across billions of years of cosmic history. "The more galaxies we get, the better we do," says Schlegel, an astronomer at Lawrence Berkeley National Laboratory (LBNL) in California and principal investigator of the massive undertaking, known as the Baryon Oscillation Spectroscopic Survey (BOSS).

But BOSS — and the method behind it — is already coming into its own. On 11 January, Schlegel and his colleagues will unveil their initial findings — based on data from 470,000 galaxies — at a meeting of the American Astronomical Society in Austin, Texas. Those data have given them a glimpse of cosmic structure by showing where galaxies clump together, like the crests of giant waves. The structure is a relic of a much younger and smaller Universe, in which acoustic waves reverberated through

concerns, but say that they are reserving judgement about whether to censor the papers until the US government provides details of how it will allow genuine researchers to obtain redacted information.

Steinbruner is frustrated that no such mechanism exists. About five years ago, he and his collaborators studied a scenario similar to the H5N1 transmissibility studies, and realized that the government would need to construct a system to disseminate such information to a limited, vetted set of users (J. Steinbruner *et al. Controlling Dangerous Pathogens: A Prototype Protective Oversight System;* Center for International and Security Studies at Maryland, Univ. Maryland, 2007). He says they notified the NSABB of their findings, to no effect.

Keim says that working out how to distribute sensitive information is "not our job", adding that it may fall to the US government's Department of Health and Human Services (DHHS), possibly with guidance from the White House Office of Science and Technology Policy.

Other biosecurity experts say that the NSABB should have been involved in reviewing the H5N1 work earlier. "The time when action is needed is at the grant-application level," says Laura Kahn, a security expert at Princeton University in New Jersey. But the NSABB can only evaluate a project when the DHHS explicitly asks it to do so. Keim and several other board members hope that in the future they will be asked to weigh in earlier.

"All along, we hoped to assist the US government in putting in place a much more far-reaching, proactive oversight system," says David Relman, a microbiologist at Stanford University in Palo Alto, California, and a member of the board since it began. But "everything of this sort comes at some cost, not just in terms of work and burden and time, but also dollars and bodies".

For security experts such as Steinbruner, the best solution is to replace the board, or to supplement its role, with a mandatory biosecurity oversight system. Even some scientists who have argued strongly in favour of self-regulation say that they have been disappointed by the NSABB's performance. "I wanted to see them do more," says Atlas.

He may get his wish. In the course of its deliberations over the H5N1 papers, the NSABB became aware of additional work on H5N1 transmissibility that was nearing publication. Keim says the board is now considering whether to recommend a voluntary moratorium on the publication of such work until the community can discuss further precautions to prevent misuse. He expects the board to vote on this in the next few weeks, and adds: "It is time for us to have a broad and global discussion." ■

D. LONG, SDSS-III

Holes in aluminium plates channel light from individual galaxies to calculate how fast each is receding.

COSMOLOGY

Survey tunes in to dark energy

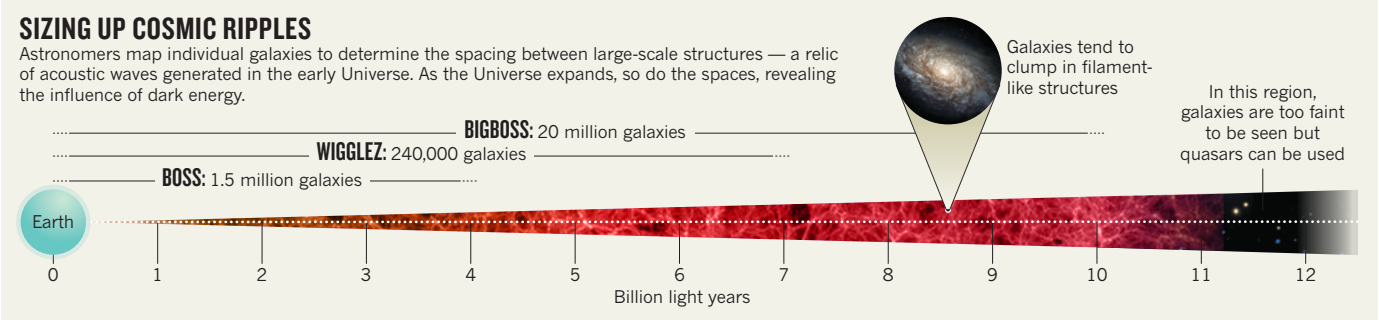
Sonic yardstick offers a measure of Universe's expansion.

BY ERIC HAND

David Schlegel's tool for exploring dark energy, one of nature's biggest mysteries, is deceptively simple. It is an aluminium plate the size of a manhole cover — or rather, 2,200 of them, each with a specific pattern of holes drilled to match the arrangement of galaxies in a particular section of the sky. Every plate is used once, for an hour, at the prime focus of the 2.5-metre telescope at Apache Point Observatory in New Mexico. When the telescope is pointing at the correct spot, light from each galaxy streams through its corresponding hole. The light is then broken up into its constituent wavelengths and used to clock how fast each galaxy is being carried away from us by the relentless expansion of space.

The study, which began in 2009, will ultimately gather data from 1.5 million galaxies. Its goal is to measure dark energy — a phenomenon thought to be driving the Universe to expand at an ever-increasing rate — and discern whether its influence has remained constant or has varied slightly across billions of years of cosmic history. "The more galaxies we get, the better we do," says Schlegel, an astronomer at Lawrence Berkeley National Laboratory (LBNL) in California and principal investigator of the massive undertaking, known as the Baryon Oscillation Spectroscopic Survey (BOSS).

But BOSS — and the method behind it — is already coming into its own. On 11 January, Schlegel and his colleagues will unveil their initial findings — based on data from 470,000 galaxies — at a meeting of the American Astronomical Society in Austin, Texas. Those data have given them a glimpse of cosmic structure by showing where galaxies clump together, like the crests of giant waves. The structure is a relic of a much younger and smaller Universe, in which acoustic waves reverberated through



the dense, hot plasma that had not yet cooled enough to form stars and galaxies. These waves, called baryon acoustic oscillations (BAOs), pushed matter into regions of high and low concentrations with fairly even spacing — a pattern that evolved, in later epochs, into the giant sheets and filaments of galaxies that are the largest structures in the Universe.

First detected in 2005, the regular spacing between these structures creates a natural cosmic yardstick[1]. In today's Universe, that spacing has grown to about 150 megaparsecs (500 million light years). By detecting deviations from this yardstick, BOSS offers the tightest constraint yet on dark energy's influence (see 'Sizing up cosmic ripples').

Surveys of powerful stellar explosions known as type Ia supernovae yielded the first clues to dark energy in 1998. These supernovae are all thought to reach roughly the same peak brightness, allowing them to be used as 'standard candles' to determine the distances to their host galaxies. When these data are combined with measurements of how fast the supernovae are receding, they reveal that the expansion of the Universe is accelerating, rather than slowing down under gravity. One explanation is that dark energy is the 'cosmological constant', a repulsive outward pressure that is innate to the vacuum of space. BOSS is expected to narrow the uncertainty bounds on this model to a few per cent. That's "a factor of a few better than anything we have today", says Schlegel.

Chris Blake, an astronomer at Swinburne University of Technology in Melbourne, Australia, says that "BOSS will be the state of the art for the next five years, there's no doubt about that". Blake was co-investigator on the WiggleZ survey, which studied BAOs by mapping almost 240,000 galaxies using the 3.9-metre Anglo-Australian Telescope at Siding Spring Observatory in Australia. That was enough for the WiggleZ team to announce last year that it had verified, independently of supernova observations, that galaxies are behaving as though dark energy exists[2,3].

But none of the probes of dark energy has been able to say with any certainty that dark energy has behaved as a constant throughout the Universe's history. To address this key question, the BOSS team has set its sights on BigBOSS, a survey that would sample a much larger volume of space, observing 20 million galaxies up to 3.1 billion parsecs away, along with 4 million quasars — the luminous cores of even more distant galaxies. This would allow it to trace the influence of dark energy across most of cosmic history and to discern whether it really has remained constant.

In place of BOSS's aluminium plates, which are laboriously positioned and connected by hand, BigBOSS will rely on an automated system that can manoeuvre fibre tips into precisely the positions needed to gather the light of distant galaxies. The collaboration has submitted a US$70-million proposal to the National Optical Astronomy Observatory to upgrade the 4-metre Mayall telescope on Kitt Peak in Arizona and begin five years of observing in 2018.

The BigBOSS survey would trace the influence of dark energy across most of cosmic history.

Schlegel says that BigBOSS would exceed the BAO-measurement capabilities of much more expensive space missions aiming to study dark energy, such as NASA's proposed $1.6-billion Wide-Field Infrared Survey Telescope. "BigBOSS scoops a lot of what we thought we had to go to space for," he says.

The rise of BAO experiments such as BOSS and WiggleZ comes just as the statistical power of supernova studies is flagging as a result of uncertainties about how consistent the explosions really are. Moreover, the BAO method has more to offer — in addition to measuring the separation between vast clusters of galaxies, it can help to tease out how gravity affects galaxies within the clusters. "We can see the flows of those galaxies into those clusters," says Blake.

This could rule out the possibility that dark energy is something other than a cosmological constant — for example, a change in general relativity at large scales. "It could be that there's something funny going on with gravity," says Josh Frieman, an astronomer at Fermilab in Batavia, Illinois.

Frieman's own tool for finding out is weak lensing, which measures how the gravity of large clumps of matter subtly distorts the shapes of more distant galaxies. Frieman is leading the Dark Energy Survey, which later this year will begin taking pictures of 300 million galaxies with a 570-megapixel camera installed at the 4-metre Blanco telescope at the Cerro Tololo Inter-American Observatory in Chile. But atmospheric blurring, which has hampered earlier efforts to map weak lensing with ground-based telescopes, could prove to be a challenge.

Because the BAO method requires only a galaxy's spectrum and not a high-resolution image, atmospheric blurring isn't as much of a problem. And the same data that BigBOSS would gather to identify the BAO effect on large scales can also be used to observe what gravity is doing around individual clusters of galaxies.

Saul Perlmutter, the LBNL astronomer who shared a Nobel prize for the discovery of cosmic acceleration, says that even if BigBOSS doesn't reveal the nature of the effect, it is bound to deliver some unexpected pay-offs. He points out that his own supernova survey — one of the two that led to the initial discovery — was originally expected to measure the deceleration of the Universe due to gravity. "We're on a roll," he says. "We have to keep pushing." ■

1. Eisenstein, D. J. *et al. Astrophys. J.* **633,** 560–574 (2005).
2. Blake, C. *et al.* preprint at http://arxiv.org/abs/1105.2862 (2011).
3. Blake, C. *et al.* preprint at http://arxiv.org/abs/1104.2948 (2011).

MORE ONLINE

TOP STORY

Citizen scientists' climate-impact survey wraps up go.nature.com/m21xtn

OTHER NEWS

- Spain cuts science ministry in government changeover go.nature.com/79c5th
- Scientists in Chile take the pulse of a shrinking glacier go.nature.com/up2kyi
- Tracking the magnetic south pole go.nature.com/hcuwgy

Z. KOCH/MÉDIA/EARTHWATCH

PROSPECTS

New year, new science

Nature *looks ahead to the key findings and events that may emerge in 2012.*

LET'S TALK ABOUT EARTH

In June, scientists, politicians and campaigners of all stripes will flock to Rio de Janeiro, Brazil, for the United Nations' fourth Earth summit, devoted to sustainable development and the green economy. The conference — undoubtedly the major environmental meeting of 2012 — comes 20 years after the UN Framework Convention on Climate Change was signed at the first UN Earth summit, also in Rio.

THE SOURCE OF MARTIAN METHANE

NASA's car-sized rover, Curiosity, is set to arrive on Mars in August. The US$2.5-billion craft will be lowered by an innovative landing system — the 'sky crane' — into Gale crater, where it will study rock strata in a bid to unpick the red planet's watery past. It will also sniff for methane in Mars's atmosphere, and could reveal whether the gas is being produced by geological processes — or by microbial martian life. Farther afield, NASA's Kepler mission surely ought to find a true extrasolar twin for Earth, with just the right size and orbit around a Sun-like star to be habitable.

ROBOTS, BRAINS OR GRAPHENE?

Six visionary research proposals will vie for huge grants from the European Commission's Future and Emerging Technologies Flagship scheme. The two winning projects, to be announced in the latter half of the year, will each receive €1 billion (US$1.3 billion) over the next decade. In the running are projects on graphene, the promising new form of carbon; robot companions for the lonely; planetary-scale modelling of human activities and their environmental impact; autonomous energy-scavenging sensors; ways to apply research data more efficiently in health care; and a supercomputer simulation of the brain.

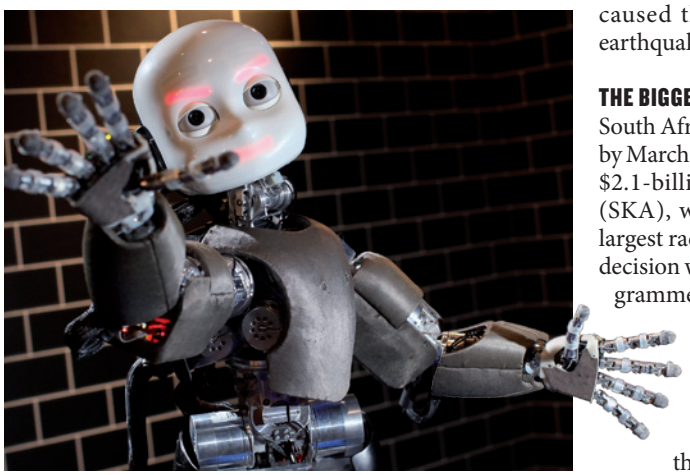
MAJORANA MYSTERY

The Large Hadron Collider, the giant particle accelerator at CERN, near Geneva in Switzerland, will gather enough data this year to either confirm or rule out the existence of the simplest form of the Higgs boson, a key part of the mechanism that is thought to confer mass on other matter. A riskier bet would be on physicists finding an example of a Majorana fermion, hypothesized to be massless, chargeless entities able to serve as their own antiparticles, which could be useful for forming stable bits in quantum computing. Experiments have suggested that in materials known as topological insulators, the collective motions of electrons create a quasiparticle that behaves like a Majorana.

➲ NATURE.COM
For a look back at the highlights of 2011, see:
go.nature.com/2011

DNA ENCYCLOPEDIA

Biologists know that much of what was once termed 'junk' DNA actually has a role. But which sequences are functional — and what do they do? The best answer so far will come with a major update from the US National Institutes of Health's ENCODE (Encyclopedia of DNA Elements) project, which aims to identify all the functional elements in the human genome.



Will iCub and his robot friends win a €1-billion grant this year?

O. SCARFF/GETTY IMAGES

PHARMACEUTICAL PROMISE

Two monoclonal antibodies to treat Alzheimer's disease — solanezumab and bapineuzumab — would be a big hit if they reported positive results from phase III trials in 2012. Both bind to the amyloid-β peptides that make up the protein plaques seen in the brains of people with the disease. Meanwhile, the US Food and Drug Administration will once again consider the thorny issue of approving obesity drugs: it rejected one last year because of worries over side effects. It will also decide whether to approve a pioneering drug for cystic fibrosis, ivacaftor, made by Vertex Pharmaceuticals of Cambridge, Massachusetts. The drug works only for people with a particular genetic mutation, but would be the first to treat the disease's underlying cause, rather than its symptoms. And blockbuster drugs will continue to lose patent protection, including the anticlotting Plavix (clopidogrel) and the antipsychotic Seroquel (quetiapine).

RAIDERS OF THE LOST LAKE

Within weeks, Russian researchers hope to finish drilling through Antarctica's ice sheet to reach Lake Vostok, a huge freshwater lake roughly 3,750 metres beneath the surface. It's a race against time: 10–50 metres of ice separate the team from its goal, which it must reach before the last aircraft of the season leaves in February. There'll be more drilling research in April, when Japan's *Chikyu* ship sets sail to bore into the underwater fault that caused the magnitude-9.0 Tohoku earthquake last year.

THE BIGGEST ARRAY

South Africa and Australia will find out by March which of them might host the $2.1-billion Square Kilometre Array (SKA), which would be the world's largest radio telescope if it is built. The decision will be made by the SKA's programme development office in Manchester, UK. Meanwhile, the Atacama Large Millimeter/Submillimeter Array in Chile's Atacama Desert should be 60% complete by the end of the year.

SPACEFLIGHT ADVANCES

In February, SpaceX of Hawthorne, California, hopes to be the first commercial firm to fly an unmanned cargo craft to the International Space Station — a milestone in private spaceflight. As for government space efforts, China, brimming with confidence after last year's docking of the unmanned spacecraft Shenzhou-8 with the experimental module Tiangong-1, hopes to send astronauts up for a manned docking manoeuvre this year.

A USEFUL SYNTHETIC GENOME

Synthetic biologists can build entire genomes from scratch, working from natural models, and they can also rewire the genetic circuitry of living things. But so far, no one has united the two approaches: Craig Venter's synthetic genome of 2010 was cribbed wholesale from a bacterium and contained no new genetic circuitry beyond a DNA watermark. Might 2012 see the first really useful artificial genome? ■

COMPILED BY RICHARD VAN NOORDEN

P. GILLOOLY/MIT NEWS

Q&A Nancy Hopkins

A passion for science without barriers

In 2005, Nancy Hopkins sparked a firestorm of controversy by walking out during a presentation by Larry Summers, then-president of Harvard University in Cambridge, Massachusetts. He had suggested that innate differences might account for the lack of women in high-achieving roles in science, and Hopkins, a well-known champion of gender equality in science, wanted to register her outrage. This May, the biologist at the Massachusetts Institute of Technology (MIT) in Cambridge, whose career has mirrored the growth and diversity of molecular genetics, will retire from laboratory research and from most public speaking on women and science.

You worked in Nobel laureate James Watson's lab at Harvard University as an undergraduate. What did you take away from that experience?
Jim would come in at teatime and talk about what he had learned that day. It was so breathtakingly exciting you couldn't ask for more. It was Jim who pushed me to pursue a career as a scientist and professor. Later, when I was at MIT and confronting obstacles, I realized how much his influence had protected me from overt discrimination because he was such a powerful person.

Are we in a new era when it comes to women advancing in science?
We are in a new era in that women now expect to have careers that are equal to men's. But are they making it to the top? Not at very high rates. It is still a man's world in terms of who is in charge. We all believe that science is a meritocracy, that being a great scientist takes brains and hard work, but that isn't the whole story.

You changed your research focus many times. What was the driving force behind those shifts?
Part of the reason is the astonishing speed at which science moves. When I started, the focus was on the secret of life, discovering how genes and their expression are regulated in viruses. By the time I finished my PhD in 1971, you could actually work on the molecular biology of cancer, so I switched. I eventually became disillusioned about how I and other women were being treated in the cancer field, and I switched again, to study the genetics of early vertebrate development, a field with more women. Different fields of biology have different levels of acceptance for women.

What motivated the landmark 1999 study that you championed on gender bias at MIT?
It was a final straw. At age 50, after 20 years as a scientist, I found that it was impossible for me to get the supplies and lab space I needed to do my work. I thought, 'I'm not going to tolerate this any more'. I told another woman, and she said she was experiencing the same thing. We talked to other women scientists at MIT and realized that we were all hitting the same roadblocks.

What was the reception like from your male colleagues when you began to speak out?
It was not too good. Some in my department were downright hostile and that was a problem for me. But suddenly I had eight women and four men on the study committee, and also Robert Birgeneau, then dean of science, whom I could actually talk to about this.

Reports like this often fall on deaf ears. What made the difference here?
We put in a huge amount of work, 5 years, just to try and understand the scope of the problem, before we wrote the report. The internal report on which the 1999 report was based was very long, and detailed enough that if you were a scientist and read it, you could see how what was happening to us would make your life as a scientist very hard. It was also a miracle that Charles Vest [then president of MIT] realized that there was a problem and was willing to publicly endorse the report.

Walking out during Larry Summers' talk was a decisive act. Are there other times you wish you had acted about gender equality but did not?
This is an issue I've been grappling with this year as I look back on my life. Why did I allow myself to be treated this way for all those early years? I think that when you are young, you just have this driving energy, you love the science so much and you find excuses to put up with it all because you want to do the science.

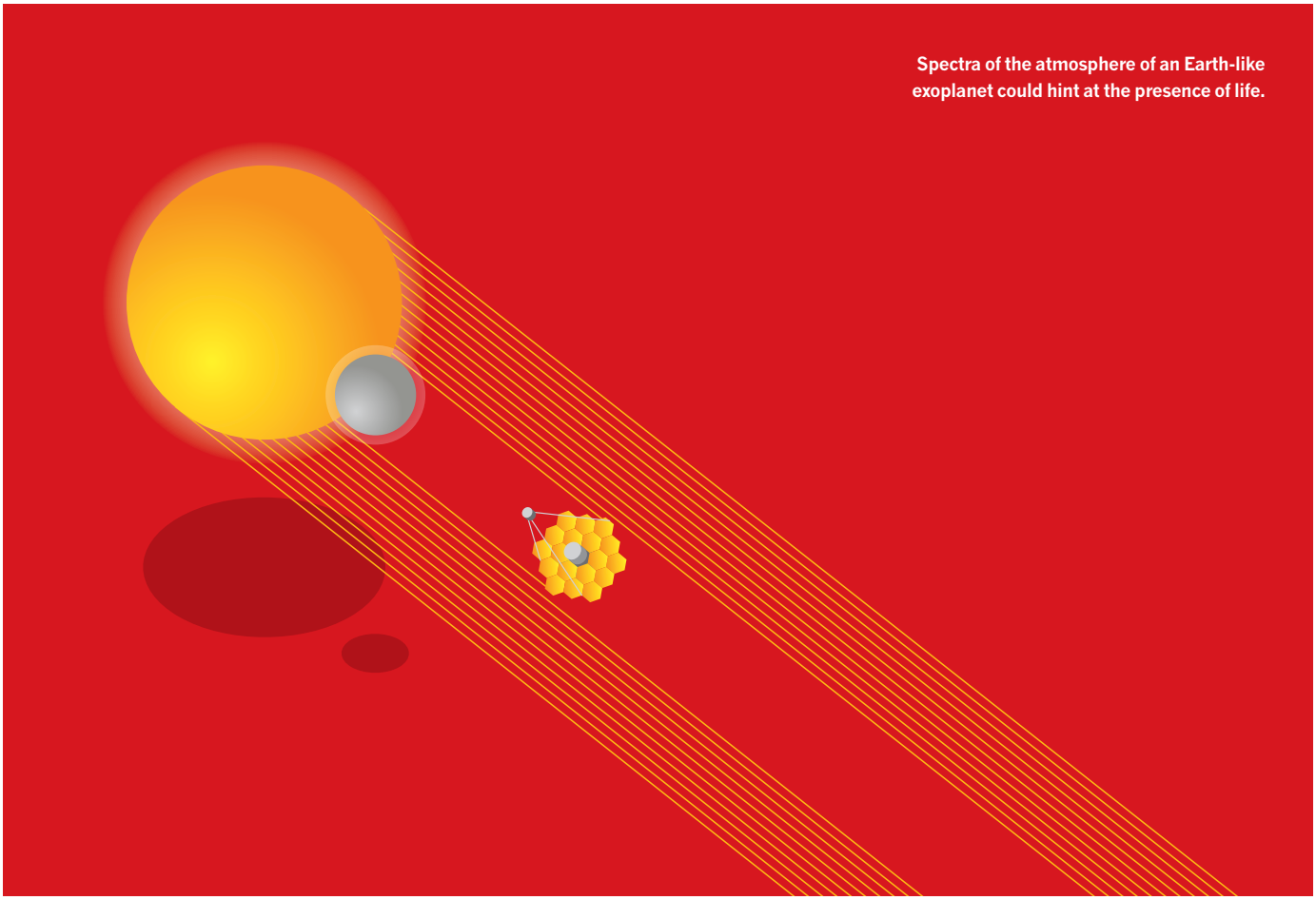
Do you think others should speak out about discrimination?
I increasingly have trouble with established women who don't speak out. At one point I asked Harvard president Drew Faust to give a talk to students about how Larry Summers' comments about women being genetically inferior were erroneous. She wouldn't do it and that bothered me. ■

INTERVIEW BY ADRIANNE APPEL

CORRECTIONS

The News Feature 'The long shot' (*Nature* **480,** 308–309; 2011) gave the wrong affiliation for Adam Burrows. He is chair of the US National Research Council's physics and astronomy board.

The News Feature 'The cell division' (*Nature* **480,** 310–312; 2011) incorrectly stated that Ireland's government does not allow research on human embryonic stem cells. In fact, the government only does not allow it to be funded with public money.



Spectra of the atmosphere of an Earth-like exoplanet could hint at the presence of life.

ILLUSTRATIONS BY THOMAS POROSTOCKY

TOUGH SCIENCE

Five experiments as hard as finding the Higgs.

BY NICOLA JONES

As the media spotlight shines on the Large Hadron Collider in Geneva and its high-profile hunt for a certain boson, other scientists are pressing forward with experiments that are just as challenging — and just as potentially transformative.

These often unsung researchers are willing to spend years or even decades getting a finicky instrument to run smoothly; setting up proper controls to minimize spurious results; beating back noise that threatens to swamp their signal; and striving for an ever more painstaking level of precision — a determination and single-mindedness that borders on heroic. Here, *Nature* describes five such quests.

SPOTTING DISTANT SIGNS OF LIFE

Back in 1999, when David Charbonneau was a graduate student at Harvard University in Cambridge, Massachusetts, he became the first person to measure the tiny dimming caused by the passage of a planet from another solar system across the face of its parent star. Today, such 'transits' are a routine way for astronomers to discover planets. The tricky part is working out what they and their atmospheres are made of. If the atmosphere turns out to contain oxygen, for instance, that could be an indication of the presence of life. But the only way to detect such elements is to find them in the spectrum of the starlight that passes through the planet's atmosphere — a signal that is ridiculously small.

To begin with, explains Charbonneau, "the fraction of light that the planet blocks is tiny". A planet the size of Jupiter passing in front of a star like the Sun would block about 1% of the light; and a smaller, Earth-size planet would block about 0.01%. "Then you look at this tiny onion skin around the planet: that's the atmosphere,"

Charbonneau says. Only the starlight that passes through that onion skin will have the spectral information that astronomers are looking for — and that's less than one photon in a million for a Sun-like star and a planet the size of Earth.

Although no telescope today has anywhere near the sensitivity required to extract a signal that small from the glare of the star itself, Jupiter-scale gas-giant planets have much bigger atmospheres than Earth-sized ones and a correspondingly bigger spectral signature, says Charbonneau. Orbital observatories such as the Hubble and Spitzer space telescopes have been able to extract atmospheric spectra for about 40 gas giants, all since 2005. Although the initial observations met with scepticism, says Charbonneau, "for the gas giants, it's now not quite commonplace, but not controversial. Now it's all about Earth-like planets, but no one has done that yet." The closest researchers have come is examining the spectra of a super-Earth — called GJ 1214b — that has a radius about 2.6 times that of Earth's and is circling a relatively small star not too far from the Sun. The first work on this planet implied that it had an atmosphere full of water vapour or clouds; observations by Charbonneau and his colleagues using Hubble confirmed this a few months ago[1].

Detecting the components of the atmosphere of an Earth-like planet around a Sun-like star — affording the best chance of detecting biological activity on another planet — requires a step-up in sensitivity. Charbonneau is crossing his fingers and hoping that NASA's long-planned and much-delayed Hubble successor, the US$8-billion James Webb Space Telescope, now scheduled for launch in 2018, will indeed reach orbit. "That would be fantastic," he says. "That would give us an honest shot at finding life on other planets."

SEEING THROUGH THE MOLECULAR MIRROR

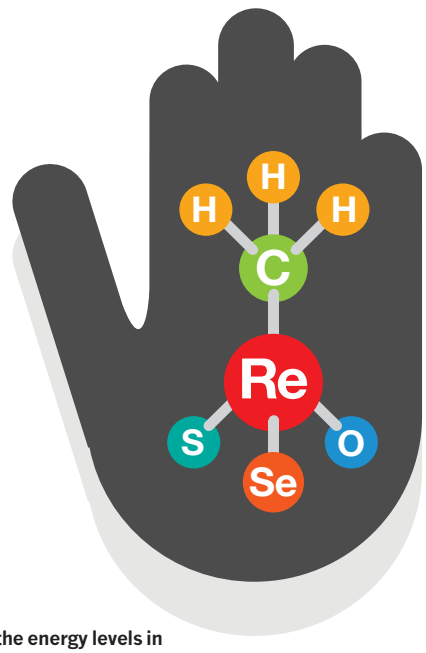
Biology has a curious lopsidedness. Many molecules are 'chiral', meaning that their atoms can be arranged in two forms that are mirror images of each other. When making such molecules in the lab, chemists typically get a mix of both forms, which, by convention, they label as right- or left-handed. But living cells are generally made from the left-handed versions only. No one knows why.

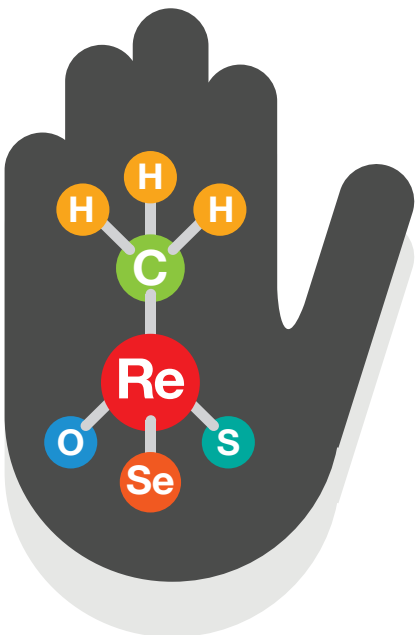
One possible explanation lies in the fact that one of the four fundamental forces in nature predicted by the standard model of particle physics — the 'weak' force that mediates certain interactions between nuclei and electrons — affects left- and right-handed molecules differently. The other forces, which include gravity, are the same in either version of a mirror universe. In theory, explains Benoît Darquié at the University of Paris 13 North in Villetaneuse, France, the weak force should cause the energy states in one form of a chiral molecule to be ever so slightly different from those in its mirror image twin — typically by just one part in 10^{15} or 10^{20}. So if one form had a vibrational frequency of, say, 30 terahertz, its partner's should differ by just a few milli- or even microhertz.

"THE HARDER IT IS, ALL THE BETTER IT FEELS WHEN YOU GET THERE."

Measuring such tiny differences could shed light on the biological lopsidedness conundrum, says Darquié, and his group is attempting to do just that. It could even fill in the values of certain parameters in the weak-force theory part of the standard model.

He and his colleagues are the only ones in the world pursuing this goal, as far as Darquié knows. Indeed, it took him a full three years to assemble the consortium of experimental physicists, quantum theoreticians and chemists he needed. They now need to crack two problems. First, they need to build extremely high-resolution spectrometers to measure the energy levels of chiral molecules. Their best instrument to date can discern energy differences as small as 5 parts in 10^{14} — about a million times better than the resolution of an off-the-shelf spectrometer. They are now building one that will be even more precise. To achieve such sensitivity, their machines need to be isolated from any external vibrations and maintained at a temperature that is steady to within 0.1 °C. And to measure the molecular vibrational frequencies with the required level of precision, Darquié's lab uses a molecular





Tiny differences in the energy levels in mirror-image molecules could hint at symmetry-breaking weak interactions.

clock linked by a fibre-optic cable to the world time standard atomic clock in Paris.

The researchers' second challenge is to create test molecules in which the asymmetrical effect is large enough to be measurable. Such a molecule needs to have a large central atom, because atomic theory says that will maximize the energy differences between the chiral forms, and must not break apart when heated to the gaseous state necessary for spectroscopy. The team is betting that the best molecule will be something like methyltrioxorhenium that has had two of its oxygen atoms replaced by sulphur and selenium, although the researchers have struggled to make that particular molecule in purely left- or right-handed forms. Even if the researchers find a molecule that works perfectly, they will still need another year to take enough measurements to bump up the signal-to-noise ratio and get a trustworthy number.

What if the experiment doesn't solve the puzzle of biological handedness? Darquié says that it won't bother him much, because the techniques they are developing will open up a lot of ways to test the theories of fundamental physics. "Most of the accurate tests are done at high energy with particles, or at lower energies with atoms," he says. "Molecules are more complex, so give access to more complex questions."

LOOKING FOR EXTRA DIMENSIONS

It is an aspect of reality so fundamental that most us can't imagine anything different: the world has precisely three spatial dimensions — left–right, forwards–backwards and up–down. But superstring theory and other attempts to devise a 'theory of everything' have led many physicists to propose that space has many more than that. These extra dimensions would presumably be curled up very tightly, and thus hidden from everyday experience. But they would affect gravity at very small scales, producing a force between two masses that differs ever so slightly from that predicted by Newton's classical law of gravity. An experiment able to detect changes in gravity at that scale might therefore be able to 'see' any other dimensions.

Eric Adelberger at the University of Washington's Center for Experimental Nuclear Physics and Astrophysics in Seattle first heard about this idea at a talk back in 1999. "Some people thought it was crazy; some thought it was really cool," he says. But he and his colleagues decided they had to test it. "What more exciting thing can you do than discover that our understanding of the dimensionality of the world has been wrong forever?" he says.

The team's tool of choice is a torsion balance — essentially an update of the equipment used by the English physicist Henry Cavendish to make the first laboratory measurement of gravity in the late 1790s. In the modern version, a metal cylinder hangs from a thread, allowing the cylinder to twist freely. Attached to the bottom of the cylinder is a disk called the detector, which has a ring of holes drilled in it. A second disk, with similarly drilled holes, sits just micrometres beneath this. When this second disk, called the attractor, is rotated, the material between its holes exerts a tiny gravitational force on the material between the holes in the detector. The force twists the thread that supports the cylinder, causing it to rotate by an amount measured in billionths of a degree.

Micrometre-scale rotations in the absence of gravity could hint at extra dimensions.

To make sure that the detector is responding to gravity and nothing else, the equipment has to be made entirely from non-magnetic materials, and all the surfaces need to be coated in gold to spread out any electrical charges on the device. The device also has to be machined to perfection and protected from all vibrations, including cars driving into the car parks outside. "We get our best data on weekends from midnight to 4 a.m.," says Adelberger. "It's frustrating. The amount of time you spend actually getting good data is very small. It's all detective work."

Tweaks to the design allow the experimenters to cancel out the force expected from Newton's law and isolate the deviations: if the detector spins anyway, they know that something funny is going on. And so far, Adelberger's group can say definitively that there are no extra dimensions larger than 44 micrometres. Two of his graduate students, as well as a handful of other groups around the world, are trying to push that limit down, he says. But how long it will take them to spot something depends on the size of the elusive dimensions. If they're curled up too tightly, he says, "the answer is never. If there is one at 30 micrometres, it'll be a year".

But Adelberger seems to thrive on the uncertainties and difficulties involved. It's like getting to the top of a mountain, he says. "The harder it is, all the better it feels when you get there."

CATCHING A GRAVITY WAVE

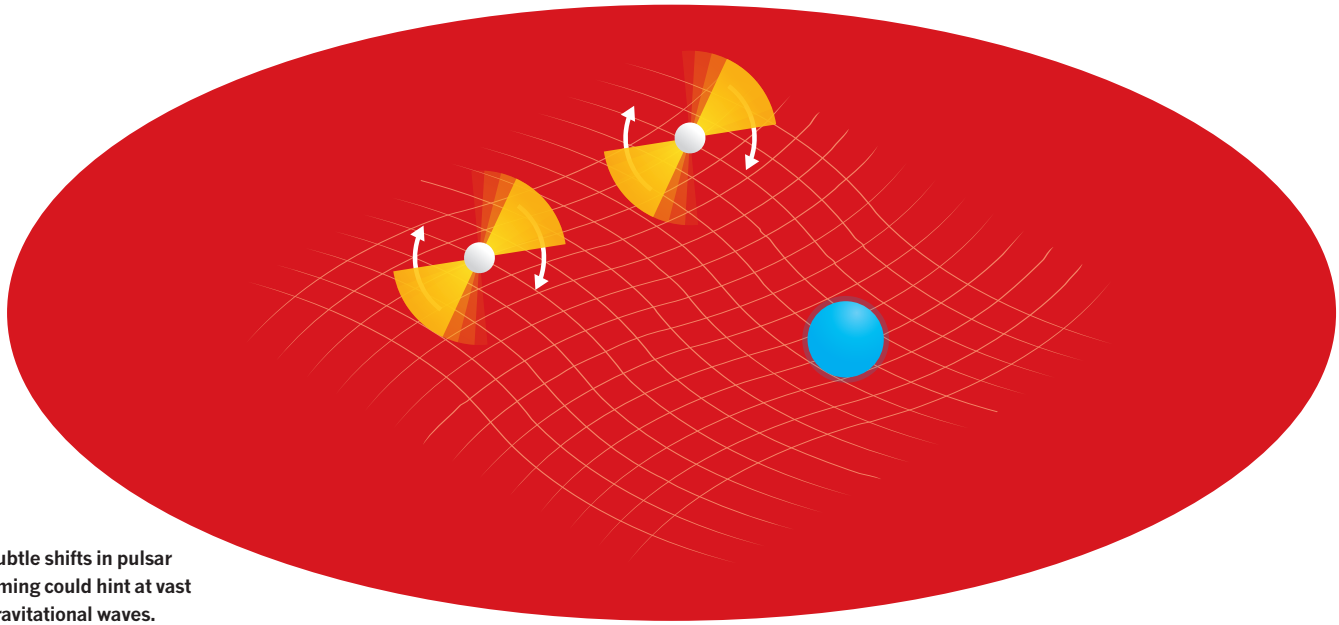
Scott Ransom has a boyish energy that seems mismatched with his subject: a project that may take a decade to produce its first result. Ransom, an astronomer at the National Radio Astronomy Observatory in Charlottesville, Virginia, uses a rapid-fire stream of words such as "awesome" and "cool" as he talks about the Galaxy's most precise natural clocks — pulsars — and how they might allow him and others to detect one of the most fundamental predictions of Einstein's general theory of relativity: gravitational waves. "It will open a whole new window on our Universe," he exclaims. "We will be able to see with mass instead of light."

According to Einstein, explains Ransom, gravity waves are ripples in the fabric of space-time caused by the movement of mass — an orbiting pair of neutron stars, for example. It's just like jiggling an electron, which causes ripples in the surrounding electric and magnetic fields to spread out as light and other forms of radiation. "When you jiggle something massive," he says, "you give off gravitational waves".

Unfortunately, even a very big gravitational wave washing over Earth would squash and expand the planet's diameter by only 10 nanometres or less. Ground-based experiments attempting to detect such tiny disturbances, such as the Laser Interferometer Gravitational wave Observatory run by the California Institute of Technology in Pasadena and the Massachusetts Institute of Technology in Cambridge, are forever trying to distinguish genuine signals from background noise caused by passing trucks, thunderstorms and even the fall of waves on a beach a hundred kilometres away.

So Ransom and his fellow enthusiasts are taking what they hope will be a cheaper path: looking at pulsars. Some of these ultra-dense stars rotate thousands of times a second, each time emitting a flash of radiation that astronomers can time to within about 100 nanoseconds. The team hopes to monitor about 20 such pulsars spread all over the sky to look for deviations in their timing caused by very-low-frequency gravity waves contracting or expanding the space-time between them and Earth. They expect that one of the strongest sources of such waves is the years-long dance of massive black holes in distant, colliding galaxies.

Ransom is one of about a dozen people devoted to this quest, which



Subtle shifts in pulsar timing could hint at vast gravitational waves.

is coordinated by the International Pulsar Timing Array consortium. The good news is they haven't needed to invent any instruments: facilities such as the Arecibo radio telescope in Puerto Rico can do the job. The bad news is that the pulsars need to be monitored for around 10 years to catch the gravitational waves from those orbiting black holes. So far, they have accurate timing measurements for about 5 years on just 6 pulsars.

Still, says an upbeat Ransom, "the cool thing is that our chance of discovery goes up dramatically with time. As long as we're patient, we will see gravitational waves."

REDEFINING THE KILOGRAM

The mass of one kilogram is meant to be an unvarying constant. Yet it actually changes, thanks to an old-fashioned way of defining it as the mass of a more-than-120-year-old cylinder of platinum and iridium that lives in a vault in the outskirts of Paris. No one knows if 'Le Grand K' is getting heavier as atoms are added to its surface, or lighter as atoms are rubbed away, but its mass is certainly drifting: copies that once had precisely the same weight now have measurably different weights.

"We need to tidy things up," says Jon Pratt, an engineer at the US National Institute of Standards and Technology (NIST) just outside Washington DC, one of a number of metrologists working on a redefinition. The kilogram is the only fundamental unit of measure still defined by a physical object, he says.

The basic idea is to pin the kilogram to a precisely measured fundamental physical constant, in much the same way that the metre is now defined in terms of the speed of light in a vacuum: it's the distance that light travels in precisely 1/299,792,458 seconds. To do this for the kilogram would mean fixing Planck's constant, h, which reflects the size of energy quanta in quantum mechanics and is famously linked to energy through the frequency of light: $E = h\nu$. Combining that equation with the even more famous $E = mc^2$ then leads to a definition of mass.

Determining a precise value for Planck's constant is fussy work, however, and the two methods currently in favour disagree with each other enough to keep the redefinition of a kilogram on hold.

One of these ways makes use of a 'watt balance'. In essence this is a simple set of scales: on one side it has a 1-kilogram mass — standardized carefully against the one in Paris — and on the other it has a current-carrying coil of wire immersed in a magnetic field. The field is tweaked until the weight of the mass is balanced by the electromagnetic force on the coil, which can then be linked through a string of equations to Planck's constant. But in practice things are not that simple. Researchers still have to measure other things — the local gravitational field, for instance, the biggest source of error — and avoid any kind of vibration.

In 2007, a Watt balance now run by Pratt produced one of the most precise measurements of Planck's constant — $6.62606891 \times 10^{-34}$ J s, with a relative uncertainty of 36 parts per billion[2]. But another instrument, built at the National Physical Laboratory (NPL) in Teddington, UK, and now located at the National Research Council's Institute for National Measurement Standards in Ottawa, Canada, has yielded a result that differs from NIST's by an amount that is small, but just outside the experimental error[3].

The other favoured approach is to count the number of atoms in a sample of isotopically pure material. That would determine the value of the Avogadro constant — the number of atoms in exactly 12 grams of carbon-12, say — which can be linked mathematically to Planck's constant through another string of equations. In 2008, scientists at the Federal Institute of Physical and Technical Affairs in Braunschweig, Germany, began working with two near-perfect 1-kilogram spheres that had been fashioned from 99.995% pure silicon-28. Since then they have been using high-precision laser interferometry to determine the spheres' volumes, and X-ray diffraction to determine their crystal structures so that they can count the atoms with ever-more accuracy. So far they have measured the Avogadro constant as $6.02214082 \times 10^{23}$ with a relative uncertainty of just 30 parts per billion[4]. The translation of that into Planck's constant agrees with the NPL's watt-balance result, but not with the NIST's.

As of 2010, the recommended value for Planck's constant is $6.62606957 \times 10^{-34}$ J s, with an uncertainty of 44 parts per billion. Some say that's good enough to use to redefine a kilogram. But others want to keep picking away at it until the numbers agree better with each other and have a smaller range of error — to within 20 parts per billion.

That could take quite a while, says Pratt. "These are hard measurements to make. That's just the way it is." ■

Nicola Jones *is a freelance reporter based near Vancouver, Canada.*

1. Berta, Z. K. *et al.* http://arxiv.org/abs/1111.5621 (2011).
2. Steiner, R. L., Williams, E. R., Liu, R. & Newell, D. B. *IEEE Trans. Instru. Meas.* **56,** 592–596 (2007).
3. Robinson, I. A. & Kibble, B. P. *Metrologia* **44,** 427–440 (2007).
4. Andreas, B. *et al. Metrologia* **48,** S1–S13 (2011).

The sounds of the stars

Data from NASA's Kepler space telescope have revolutionized the search for planets outside the Solar System — and are now doing the same for asteroseismology.

BY RON COWEN

SHUTTERSTOCK

Most astronomers gaze at the heavens and see stars. William Chaplin hears an orchestra — a celestial symphony in which the smallest stars are flutes, the medium-sized ones are trombones and the giants are reverberating tubas.

The sounds are internal vibrations that reveal themselves as a subtle, rhythmic brightening and dimming of a star, explains Chaplin, an astrophysicist at the University of Birmingham, UK, and a specialist in asteroseismology. These waves provide information that astronomers can't get in any other way: triggered by the turbulent rise and fall of hot gases on the star's surface, the vibrations penetrate deep into the stellar interior and become resonating tones that reveal the star's size, composition and mass (see 'Celestial music'). So by watching for the characteristic fluctuations in brightness, says Chaplin, "we can literally build up a picture of what the inside of a star looks like".

Better still, he adds, asteroseismologists are now hauling in the data wholesale. After years of being hampered by Earth's turbulent atmosphere, which obscures the view of the Universe and has limited asteroseismology to about 20 of the brightest nearby stars, researchers have been astonished by the trove of information coming from a new generation of space observatories. Thanks to the French-led Convection, Rotation and Planetary Transits (COROT) space telescope, launched in 2006, and NASA's Kepler space telescope, launched in 2009, they can now listen in on hundreds of stars at a time.

"We are in a golden age for the study of stellar structure and evolution," says Hans Kjeldsen, an astronomer at Aarhus University in Denmark.

"Nature seems to have been kind to us," agrees Ronald Gilliland, an astronomer at Pennsylvania State University in University Park. "The stars seem not to be shy about showing us lots of oscillations that will allow us to reveal their innermost secrets." The flood of data has shed light on the interior of red-giant stars, and forced astronomers to question their understanding of how stars and galaxies form.

Stellar serendipity

Asteroseismology isn't the main mission of either COROT or Kepler: they are intended to hunt for planets outside the Solar System (exoplanets) that have roughly the size and orbital radius of Earth. But because they both look for the tiny dip in brightness caused when a planet transits, or passes in front of, its parent star, they both have to record a drop in stellar brightness of no more than 1 part in 1,000. And that, in theory, makes them able to detect the effects of the stellar sound waves.

NATURE.COM
More about Kepler's exoplanet search:
go.nature.com/ooqztr

Before launch, no one could say whether the satellites would make good on this. Kepler's exoplanet search has, in fact, been hindered by stellar oscillations that obscure transits, but are caused by magnetic activity[1], so are unrelated to sound waves. Acoustic oscillations and transits don't interfere with each other: sound waves cause the brightness of Sun-like stars to vary on time scales of 5–15 minutes, whereas planetary transits last for hours. So the planners for both COROT and Kepler were happy to include asteroseismologists in their mission teams. "We are riding on the back of the planet hunters," says Douglas Gough, an asteroseismologist at the University of Cambridge, UK.

As it turned out, the sound-wave data came down in an avalanche — especially from Kepler, which has a 0.95-metre-aperture telescope — nine times the sensitivity of COROT's — plus the ability to look at a larger group of stars for a longer period of time than COROT.

"Everything came together marvellously well," says Gilliland.

Last April, Chaplin and his colleagues published their analysis[2] of acoustic oscillations observed by Kepler in 500 Sun-like stars. The frequency and amplitude of the oscillations revealed that the stars have roughly the sizes predicted by established theories of astrophysics,

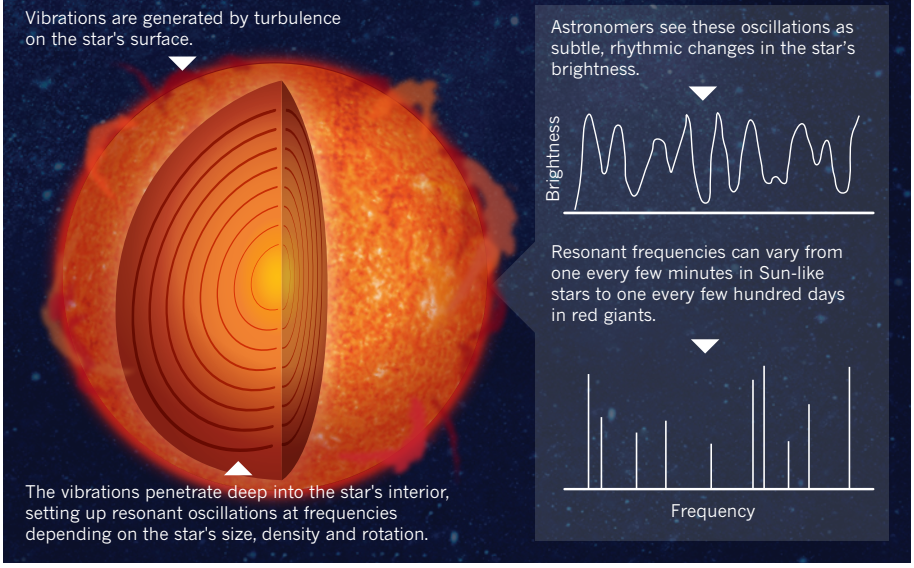
but the distribution of their masses turned out to be significantly lower than expected.

Chaplin isn't yet sure what to make of these findings. But if further observations of the same stars continue to show masses lower than estimated, theorists may have to rethink models of star formation and galaxy assembly. "We didn't have a way of testing these models until we began doing the asteroseismology with Kepler," says Chaplin. And getting them right is crucial: not only do stellar masses not only underlie theories of galaxy formation, but they are also essential for understanding how thermonuclear reactions in stars have produced heavy elements throughout the history of the galaxy — heavy elements that eventually formed planets including Earth.

"This is amazing to witness," says Kjeldsen. With the latest data, he adds "we can test our assumptions, ideas, theories and models in great detail. And we can correct all our errors too."

Celestial music

In the same way as a sound wave resonates inside an organ pipe to create a musical tone, sound waves on a far vaster scale can resonate inside a star. By measuring the frequencies of these waves, astronomers can learn about the star's internal structure.



Secrets of the giants

Some of Kepler's biggest surprises have been in its sounding out of red giants. These are Sun-like stars that have exhausted the hydrogen at their cores, causing a fuel crisis that paradoxically leads them to swell up to more than 100 times their original diameters. In about 5 billion years' time, for example, the Sun will become a red giant big enough devour the innermost planets of the Solar System.

Astronomers would like to be able to distinguish between two phases of red-giant evolution: an early stage, in which the giant is still fuelled by hydrogen in a thin shell around a dense core only a few times bigger than Earth; and a later stage, in which the star has begun burning the helium at its core. Knowing the difference would help them to determine the red giants' ages, how quickly they evolve and the amount of gas and heavy elements that they shed into interstellar space during each phase.

That was impossible until Kepler: from the outside, a red giant looks the same regardless of what it is burning. But last March, Timothy Bedding, an astronomer at the University of Sydney in Australia, and his colleagues reported[3] that Kepler oscillation data allowed for a clear distinction.

"It's difficult not to be fascinated by an ability to learn about the properties of the tiny core of these huge stars from oscillations on their surface," says Jørgen Christensen-Dalsgaard, an astronomer at Aarhus University and a co-author of the study.

Going further, the same researchers reported in December[4] that they had measured the rotation rate of the core region of a red giant for the first time, and discovered that it whips around about ten times faster than the surface.

This finding confirms the standard model of red-giant formation — Sun-like stars flinging their shallower layers outwards while their cores contract. Basic physics demands that angular momentum is conserved, so the outer layers must slow their rotation and the contracting core must speed up, just as observed.

Mission not yet accomplished

Many astronomers have called for an extension to Kepler's mission, which is currently slated to end in November. It is unclear whether NASA will be able to heed them; funding is tight, and other missions need money, too. But asteroseismologists are helping to make the case. They point out, for example, that acoustic oscillations in the Sun shift their frequency ever so slightly — a change of about 1 part in 10,000 — over the course of the Sun's 11-year magnetic cycle. The shift provides a new way to measure the length of the cycle, in which changes in the Sun's magnetic field drive sunspots, flares and other fluctuations in energy that can wreak havoc on Earth's satellites and communication systems.

Astronomers would now like to compare the Sun's magnetic-activity cycle with those of a slew of similar stars. If the other stars have cycles extending over many years, says Gilliland, Kepler's baseline mission will not be able to track them. "But with observations extending to 7–8 years, or even more — the spacecraft seems good to allow 11 — we will be able to probe many stellar activity cycles. It would be profoundly more powerful," he adds.

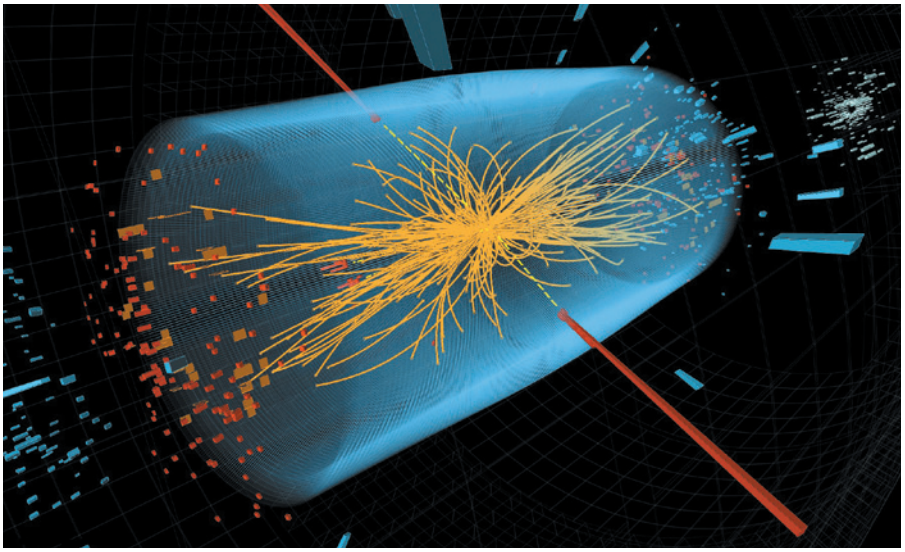
An extended mission could also allow astronomers to learn more about a different class of oscillation that originates deep in a red giant's core, and could tell them a great deal about the core's structure and density. These oscillations have a very small amplitude by the time they make their presence known at the surface, but their reverberations are persistent, like those of a heavy bell, lasting for months or years. "We've just begun to wring the interesting astrophysics out of these results," says Gilliland, and the ability to take data over many years would be an immense help.

Asteroseismology may even help in Kepler's primary mission of finding Earth-sized exoplanets orbiting in the habitable zone around their stars, notes Chaplin. Because the craft can detect an exoplanet only by the amount of light that it blocks as it passes in front of its host star, and can measure it only in relation to the host, the radius of a planet is known only as accurately as the radius of the star. But sound-wave oscillations recorded from the parent star can pin down its size very accurately[5]. Such measurements are possible for only the brightest of the stars in Kepler's field of view, but they could make a huge difference in the researchers' confidence in their data as they begin to report detection of planets approaching Earth's size[6,7].

Johannes Kepler, the seventeenth-century astronomer after whom the spacecraft was named, theorized that Earth and all the other known planets made their own sounds — an arrangement that he called the music of the spheres. It would be only fitting if celestial music, of a kind, had a key role in the Kepler space telescope's most prized discovery. ■

Ron Cowen *is a freelance writer based in Silver Spring, Maryland.*

1. Cowen, R. *Nature* **477,** 142–143 (2011).
2. Chaplin, W. J. *et al. Science* **332,** 213–216 (2011).
3. Bedding, T. R. *et al. Nature* **471,** 608–611 (2011).
4. Beck, P. G. *et al. Nature* http://dx.doi.org/10.1038/nature10612 (2011).
5. Howell, S. B. *et al.* Preprint at http://arxiv.org/abs/1112.2165 (2011).
6. Hand, E. *Nature* **480,** 302 (2011).
7. Fressin, F. *et al. Nature* http://dx.doi.org/10.1038/nature10780 (2011).



T. MCCAULEY/L. TAYLOR/CERN

Hint of the Higgs: photons (red) and particles (yellow) result from a proton–proton collision at CERN.

The need for new physics

Regardless of whether the Higgs detection is confirmed, the standard model is incomplete, says **John Ellis**.

At the time of writing, physicists are excited by the possible detection of a Higgs boson. After analysing data from the Large Hadron Collider (LHC) at CERN, Europe's high-energy physics lab near Geneva, Switzerland, signals of the particle — thought to give other particles mass — have been seen in the region of 125 gigaelectronvolts (GeV).

It is too early to say whether these promising hints will be confirmed, but if they are, many people would take this to be a validation of the standard model of particle physics. There have been previous indirect signs from other data that the Higgs boson probably weighs less than 150 GeV, and CERN's possible observation would be in line with that. But I am a contrarian. I argue that whether or not the Higgs boson exists, we already know that there must be physics beyond the standard model. Let us consider the possibilities.

AVERTING COLLAPSE

One option is that the evidence from the LHC will be confirmed, and a standard-model Higgs boson exists in the low-mass range below 130 GeV. Experimentalists would say that this is where theorists were telling them to look all along, which is true. But there is a catch. Within the standard model, it is possible to calculate the lowest energy state of the Universe. If the Higgs is light, this calculation predicts a lowest energy state totally unlike our current Universe. It implies that our Universe is in some other, unstable state that will eventually flip over to its lowest energy condition — next week, or in a few billion years, we could go down the cosmological tubes. Some physicists say that this unsettling conclusion is the result of taking theory too seriously, and that we should simply find the Higgs and cross this bridge when we come to it. But it is worrying. The only way to produce a more stable model would be to add new particles to the standard model, in this case making it look as predicted by theories of supersymmetry.

AVERTING A BLOW-UP

What if the LHC evidence is not confirmed? In principle, there are other options. For example, there is another range of possible Higgs mass, above 600 GeV, where CERN has yet to look. In this case, we must go beyond the standard model for two reasons. First, physicists have found and measured the masses and other properties of W and Z bosons, which mediate the fundamental weak nuclear forces. The theoretical interpretation of these measurements is what hints that the Higgs should weigh less than 150 GeV. If the Higgs is actually much heavier, then this can be reconciled only by adding a host of new interactions between the known particles. Second, the interactions between heavy Higgs bosons become seemingly infinite at high energies. Clearly this blow-up must be averted, but understanding how also lies beyond the standard model.

➲ NATURE.COM
Detectors home in on Higgs boson:
go.nature.com/rvar9q

Could there be a lighter Higgs boson, hidden between 130 GeV and 600 GeV? The LHC has excluded a standard Higgs boson in this region, but it is still possible that the Higgs lurks there, if its interactions with other particles are not as the standard model predicts. One possibility might be that it has some additional, invisible mode of decaying into particles, such as via pairs of weakly interacting particles of dark matter. Another possibility is that its decays are normal, but the rate at which other particles collide to produce Higgs bosons is suppressed. Make such an adjustment, however, and the standard model is no longer a calculable theory; and the whole point of the Higgs is to make the standard model calculable. The only way to fix this would be to add new particles to the mix, such as massive spin-one particles, to finish the job that the Higgs was supposed to do.

These problems would be particularly severe in the absence of a Higgs boson. Physicists would have to explain, in a way that is both calculable and consistent with existing measurements, how and why the symmetry between different species of standard-model particles is broken — only some particles in the standard model have mass. The Higgs is assumed to be what forces the symmetric standard-model equations to have these asymmetric solutions. Breaking the intrinsic symmetry of the equations themselves would be inadvisable, because it is not known how to calculate sensible results in such a case.

Alternatively, one could break the symmetry of the solutions to the standard-model equations by setting specific boundary conditions. Just as a physicist's description of how a washing line hangs between two walls is dependent not just on the properties of the line, but also on its attachment to the walls, so too a description of bosons depends in part on their behaviour at the edges of the Universe. Our current description of a four-dimensional Universe doesn't have edges in this sense, but additional dimensions, presumably curled up into tiny spaces so that no experiment has yet been able to detect them, might provide suitable boundaries: physics beyond the standard model indeed.

Regardless of whether and where the Higgs boson is found, new physics is needed. ■

John Ellis *is a theoretical physicist at King's College London, London WC2R 2LS, UK, and a guest professor at CERN.*
e-mail: john.ellis@cern.ch

Keep standards high

As models of authorship and collaboration change in the digital age, we must rely on trust to filter the products of research, says **Jerome Ravetz**.

Science is unique among areas of organized activity and production in that it has an informal quality-assurance system: peer review, publication and replication. The system has worked well since its inception in the seventeenth century, when the scientific journal came into being. But it is now being challenged as technology changes social practices of science. How might it evolve?

Some trends are apparent. The rise of digital media has revolutionized the management of information and created opportunities for broader involvement in science's production. Collaborations are growing ever larger, transforming the concept of authorship. Prepublication discussions of research on blogs dilute a principal author's claim to discovery. And the public is increasingly involved.

Amateurs are returning to mainstream research after an absence of generations. By completing online tasks, from classifying galaxies (http://galaxyzoo.org) to solving complex protein-folding problems (http://folding.stanford.edu), anyone can become a co-creator of scientific knowledge. Such a widening of participation might be liberating, but it also risks lowering standards. Not everyone shares the ideal that intellectual integrity comes before personal gain.

As a result of these developments, the product of research is becoming more fluid. The journal is losing its status as the sole gatekeeper — simultaneous guarantor of quality, certifier of property, medium of communication and also archive. Other means of sharing material, assessing quality and screening out the incompetent or fraudulent are emerging to fill the gap, but ultimately the professional monopoly on quality assurance of science will have to be modified.

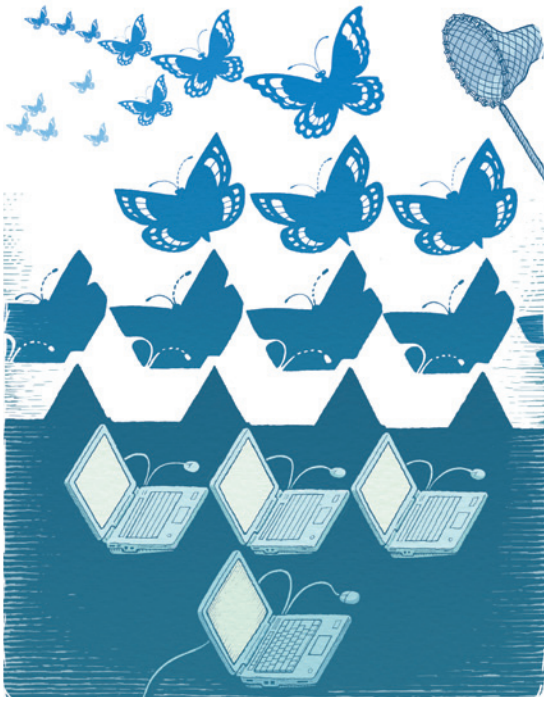


ILLUSTRATION BY DAVID PARKINS

NEW GATEKEEPERS

In response to these trends, some individuals are becoming self-appointed gatekeepers. During the polarized 'climategate' debates in 2010, for example, climate scientists stepped in to defend the work of a reputable colleague from criticism by a 'mere' mining engineer. That critic, Steve McIntyre, claimed on his blog simply to be applying the standards of the business world to climate data.

Although scientific expertise presents a bar to interference, concerned outsiders have a legitimate and useful role. The setting of policy priorities is one such example. Institutions such as the Cochrane Collaboration employ thousands of 'expert patients' across the globe to review the quality of research papers on treatments.

Whistle-blowing is another vital form of intervention. As technological advances bring new perils, people can provide early warnings that prevent real harm. Here, the blogosphere holds great promise for free information sharing. But with its lack of formal quality assurance, some see it as a possible vehicle for a scientific demagogue like Soviet biologist Trofim Lysenko.

In some circles, gatekeepers are being done away with. Many high-technology sectors already operate along communitarian lines. For example, 'open source' and 'creative commons' enterprises handle intellectual property collaboratively. Signatories publish their specifications freely, allowing others to copy and adapt the work, as long as people credit it and there is mutual access to ideas. These sound like the ideals of science as based on the work of US sociologist Robert Merton: communalism, universalism, disinterestedness, originality and scepticism.

It would be ironic but welcome if the technology sector, traditionally managed through the private-property system of patents, were to lead the way in open science. Within a more informal system, the current excessively quantitative systems of quality-assessment — the numbers of papers and citations — could be relaxed and made more personal and communal. Quality assurance could be achieved by a societal consensus and professional gatekeepers.

For this to happen, barriers to sharing scientific information with the public, such as journal paywalls, should come down. And better online discussion forums must be developed. The presentation and archiving of blogs and other forms of internet communications should be improved, so that ideas can be debated and added to over time.

As more people become involved in online debates, quality need not fall by the wayside. It is encouraging to see that well-conducted discussions of points of contention between the scientific mainstream and critics are emerging, as the Berkeley Earth Surface Temperature study demonstrates (see *Nature* **478,** 428; 2011).

Ultimately, effective quality assurance depends on trust. And science relies on trust more than most institutions. As Steven Shapin, a historian of science at Harvard University in Cambridge, Massachusetts, showed in his 1994 book *A Social History of Truth*, trust is achieved and maintained only by mutual respect and civility of discourse. In a digital age, civility should be extended to, and reciprocated by, the extended peer community.

Scientists have a special responsibility, but also a special difficulty. When their training has been restricted to puzzles with just one right answer, scientists may find it hard to comprehend honest error, and may condemn those who persist in apparently wrong beliefs. But amid all the uncertainties of science in the digital age, if quality assurance is to be effective, this lesson of civility will need to be learned by us all. ■

Jerome Ravetz *is at the Institute for Science, Innovation and Society at the University of Oxford, Oxford OX1 1HP, UK. He is the author of* Scientific Knowledge and its Social Problems *(1971, 1996).*
e-mail: jerome-ravetz@ntlworld.com

S. NELSON/WPN/PHOTOSHOT

Military surveillance drones must be designed in accordance with laws governing warfare, such as the avoidance of civilian casualties.

ROBOTICS

Morals and machines

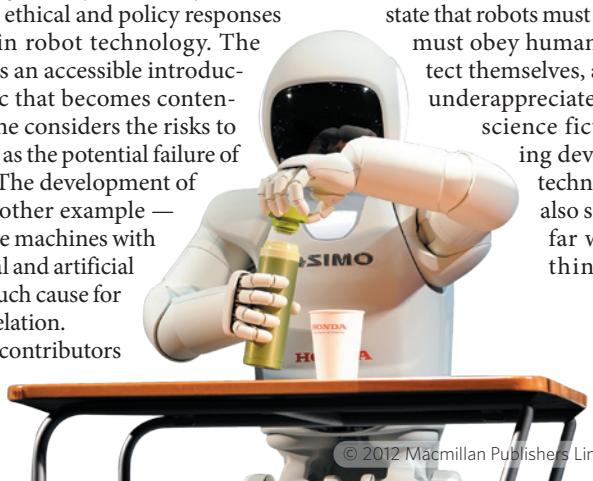
A view of robotics reveals ethics has not kept up with technology, finds **Braden Allenby**.

Fantastically complex technological systems now criss-cross Earth, harnessing nanotechnology, biotechnology, information technology, applied cognitive science and, of course, robotics. How do we regulate such systems? Adaptively, according to a book that considers the evolving and unpredictable nature of the robotics frontier.

Robot Ethics, edited by philosophers Patrick Lin and Keith Abney, and computer scientist George Bekey, is a timely round-up of sensible ethical and policy responses to advances in robot technology. The book provides an accessible introduction to a topic that becomes contentious when one considers the risks to humans, such as the potential failure of robotic cars. The development of cyborgs is another example — for some, these machines with both biological and artificial parts are as much cause for unease as for elation.

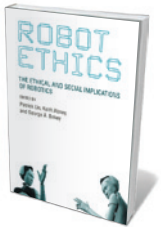
The book's contributors address a wide spectrum of concerns: from robotics expert Noel Sharkey on military applications and philosopher Peter Asaro on the legal perspectives, to ethicists Jason Borenstein and Yvette Pearson on robotic carers in ageing societies. Robotics aimed at childcare, medicine, surgery and even recreation also get a look-in.

Several chapters touch on Isaac Asimov's 'Three Laws of Robotics', which first appeared in his 1942 science-fiction story *Runaround*, and still spark debate. The laws state that robots must not hurt humans, must obey humans and must protect themselves, and illustrate the underappreciated importance of science fiction as a visioning device for emerging technologies. The laws also serve to show how far we have come in thinking about the ethics of robotics — a good thing, given that Asimov's stories mainly highlighted the laws' inadequacy.



Asimo is designed to aid people who lack mobility.

T. YAMANAKA/AFP/GETTY



Robot Ethics: The Ethical and Social Implications of Robotics
EDITED BY PATRICK LIN, KEITH ABNEY AND GEORGE A. BEKEY
MIT Press: 2011. 400 pp. $45, £31.95

Roboticists are now integrating technology with humans in ways that challenge fundamental ethical and cultural ideas. As prosthetics become more powerful and computer–brain interfaces move ever closer to cyborg territory, the definition of what it means to be human becomes unpredictable in new ways. Yet as this book makes clear, ethics has not kept pace with complex techno-human systems, such as 'augmented cognition' networks that boost a person's capacity to handle multiple streams of incoming data. For example, who is responsible if an autonomous military robot kills a group of civilians?

The manufacturer, the commander, the operator or the robot itself? The time for ethical considerations is now — once the technologies are embedded, they will be much harder to change.

The book touches on another hotly debated issue in robotics: military applications and their impacts. These include the operational, policy, governance and ethical implications of unmanned aerial vehicles or insect-sized mobile surveillance devices. Greater challenges will arise when such technologies migrate into civil society; imagine what divorce lawyers, or internal security organizations, could do with tiny mobile surveillance cyborgs.

"The time for ethical considerations is now — once the technologies are embedded, they will be much harder to change."

Robot Ethics succeeds as a stand-alone text, with its varied contributors striving for objectivity and avoiding hyperbole. The broad spread of applications discussed is key because the ethics differ depending on the use. Military robots, for instance, must be designed to obey the laws that govern warfare. Carer robots must be capable of interacting with patients, who may give them trust and even affection.

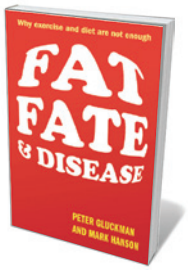
However, there are omissions. The lack of a contribution from robotics expert Ronald Arkin at the Georgia Institute of Technology in Atlanta is glaring, because several chapters cite and challenge his work on autonomous lethal military robots. A chapter summarizing the state of the technological art in robotics would have been valuable, so that a novice reader could distinguish between existing technologies — which raise concerns that might need to be addressed immediately — and hypothetical inventions.

Similarly useful would have been a summary of the underlying economic and cultural factors driving robotic technologies. That would help to distinguish likely scenarios from economically or politically impractical ones. For example, military robots are likely to be implemented because they save soldiers' lives. So outright bans on robots might be politically impractical, suggesting that regulation is more realistic.

By portraying robots as real-world experiments in ethics, *Robot Ethics* conveys an important lesson for our technological era: we must develop responses to emerging technologies in real time, rather than simply reacting to them using existing ethical frameworks. ■

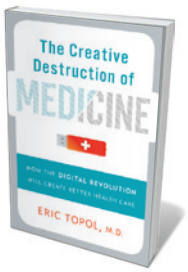
Braden Allenby *is a professor of engineering and of law at Arizona State University, Tempe, Arizona 85287, USA.*
e-mail: braden.allenby@asu.edu

Books in brief

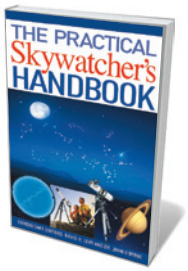


Fat, Fate and Disease: Why Exercise and Diet Are Not Enough
Peter Gluckman and Mark Hanson OXFORD UNIVERSITY PRESS *304 pp. £16.99 (2012)*
Today's 'epidemics' of heart disease, obesity and diabetes could go global just as the squeeze on planetary resources tightens. Disease-development experts Peter Gluckman and Mark Hanson probe why we are losing the battle. Arguing for the inclusion of cultural, social and biological realities, they explore how fat and disease are linked; lifestyle shifts in the developing world; and genetics and fetal development. The solution, they say, lies in focusing on the health of mothers and children.

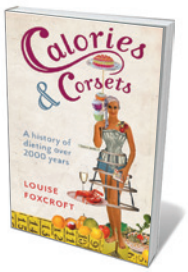
Design in Nature: How the Constructal Law Governs Evolution in Biology, Physics, Technology, and Social Organization
Adrian Bejan and J. Peder Zane DOUBLEDAY *288 pp. $27.95 (2012)*
What links the alveoli of the human lung and a bolt of lightning? The answer is branching — a configuration that is optimal for the flow of oxygen, electricity and much more. Mechanical engineer Adrian Bejan and writer J. Peder Zane make a fascinating case for how a single law of physics governs shape and structure in everything, animate or inanimate. By reframing things as flow systems, they reveal how function determines form in everything from corporate hierarchies to Canada geese.



The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care
Eric Topol BASIC BOOKS *304 pp. $26.99 (2012)*
The digital age opens up the possibility of a new type of medicine in which an individual's health data are digitized using wearable sensors, smartphone apps and genome information, writes geneticist and cardiologist Eric Topol. With this wealth of data, medical interventions could be tailored to our uniqueness. Topol covers failures in patient information; what might happen if genomics, imaging, sensors and better health information were to converge; and the potential pitfalls of this brave new medical world.



The Practical Skywatcher's Handbook
Edited by David H. Levy and John O'Byrne A&C BLACK *480 pp. £16.99 (2012)*
Astronomer David Levy and physicist John O'Byrne provide zealous amateur stargazers — and serious sailors who need a backup to the Global Positioning System — with a comprehensive 'field guide' to the heavens. Every constellation in the Northern and Southern Hemispheres is mapped across 150 sky charts, with easy methods for finding the brightest stars and working out the rest. Levy (co-discoverer of Comet Shoemaker–Levy) and O'Byrne also give the lowdown on which stars are visible with the naked eye.



Calories and Corsets: A History of Dieting Over 2,000 Years
Louise Foxcroft PROFILE *240 pp. £14.99 (2012)*
Medical historian Louise Foxcroft's astutely researched chronicle of dieting is amusing, alarming and poignant by turns. She recounts how Hippocrates advocated vomiting and naked jogging; how the Royal Society experimented with the potato as a hunger suppressant in the seventeenth century; and how 'remedies' from metal corsets to industrial diet drinks have sometimes proved lethal. Foxcroft sees *diaita* — an ancient Greek philosophy centred on all-round health — as more sensible than a fixed focus on fat.

The manufacturer, the commander, the operator or the robot itself? The time for ethical considerations is now — once the technologies are embedded, they will be much harder to change.

The book touches on another hotly debated issue in robotics: military applications and their impacts. These include the operational, policy, governance and ethical implications of unmanned aerial vehicles or insect-sized mobile surveillance devices. Greater challenges will arise when such technologies migrate into civil society; imagine what divorce lawyers, or internal security organizations, could do with tiny mobile surveillance cyborgs.

"The time for ethical considerations is now — once the technologies are embedded, they will be much harder to change."

Robot Ethics succeeds as a stand-alone text, with its varied contributors striving for objectivity and avoiding hyperbole. The broad spread of applications discussed is key because the ethics differ depending on the use. Military robots, for instance, must be designed to obey the laws that govern warfare. Carer robots must be capable of interacting with patients, who may give them trust and even affection.

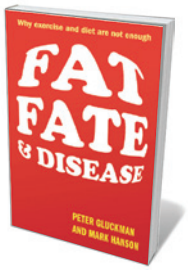
However, there are omissions. The lack of a contribution from robotics expert Ronald Arkin at the Georgia Institute of Technology in Atlanta is glaring, because several chapters cite and challenge his work on autonomous lethal military robots. A chapter summarizing the state of the technological art in robotics would have been valuable, so that a novice reader could distinguish between existing technologies — which raise concerns that might need to be addressed immediately — and hypothetical inventions.

Similarly useful would have been a summary of the underlying economic and cultural factors driving robotic technologies. That would help to distinguish likely scenarios from economically or politically impractical ones. For example, military robots are likely to be implemented because they save soldiers' lives. So outright bans on robots might be politically impractical, suggesting that regulation is more realistic.

By portraying robots as real-world experiments in ethics, *Robot Ethics* conveys an important lesson for our technological era: we must develop responses to emerging technologies in real time, rather than simply reacting to them using existing ethical frameworks. ■

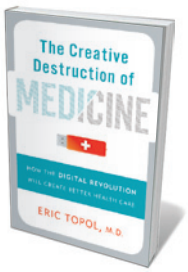
Braden Allenby *is a professor of engineering and of law at Arizona State University, Tempe, Arizona 85287, USA.*
e-mail: braden.allenby@asu.edu

Books in brief

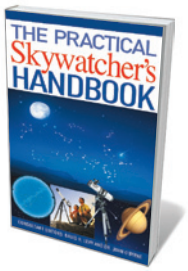


Fat, Fate and Disease: Why Exercise and Diet Are Not Enough
Peter Gluckman and Mark Hanson OXFORD UNIVERSITY PRESS *304 pp. £16.99 (2012)*
Today's 'epidemics' of heart disease, obesity and diabetes could go global just as the squeeze on planetary resources tightens. Disease-development experts Peter Gluckman and Mark Hanson probe why we are losing the battle. Arguing for the inclusion of cultural, social and biological realities, they explore how fat and disease are linked; lifestyle shifts in the developing world; and genetics and fetal development. The solution, they say, lies in focusing on the health of mothers and children.

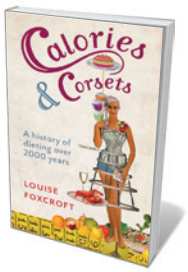
Design in Nature: How the Constructal Law Governs Evolution in Biology, Physics, Technology, and Social Organization
Adrian Bejan and J. Peder Zane DOUBLEDAY *288 pp. $27.95 (2012)*
What links the alveoli of the human lung and a bolt of lightning? The answer is branching — a configuration that is optimal for the flow of oxygen, electricity and much more. Mechanical engineer Adrian Bejan and writer J. Peder Zane make a fascinating case for how a single law of physics governs shape and structure in everything, animate or inanimate. By reframing things as flow systems, they reveal how function determines form in everything from corporate hierarchies to Canada geese.



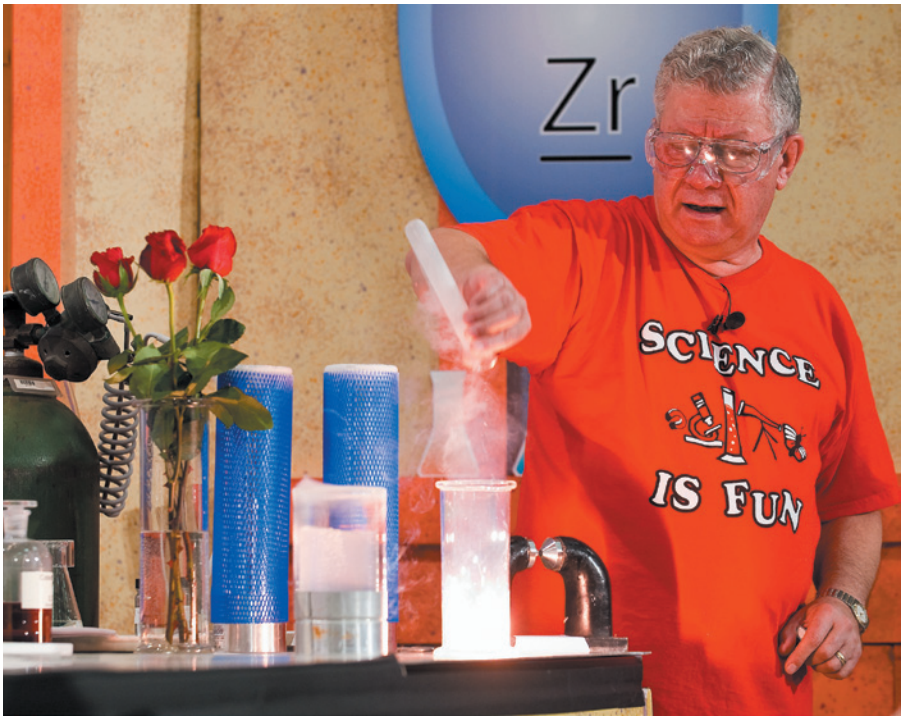
The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care
Eric Topol BASIC BOOKS *304 pp. $26.99 (2012)*
The digital age opens up the possibility of a new type of medicine in which an individual's health data are digitized using wearable sensors, smartphone apps and genome information, writes geneticist and cardiologist Eric Topol. With this wealth of data, medical interventions could be tailored to our uniqueness. Topol covers failures in patient information; what might happen if genomics, imaging, sensors and better health information were to converge; and the potential pitfalls of this brave new medical world.



The Practical Skywatcher's Handbook
Edited by David H. Levy and John O'Byrne A&C BLACK *480 pp. £16.99 (2012)*
Astronomer David Levy and physicist John O'Byrne provide zealous amateur stargazers — and serious sailors who need a backup to the Global Positioning System — with a comprehensive 'field guide' to the heavens. Every constellation in the Northern and Southern Hemispheres is mapped across 150 sky charts, with easy methods for finding the brightest stars and working out the rest. Levy (co-discoverer of Comet Shoemaker–Levy) and O'Byrne also give the lowdown on which stars are visible with the naked eye.



Calories and Corsets: A History of Dieting Over 2,000 Years
Louise Foxcroft PROFILE *240 pp. £14.99 (2012)*
Medical historian Louise Foxcroft's astutely researched chronicle of dieting is amusing, alarming and poignant by turns. She recounts how Hippocrates advocated vomiting and naked jogging; how the Royal Society experimented with the potato as a hunger suppressant in the seventeenth century; and how 'remedies' from metal corsets to industrial diet drinks have sometimes proved lethal. Foxcroft sees *diaita* — an ancient Greek philosophy centred on all-round health — as more sensible than a fixed focus on fat.



B. RICHTER

Q&A Bassam Shakhashiri
The science showman

Chemist and educator Bassam Shakhashiri is known for his dramatic live demonstrations of chemistry in action. As he takes the helm as president of the American Chemical Society (ACS) this week, he talks about the state of science education and how to engage people in chemistry through the wonders of the lab.

What's your take on sensationalized media reports on science in the United States?
Science has a social contract: society expects great intellectual achievements to benefit the human condition and protect the environment. My focus is not on countering sensational media reports, but on addressing grand challenges to society. Along with population growth, finite resources and climate change, those challenges include the denial of human rights — especially the right to benefit from scientific and technological progress.

What are your goals as 2012 president of the ACS?
I am deeply committed to the mission of the ACS, which is "to advance the broader chemistry enterprise and its practitioners for the benefit of Earth and its people". As a learned society, the ACS must show the public that the chemical sciences are major drivers of our economy and contribute hugely to prosperity, fairness and justice. I have appointed a working group on climate science to develop a web-based toolkit as a resource for ACS members. Through this they will be able to gain expertise in climate issues, and formulate strategies for disseminating their knowledge to all levels and areas of society — from academics to industrial scientists, civic and religious groups and government. Another of my initiatives is a new commission that will examine US graduate education and research in chemistry. It will aim to recommend radical changes in how we use our country's vast educational, industrial and government resources to prepare students for careers that can adapt to changing human needs over the next 50 years.

How did you start out as a scientist?
When I was growing up in Lebanon, my mother knitted me a yellow sweater, and I asked her what made it yellow. Much later, as a chemistry major at Boston University in Massachusetts, I found answers to my questions. In graduate school at the University of Maryland in College Park, I was still fascinated by colour changes and studied electron-transfer reactions involving colourful metal-ion complexes in aqueous solutions. My studies of how light and matter interact helped me to get satisfying answers for my youthful curiosities, which led to more questions. As a university professor, I was struck by the profound impression made on my students by the deep red glow produced by the decay of singlet molecular oxygen or by the blue colour of paramagnetic liquid oxygen.

What is your approach to science education?
Science education should aim to share the beauty, challenges and rewards of open enquiry and help people to avoid sham, quackery and unproven conjecture. Interacting with students deepens my own understanding of science and of the process of learning science. When I joined the University of Wisconsin–Madison as a faculty member in 1970, my mission was to improve undergraduate chemistry education for all students, not just for science majors. In 1984, I became the assistant director for science and engineering education at the US National Science Foundation, after those programmes were almost phased out early in the administration of President Ronald Reagan. I rebuilt the programmes and created new ones. When I returned to the University of Wisconsin–Madison in 1990, I worked on science-literacy initiatives that focused on classroom instruction and the public appreciation of science.

How do you engage the public with science?
Scientists do what we do because it interests us, it satisfies our curiosity and we enjoy it. I want to share this enjoyment with everyone. So I aim to engage people in meaningful and thoughtful explorations of science, to share core values, to develop a deeper public understanding of science and to influence attitudes. The real gauge is people's behaviour in society. We need more science fans supporting and talking about science. It's about developing an attitude. On my website, Science is Fun [www.scifun.org], you can download instructions for home experiments that use ordinary chemicals. My 'Science is Fun' badge prompts smiles and questions. If anyone asks what it is about, I tell them that science contributes to the quality of life in modern society, and that it is intellectually stimulating and emotionally rewarding. They always want to know how to get a badge — and effectively spread the word.

How do you make science fun?
I'm a strong advocate of demonstrations — exploding balloons, glowing liquids, spectacular chemical transformations. Connectivity is key. My ultimate purpose is to trigger cerebral and emotional reactions to heighten an audience's joy of learning. I want them to be sceptical about what they see and hear. That's how science works. ■

INTERVIEW BY JASCHA HOFFMAN

Correspondence

Call to veto Brazil's forest-code revisions

At last month's COP17 climate-change negotiations in Durban, South Africa, Brazilian president Dilma Rousseff called for renewal of the Kyoto Protocol's binding targets for reducing greenhouse-gas emissions. We look to her to reaffirm Brazil's commitment to sustainable development by vetoing the proposed changes to the country's Forest Code when it comes before the Brazilian National Assembly for ratification in March.

The Senate has already approved revisions to the code, which would provide amnesty for previously illegal forest clearance and undermine efforts to limit deforestation.

Brazil should be setting an example as a green economy, particularly as it will be hosting the Rio+20 conference this year — the twentieth anniversary of the 1992 United Nations Conference on Environment and Development.

The adverse environmental effects that would be generated by the Forest Code revisions threaten to be made worse by the country's poor record for enforcing environmental law. For example, appeals against fines for environmental crimes are almost always successful (see go.nature.com/2xvs2o), amounting to nearly US$8 billion in unpaid penalties between 2005 and 2010.

Brazil's forestry regulations must be tailored to ensure the protection of forests alongside sustainable development. The new Forest Code would do neither.

Alison G. Nazareno *Federal University of Santa Catarina, Florianópolis, Brazil.*
alison_nazareno@yahoo.com.br

Mauritius is putting conservation at risk

A government policy for the national parks of Mauritius is threatening important research into conservation and undermining the ownership and sustainability of conservation projects. Appeals to modify this policy have remained unanswered.

Mauritius is known for its conservation successes — including that for the Mauritian kestrel, *Falco punctatus* — thanks to the development of innovative techniques. Its national parks provide ideal terrain for conservation research because they harbour most of the country's highly threatened biodiversity and provide a variety of experimental settings for study.

The new policy makes it difficult for biologists and students from the local university to access field camps, restricting their fieldwork to office hours on weekdays. Fieldwork outside these periods is rarely allowed: applications typically take at least three months to process, and payment to the National Parks and Conservation Service is necessary if the request is successful.

The Mauritian government's policy is insensitive to the temporal expression of biological phenomena. Its justification of concern for biologists' safety is weak, given that no significant security scares have ever been recorded. Worse, many foreign biologists are allowed to work at the same sites, unimpeded by the restrictions facing local academics.

The Mauritian authorities should note that the local university has been spearheading capacity-building to support native conservation efforts since the 1990s, with more than 50 research projects completed by its graduate and undergraduate students. With the help of the international community, the government should be persuaded to return to a stance that is less harmful to conservation.

F. B. Vincent Florens *University of Mauritius, Réduit, Mauritius.*
vin.florens@uom.ac.mu

Phoenix first to see silt grains on Mars

The successful launch of NASA's Mars Science Laboratory, or Curiosity, offers a chance to correct a misconception about its imaging systems. You say that these have sufficient magnification and resolution "to see, for the first time, tiny grains of silt … particles that would bear the history of billions of years of erosion from wind and water" (*Nature* **479,** 446; 2011).

Curiosity's imager will be the first to see silt-sized particles *in situ*, on the undisturbed surfaces of rocks and soils. But bragging rights for the first images of silt grains (4–62.5 micrometres in diameter) and even some clay-sized particles (smaller than 4 micrometres in diameter) go to the Phoenix Mars Lander, which accomplished the feat in 2008 (P. H. Smith *et al. Science* **325,** 58–61; 2009).

Soil samples from trenches on the periglacial landforms surrounding Phoenix were imaged by an optical microscope in the Microscopy, Electrochemistry, and Conductivity Analyzer (MECA), which returned hundreds of colour images of grains as fine as 4 micrometres in diameter (W. Goetz *et al. J. Geophys. Res.* **115,** E00E22; 2010). Imaging grains from diverse substrates enables the microscopic characterization of grain magnetic properties. The MECA atomic-force microscope returned three-dimensional images of particles as small as 0.1 micrometres (M. A. Velbel and A. I. Losiak *J. Sed. Res.* **80,** 771–780; 2010).

When it lands, Curiosity will comprise the most capable suite of instruments delivered to the surface of another planet. But Phoenix has already given us a close look at Mars.

Michael Velbel *Michigan State University, East Lansing, Michigan, USA.*
velbel@msu.edu

Science and politics need more empathy

Some important pointers for improving communication between scientists and politicians (*Nature* **480,** 153; 2011) emerged from a meeting last year between the two groups, organized by the International Risk Governance Council.

Support for fundamental research is essential, but scientists shouldn't be tempted to overstate their claims. Specifically, requests for funding for basic research into a particular question shouldn't imply that the research will definitely provide the answer. When basic questions lead to technological advances, these typically emerge in another, unexpected area — take quantum mechanics and the transistor, for example, or the study of gas conductivity that led to the discovery of X-rays.

The best bets for answering immediate, focused questions are likely to be the development of existing technologies and the novel juxtaposition of established ideas from different areas. Governments should therefore put more faith in interdisciplinary scientists who have the vision and ability to bring apparently disparate fields together, rather than turning to subject specialists, who often have their own agendas.

For their part, scientists must recognize and respect the need of politicians to win votes. Without power, politicians can't carry through a long-term, scientifically based policy. Scientists who promote these policies will fare better if they can identify and suggest short-term, intermediate benefits.

Len Fisher *University of Bristol, Bristol, UK.*
len.fisher@bristol.ac.uk

CONTRIBUTIONS
Please consult the author guidelines for submitting Correspondence at **go.nature.com/cmchno**.

CAREERS

TURNING POINT UK postdoc hopes that prizewinning essay will open doors **p.105**

ADVICE FORUM Get expert tips on scientific career issues **go.nature.com/lm1x4t**

NATUREJOBS For the latest career listings and advice **www.naturejobs.com**

J. BERKELEY

WOMEN IN BUSINESS

Finding a way in

Female scientists hoping to become entrepreneurs face obstacles — but there are organizations that can help.

BY VIRGINIA GEWIN

When her adviser left, Mona Jhaveri was two years into her biochemistry postdoc at the US National Cancer Institute in Bethesda, Maryland. But for Jhaveri, it was as much an opportunity as a setback.

Jhaveri had helped to discover that a small DNA sequence could be designed to knock down levels of the messenger RNA that codes for the folate-receptor protein, which is overexpressed in ovarian cancer cells. When her adviser left in 2000, Jhaveri secured all patent rights to the sequence. In 2006, intellectual property in hand, she launched FOLIGO Therapeutics, a biotechnology company in Rockville, Maryland, that focused on developing DNA-based medicines and molecular diagnostics for ovarian cancer. She raised US$500,000 in capital through social connections.

Then the money ran out. In 2010, Jhaveri shut down operations. She is now attempting to raise the $5 million to $6 million that she needs to get a compound through clinical trials. Jhaveri acknowledges that the recession and the dismal state of the biotech industry have made it difficult for everyone to secure funding, but she suggests that women still have a tougher time than men — including when it comes to accessing capital. "Deals are still made on the golf course," she says.

Jhaveri's struggles are typical. In the United States and Europe, women make up only a small fraction of business leaders. They find it hard to raise funds, make contacts and pitch ideas, and they are less confident than men about finding and creating opportunities.

Still, the number of women entrepreneurs could be set to rise, thanks to a host of resources available to help women to overcome the apparent crises of confidence and capital. Female-focused mentoring and training programmes promise participants the support and advice that they need to become investment-ready — but raising the funds requires a unisex approach.

GENDER BIAS

Women account for just 35.3% of total US entrepreneurship, according to the Index of Entrepreneurial Activity published by the Kauffman Foundation, a non-profit organization focused on entrepreneurship and based in Kansas City, Missouri. Of the US-based companies that received venture-capital financing in 2010, only 10% had ever had a female founder or chief executive, according to Dow Jones VentureSource. And fear of financial woes is among the main reasons that women opt out of becoming entrepreneurs, says Jeffrey Sohl, director of the Center for Venture Research at the University of New Hampshire in Durham.

But the paucity of female entrepreneurs is a result of more than a lack of capital. Surveying 59 economies, the *Global Entrepreneurship Monitor 2010 Women's Report*, published last month by Babson College in Babson Park, Massachusetts, found that women are less likely than men to believe that they have the opportunities to start a business. They are also more likely to let fear of failure dissuade them from entrepreneurship. According to *Overcoming the Gender Gap: Women Entrepreneurs as Economic Drivers*, a report released in September by the Kauffman Foundation, women entrepreneurs often neglect to take the steps necessary to launch a high-growth business, such as patenting their research or making connections outside academia.

For female would-be entrepreneurs, these challenges make it difficult to pursue a ▶

potentially rewarding career path. In the United Kingdom, for example, 15% of businesses are led by women, yet women account for only 7% of entrepreneurs in science, engineering or technology fields.

"Women don't ask for opportunities and they undersell their abilities and expertise," says Sharon Vosmek, chief executive of Astia, a non-profit organization in San Francisco, California, that supports women-led, high-growth companies in technology and the life sciences around the world. Often, she adds, women miss opportunities because they don't know how to take advantage of their scientific credentials.

KAUFFMAN FOUNDATION

Women, especially scientists and engineers, don't have the necessary networks in the private sector.
Lesa Mitchell

When it comes to pitching business ideas, women are often less aggressive and more cautious than men — which can be interpreted as a lack of confidence, undermining the pitch. Women are also "more open to a discussion about the cons as well as the pros of a potential business — which can make an idea look less attractive to an investor", says Joanna Horobin, president and chief executive of Syndax Pharmaceuticals in Waltham, Massachusetts.

FOLLOW THE LEADERS

Experts in entrepreneurialism agree that finding mentors — preferably ones who have formed successful companies and are willing to walk newcomers through the process — is a must for women and men alike. But Lydia Villa-Komaroff, chief scientific officer of CytonomeST in Boston, Massachusetts, says that women require extra support. "Women need advocates — people willing to make introductions and vouch that you are a known quantity," she says. A scheme announced on 4 November by Theresa May, the UK women's and equalities minister, will spend £2 million (US$3.14 million) to train 5,000 volunteer business mentors for women.

And a number of organizations have popped up in recent years to bolster women's entrepreneurial activity. ACTiVATE, a programme that was originally funded by the US National Science Foundation (NSF), teaches women how to create technology companies. It lasts for 10–12 months and trains participants in how to assess a market for their idea, develop a business plan and form a company. The 130 programme graduates have together created more than 40 businesses. "Even though all of them won't be chief executives, understanding the process makes them more valuable members of a start-up team, such as a chief scientific officer," says Julie Lenzer Kirk, co-founder of the Path Forward Center for Innovation and Entrepreneurship, a non-profit group in Germantown, Maryland, that licensed ACTiVATE from the NSF.

ACTiVATE and other programmes, such as those sponsored by Women Entrepreneurs in Science and Technology in Cambridge, Massachusetts, or the Association for Women in Science and Engineering, a group of female academic and industrial scientists in Cambridge, UK, are great starting points for learning about businesses and finding mentors, say participants. Still, their reach can go only so far. "ACTiVATE gave me the confidence and a path that I wouldn't have formed on my own," says Jhaveri. But it did not give her access to capital.

A DEARTH OF FUNDS

Women tend to limit themselves in how they search for funding, according to research from Sohl. Angel investors — affluent individuals who invest their own funds — are generally the largest source of funding for early-stage, high-growth companies, yet women pursue such investment at substantially lower rates than men. And women most often seek money from other women. This presents a problem, given that only 10–15% of angels are women. Yet Sohl's research shows that once women overcome barriers and submit proposals for angel funding, they have a roughly 14% success rate, about the same as men.

Even so, women have problems securing more lucrative, later-stage venture-capital deals. In the United States, just 2% of venture capital went to companies with women chief executives and founders in 1998, according to the Diana Project, a longitudinal study on women in business, based at Babson College. The proportion had increased to 10% by 2010, finds Dow Jones, but that is still a small total — especially considering that women make up half of the nation's MDs and half of PhDs in some fields. Women, especially scientists and engineers, don't have the necessary networks in the private sector, says Lesa Mitchell, a vice-president with the Kauffman Foundation.

Women looking for funding might consider organizations such as Golden Seeds, an angel-investment network based in New York City. Founded in 2004, it has invested $30 million in dozens of companies, many of which focus on science and technology — and all of which have at least one woman in a leadership role who has equity in the company. They typically see 300 applications each year, and fund between 10 and 12 of those companies, according to Nadia Jain, Golden Seeds managing partner. She adds that the network often offers guidance to entrepreneurs in whom it does not actually invest.

Too frequently, budding women entrepreneurs are content to scrape by rather than make a bold request. "Women tend to think they can 'get by' with $250,000 when they should be asking for $1 million," says Susan Windham-Bannister, president and chief executive of Massachusetts Life Sciences Center in Waltham. That strategy effectively shoots an entrepreneur in the foot twice, she points out — not only does the fund-seeker fail to secure the necessary support, but she also sends a signal that she doesn't understand what it takes to develop a product.

There are few women-centric organizations focused on raising capital, but they do exist. Springboard Enterprises, a non-profit organization in Washington DC, has helped to raise more than $5 billion for more than 400 women-led companies. Astia screens and selects women entrepreneurs with fundable high-growth companies, and fosters networking opportunities to get the capital to succeed. Now serving 307 companies globally, Astia began tracking a cohort of women-led start-ups in 2005; the group of women eligible for Astia support has since grown from 50 to 500.

Exposure to the venture side of business can be valuable. Horobin started her career in clinical drug development and served as vice-president of oncology at Rhône-Poulenc Rorer (now Sanofi) in Paris, before moving on to smaller start-ups and deciding that she wanted to become an entrepreneur. But she needed a better understanding of how to raise capital. So she joined a venture-capital firm for one year as an entrepreneur-in-residence.

M. HAKOLA

"Women tend to think they can 'get by' with $250,000 when they should be asking for $1 million."
Susan Windham-Bannister.

"If I'm going to run a venture-backed company, I need to see how things work from the other side of the table," says Horobin. The most important thing she learned, she says, is to research thoroughly which investors will be the best fit for a specific company: exchanging company equity for capital requires trust and like-mindedness. Details such as the size and age of the investors' fund, what investments they've made in the past and what they hope to do in the future can be important indicators of the potential success of the partnership.

Vosmek and other entrepreneurial experts say that women should be careful not to network and seek resources only from women's groups. Doing so often separates the would-be entrepreneur from sources of money controlled by men, points out Barbara Fox, chief executive of Avaxia Biologics, a biotechnology company in Lexington, Massachusetts.

Vosmek encourages women to make their way into the ownership structure of the company. For example, scientific advisory board members are often compensated in stock. "You don't have to be an entrepreneur to benefit from the financial upside of your science," says Vosmek. But only 6.5% of scientific advisory board members of US life-sciences firms are women. In the United Kingdom, 14.2% of board directors of firms on the London Stock Exchange's FTSE 100 index are women. That is up almost 2% since 2010, following the February 2011 publication of *Women on Boards*, a report by Mervyn Davies, the former UK trade and industry minister, which called for a minimum of 25% female board representation by 2015.

CULTIVATING CONTACTS

Learning about the entrepreneurial ecosystem is a crucial step towards breaking out of the female niche. One of the most efficient ways for a woman to network may be to join the lab of a principal investigator who has established industry contacts. Villa-Komaroff took a postdoc with Walter Gilbert, a molecular biologist at Harvard University in Cambridge, Massachusetts, who would go on to be a co-winner of the 1980 Nobel Prize in Chemistry. Unbeknown to Villa-Komaroff, Gilbert was in the middle of co-founding Biogen, one of the first biotechnology companies. Although she turned down an offer to join Biogen (now Biogen Idec, based in Weston, Massachusetts), Villa-Komaroff's interest in industry was sparked once she began to attend board meetings as a consultant to the company. She went on to serve on other company boards, until John Gilbert, Walter's son, approached her to join CytonomeST. Villa-Komaroff served as the company's chief executive before taking on her current post.

With confidence and mentors in hand, women still face the same challenges as all entrepreneurs: identifying a good idea, coming up with a business plan that gauges a product's market and attracting investment. Even as she struggles to gain a foothold, Jhaveri is certain that there is a market for the detection and treatment of ovarian cancer. But given the early stage of her work, and with no return in sight, she realizes the risk for investors. So she is turning to her networks, including contacts in philanthropy and entertainment, to see whether fund-raising events involving leading comedians will provide the money she needs. "The best entrepreneurs," says Vosmek, "are innovating the business model as well as the science." ■

Virginia Gewin *is a freelance writer in Portland, Oregon.*

TURNING POINT
Tiago Branco

Tiago Branco, a postdoc at University College London (UCL), received the 2011 Eppendorf and Science *Prize for Neurobiology on 12 November at the Society for Neuroscience annual meeting in Washington DC.*

What do you consider your most pivotal career decision?

I was in my last year of medical school at the University of Lisbon in 2002, and I had to decide whether to go on for two years of residency or do a PhD. Given my interest in research, my adviser at the time encouraged me to apply to UK graduate programmes. I decided to attend UCL. It was all very quick: I had to finish medical school on a Friday and start at UCL on the Monday. I really wanted to do this programme because I didn't have a background in neuroscience and it offered one year of lectures and experience in different labs.

Describe your PhD research.

Transmission of signals between neurons fails most of the time. My PhD studied how this is regulated — and why some junctions, or synapses, are reliable and others are not. We found that the neuron receiving the connection talks to the neurons sending the transmitter, and regulates the reliability of the synapse, so that it is not too excited or silent — it maintains a balance.

Is there something you would like to have done differently during your PhD?

The PhD is an ideal time to try high-risk research, but I naively didn't realize how crucial it is to have publications by the end of it. If I were to do things differently, I would conduct experiments that are sure to produce data, as well as try riskier things. I didn't account for the time to publication, which can be problematic under Britain's three-year PhD programme. If it takes, on average, a year to get a paper published and a year to get your technique up and running, you basically have one year to generate publishable data. My advice for PhD students, especially in Britain, is that it is important to determine the work that will define your PhD research as early as possible. It will make life much easier.

How did the lack of publications affect your career progression?

I had a paper accepted for publication one year into my postdoc. Luckily, I was able to start a postdoc at UCL without a first-authored paper. But if I hadn't had that opportunity, I don't know what I would have done. The lack of publications at the end of my PhD did knock me out of competition for a postdoc fellowship. I found out that my paper had been accepted by *Neuron* two months after I was declined for the fellowship. Timing can be as narrow as that.

You won the Eppendorf prize with an essay on how dendrites affect neurotransmission. What inspired it?

I wanted to write about why I'm motivated to study neurobiology — why I think that tracking single neurons is a good way to investigate how the brain controls behaviour. I wanted to write about my research without the constraints of a scientific publication or the worry that the paper might be shot down. It was a great exercise.

Do you think the award will help your job search?

It will increase my visibility, and might help me to get past the first round of eliminations. Selection committees are looking for something to make you stand out. I'm hoping that this increases awareness of who I am and what I've done.

How would you describe the job scene?

In Britain, the recession is definitely affecting jobs. Most universities have a freeze on hiring — which means that your options for starting a career, or a lab, are limited. The main way of starting a job in academia is to get a lectureship and apply for grants for research money. That's hardly an option at the moment because so few lectureships are available. My only chance of staying at UCL is to get a fellowship. But the odds of that are minute. And the success rate in terms of grant funding has decreased. It feels as if you have to get every career decision right or you might end up in a bad spot. ■

INTERVIEW BY VIRGINIA GEWIN

Vosmek encourages women to make their way into the ownership structure of the company. For example, scientific advisory board members are often compensated in stock. "You don't have to be an entrepreneur to benefit from the financial upside of your science," says Vosmek. But only 6.5% of scientific advisory board members of US life-sciences firms are women. In the United Kingdom, 14.2% of board directors of firms on the London Stock Exchange's FTSE 100 index are women. That is up almost 2% since 2010, following the February 2011 publication of *Women on Boards*, a report by Mervyn Davies, the former UK trade and industry minister, which called for a minimum of 25% female board representation by 2015.

CULTIVATING CONTACTS

Learning about the entrepreneurial ecosystem is a crucial step towards breaking out of the female niche. One of the most efficient ways for a woman to network may be to join the lab of a principal investigator who has established industry contacts. Villa-Komaroff took a postdoc with Walter Gilbert, a molecular biologist at Harvard University in Cambridge, Massachusetts, who would go on to be a co-winner of the 1980 Nobel Prize in Chemistry. Unbeknown to Villa-Komaroff, Gilbert was in the middle of co-founding Biogen, one of the first biotechnology companies. Although she turned down an offer to join Biogen (now Biogen Idec, based in Weston, Massachusetts), Villa-Komaroff's interest in industry was sparked once she began to attend board meetings as a consultant to the company. She went on to serve on other company boards, until John Gilbert, Walter's son, approached her to join CytonomeST. Villa-Komaroff served as the company's chief executive before taking on her current post.

With confidence and mentors in hand, women still face the same challenges as all entrepreneurs: identifying a good idea, coming up with a business plan that gauges a product's market and attracting investment. Even as she struggles to gain a foothold, Jhaveri is certain that there is a market for the detection and treatment of ovarian cancer. But given the early stage of her work, and with no return in sight, she realizes the risk for investors. So she is turning to her networks, including contacts in philanthropy and entertainment, to see whether fund-raising events involving leading comedians will provide the money she needs. "The best entrepreneurs," says Vosmek, "are innovating the business model as well as the science." ■

Virginia Gewin *is a freelance writer in Portland, Oregon.*

TURNING POINT
Tiago Branco

Tiago Branco, a postdoc at University College London (UCL), received the 2011 Eppendorf and Science *Prize for Neurobiology on 12 November at the Society for Neuroscience annual meeting in Washington DC.*

What do you consider your most pivotal career decision?

I was in my last year of medical school at the University of Lisbon in 2002, and I had to decide whether to go on for two years of residency or do a PhD. Given my interest in research, my adviser at the time encouraged me to apply to UK graduate programmes. I decided to attend UCL. It was all very quick: I had to finish medical school on a Friday and start at UCL on the Monday. I really wanted to do this programme because I didn't have a background in neuroscience and it offered one year of lectures and experience in different labs.

Describe your PhD research.

Transmission of signals between neurons fails most of the time. My PhD studied how this is regulated — and why some junctions, or synapses, are reliable and others are not. We found that the neuron receiving the connection talks to the neurons sending the transmitter, and regulates the reliability of the synapse, so that it is not too excited or silent — it maintains a balance.

Is there something you would like to have done differently during your PhD?

The PhD is an ideal time to try high-risk research, but I naively didn't realize how crucial it is to have publications by the end of it. If I were to do things differently, I would conduct experiments that are sure to produce data, as well as try riskier things. I didn't account for the time to publication, which can be problematic under Britain's three-year PhD programme. If it takes, on average, a year to get a paper published and a year to get your technique up and running, you basically have one year to generate publishable data. My advice for PhD students, especially in Britain, is that it is important to determine the work that will define your PhD research as early as possible. It will make life much easier.

How did the lack of publications affect your career progression?

I had a paper accepted for publication one year into my postdoc. Luckily, I was able to start a postdoc at UCL without a first-authored paper. But if I hadn't had that opportunity, I don't know what I would have done. The lack of publications at the end of my PhD did knock me out of competition for a postdoc fellowship. I found out that my paper had been accepted by *Neuron* two months after I was declined for the fellowship. Timing can be as narrow as that.

You won the Eppendorf prize with an essay on how dendrites affect neurotransmission. What inspired it?

I wanted to write about why I'm motivated to study neurobiology — why I think that tracking single neurons is a good way to investigate how the brain controls behaviour. I wanted to write about my research without the constraints of a scientific publication or the worry that the paper might be shot down. It was a great exercise.

Do you think the award will help your job search?

It will increase my visibility, and might help me to get past the first round of eliminations. Selection committees are looking for something to make you stand out. I'm hoping that this increases awareness of who I am and what I've done.

How would you describe the job scene?

In Britain, the recession is definitely affecting jobs. Most universities have a freeze on hiring — which means that your options for starting a career, or a lab, are limited. The main way of starting a job in academia is to get a lectureship and apply for grants for research money. That's hardly an option at the moment because so few lectureships are available. My only chance of staying at UCL is to get a fellowship. But the odds of that are minute. And the success rate in terms of grant funding has decreased. It feels as if you have to get every career decision right or you might end up in a bad spot. ■

INTERVIEW BY VIRGINIA GEWIN

1-9-4-BLUE-3-7-2-6-GAMMA-TETRAHEDRON

How to play the futures market.

BY IAN RANDAL STROCK

I always knew I was destined for great things, even as a child. It was only when I started growing up that I learned how the world worked, and realized that great wealth would make those great things much easier to attain. Unfortunately, attaining great wealth wasn't quite so easy.

I didn't find my fortune on Wall Street. My writing career looked to be pleasant, but not blockbuster. Lottery wins were hard to come by. And I didn't even have a big enough stake to take the poker world by storm.

Then, one day, I hit upon the solution. It was so simple I almost laughed with joy at it.

I didn't have to find my fortune all by myself.

As long as the Universe would allow time travel, my future self already knew how I'd made my fortune. With time travel, of course, he'd come back to tell me how to do it, ensuring that he would have that fortune when the time came.

So all I had to do was be prepared for my visit from my future self. Chance favours the prepared mind, and I was going to be prepared. I needed a foolproof way of recognizing my future self, because he might have only a moment to give me what I needed. He might not look like me any more. I needed something. A recognition code. Something I would know, he would remember, and no one else would ever even think of.

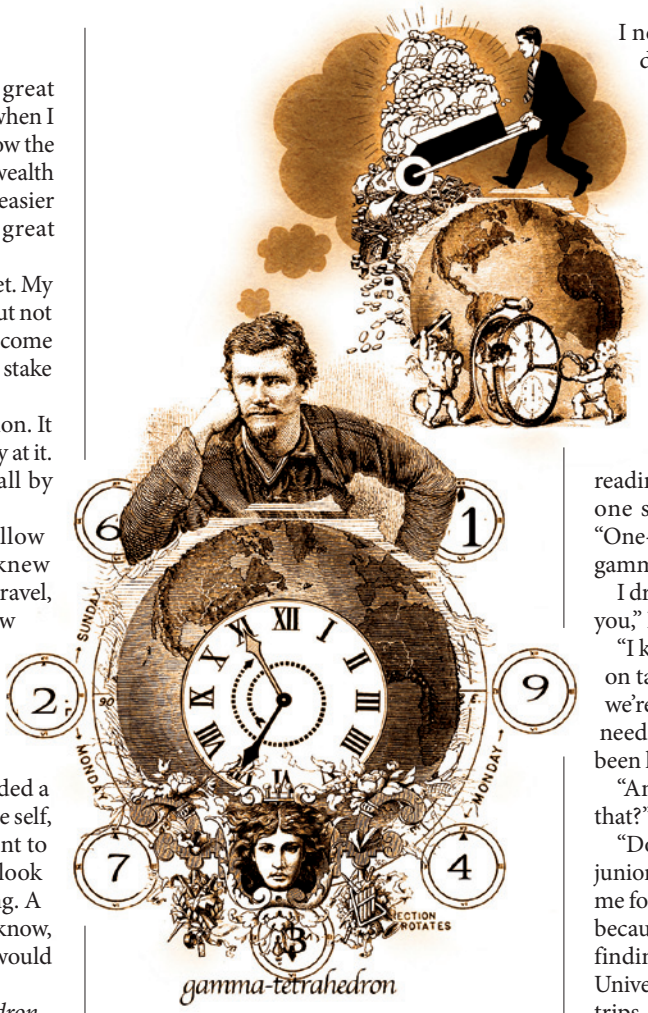
1-9-4-blue-3-7-2-6-gamma-tetrahedron.

A code. A code I repeated to myself nearly constantly at first, until it became ingrained in my brain. And then only regularly, to keep it fresh, so I would recognize it instantly. Who knew? My contact with my future self might be only a few seconds. I would need to be able to hear (or see) that code and know it immediately, so as not to waste whatever brief time interval we might have together.

1-9-4-blue-3-7-2-6-gamma-tetrahedron.

I kept plugging away, trying to write that best-selling novel (no luck yet); on Wall Street, everyone seemed to be making money but me; I hadn't yet hit a winning lottery combination. But I knew my destiny was assured. Somewhere out there was the future me who had the answers; who knew how I would make my fortune. And he'd be back to tell me. After all, he needed to tell me how to do it, so that he would have that fortune.

1-9-4-blue-3-7-2-6-gamma-tetrahedron.



JACEY

I went to work for an Internet start-up company, but it didn't make it out of the gate. I tried my hand at poker, but was only a fair player, and without a large enough bankroll, the big-money tournaments were well out of my reach. I even started several businesses, on my own and with friends, but they all came to nought.

1-9-4-blue-3-7-2-6-gamma-tetrahedron.

There were times when I was depressed, thinking it was all a cosmic joke on me, that of course there was no way my future self would be able to tell me anything. But at my core, I held fast to that feeling; it felt so right, it made the Universe make sense. It had to be. My future self would tell me how to do it.

1-9-4-blue-3-7-2-6-gamma-tetrahedron.

I tried to imagine what would be the best place, the right time for my future self to visit, to share the words or data I needed to know, and then I realized it didn't matter. My future self already knew when and where we would meet. After that meeting, I'd know it, too, and then I could remember it for us.

ONE-NINE-FOUR-BLUE-THREE-SEVEN-TWO-SIX-GAMMA-TETRAHEDRON.

It was my code. I never wrote it down; never told it to anyone; never even told anyone my code existed. It was going to work; it had to work.

1-9-4-blue-3-7-2-6-gamma-tetrahedron.

One day, I was sitting in the park, reading a book, taking a break, when someone sat down on the bench beside me. "One-nine-four-blue-three-seven-two-six-gamma-tetrahedron," he said.

I dropped my book. "I've been waiting for you," I said.

"I know," he said, in a voice I'd only heard on tape. "I'm sorry to disappoint you, but we're not going to be rich anytime soon. I need a sample of your blood, to prove I've been here."

"And then what? Will we be rich after that?"

"Doubtful," the future me said. "I'm just a junior member of the team. They only chose me for the trip because I'm expendable, and because I told them I had a foolproof way of finding my earlier self. But the fabric of the Universe won't allow more than three or four trips, so this is a proof-of-concept trip that will probably never be repeated."

"But what about the stock market, or the lottery, or —"

"I'm not rich in my time, so I can't help you, there. But can I have a sample of your blood?"

I sat there, stunned, while he drew a sample. Then he walked away. I didn't even pay attention.

8-5-omega-0-3-3-orange. 8-5-omega-0-3-3-orange.

There must be an alternate universe, one in which I find my fortune, and can travel between universes to tell myself how to do it … ■

Ian Randal Strock *is the editor of SFScope.com, the publisher of Fantastic Books, and a graduate of the Launch Pad Astronomy Workshop for Writers. His fiction has appeared in* Analog *(from which he won two AnLab Awards) and once in* Nature.

NATURE.COM
Follow Futures on Facebook at:
go.nature.com/mtoodm

ASTROPHYSICS

The final plunge

A gas cloud has been spotted approaching the Milky Way's central black hole. Observations of its closest approach, expected to occur in mid-2013, may offer insight into the black hole's immediate surroundings. SEE LETTER P.51

MARK MORRIS

From observations made at radio, infrared and X-ray wavelengths, we know that matter is continuously falling into the supermassive black hole that lies at the centre of the Milky Way. As that gaseous matter plunges towards the black hole, it heats up and emits copious electromagnetic radiation, allowing us to 'see' the black hole's immediate surroundings as a source known as Sagittarius A* (or Sgr A*). During the decade or so over which researchers have followed this emission, Sgr A* has flickered on timescales of minutes to hours — presumably because of inhomogeneities or instabilities in its accretion flow. But on longer timescales its average emission has been fairly steady[1]. This may not be the case for much longer: on page 51 of this issue, Gillessen *et al.*[2] report the discovery of a small, coherent blob of ionized gas and dust accelerating along a trajectory that they project will carry it close to the black hole[†].

Gillessen and colleagues estimate that the gas blob's closest approach to the black hole — about 260 times the mean Earth–Sun distance — will occur in mid-2013. By that time, and perhaps well before that, the blob will be disrupted by a combination of dynamical instabilities and the black hole's tidal forces. Consequently, according to the authors, the gas will be dispersed over a broad range of orbital trajectories (Fig. 1), and some of it could accrete onto the black hole in mid-to-late 2013, possibly leading to a marked, long-duration brightening of Sgr A*.

The cautiously predicted[2] flare-up would be a valuable probe of the environment immediately surrounding the black hole, about which little is known. What we do know is that the average rate of mass accretion driving the emission from Sgr A* is exceptionally small, thus explaining why our Galactic black hole is extremely dim. That we can detect it at all is due largely to its proximity — it is about 100 times closer to Earth than the nearest supermassive black hole in another galactic nucleus, that of the Andromeda galaxy. We also know that the accreting gas is heated so much that the individual gas particles almost all approach the speed of light, and that the observed emission emanates from a magnetohydrodynamic maelstrom[3]. But the basic parameters of this small-accretion-rate regime are still being sought[4]: does the accreting gas form a well-defined disk or an outflowing jet, as it often does in the more prominent cases observed in many other galaxies? Or is it better described as a more isotropic inflow comprising stochastically distributed density fluctuations?

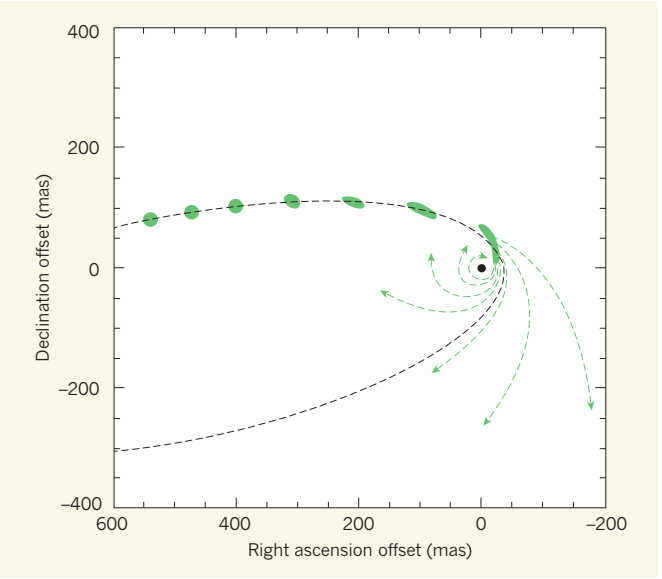


Figure 1 | On its way to the Galactic black hole. The sequence of green blobs, which are not drawn to scale, represents the successive locations of the cloud of gas and dust reported by Gillessen *et al.*[2] as it approaches the Milky Way's black hole (black spot). The dashed black line is the trajectory that a point mass would follow in the absence of tidal effects. The cloud becomes elongated because of the tidal force acting on it, especially as it nears the point of closest approach, where it is violently sheared and the trajectories of different parts of the cloud diverge (dashed green lines). The authors argue that some of the gas will circle the black hole and provoke an accretion event that might cause a brightening of Sgr A*, the currently rather dim source associated with the black hole, in mid-to-late 2013. The angular offsets from the black hole in right ascension and declination are shown in units of milliarcseconds (mas).

Gillessen and colleagues' projected 2013 approach of a gas cloud with known dynamical characteristics into this arena could offer a fresh perspective on the geometry and energetics of the accretion flow. It will be a challenging investigation; the emission that can be observed in several parts of the electromagnetic spectrum will provide a rather narrow window into the ferment of complex physics taking place there.

The scale of the incoming cloud of gas and dust is very modest: Gillessen *et al.*[2] estimate it to be about the size of the Solar System (with a radius about 120 times the mean Earth–Sun distance), and to have a mass only about three times that of Earth. It therefore has the status of a tiny fragment of a dense interstellar cloud. This raises significant questions about whether it will remain coherent as it undergoes its presumably terminal plunge through the violent, hot, low-density medium around Sgr A*.

Such a small cloud cannot possibly hold itself together with its own gravity. External gas pressure and the compression resulting from the impact of the ambient gas through which the cloud is moving, perhaps assisted by an enveloping magnetic field, might promote its confinement. But the inevitable doom of such a blob of gas is its inexorable tendency towards fragmentation, as dynamical instabilities, such as the Rayleigh–Taylor and Kelvin–Helmholtz instabilities, cleave and pare the cloud into a cluster of smaller and more ill-defined fragments, like a disintegrating satellite entering Earth's atmosphere. The fragmentation of a disintegrating gas cloud would continue on ever shorter scales of size and time until little was left but a faint, broad breeze representing the initial momentum of the incoming cloud. One might wonder why this hasn't already happened to this tiny cloud.

Gillessen *et al.*[2] have undertaken an initial numerical investigation of how the black

[†]This article and the paper[2] under discussion were published online on 14 December 2011.

hole's tidal force would disrupt a small cloud, and their results seem to be consistent with their observation of a progressive elongation of the blob, and perhaps with a gaseous tail following the blob in its orbit. But it will be necessary to add numerical hydrodynamic calculations of this phenomenon in order to assess whether the tiny cloud that the authors have found could survive the violent plunge long enough to deposit a significant amount of matter near the black hole. If not, one might consider the possibility that a much larger mass than the cloud — a relatively faint star or a stellar-mass black hole — binds the dusty gas that Gillessen *et al.* have observed in the form of a circumstellar disk as the ensemble orbits towards Sgr A*. Even then, the black hole's tidal force near closest approach may be able to wrench some of the gas from the disk, and fuel an accretion-induced brightening of the Galactic black hole. Many telescopes are likely to be watching. ■

Mark Morris *is in the Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095–1547, USA.*
e-mail: morris@astro.ucla.edu

1. Meyer, L. *et al. Astrophys. J.* **694,** L87–L91 (2009).
2. Gillessen, S. *et al. Nature* **481,** 51–54 (2012).
3. Melia, F. & Falcke, H. *Annu. Rev. Astron. Astrophys.* **39,** 309–352 (2001).
4. Genzel, R., Eisenhauer, F. & Gillessen, S. *Rev. Mod. Phys.* **82,** 3121–3195 (2010).

SYNTHETIC BIOLOGY

Bacteria collaborate to sense arsenic

A method developed to allow rapid communication between bacterial cells across long distances enables the cells to detect arsenic collectively, and to report it as an oscillatory output. **See Article p.39**

CHRISTOPHER A. VOIGT

Living cells can be exploited to sense and process environmental stimuli, including poorly defined microenvironments, biological markers of disease, defects in materials and complex small molecules. But obtaining a reliable signal from individual cells has proved a challenge. On page 39 of this issue, Prindle *et al.*[1] report a solution to this problem: a sensor composed of millions of bacterial cells that communicate with each other over long distances (up to 2.4 centimetres). The cells respond to the presence of arsenic by altering the rate at which they produce synchronized pulses of fluorescence.

Cells have been engineered to sense many environmental signals, including light, chemicals, touch, metal ions and pH. For example, sensors have been made in which human olfactory receptors are expressed in yeast cells[2]. Most cellular sensors are based on a protein or messenger RNA that responds to a signal by causing the expression of a gene. Such genetic sensors often suffer from low dynamic range (that is, there is little change in output between the absence and presence of a signal) and nonspecificity (they are activated by multiple signals). Furthermore, because cells are living systems, individual responses may vary because of stochastic effects or differences in growth states.

Prindle *et al.*[1] have addressed the problem of dynamic range by applying the principles of signal processing to a biosensor based on genetic circuits. Such circuits use biochemical interactions to produce functions analogous to their electronic counterparts. Previously, the same group built a robust genetic oscillator[3] — a network of genes and proteins that produced regular pulses of molecules — and used this as a time-keeping mechanism to control cell–cell communication between bacteria. This yielded populations of bacteria that expressed a fluorescent protein in unison, and so produced synchronized pulses of light.

Theoretically, such an oscillator would enable a sensor to use the frequency of oscillations as signals, making the sensor less sensitive to environmental noise and exposure time than systems based on steady-state signals. A problem with the previous oscillator[3], however, was that cell–cell coupling relied on the diffusion of a small molecule through cellular media, a process that is too slow to allow rapid, long-range coupling of millions of cells. Molecular diffusion in the gas phase is much faster, so Prindle *et al.*[1] used this mechanism to accelerate the coupling between separate colonies of bacteria.

In this way, the authors were able to couple 2.5 million cells of the bacterium *Escherichia coli*, which were arranged as an array of colonies across a distance of 5 millimetres. As in the previously reported oscillator[3], the output of the system[1] was the coordinated, oscillating expression of a fluorescent protein, which the authors detected using a microscope. The period of the oscillations was quite long (more than an hour), but the degree of synchronization was high — the colonies produced light pulses within 2 minutes of each other.

To demonstrate a potential application of their system, Prindle *et al.* 'rewired' their network to incorporate elements that respond to arsenic. The resulting system acted as an arsenic sensor: once the concentration of arsenic reached a threshold value, the amplitude and period of the oscillations increased significantly. This produced a device with a large dynamic range. What's more, because the device averaged the outputs of a population of cells, noise was reduced and the sensor's response was decoupled from the growth state of individual cells. The authors scaled up their device so that it included more than 12,000 communicating bacterial colonies, covering an area of 2.4 × 1.2 centimetres.

Several advances, yet to be achieved, would improve the ability to connect this living sensor to an electronic system. One problem is that the minimum response time of the sensor to an input signal is slow, because gene expression — which takes about 20 minutes to occur — is required. Another issue is that the output involves fluorescence, which is awkward for electronic devices to use; the ideal output would be a direct electrical signal. To this end, cells have been metabolically engineered so that they can be induced to release electrons, which can then be read by an electronic sensor[4]. The electron-transport system found in bacterial nanowires (extracellular appendages that conduct electricity[5]) has also been harnessed[6] to link cells to an electronic system. Nevertheless, these strategies still require the expression of a gene that triggers electron flux, so the resulting sensors are relatively slow to respond to signals.

More broadly, there are several collaborative research efforts aiming to develop better toolboxes for building interfaces between cellular and electronic components. One such project is to build a millimetre-scale robot that swims like a lamprey, using a combination of human muscle cells, yeast-based sensors, an electronic brain and flexible materials (nicknamed 'cyberplasm')[7]. Another project is to develop genetic sensors, along with genetic circuits to apply signal processing within the cell[8] and new approaches to link cellular outputs to an electronic system, with the ultimate objective of controlling robots[9]. As the integration of cellular and electronic systems matures, it will be interesting to see how circuitry in future devices is divided between biological and electronic components. ■

Christopher A. Voigt *is in the Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.*
e-mail: cavoigt@gmail.com

hole's tidal force would disrupt a small cloud, and their results seem to be consistent with their observation of a progressive elongation of the blob, and perhaps with a gaseous tail following the blob in its orbit. But it will be necessary to add numerical hydrodynamic calculations of this phenomenon in order to assess whether the tiny cloud that the authors have found could survive the violent plunge long enough to deposit a significant amount of matter near the black hole. If not, one might consider the possibility that a much larger mass than the cloud — a relatively faint star or a stellar-mass black hole — binds the dusty gas that Gillessen *et al.* have observed in the form of a circumstellar disk as the ensemble orbits towards Sgr A*. Even then, the black hole's tidal force near closest approach may be able to wrench some of the gas from the disk, and fuel an accretion-induced brightening of the Galactic black hole. Many telescopes are likely to be watching. ■

Mark Morris *is in the Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095–1547, USA.*
e-mail: morris@astro.ucla.edu

1. Meyer, L. *et al. Astrophys. J.* **694,** L87–L91 (2009).
2. Gillessen, S. *et al. Nature* **481,** 51–54 (2012).
3. Melia, F. & Falcke, H. *Annu. Rev. Astron. Astrophys.* **39,** 309–352 (2001).
4. Genzel, R., Eisenhauer, F. & Gillessen, S. *Rev. Mod. Phys.* **82,** 3121–3195 (2010).

SYNTHETIC BIOLOGY

Bacteria collaborate to sense arsenic

A method developed to allow rapid communication between bacterial cells across long distances enables the cells to detect arsenic collectively, and to report it as an oscillatory output. **See Article p.39**

CHRISTOPHER A. VOIGT

Living cells can be exploited to sense and process environmental stimuli, including poorly defined microenvironments, biological markers of disease, defects in materials and complex small molecules. But obtaining a reliable signal from individual cells has proved a challenge. On page 39 of this issue, Prindle *et al.*[1] report a solution to this problem: a sensor composed of millions of bacterial cells that communicate with each other over long distances (up to 2.4 centimetres). The cells respond to the presence of arsenic by altering the rate at which they produce synchronized pulses of fluorescence.

Cells have been engineered to sense many environmental signals, including light, chemicals, touch, metal ions and pH. For example, sensors have been made in which human olfactory receptors are expressed in yeast cells[2]. Most cellular sensors are based on a protein or messenger RNA that responds to a signal by causing the expression of a gene. Such genetic sensors often suffer from low dynamic range (that is, there is little change in output between the absence and presence of a signal) and nonspecificity (they are activated by multiple signals). Furthermore, because cells are living systems, individual responses may vary because of stochastic effects or differences in growth states.

Prindle *et al.*[1] have addressed the problem of dynamic range by applying the principles of signal processing to a biosensor based on genetic circuits. Such circuits use biochemical interactions to produce functions analogous to their electronic counterparts. Previously, the same group built a robust genetic oscillator[3] — a network of genes and proteins that produced regular pulses of molecules — and used this as a time-keeping mechanism to control cell–cell communication between bacteria. This yielded populations of bacteria that expressed a fluorescent protein in unison, and so produced synchronized pulses of light.

Theoretically, such an oscillator would enable a sensor to use the frequency of oscillations as signals, making the sensor less sensitive to environmental noise and exposure time than systems based on steady-state signals. A problem with the previous oscillator[3], however, was that cell–cell coupling relied on the diffusion of a small molecule through cellular media, a process that is too slow to allow rapid, long-range coupling of millions of cells. Molecular diffusion in the gas phase is much faster, so Prindle *et al.*[1] used this mechanism to accelerate the coupling between separate colonies of bacteria.

In this way, the authors were able to couple 2.5 million cells of the bacterium *Escherichia coli*, which were arranged as an array of colonies across a distance of 5 millimetres. As in the previously reported oscillator[3], the output of the system[1] was the coordinated, oscillating expression of a fluorescent protein, which the authors detected using a microscope. The period of the oscillations was quite long (more than an hour), but the degree of synchronization was high — the colonies produced light pulses within 2 minutes of each other.

To demonstrate a potential application of their system, Prindle *et al.* 'rewired' their network to incorporate elements that respond to arsenic. The resulting system acted as an arsenic sensor: once the concentration of arsenic reached a threshold value, the amplitude and period of the oscillations increased significantly. This produced a device with a large dynamic range. What's more, because the device averaged the outputs of a population of cells, noise was reduced and the sensor's response was decoupled from the growth state of individual cells. The authors scaled up their device so that it included more than 12,000 communicating bacterial colonies, covering an area of 2.4 × 1.2 centimetres.

Several advances, yet to be achieved, would improve the ability to connect this living sensor to an electronic system. One problem is that the minimum response time of the sensor to an input signal is slow, because gene expression — which takes about 20 minutes to occur — is required. Another issue is that the output involves fluorescence, which is awkward for electronic devices to use; the ideal output would be a direct electrical signal. To this end, cells have been metabolically engineered so that they can be induced to release electrons, which can then be read by an electronic sensor[4]. The electron-transport system found in bacterial nanowires (extracellular appendages that conduct electricity[5]) has also been harnessed[6] to link cells to an electronic system. Nevertheless, these strategies still require the expression of a gene that triggers electron flux, so the resulting sensors are relatively slow to respond to signals.

More broadly, there are several collaborative research efforts aiming to develop better toolboxes for building interfaces between cellular and electronic components. One such project is to build a millimetre-scale robot that swims like a lamprey, using a combination of human muscle cells, yeast-based sensors, an electronic brain and flexible materials (nicknamed 'cyberplasm')[7]. Another project is to develop genetic sensors, along with genetic circuits to apply signal processing within the cell[8] and new approaches to link cellular outputs to an electronic system, with the ultimate objective of controlling robots[9]. As the integration of cellular and electronic systems matures, it will be interesting to see how circuitry in future devices is divided between biological and electronic components. ■

Christopher A. Voigt *is in the Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.*
e-mail: cavoigt@gmail.com

1. Prindle, A. *et al. Nature* **481,** 39–44 (2012).
2. Radhika, V. *et al. Nature Chem. Biol.* **3,** 325–330 (2007).
3. Danino, T. *et al. Nature* **463,** 326–330 (2010).
4. Weber, W. *et al. Nucleic Acids Res.* **37,** e33 (2009).
5. El-Naggar, M. Y. *et al. Proc. Natl Acad. Sci. USA* **107,** 18127–18131 (2010).
6. Jensen, H. M. *Proc. Natl Acad. Sci. USA* **107,** 19213–19218 (2010).
7. Mervis, J. *Science* **324,** 1128–1129 (2009).
8. Khalil, A. S. & Collins, J. J. *Nature Rev. Genet.* **11,** 367–379 (2010).
9. Sakar, M. S. *et al. Int. J. Robotics Res.* **30,** 647–658 (2011).

OPTICAL PHYSICS

How to hide in time

As if the idea of a device that makes an object seem invisible was not mind-boggling enough, researchers have now demonstrated a system that can conceal an event in time. **See Letter p.62**

ROBERT W. BOYD & ZHIMIN SHI

An exciting development in optical physics has been the proposal[1,2] and subsequent demonstration[3–6] of a spatial cloak, a structure that can render invisible any object placed in a specific region of space. Writing in this issue, Fridman *et al.*[7] (page 62) extend this concept by demonstrating a temporal cloak — a device that hides events occurring during a specific time window.

Let us first describe the operation of a spatial cloak. One example of such a device consists of a shell that surrounds the object to be hidden[2]. Using a method known as transformation optics, the way in which the refractive index changes across the material that constitutes the shell is set such that any light ray incident on the shell is deflected so as to miss the object to be hidden. The ray is redirected so that, when it leaves the shell, it is travelling in the same direction as if both the shell and the object hidden inside had not been present at all.

The experimental realization of spatial cloaking is intimately related to the development of optical metamaterials[8]. These are artificial materials with highly controllable optical properties that can be very different from those of naturally occurring materials. A prime example is a metamaterial designed to have a negative refractive index so that it bends light rays in the opposite direction to that in which conventional materials do. So far, spatial cloaking has been realized in, for example, a cylindrical geometry at radio frequency[3] and 'carpet' geometries at infrared[4,5] and visible[6] wavelengths.

The concept of cloaking has been extended to cloaking in time by a recent theoretical treatment[9]. This work showed that a time gap can be opened in an optical wave by locally manipulating the speed of light such that the front and rear parts of the wave get accelerated and slowed down, respectively. Any event that occurs within the resulting time gap — in which no light is present — would be rendered invisible to someone monitoring the transmitted light wave. However, the presence of this time gap in the light intensity would be a clear indication that someone had tampered with the time history of the system. The gap can be closed by subsequently reversing the modification of the light's speed as it leaves the 'interrogation region' that is to be cloaked. In this way, the previously accelerated light gets slowed down and the previously slowed-down light gets accelerated. When the restored light reaches an observer, a continuous, uniform light field is observed, and there is no indication that some event has occurred.

In their experimental study of temporal cloaking, Fridman *et al.*[7] made use of time-lenses and dispersive media[10]. To understand the principle of a time-lens, we should recall that a conventional optical lens is a device that can cause an incident light beam to converge or diverge spatially. From a mathematical perspective, the spatial and temporal evolution of light are quite similar, and therefore the principle of a lens can be extended to a time-lens.

A time-lens modifies a light field's temporal, rather than spatial, distribution. An ideal time-lens changes the colour of the light field at different moments in time. This modified light field is then passed through a dispersive medium in which different colours of light travel at different velocities and therefore emerge from the medium with different time delays. When the system is properly designed, all the colours can be made to arrive at a given spatial point at the same time, or, by analogy with a conventional lens, they can all be 'focused' to the same point in time.

In their work, Fridman and colleagues used a split time-lens, which is a slight modification of a time-lens. This lens is composed of two half time-lenses, which are connected at their tips. The light passing through the first half of the split time-lens experiences a colour change in the opposite direction to that passing through the second half: the first half makes the light bluer and the second half makes it redder. Then, after passing through a dispersive medium — an optical fibre in the

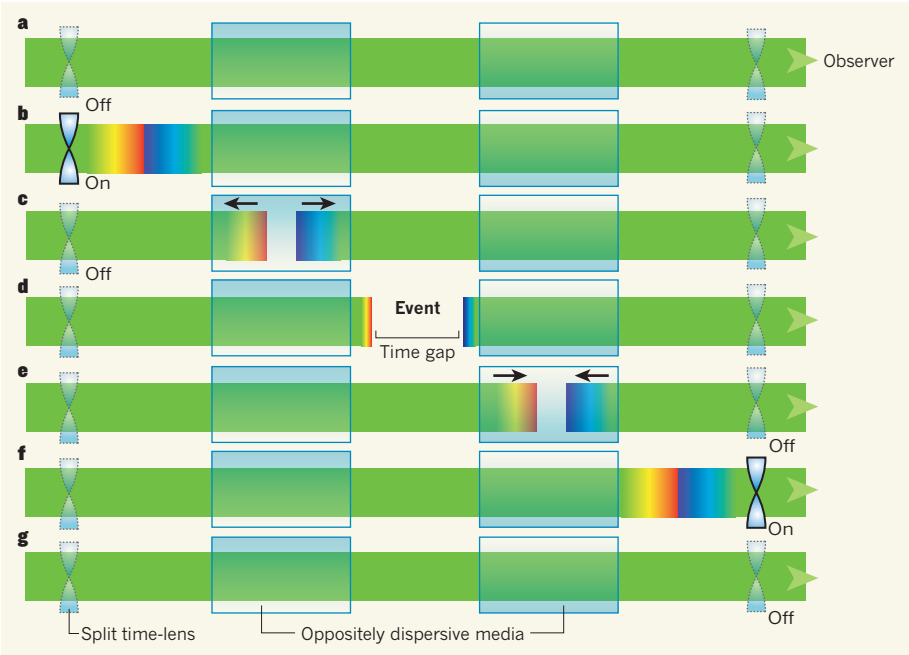


Figure 1 | Schematics of the temporal cloaking system of Fridman and colleagues[7] at different times. **a**, A continuous stream of green light passes through the system from left to right. **b**, The first time-lens is turned on and the light's colour changes as a function of time. **c**, The modified light travels through a dispersive medium. Because in this medium the blue-shifted light travels faster than the green, and the red-shifted light slower, a time-gap gradually opens up. **d**, The time gap is maximally open, and in its centre an event occurs in the form of a light pulse (not shown). **e**, The time gap gradually closes as the light passes through an oppositely dispersive medium from the first one. **f**, After the time gap is completely closed, a second time-lens is turned on such that all the colours are changed back to green. **g**, All the observer sees is basically a continuous green light as if the event in **d** never occurred.

1. Prindle, A. *et al. Nature* **481,** 39–44 (2012).
2. Radhika, V. *et al. Nature Chem. Biol.* **3,** 325–330 (2007).
3. Danino, T. *et al. Nature* **463,** 326–330 (2010).
4. Weber, W. *et al. Nucleic Acids Res.* **37,** e33 (2009).
5. El-Naggar, M. Y. *et al. Proc. Natl Acad. Sci. USA* **107,** 18127–18131 (2010).
6. Jensen, H. M. *Proc. Natl Acad. Sci. USA* **107,** 19213–19218 (2010).
7. Mervis, J. *Science* **324,** 1128–1129 (2009).
8. Khalil, A. S. & Collins, J. J. *Nature Rev. Genet.* **11,** 367–379 (2010).
9. Sakar, M. S. *et al. Int. J. Robotics Res.* **30,** 647–658 (2011).

OPTICAL PHYSICS

How to hide in time

As if the idea of a device that makes an object seem invisible was not mind-boggling enough, researchers have now demonstrated a system that can conceal an event in time. **See Letter p.62**

ROBERT W. BOYD & ZHIMIN SHI

An exciting development in optical physics has been the proposal[1,2] and subsequent demonstration[3–6] of a spatial cloak, a structure that can render invisible any object placed in a specific region of space. Writing in this issue, Fridman *et al.*[7] (page 62) extend this concept by demonstrating a temporal cloak — a device that hides events occurring during a specific time window.

Let us first describe the operation of a spatial cloak. One example of such a device consists of a shell that surrounds the object to be hidden[2]. Using a method known as transformation optics, the way in which the refractive index changes across the material that constitutes the shell is set such that any light ray incident on the shell is deflected so as to miss the object to be hidden. The ray is redirected so that, when it leaves the shell, it is travelling in the same direction as if both the shell and the object hidden inside had not been present at all.

The experimental realization of spatial cloaking is intimately related to the development of optical metamaterials[8]. These are artificial materials with highly controllable optical properties that can be very different from those of naturally occurring materials. A prime example is a metamaterial designed to have a negative refractive index so that it bends light rays in the opposite direction to that in which conventional materials do. So far, spatial cloaking has been realized in, for example, a cylindrical geometry at radio frequency[3] and 'carpet' geometries at infrared[4,5] and visible[6] wavelengths.

The concept of cloaking has been extended to cloaking in time by a recent theoretical treatment[9]. This work showed that a time gap can be opened in an optical wave by locally manipulating the speed of light such that the front and rear parts of the wave get accelerated and slowed down, respectively. Any event that occurs within the resulting time gap — in which no light is present — would be rendered invisible to someone monitoring the transmitted light wave. However, the presence of this time gap in the light intensity would be a clear indication that someone had tampered with the time history of the system. The gap can be closed by subsequently reversing the modification of the light's speed as it leaves the 'interrogation region' that is to be cloaked. In this way, the previously accelerated light gets slowed down and the previously slowed-down light gets accelerated. When the restored light reaches an observer, a continuous, uniform light field is observed, and there is no indication that some event has occurred.

In their experimental study of temporal cloaking, Fridman *et al.*[7] made use of time-lenses and dispersive media[10]. To understand the principle of a time-lens, we should recall that a conventional optical lens is a device that can cause an incident light beam to converge or diverge spatially. From a mathematical perspective, the spatial and temporal evolution of light are quite similar, and therefore the principle of a lens can be extended to a time-lens.

A time-lens modifies a light field's temporal, rather than spatial, distribution. An ideal time-lens changes the colour of the light field at different moments in time. This modified light field is then passed through a dispersive medium in which different colours of light travel at different velocities and therefore emerge from the medium with different time delays. When the system is properly designed, all the colours can be made to arrive at a given spatial point at the same time, or, by analogy with a conventional lens, they can all be 'focused' to the same point in time.

In their work, Fridman and colleagues used a split time-lens, which is a slight modification of a time-lens. This lens is composed of two half time-lenses, which are connected at their tips. The light passing through the first half of the split time-lens experiences a colour change in the opposite direction to that passing through the second half: the first half makes the light bluer and the second half makes it redder. Then, after passing through a dispersive medium — an optical fibre in the

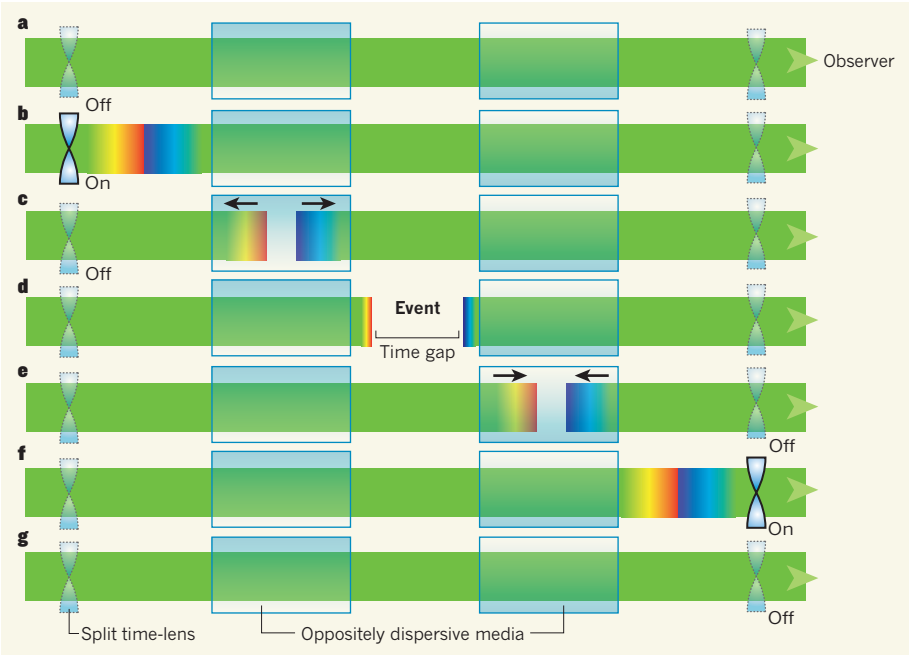


Figure 1 | Schematics of the temporal cloaking system of Fridman and colleagues[7] at different times. **a**, A continuous stream of green light passes through the system from left to right. **b**, The first time-lens is turned on and the light's colour changes as a function of time. **c**, The modified light travels through a dispersive medium. Because in this medium the blue-shifted light travels faster than the green, and the red-shifted light slower, a time-gap gradually opens up. **d**, The time gap is maximally open, and in its centre an event occurs in the form of a light pulse (not shown). **e**, The time gap gradually closes as the light passes through an oppositely dispersive medium from the first one. **f**, After the time gap is completely closed, a second time-lens is turned on such that all the colours are changed back to green. **g**, All the observer sees is basically a continuous green light as if the event in **d** never occurred.

authors' study — the light from the first half experiences a negative time delay (it accelerates) compared with the original green light, whereas that from the second half experiences a positive time delay (it slows down). This opens up a time gap of approximately 50 picoseconds in the transmitted light intensity. Afterwards, the time gap is closed seamlessly using similar techniques involving an oppositely dispersive medium from the first one and a second split time-lens (Fig. 1).

To demonstrate temporal cloaking in this system, Fridman *et al.*[7] created an 'event' in the form of a light pulse, at the centre of the time gap, that has a different frequency from that of the light passing through the system. The temporal cloaking is turned on or off by controlling the operation of the split time-lenses using additional laser light. The authors found that the detected signal associated with this event becomes more than tenfold weaker than the event's original signal. This result demonstrates that the event has been cloaked.

The distinction between temporal and spatial cloaking can be understood in terms of a metaphor involving automobile traffic. A spatial cloak acts like a junction in the form of a 'cloverleaf' interchange or flyover, in which the traffic is guided (by slip roads) to bend around a certain region of space. After passing through the junction, the traffic continues in the same direction as if the junction did not exist. By contrast, a temporal cloak behaves like a railway crossing. Traffic is stopped when a train passes, forming a gap in the traffic. After the train has passed the crossing, the stopped cars speed up until they catch up with the traffic in front of them, and the fact that a train has crossed the intersection cannot be deduced by observing the traffic flow.

Because spatial and temporal cloaking work in different physical dimensions — space and time, respectively — there is no fundamental reason why the two techniques cannot be combined so that full spatial–temporal cloaking could be turned on or off at will. Nonetheless, what Fridman *et al.* have demonstrated as a first unidirectional temporal cloaking device could already be useful in some applications, such as enhancing the security of communication in fibre-optic systems. Future directions may include increasing the cloaking time towards the order of microseconds to milliseconds, and building a device that can work simultaneously for incident light coming from different directions. ■

Robert W. Boyd *and* **Zhimin Shi** *are at the Institute of Optics, University of Rochester, Rochester, New York 14627, USA.* **R.W.B.** *is also in the Department of Physics and the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario K2P 1A3, Canada.*
e-mails: boyd@optics.rochester.edu; zshi@optics.rochester.edu

1. Leonhardt, U. *Science* **312,** 1777–1780 (2006).
2. Pendry, J. B., Schurig, D. & Smith, D. R. *Science* **312,** 1780–1782 (2006).
3. Schurig, D. *et al. Science* **314,** 977–980 (2006).
4. Gabrielli, L. H., Cardenas, J., Poitras, C. B. & Lipson, M. *Nature Photon.* **3,** 461–463 (2009).
5. Valentine, J., Li, J., Zentgraf, T., Bartal, G. & Zhang, X. *Nature Mater.* **8,** 568–571 (2009).
6. Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. & Wegener, M. *Science* **328,** 337–339 (2010).
7. Fridman, M., Farsi, A., Okawachi, Y. & Gaeta, A. L. *Nature* **481,** 62–65 (2012).
8. Cai, W. & Shalaev, V. *Optical Metamaterials: Fundamentals and Applications* (Springer, 2009).
9. McCall, M. W., Favaro, A., Kinsler, P. & Boardman, A. *J. Opt.* **13,** 024003 (2011).
10. Kolner, B. H. & Nazarathy, M. *Opt. Lett.* **14,** 630–632 (1989).

NEUROSCIENCE

Behavioural effects of cocaine reversed

Cocaine use causes lasting changes in behaviour by altering the strength of connections between neurons. The finding that these changes can be reversed in mice suggests strategies that could be used to treat drug addiction. **SEE LETTER P.71**

MARINA E. WOLF

Synaptic plasticity — the process by which connections (synapses) between nerve cells grow stronger or weaker depending on their activity level — is essential to normal development and learning. But synaptic plasticity also has a role in brain disease, including that resulting from drug abuse. Understanding this role is a challenging problem. Over the past decade, drug-addiction researchers have made progress towards this goal, aided by the fact that different facets of addiction can be modelled in animals and involve well-characterized brain circuits. From these studies, we know that drugs of abuse produce synaptic plasticity in the brain's 'reward circuitry' and that this contributes to addiction-related behaviours. On page 71 of this issue, Pascoli *et al.*[1] report that reversal of cocaine-induced synaptic plasticity in mice resets such behaviours to the pre-cocaine baseline.

Pascoli *et al.* studied synaptic plasticity associated with cocaine-induced behavioural sensitization — the increased behavioural response to a drug that occurs over the course of repeated administration and which persists long after drug exposure is discontinued. Even a single cocaine exposure in mice can cause sensitization to the drug's locomotor stimulatory effects (hyperactivity), thereby enhancing the locomotor response to a subsequent 'challenge' injection of cocaine. Opinions differ about the clinical relevance of sensitization, but according to one influential addiction theory, cocaine's incentive motivational properties (which make users want it) undergo sensitization[2].

The authors focused on part of the brain known as the nucleus accumbens and its major cell type, the medium spiny neuron (MSN). These neurons receive and integrate input signals — in the form of glutamate molecules — from cortical and limbic brain regions that control motivated behaviours, and then signal the motor circuitry to trigger a behavioural response (Fig. 1). There is evidence that glutamate synapses to MSNs are strengthened in cocaine-sensitized rodents[3], but exactly which synapses are strengthened, and how this relates to sensitization, is controversial.

To address these fundamental questions, Pascoli *et al.* gave mice a single injection of saline or cocaine, using enough of the drug to produce 'one-shot' locomotor sensitization. They took brain slices from the mice 7 days later, and used high-frequency stimulation (HFS; a series of electrical pulses) to produce long-term potentiation of glutamate synapses onto MSNs of the nucleus accumbens in the slices. Long-term potentiation (LTP) is a form of synaptic plasticity in which increased activity strengthens glutamate synapses, often through insertion into the neuronal membrane of additional glutamate receptors known as AMPA receptors. The authors found that the magnitude of HFS-induced LTP produced in cocaine-exposed neurons was approximately half of that observed in saline-treated controls. This could be explained if cocaine selectively eliminates HFS-induced LTP in a subpopulation of MSNs. But which one?

MSNs can be classified according to whether they express the D1 or D2 subtype of dopamine receptor (D1R or D2R). These subpopulations generally have distinct projection targets and different functions, although the distinctions are less clear in the nucleus accumbens than in the neighbouring dorsal striatum[4]. To distinguish between these subpopulations, Pascoli *et al.*[1] used transgenic mice that express green fluorescent protein in either D1R- or D2R-expressing MSNs. They thus observed that a single exposure to cocaine abolished HFS-induced LTP selectively in D1R neurons.

Crucially, the authors found that the abolition was not caused by impairment of mechanisms

authors' study — the light from the first half experiences a negative time delay (it accelerates) compared with the original green light, whereas that from the second half experiences a positive time delay (it slows down). This opens up a time gap of approximately 50 picoseconds in the transmitted light intensity. Afterwards, the time gap is closed seamlessly using similar techniques involving an oppositely dispersive medium from the first one and a second split time-lens (Fig. 1).

To demonstrate temporal cloaking in this system, Fridman *et al.*[7] created an 'event' in the form of a light pulse, at the centre of the time gap, that has a different frequency from that of the light passing through the system. The temporal cloaking is turned on or off by controlling the operation of the split time-lenses using additional laser light. The authors found that the detected signal associated with this event becomes more than tenfold weaker than the event's original signal. This result demonstrates that the event has been cloaked.

The distinction between temporal and spatial cloaking can be understood in terms of a metaphor involving automobile traffic. A spatial cloak acts like a junction in the form of a 'cloverleaf' interchange or flyover, in which the traffic is guided (by slip roads) to bend around a certain region of space. After passing through the junction, the traffic continues in the same direction as if the junction did not exist. By contrast, a temporal cloak behaves like a railway crossing. Traffic is stopped when a train passes, forming a gap in the traffic. After the train has passed the crossing, the stopped cars speed up until they catch up with the traffic in front of them, and the fact that a train has crossed the intersection cannot be deduced by observing the traffic flow.

Because spatial and temporal cloaking work in different physical dimensions — space and time, respectively — there is no fundamental reason why the two techniques cannot be combined so that full spatial–temporal cloaking could be turned on or off at will. Nonetheless, what Fridman *et al.* have demonstrated as a first unidirectional temporal cloaking device could already be useful in some applications, such as enhancing the security of communication in fibre-optic systems. Future directions may include increasing the cloaking time towards the order of microseconds to milliseconds, and building a device that can work simultaneously for incident light coming from different directions. ■

Robert W. Boyd *and* **Zhimin Shi** *are at the Institute of Optics, University of Rochester, Rochester, New York 14627, USA.* **R.W.B.** *is also in the Department of Physics and the School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario K2P 1A3, Canada.*
e-mails: boyd@optics.rochester.edu; zshi@optics.rochester.edu

1. Leonhardt, U. *Science* **312,** 1777–1780 (2006).
2. Pendry, J. B., Schurig, D. & Smith, D. R. *Science* **312,** 1780–1782 (2006).
3. Schurig, D. *et al. Science* **314,** 977–980 (2006).
4. Gabrielli, L. H., Cardenas, J., Poitras, C. B. & Lipson, M. *Nature Photon.* **3,** 461–463 (2009).
5. Valentine, J., Li, J., Zentgraf, T., Bartal, G. & Zhang, X. *Nature Mater.* **8,** 568–571 (2009).
6. Ergin, T., Stenger, N., Brenner, P., Pendry, J. B. & Wegener, M. *Science* **328,** 337–339 (2010).
7. Fridman, M., Farsi, A., Okawachi, Y. & Gaeta, A. L. *Nature* **481,** 62–65 (2012).
8. Cai, W. & Shalaev, V. *Optical Metamaterials: Fundamentals and Applications* (Springer, 2009).
9. McCall, M. W., Favaro, A., Kinsler, P. & Boardman, A. *J. Opt.* **13,** 024003 (2011).
10. Kolner, B. H. & Nazarathy, M. *Opt. Lett.* **14,** 630–632 (1989).

NEUROSCIENCE

Behavioural effects of cocaine reversed

Cocaine use causes lasting changes in behaviour by altering the strength of connections between neurons. The finding that these changes can be reversed in mice suggests strategies that could be used to treat drug addiction. **SEE LETTER P.71**

MARINA E. WOLF

Synaptic plasticity — the process by which connections (synapses) between nerve cells grow stronger or weaker depending on their activity level — is essential to normal development and learning. But synaptic plasticity also has a role in brain disease, including that resulting from drug abuse. Understanding this role is a challenging problem. Over the past decade, drug-addiction researchers have made progress towards this goal, aided by the fact that different facets of addiction can be modelled in animals and involve well-characterized brain circuits. From these studies, we know that drugs of abuse produce synaptic plasticity in the brain's 'reward circuitry' and that this contributes to addiction-related behaviours. On page 71 of this issue, Pascoli *et al.*[1] report that reversal of cocaine-induced synaptic plasticity in mice resets such behaviours to the pre-cocaine baseline.

Pascoli *et al.* studied synaptic plasticity associated with cocaine-induced behavioural sensitization — the increased behavioural response to a drug that occurs over the course of repeated administration and which persists long after drug exposure is discontinued. Even a single cocaine exposure in mice can cause sensitization to the drug's locomotor stimulatory effects (hyperactivity), thereby enhancing the locomotor response to a subsequent 'challenge' injection of cocaine. Opinions differ about the clinical relevance of sensitization, but according to one influential addiction theory, cocaine's incentive motivational properties (which make users want it) undergo sensitization[2].

The authors focused on part of the brain known as the nucleus accumbens and its major cell type, the medium spiny neuron (MSN). These neurons receive and integrate input signals — in the form of glutamate molecules — from cortical and limbic brain regions that control motivated behaviours, and then signal the motor circuitry to trigger a behavioural response (Fig. 1). There is evidence that glutamate synapses to MSNs are strengthened in cocaine-sensitized rodents[3], but exactly which synapses are strengthened, and how this relates to sensitization, is controversial.

To address these fundamental questions, Pascoli *et al.* gave mice a single injection of saline or cocaine, using enough of the drug to produce 'one-shot' locomotor sensitization. They took brain slices from the mice 7 days later, and used high-frequency stimulation (HFS; a series of electrical pulses) to produce long-term potentiation of glutamate synapses onto MSNs of the nucleus accumbens in the slices. Long-term potentiation (LTP) is a form of synaptic plasticity in which increased activity strengthens glutamate synapses, often through insertion into the neuronal membrane of additional glutamate receptors known as AMPA receptors. The authors found that the magnitude of HFS-induced LTP produced in cocaine-exposed neurons was approximately half of that observed in saline-treated controls. This could be explained if cocaine selectively eliminates HFS-induced LTP in a subpopulation of MSNs. But which one?

MSNs can be classified according to whether they express the D1 or D2 subtype of dopamine receptor (D1R or D2R). These subpopulations generally have distinct projection targets and different functions, although the distinctions are less clear in the nucleus accumbens than in the neighbouring dorsal striatum[4]. To distinguish between these subpopulations, Pascoli *et al.*[1] used transgenic mice that express green fluorescent protein in either D1R- or D2R-expressing MSNs. They thus observed that a single exposure to cocaine abolished HFS-induced LTP selectively in D1R neurons.

Crucially, the authors found that the abolition was not caused by impairment of mechanisms

that elicit LTP. Rather, they showed that synapses in the D1R neurons had been strengthened by cocaine, and therefore could not be further potentiated by HFS. This cocaine-induced potentiation depended on the same mechanisms that underlie HFS-induced LTP in the nucleus accumbens, namely activation of a subgroup of glutamate receptors called NMDA receptors, and the ERK signalling pathway. So, when Pascoli *et al.* gave mice an inhibitor of the ERK pathway before cocaine injection, no cocaine-induced potentiation occurred, allowing HFS-induced LTP to be elicited.

The authors also observed that the time course of cocaine-induced potentiation in D1R MSNs mirrored that of locomotor sensitization: both sensitization and synaptic potentiation were observed a week after cocaine injection, but were absent after a month. The million-dollar question is whether synaptic potentiation is causally involved in sensitization. To find out, Pascoli *et al.* tested whether reversal of synaptic potentiation — a process known as depotentiation — would also reverse locomotor sensitization. To do this, they used an optogenetic method in which nerve terminals projecting from the infralimbic cortex to the nucleus accumbens in mice were modified so that they could be depolarized by light pulses. The authors gave these mice a single injection of cocaine, and applied light pulses 7 days later using a protocol that produces depotentiation at MSN synapses. Remarkably, when they gave the mice a challenge injection of cocaine 45 minutes after depotentiation, the 'depotentiated' mice did not exhibit locomotor sensitization.

Pascoli *et al.* confirmed these findings by using a different protocol (consisting of five daily cocaine injections) that produces more robust sensitization in mice. They found that when optogenetic depotentiation was performed on withdrawal day 10, just before a cocaine challenge, the expression of locomotor sensitization in the animals was prevented. In a key experiment, the authors then showed that this effect was long-lasting — a challenge injection of cocaine given to the mice 5 days after depotentiation was still unable to elicit a sensitized locomotor response.

These results are exciting because they offer hope that some of the neuroadaptations associated with cocaine exposure are reversible. This would not necessarily have been predicted, given that addicts exhibit extremely persistent changes in the brain and a lifelong vulnerability to relapse[5]. Furthermore, Pascoli and colleagues' approach differs in a subtle but important way from previous studies that have targeted plasticity mechanisms to inhibit cocaine seeking. Typically, a particular abnormality at a cocaine-exposed synapse has been identified and then normalized using an exogenous reagent (see, for example, refs 6, 7). By contrast, the current approach[1] normalized nucleus-accumbens synapses by altering their level of synaptic activation — that is, through

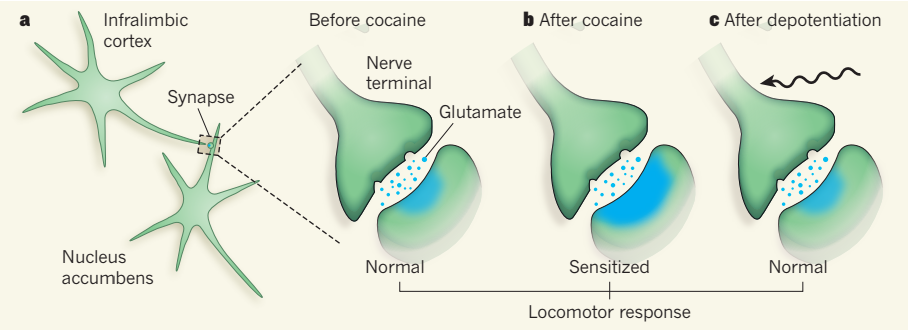


Figure 1 | Normalizing the locomotor response to cocaine in mice. **a**, Glutamate-releasing nerve terminals originating in the infralimbic cortex make excitatory connections (synapses) with medium spiny neurons that express D1 dopamine receptors in the shell subregion of the nucleus accumbens. These synapses regulate locomotor behaviour in mice. The blue region represents the postsynaptic density — the location of AMPA-type receptors that respond to glutamate released by the infralimbic nerve terminals. **b**, Cocaine injections strengthen these synapses (represented by a larger blue area). This synaptic potentiation is associated with a sensitized locomotor response — increased hyperactivity — in mice that are given a subsequent 'challenge' injection of cocaine. **c**, Pascoli *et al.*[1] report that depotentiation of these synapses using an optogenetic strategy (represented by the wavy arrow) resets the synaptic strength and normalizes the locomotor response to cocaine.

the same physiological mechanism that the brain uses to regulate synaptic strength. Harnessing this physiological mechanism may make it possible to produce a more global and lasting normalization of synapses and behaviour. In humans, this might be accomplished using deep brain stimulation or transcranial magnetic stimulation, rather than optogenetics. Indeed, these approaches have shown some promise in animal studies[8].

The next step should be evaluation of the depotentiation strategy in more sophisticated animal models that involve voluntary drug self-administration. These models measure an animal's motivation to obtain a drug based on the effort it is willing to expend — or even its willingness to tolerate punishment — in order to do so. Complexities in the neuroanatomy of the reward system will no doubt be important. For example, neurons in the nucleus accumbens receive glutamate projections from regions other than the infralimbic cortex, regions that also have critical roles in drug seeking[5]. So, although depotentiating a single pathway is sufficient to block locomotor sensitization[1], more complex approaches may be needed in self-administration experiments.

Similarly, although differentiating between D1R- and D2R-expressing cells is a useful starting point for studying functionally distinct MSN subpopulations in the nucleus accumbens, the issue of dopamine-receptor segregation is complex[4]. For example, the target regions of these subpopulations overlap[4], and many studies have uncovered cooperative effects of D1Rs and D2Rs on nucleus-accumbens firing and related behaviours[9,10]. Furthermore, within any neuronal population defined by the expression of a particular protein marker (such as the D1R), only a few neurons may actually contribute to behavioural output in a particular situation[11]. Finally, given that LTP in the nucleus accumbens is presumably caused by an increase in the number of synaptic AMPA receptors, it is logical to infer that depotentiation and reversal of sensitization reflect AMPA-receptor removal. Yet this inference is difficult to reconcile with previous studies[3] in cocaine-sensitized rodents, which showed that locomotor sensitization can be expressed under conditions in which AMPA-receptor levels are not elevated — at early withdrawal times, for example, before there is any increase in AMPA-receptor number.

Nevertheless, Pascoli and colleagues' paper[1] is a landmark in our understanding of cocaine-related plasticity and of how such plasticity might be harnessed in the development of addiction therapies. Their experimental approach provides a strategy for future studies: targeting a particular synapse, determining the stimulation parameters for its potentiation or depotentiation, and then developing a method to accomplish a desired adjustment in an intact animal. ■

Marina E. Wolf *is in the Department of Neuroscience, Chicago Medical School, Rosalind Franklin University of Medicine and Science, North Chicago, Illinois 60064-3095, USA.*
e-mail: marina.wolf@rosalindfranklin.edu

1. Pascoli, V. *et al. Nature* **481,** 71–75 (2012).
2. Robinson, T. E. & Berridge, K. C. *Phil. Trans. R. Soc. Lond. B* **363,** 3137–3146 (2008).
3. Wolf, M. E. & Ferrario, C. R. *Neurosci. Biobehav. Rev.* **35,** 185–211 (2010).
4. Sesack, S. R. & Grace, A. A. *Neuropsychopharmacology* **35,** 27–47 (2010).
5. Koob, G. F. & Volkow, N. D. *Neuropsychopharmacology* **35,** 217–238 (2010).
6. Conrad, K. L. *et al. Nature* **454,** 118–121 (2008).
7. Moussawi, K. *et al. Proc. Natl Acad. Sci. USA* **108,** 385–390 (2011).
8. Feil, J. & Zangen, A. *Neurosci. Biobehav. Rev.* **34,** 559–574 (2010).
9. Clark, D. & White, F. J. *Synapse* **1,** 347–388 (1987).
10. Seif, T., Makriyannis, A., Kunos, G., Bonci, A. & Hopf, F. W. *Neuroscience* **193,** 21–23 (2011).
11. Koya, E. *et al. Nature Neurosci.* **12,** 1069–1073 (2009).

doi:10.1038/nature10722

A sensing array of radically coupled genetic 'biopixels'

Arthur Prindle[1]*, Phillip Samayoa[2]*, Ivan Razinkov[1], Tal Danino[1], Lev S. Tsimring[3] & Jeff Hasty[1,2,3,4]

Although there has been considerable progress in the development of engineering principles for synthetic biology, a substantial challenge is the construction of robust circuits in a noisy cellular environment. Such an environment leads to considerable intercellular variability in circuit behaviour, which can hinder functionality at the colony level. Here we engineer the synchronization of thousands of oscillating colony 'biopixels' over centimetre-length scales through the use of synergistic intercellular coupling involving quorum sensing within a colony and gas-phase redox signalling between colonies. We use this platform to construct a liquid crystal display (LCD)-like macroscopic clock that can be used to sense arsenic via modulation of the oscillatory period. Given the repertoire of sensing capabilities of bacteria such as *Escherichia coli*, the ability to coordinate their behaviour over large length scales sets the stage for the construction of low cost genetic biosensors that are capable of detecting heavy metals and pathogens in the field.

Synthetic biology can be broadly broken down into the 'top-down' synthesis of genomes[1] and the 'bottom-up' engineering of relatively small genetic circuits[2–10]. In the field of genetic circuits, toggle switches[11] and oscillators[12] have progressed into triggers[13], counters[14] and synchronized clocks[15]. Sensors have arisen as a major focus in the context of biotechnology[6,16,17], while oscillators have provided insights into the basic-science functionality of cyclic regulatory processes[18–20]. A common theme is the concurrent development of mathematical modelling that can be used for experimental design and characterization, as in physics and the engineering disciplines.

The synchronization of genetic clocks provides a particularly attractive avenue for synthetic biology applications. Oscillations permeate science and technology in a number of disciplines, with familiar examples including alternating current (AC) power[21], the global positioning system (GPS)[22] and lasers[23]. These technologies have demonstrated that operating in the frequency domain can offer considerable advantages over steady-state designs in terms of information gathering and transmission. In particular, oscillatory sensors confer a number of advantages to traditional ones[24], as frequency is easily digitized and can be quickly updated with repeated measurements. For sensors that use optical reporters, measurements of frequency are less sensitive to experimental factors such as beam power and exposure time than intensity measurements, which must be normalized and calibrated.

Although the bottom-up approach to synthetic biology is increasingly benefiting from DNA synthesis technologies, the general design principles are still evolving. In this context, a substantial challenge is the construction of robust circuits in a cellular environment that is governed by noisy processes such as random bursts of transcription and translation[25–29]. Such an environment leads to considerable intercellular variability in circuit behaviour, which can impede coherent functionality at the colony level. An ideal design strategy for reducing variability across a cellular population would involve both strong and long-range coupling that would instantaneously synchronize the response of millions of cells. Quorum sensing typically involves strong intercellular coupling over tens of micrometres[8,15,30], yet the relatively slow diffusion time of molecular communication through cellular media leads to signalling delays over millimetre scales. Faster communication mechanisms, such as those mediated in the gas phase, may increase the length scale for instantaneous communication, but are comparatively weak and short lived because the vapour species more readily disperse.

Synergistic synchronization

To develop a frequency-modulated biosensor, we designed a gene network capable of synchronizing genetic oscillations across multiple scales (Fig. 1a and Supplementary Fig. 1). We constructed an LCD-like microfluidic[31] array that allows many separate colonies of sensing bacteria to grow and communicate rapidly by gas exchange (Fig. 1b, c and Supplementary Fig. 9). As previous work[15] has demonstrated that coupling through quorum sensing leads to incoherent oscillations at the millimetre scale, this mode of cellular communication is too slow for the generation of macroscopic synchronized oscillations. However, the slower quorum sensing can be used to synchronize small local colonies, provided there is a second level of design that involves faster communication for coordination between the colonies. Therefore, rather than attempting to engineer a sensor from a single large-colony oscillator, we wired together thousands of small oscillating colonies, or 'biopixels', in a microfluidic array. Coupling between biopixels involves redox signalling by hydrogen peroxide (H_2O_2) and the native redox sensing machineries of *E. coli*. The two coupling mechanisms act synergistically in the sense that the stronger, yet short-range, quorum sensing is necessary to coherently synchronize the weaker, yet long-range, redox signalling. Using this method we demonstrate synchronization of approximately 2.5 million cells across a distance of 5 mm, over 1,000 times the length of an individual cell (Fig. 1c, d and Supplementary Movies 1 and 2). This degree of synchronization yields extremely consistent oscillations, with a temporal accuracy of about 2 min compared to 5–10 min for a single oscillator[15] (Fig. 1d).

The global synchronization mechanism is comprised of two modes of communication that work on different scales. The quorum-sensing

[1]Department of Bioengineering, University of California, San Diego, La Jolla, California 92093, USA. [2]Bioinformatics Program, University of California, San Diego, La Jolla, California 92093, USA. [3]BioCircuits Institute, University of California, San Diego, La Jolla, California 92093, USA. [4]Molecular Biology Section, Division of Biological Science, University of California, San Diego, La Jolla, California 92093, USA.
*These authors contributed equally to this work.

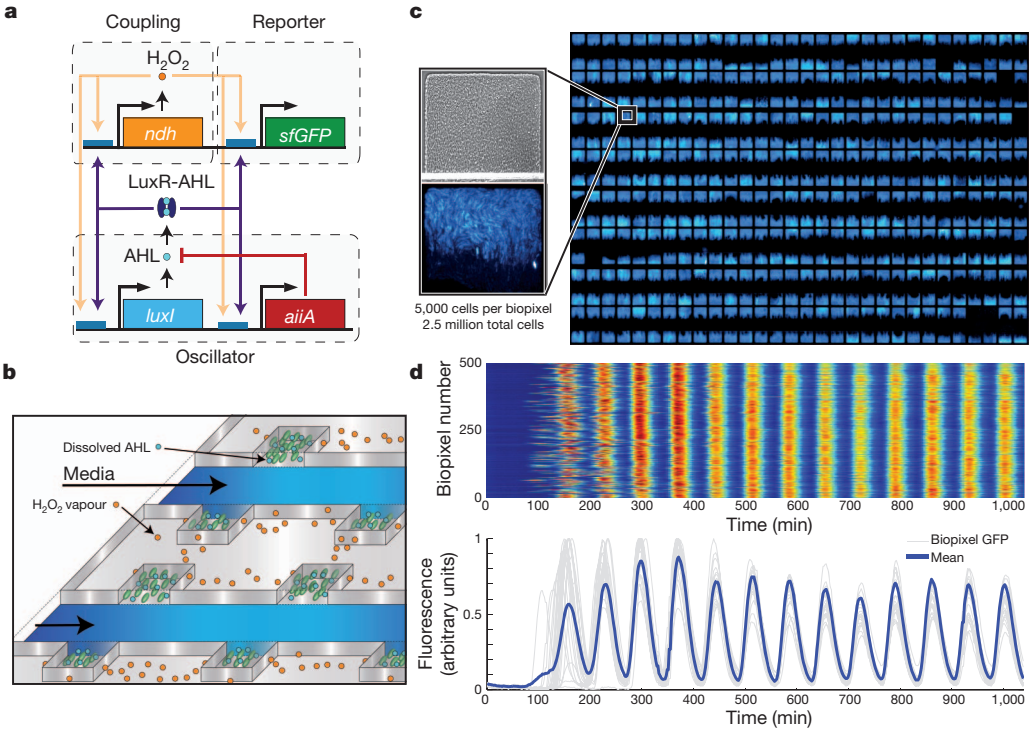


Figure 1 | Sensing array of radically coupled genetic biopixels. **a**, Network diagram. The *luxI* promoter drives expression of *luxI*, *aiiA*, *ndh* and *sfGFP* (superfolder variant of GFP) in four identical transcription modules. The quorum-sensing genes *luxI* and *aiiA* generate synchronized oscillations within a colony via AHL. The *ndh* gene codes for NDH-2, an enzyme that generates H_2O_2 vapour, which is an additional activator of the *luxI* promoter. H_2O_2 is capable of migrating between colonies and synchronizing them. **b**, Conceptual design of the sensing array. AHL diffuses within colonies while H_2O_2 migrates between adjacent colonies through the PDMS. Arsenite-containing media is passed in through the parallel feeding channels. **c**, Fluorescent image of an array of 500 *E. coli* biopixels containing about 2.5 million cells. Inset, bright-field and fluorescent images display a biopixel of 5,000 cells. **d**, Heat map and trajectories depicting time-lapse output of 500 individual biopixels undergoing rapid synchronization. Sampling time is 2 min.

machinery (LuxI, AiiA) uses an acyl-homoserine lactone (AHL) to mediate intracolony synchronization. In our device, the degree to which neighbouring colonies are able to influence each other via AHL diffusion is negligible owing to the high media channel flow rates. Instead, we engineered the cells to communicate via gas exchange by placing a copy of the gene coding for NADH dehydrogenase II (*ndh*) under the control of an additional *lux* promoter. NDH-2 is a membrane-bound respiratory enzyme that produces low levels of H_2O_2 and superoxide (O_2^-)[32]. As H_2O_2 vapour is able to pass through the 25-μm oxygen-permeable polydimethylsiloxane (PDMS) walls that separate adjacent colonies, periodic production of NDH-2 yields periodic exchange of H_2O_2 between biopixels. When H_2O_2 enters the cell, it transiently changes its redox state, interacting with our synthetic circuit through the native aerobic response control systems, including ArcAB, which has a binding site in the *lux* promoter region[33,34]. Under normal conditions, ArcAB is partially active so *lux* is partially repressed. In contrast, oxidizing conditions triggered by H_2O_2 inactivate ArcAB, relieving this repression. Each oscillatory burst promotes firing in neighbouring colonies by relieving repression on the *lux* promoter. This constitutes an additional positive feedback that rapidly synchronizes the population (Supplementary Fig. 2 and Supplementary Movie 1).

We investigated the effects of catalase and superoxide dismutase (SOD) to probe the nature of H_2O_2 communication. When a population of synchronized colonies was exposed to a step increase of 200 U ml^{-1} catalase, an enzyme that rapidly degrades extracellular H_2O_2[35], synchronization was broken and colonies continued to oscillate individually (Supplementary Fig. 3). As the cell membrane is impermeable to catalase, asynchronous colony oscillations confirm that communication between colonies depends on external H_2O_2 whereas oscillations within a colony do not. Conversely, when we enhanced the rate of superoxide conversion to H_2O_2 by expressing *sodA*[36,37] from an additional *lux* promoter, colonies quickly fired in a spatial wave and failed to oscillate further despite no changes to growth rate or cell viability (Supplementary Fig. 4). Because H_2O_2 is produced internal to the cell, this confirms that H_2O_2 is capable of escaping the cell and activating *lux*-regulated genes in neighbouring colonies via diffusion. The apparent higher output of H_2O_2 by SOD as compared to NDH-2 is probably due to its very high catalytic efficiency[38]. Lastly, we observed synchronization between arrays of traps even when they were fluidically isolated but held in close proximity (Supplementary Fig. 5). These devices share no common fluid sources or channels, making communication by dissolved molecules like AHL impossible. Taken together, these results confirm that gaseous H_2O_2 is the mode of communication between oscillating colonies.

On the basis of our understanding of the mechanism for global synchronization, we expected that we could simplify the circuitry by eliminating *ndh* and achieve the same effect with intermittent bursts of high-intensity blue light. In this design, the GFP molecule acts as a photosensitizer, releasing free radicals upon exposure that produce reactive oxygen species (ROS) including H_2O_2[39]. At the peak of oscillation, considerable vapour-phase H_2O_2 is produced by exposing GFP-containing cells to fluorescent light. Conversely, at the trough of oscillation, cells contain almost no GFP, and therefore produce very little H_2O_2 upon fluorescing. Bursts of light thus generate bursts of H_2O_2 vapour whose concentration depends on the oscillating GFP level, just as periodic production of NDH-2 did previously. Indeed, this strategy was similarly able to synchronize our sensor array (Fig. 1d and Supplementary Movie 2). Numerous controls were performed to ensure that synchronized oscillations did not occur at low fluorescence intensities (Supplementary Fig. 6 and Supplementary Movie 9).

To probe this mode of synchronization, we investigated the effects of thiourea and the antibiotics ampicillin and kanamycin. When a synchronized population of colonies was exposed to 35 mM thiourea, a

potent radical quencher[40,41], we observed sharply decaying synchronized oscillations whereas growth rate and cell viability were unaffected (Supplementary Fig. 7). This suggests that without radical species, oscillations cannot be produced. Next, we ran a series of experiments switching the antibiotic resistance genes on our plasmids. We noted that radical-producing antibiotics[42], particularly ampicillin, significantly reduced the degree of synchronization, showing that an excess of radical species also hinders communication (Supplementary Fig. 8). As our final constructs included a plasmid with kanamycin resistance, which was also found to produce some radicals, we used full (50 μg ml^{-1}) selection when growing up the cells but very low (5 μg ml^{-1}) selection during the experimental run. Persistence of oscillations, sequencing, and subsequent growth in full selection following the run confirmed the presence of all three plasmids despite this low experimental selection. Catalase and SOD results were identical to those with NDH-2 synchronization (not shown). These results show that fluorescence-mediated synchronization involves the production of radical species after fluorescence exposure and communication via H_2O_2.

Sensing array of biopixels

With a platform for generating consistent and readily detectable oscillations, we sought to use the circuit to engineer an arsenic-sensing macroscopic biosensor. We rewired the network to include an extra copy of the positive-feedback element, the AHL synthase LuxI, under the control of a native arsenite-responsive promoter that is repressed by ArsR in the absence of arsenite (Fig. 2a, right). When arsenite is not present in the media, supplementary *luxI* is not transcribed and the circuit functions normally, generating baseline oscillations. However, the addition of trace amounts of arsenite relieves this repression and allows supplementary *luxI* to be transcribed, increasing the oscillatory amplitude and period. Tuning the level of LuxI by varying arsenite concentration results in clear changes to the oscillatory period (Fig. 2b and Supplementary Movie 2). To determine the range of detection, we swept arsenite concentrations from 0–1 μM and measured the oscillatory period (Fig. 2c, top). Using statistical methods (Supplementary Methods), we generated a sensor calibration curve (Fig. 2c, bottom) that depicts the maximum possible arsenite concentration present ($\alpha = 95\%$) for a given measured period. This curve is an illustration of how data generated by our array would be used to measure arsenite concentrations in an unknown sample using our device. Our system was able to reliably quantify arsenite levels as low as 0.2 μM, below the 0.5 μM World Health Organization-recommended level for developing nations[43].

As an alternative sensing strategy, we rewired the network to include a copy of the *luxR* gene controlled by an arsenite-responsive promoter while removing it from the rest of the circuit (Fig. 2a, left). Because the LuxR–AHL complex must be present to activate the *lux* promoter[30], cells produce no LuxR when the media is free of arsenite, generating no fluorescence or oscillations. The addition of arsenite stimulates the production of LuxR, restoring circuit function and producing clear, synchronized oscillations (Fig. 2d and Supplementary Movie 3). This ON/OFF detection system has a threshold of 0.25 μM, a detection limit that can be adjusted by changing the copy number, ribosome binding site (RBS) strength, or promoter strength of the sensing plasmid (Supplementary Methods).

The sensing array is also capable of producing complex behaviours arising from the dynamic interaction of cellular colonies. By making modifications to the size, number and arrangement of biopixels in the device, we are able to markedly alter the output waveforms. For

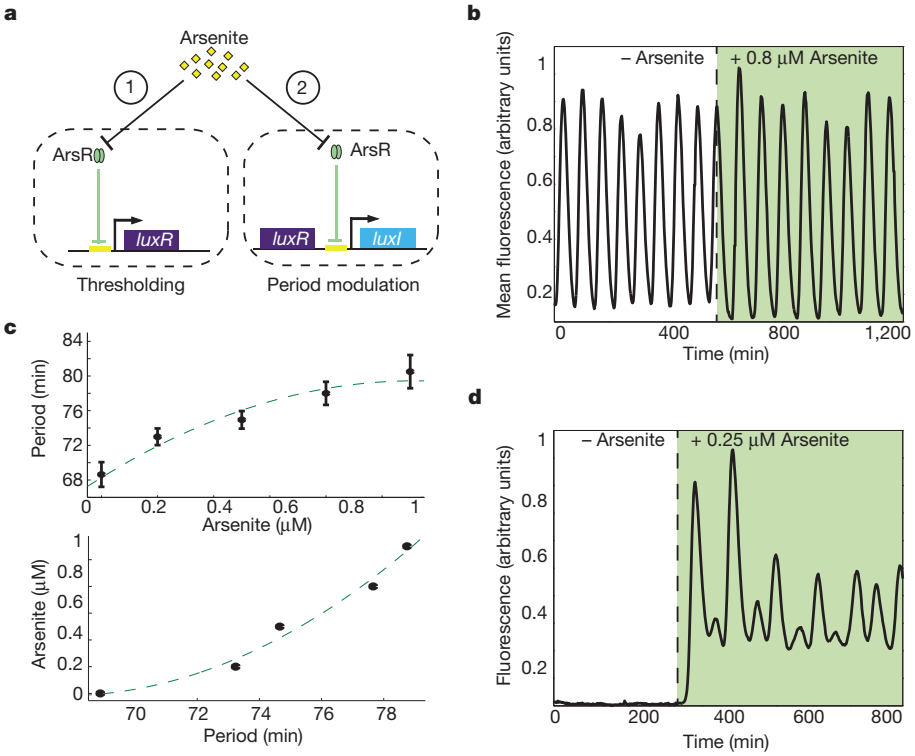


Figure 2 | Frequency-modulated genetic biosensor. **a**, Network diagrams depicting two constructed sensing modules. In thresholding (1), the *luxR* gene is removed from the oscillator network and supplemented by a new copy driven by an arsenite-responsive promoter. In period modulation (2), a supplementary *luxI* gene tagged for increased degradation is driven by the arsenic-responsive promoter, which affects the period of oscillation. **b**, A sample period modulation sensor output following a step increase of 0.8 μM arsenite. Oscillatory period increases from 69 min to 79 min. **c**, Top, period versus arsenite concentration for the sensor array. Error bars indicate ± 1 standard deviation averaged over 500 biopixel trajectories. Dotted line represents model-predicted curve. Bottom, sensor calibration curve generated from experimental data. Points indicate the maximum arsenite level with 95% certainty for a given measured period as determined statistically from experimental data. **d**, Thresholder output following a step increase of 0.25 μM arsenite. A marked shift from rest to oscillatory behaviour is observed within 20 min after the addition of arsenite.

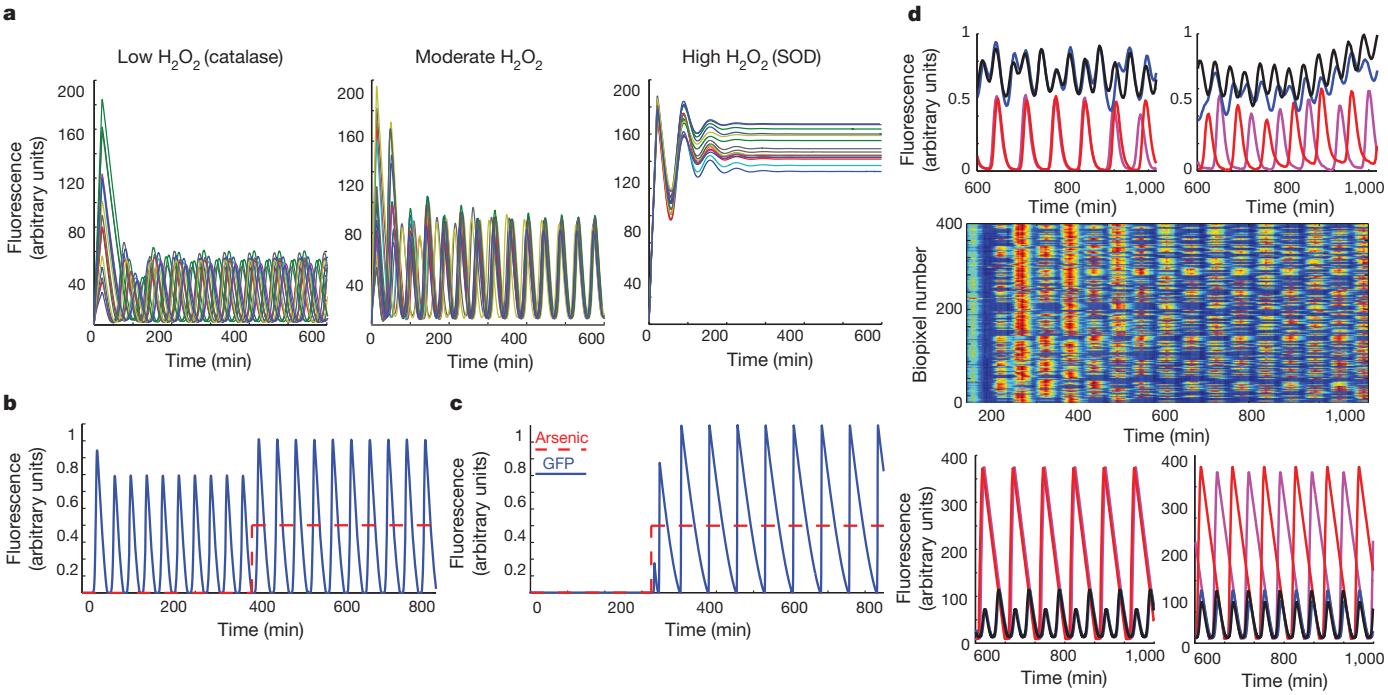


Figure 3 | Computational modelling of radical synchronization and biosensing. **a**, Time series of a population of biopixels producing varying amounts of H_2O_2 vapour. Synchronization occurs only for moderate levels whereas high levels lock ON and low levels oscillate asynchronously. **b**, A typical time series for our period modulation sensor undergoing a step increase of arsenite. Oscillations increase in both amplitude and period. **c**, A typical time series output for the thresholding sensor. Oscillations arise after the addition of arsenite. **d**, Experimental and computational output depicting complex dynamic behaviours between neighbouring traps. Top, 1:2 resonance and anti-phase synchronization observed when trap size (left, black/blue = 95 μm depth and red/magenta = 85 μm depth) and separation distance (right, same colours) are modified experimentally. Middle, scaled-up array experimental data for increased trap separation experiments demonstrating anti-phase synchronization. Bottom, computational model trajectories depicting 1:2 resonance and anti-phase synchronization when trap size (same colours as experimental data) and separation distance are changed.

example, when we constructed a device in which trap separation distance is increased (45 μm versus 25 μm), we observed local anti-phase synchronization between neighbouring colonies (Fig. 3d, top right). To explore this phenomenon on a larger scale, we constructed a device that contains an array of 416 traps constructed according to the specifications above. In these experiments, we observe initial global synchronization that gradually falls into local anti-phase synchronization across the array (Fig. 3d, middle. and Supplementary Movie 4). Phase alignment is maintained over at least 48 h, with patches of synchronization typically 3–6 colonies in size. Alternatively, by changing dimensions such that the array contains traps of two slightly different sizes, we observe a 1:2 resonance synchronization where larger traps pulse at double the frequency of smaller traps while maintaining synchronization (Fig. 3d, top left and Supplementary Movie 7). Lastly, when LuxR is limited, as in the thresholding scheme, we observe synchronized oscillations of alternating large and small peaks in both experiment and model (Supplementary Fig. 12). Our computational model (Box 1) captures these effects (Fig. 3d, bottom, and Supplementary Figs 11 and 12) and indicates that further array manipulation will yield new, richer dynamics that could not be produced directly by changing circuit structure.

Although our sensor array is capable of performing a variety of complex functions in the laboratory, adapting this technology to a real-world device will require the elimination of the expensive and bulky microscopy equipment. However, measuring genetic oscillations in the absence of any magnification or powerful illumination will require an even further increased signal. Using this mechanism of global synchronization, we were able to scale up to a 24 mm × 12 mm array that houses over 12,000 communicating biopixels (Fig. 4a). Synchronization is maintained across the entire array, a distance over 5,000 times the length of an individual cell, using an inexpensive light-emitting diode (LED; Fig. 4b, c and Supplementary Movie 5). The signal strength generated by the large number of cells in the array (about 50 million) will allow us to adapt the device to function as a handheld sensor. In our conceptual design (Fig. 4d), the sensor will

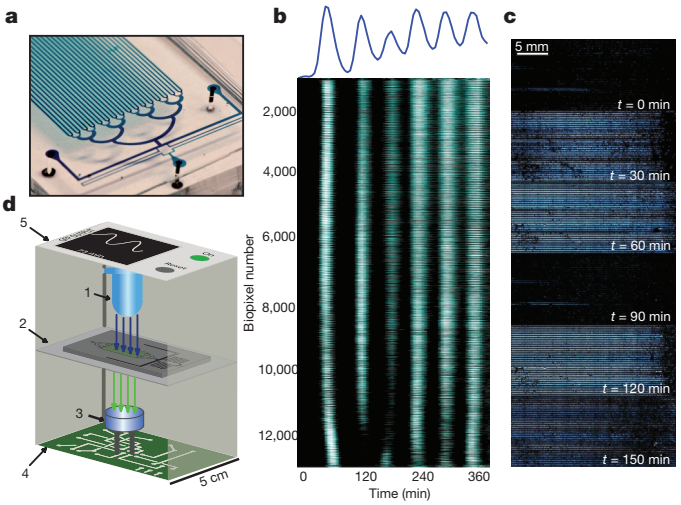


Figure 4 | Radical synchronization on a macroscopic scale. **a**, The scaled-up array is 24 mm × 12 mm and houses over 12,000 biopixels that contain approximately 50 million total cells when filled. **b**, Global synchronization is maintained across the array. Heat map of individual trajectories of all 12,224 oscillating biopixels. **c**, Image series depicting global synchronization and oscillation for the macroscopic array. Each image is produced by stitching 72 fields of view imaged at ×4 magnification. **d**, Schematic diagram illustrating our design for a handheld device using the sensing array. An LED (1) excites the array (2) and emitted light is collected by a photodetector (3), analysed by an onboard processor (4), and displayed graphically (5).

BOX 1

Computational modelling

Our model of the frequency-modulated biosensor is based on a previously described model for the quorum-sensing synchronized oscillator[15]. In addition to the reactions reflected in that model, we include the arsenite-induced production and degradation of LuxI and/or LuxR. From the biochemical reactions, we derived a set of delay-differential equations to be used as our model. These delayed reactions mimic the complex cascade of processes (transcription, translation, maturation, etc.) leading to formation of functional proteins. As expected, our model predicts oscillations that change frequency when changes in arsenite occur (Figs 2c and 3b). The amplitude and period of the oscillations both depend on the concentrations of the toxin. We then modified the model to describe the LuxR-based detection system. Our model predicts a marked transition from rest to oscillations upon addition of arsenite, consistent with experimental observations (Fig. 3c).

The multi-scale nature of communication in our array allows us to treat colony and array-level dynamics separately; in the latter, arsenite affects the quorum-sensing machinery of a colony, producing changes to oscillatory period that propagate between biopixels in the array. To describe quantitatively the mechanisms driving synchronization at the array level, we treat each colony as a single oscillator that acts according to degrade-and-fire kinetics[47]. We also include the production of H_2O_2 and its interaction with neighbouring colonies by two-dimensional diffusion. Using this model we identified three regimes that correlate well with experimental observations (Fig. 3a). When the effective production of H_2O_2 is low, as with catalase, we observe unsynchronized oscillations owing to constant, mild repression of the *lux* promoter via ArcAB (Fig. 3a, left). In contrast, when H_2O_2 production is very high, neighbouring colonies rapidly fire in succession and remain on because of the permanent activation of the *lux* promoter, consistent with the SOD experiment (Fig. 3a, right). Finally, at intermediate H_2O_2, we observe globally synchronized oscillations (Fig. 3a, middle). As colonies are moved further apart, synchronicity breaks owing to slowed migration of H_2O_2 (Supplementary Fig. 10).

continuously read the oscillatory frequency using off-the-shelf electronic components costing less than 50 dollars.

There have been many examples of bacteria-based biosensors[44–46], usually involving an optical reporter driven by a toxin-responsive promoter. Because optical intensity readings are sensitive to imaging conditions like beam power and exposure time, measurements must typically be normalized and calibrated. Measuring the period of oscillation allows us to avoid these issues because peak-to-peak time does not depend on individual peak intensity. Also, oscillations produced at the colony level effectively decouple the signal from the growth state of individual cells, which can also affect fluorescence intensity. By using a dynamic readout that depends on communication between biopixels, we scan and tune potential output signals by changing device parameters rather than redesigning the underlying circuit. For example, we might design a new sensing scheme in which oscillations synchronize with the addition of some toxin and shift to anti-phase or resonant synchronization when critical toxin levels are present.

Scaling up synthetic biology

By nesting two modes of communication we are able to expand the scale over which individual cells are coordinated and increase the complexity of their interaction. Indeed, there are many familiar examples of hierarchical systems. Airline routes are often designed such that small airports are connected locally to larger hubs that are connected internationally. It would neither be feasible nor desirable to connect every airport together. Similarly, individual cells communicate locally by one method, generating impulses large enough to enable colonies to communicate globally by another. Nesting communication mechanisms in this way may allow us to better scale up synthetic circuits of different types, such as switches and logic gates, paving the way for the next generation of synthetic biology pursuits.

METHODS SUMMARY

Strains and plasmids. The plasmids were constructed using a PCR-based cloning strategy[48] in which the origin of replication, antibiotic resistance, and circuit genes were assembled in different combinations. The *ndh* and *sodA* genes were amplified directly from the native *E. coli* genome by PCR. Various arsenite-responsive promoters were tested, including a recently reported synthetic version[49], but the final design uses the native *E. coli* version. Promoter output was tuned by changing the RBS sequence and quantified using flow cytometry. All circuit components except *luxR* were tagged by PCR with a carboxy-terminal *ssrA* tag (AANDENYALAA)[50] for fast degradation.

Microfluidics and microscopy. Image acquisition was performed on a Nikon Eclipse TI epifluorescent inverted microscope outfitted with fluorescence filter cubes optimized for GFP imaging and a phase-contrast-based autofocus algorithm. Images were acquired using an Andor Clara cooled CCD camera or Andor DU-897 EMCCD camera, both controlled by Nikon Elements software. Images were acquired every 2 min in phase contrast and fluorescence. The cells were imaged inside a microfluidic device with an upstream switch, with the ability to mix or switch between two different media sources. A custom application written in LabVIEW (National Instruments) controlled linear actuators, to which two reservoirs of arsenite-containing and pure medium were attached. Using this algorithm, arsenite concentration was dynamically varied to probe sensor output.

The microfluidic experiments were performed as previously described[15]. Briefly, 50 μl of an overnight culture was diluted in 50 ml of LB medium (Difco) plus antibiotics the day of the experiment. When cells reached an optical density ($OD_{600\,nm}$) of 0.1, cells were spun down and resuspended in 5 ml of fresh media and loaded into the device.

Received 2 September; accepted 15 November 2011.
Published online 18 December 2011.

1. Gibson, D. *et al.* Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* **329,** 52–56 (2010).
2. Hasty, J., McMillen, D. & Collins, J. J. Engineered gene circuits. *Nature* **420,** 224–230 (2002).
3. Sprinzak, D. & Elowitz, M. B. Reconstruction of genetic circuits. *Nature* **438,** 443–448 (2005).
4. Endy, D. Foundations for engineering biology. *Nature* **438,** 449–453 (2005).
5. Ellis, T., Wang, X. & Collins, J. Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nature Biotechnol.* **27,** 465–471 (2009).
6. Kobayashi, H. *et al.* Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl Acad. Sci. USA* **101,** 8414–8419 (2004).
7. You, L., Cox, R. S. III, Weiss, R. & Arnold, F. H. Programmed population control by cell–cell communication and regulated killing. *Nature* **428,** 868–871 (2004).
8. Basu, S., Gerchman, Y., Collins, C. H., Arnold, F. H. & Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* **434,** 1130–1134 (2005).
9. Mukherji, S. & Van Oudenaarden, A. Synthetic biology: understanding biological design from synthetic circuits. *Nature Rev. Genet.* **10,** 859–871 (2009).
10. Grilly, C., Stricker, J., Pang, W., Bennett, M. & Hasty, J. A synthetic gene network for tuning protein degradation in *Saccharomyces cerevisiae. Mol. Syst. Biol.* **3,** 127 (2007).
11. Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in *Escherichia coli. Nature* **403,** 339–342 (2000).
12. Elowitz, M. B. & Leibler, S. A synthetic oscillatory network of transcriptional regulators. *Nature* **403,** 335–338 (2000).
13. Lu, T. & Collins, J. Dispersing biofilms with engineered enzymatic bacteriophage. *Proc. Natl Acad. Sci. USA* **104,** 11197–11202 (2007).
14. Friedland, A. *et al.* Synthetic gene networks that count. *Science* **324,** 1199–1202 (2009).
15. Danino, T., Mondragon-Palomino, O., Tsimring, L. & Hasty, J. A synchronized quorum of genetic clocks. *Nature* **463,** 326–330 (2010).
16. Tamsir, A., Tabor, J. & Voigt, C. Robust multicellular computing using genetically encoded NOR gates and chemical 'wires'. *Nature* **469,** 212–215 (2011).
17. Tabor, J. *et al.* A synthetic genetic edge detection program. *Cell* **137,** 1272–1281 (2009).
18. Stricker, J. *et al.* A fast, robust and tunable synthetic gene oscillator. *Nature* **456,** 516–519 (2008).
19. Mondragon-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L. & Hasty, J. Entrainment of a population of synthetic genetic oscillators. *Science* **333,** 1315–1319 (2011).
20. Tigges, M., Marquez-Lago, T., Stelling, J. & Fussenegger, M. A tunable synthetic mammalian oscillator. *Nature* **457,** 309–312 (2009).

21. Westinghouse, G. System of electrical distribution. US patent 373 035 (1887).
22. Lewandowski, W., Azoubib, J. & Klepczynski, W. GPS: primary tool for time transfer. *Proc. IEEE* **87,** 163–172 (1999).
23. Vladimirov, A., Kozyreff, G. & Mandel, P. Synchronization of weakly stable oscillators and semiconductor laser arrays. *Europhys. Lett.* **61,** 613 (2003).
24. Gast, T. Sensors with oscillating elements. *J. Phys. E: Sci. Instrum.* **18,** 783 (1985).
25. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31,** 69–73 (2002).
26. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic gene expression in a single cell. *Science* **297,** 1183–1186 (2002).
27. Golding, I., Paulsson, J., Zawilski, S. & Cox, E. Real-time kinetics of gene activity in individual bacteria. *Cell* **123,** 1025–1036 (2005).
28. Blake, W. *et al.* Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell* **24,** 853–865 (2006).
29. Austin, D. *et al.* Gene network shaping of inherent noise spectra. *Nature* **439,** 608–611 (2006).
30. Waters, C. & Bassler, B. Quorum sensing: cell-to-cell communication in bacteria. *Annu. Rev. Cell Dev. Biol.* **21,** 319–346 (2005).
31. Ferry, M., Razinkov, I. & Hasty, J. Microfluidics for synthetic biology from design to execution. *Methods Enzymol.* **497,** 295 (2011).
32. Messner, K. & Imlay, J. The identification of primary sites of superoxide and hydrogen peroxide formation in the aerobic respiratory chain and sulfite reductase complex of *Escherichia coli*. *J. Biol. Chem.* **274,** 10119–10128 (1999).
33. Bose, J. L. *et al.* Bioluminescence in *Vibrio fischeri* is controlled by the redox-responsive regulator arca. *Mol. Microbiol.* **65,** 538–553 (2007).
34. Georgellis, D., Kwon, O. & Lin, E. Quinones as the redox signal for the arc two-component system of bacteria. *Science* **292,** 2314–2316 (2001).
35. Seaver, L. & Imlay, J. Hydrogen peroxide fluxes and compartmentalization inside growing *Escherichia coli*. *J. Bacteriol.* **183,** 7182–7189 (2001).
36. Fridovich, I. The biology of oxygen radicals. *Science* **201,** 875–880 (1978).
37. McCord, J. & Fridovich, I. Superoxide dismutase. *J. Biol. Chem.* **244,** 6049–6055 (1969).
38. Berg, J., Tymoczko, J. L. & Stryer, L. *Biochemistry* (W.H. Freeman, 2006).
39. Remington, S. Fluorescent proteins: maturation, photochemistry and photophysics. *Curr. Opin. Struct. Biol.* **16,** 714–721 (2006).
40. Kelner, M., Bagnell, R. & Welch, K. Thioureas react with superoxide radicals to yield a sulfhydryl compound. explanation for protective effect against paraquat. *J. Biol. Chem.* **265,** 1306–1311 (1990).
41. Touati, D., Jacques, M., Tardat, B., Bouchard, L. & Despied, S. Lethal oxidative damage and mutagenesis are generated by iron in delta fur mutants of *Escherichia coli*: protective role of superoxide dismutase. *J. Bacteriol.* **177,** 2305–2314 (1995).
42. Kohanski, M. A., DePristo, M. A. & Collins, J. J. Sublethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol. Cell* **37,** 311–320 (2010).
43. Nordstrom, D. Worldwide occurrences of arsenic in ground water. *Science* **296,** 2143 (2002).
44. van der Meer, J. & Belkin, S. Where microbiology meets microengineering: design and applications of reporter bacteria. *Nature Rev. Microbiol.* **8,** 511–522 (2010).
45. Daunert, S. *et al.* Genetically engineered whole-cell sensing systems: coupling biological recognition with reporter genes. *Chem. Rev.* **100,** 2705–2738 (2000).
46. Leveau, J. & Lindow, S. Bioreporters in microbial ecology. *Curr. Opin. Microbiol.* **5,** 259–265 (2002).
47. Mather, W., Bennett, M., Hasty, J. & Tsimring, L. Delay-induced degrade-and-fire oscillations in small genetic circuits. *Phys. Rev. Lett.* **102,** 068105 (2009).
48. Quan, J. & Tian, J. Circular polymerase extension cloning of complex gene libraries and pathways. *PLoS ONE* **4,** e6441 (2009).
49. Stocker, J. *et al.* Development of a set of simple bacterial biosensors for quantitative and rapid measurements of arsenite and arsenate in potable water. *Environ. Sci. Technol.* **37,** 4743–4750 (2003).
50. Keiler, K., Waller, P. & Sauer, R. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271,** 990–993 (1996).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the National Institutes of Health and General Medicine (R01GM69811), the San Diego Center for Systems Biology (P50GM085764), the DoD NDSEG (A.P.) and NSF Graduate Research (P.S.) Fellowship Programs. L.T. was supported, in part, by the Office of Naval Research (MURI N00014-07-0741). We would like to thank J. Imlay, S. Ali and J. Collins for helpful discussions and J. Hickman, B. Taylor and K. Lomax for their help with illustrations.

Author Contributions All authors contributed extensively to the work presented in this paper. A.P. and P.S. are equally contributing first authors.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.H. (hasty@bioeng.ucsd.edu).

doi:10.1038/nature10703

Antiparallel EmrE exports drugs by exchanging between asymmetric structures

Emma A. Morrison[1]*, Gregory T. DeKoster[1]*, Supratik Dutta[1], Reza Vafabakhsh[2], Michael W. Clarkson[3], Arjun Bahl[1], Dorothee Kern[3], Taekjip Ha[2] & Katherine A. Henzler-Wildman[1]

Small multidrug resistance transporters provide an ideal system to study the minimal requirements for active transport. EmrE is one such transporter in *Escherichia coli*. It exports a broad class of polyaromatic cation substrates, thus conferring resistance to drug compounds matching this chemical description. However, a great deal of controversy has surrounded the topology of the EmrE homodimer. Here we show that asymmetric antiparallel EmrE exchanges between inward- and outward-facing states that are identical except that they have opposite orientation in the membrane. We quantitatively measure the global conformational exchange between these two states for substrate-bound EmrE in bicelles using solution NMR dynamics experiments. Förster resonance energy transfer reveals that the monomers within each dimer are antiparallel, and paramagnetic relaxation enhancement NMR experiments demonstrate differential water accessibility of the two monomers within each dimer. Our experiments reveal a 'dynamic symmetry' that reconciles the asymmetric EmrE structure with the functional symmetry of residues in the active site.

EmrE is a secondary active antiporter, driving uphill transport of each polyaromatic cation substrate against its concentration gradient by coupling it to downhill import of two protons across the inner membrane. As one of the smallest known active transporters, with only 110 amino acids and four transmembrane helices per monomer, EmrE would seem to be an ideal model system[1–3]. There is broad agreement that the minimal functional unit is a dimer[4–7], but its structure and topology remain controversial[3,8–13], perhaps because the membrane topology and oligomeric state are exquisitely sensitive to sequence alteration and environment[14–17].

EmrE is proposed to function through a single-site alternating access model[18–20], as shown in Fig. 1a. In this well-established model[21,22], transporters are inherently dynamic proteins, converting between inward- and outward-facing conformations to move substrates across a membrane barrier. To achieve coupled antiport, both substrates share a single binding site and conformational exchange only occurs when substrate (two protons or one polyaromatic cation) is bound. Thus, saturating EmrE with its polyaromatic cation substrate, tetraphenylphosphonium (TPP^+), should drive EmrE into a two-state equilibrium (Fig. 1a, b) suitable for direct observation and quantitative analysis of conformational exchange with site-specific resolution by solution NMR spectroscopy.

To take advantage of well-developed high-resolution solution NMR dynamics methods, we solubilized EmrE in isotropic bicelles, which surround the protein with a more native-like lipid environment than detergents while preserving the fast tumbling needed for solution NMR[23–25]. Several features of the ^1H–^{15}N transverse relaxation-optimized spectroscopy (TROSY) spectrum of TPP^+-bound $^2H/^{15}N$-EmrE in isotropic bicelles are worth noting (Fig. 1c). First, the chemical shift dispersion is much greater than expected for a helical bundle membrane protein. This dispersion is probably enhanced by ring currents from the bound polyaromatic substrate and the many aromatic residues in the binding site. Second, there are twice as many peaks (approximately 210 resolved peaks) as expected for a monomer of EmrE (105 non-proline residues). A solid-state NMR study of EmrE specifically labelled at Glu 14 (ref. 26) also noted such a peak doubling for this active-site residue in the TPP^+-bound state.

There are two possible explanations for this peak doubling: (1) an asymmetric dimer, as observed by cryo-electron microscopy and X-ray crystallography[4,7,27], where each monomer has a distinct structure and thus different chemical shifts or (2) exchange between two symmetric (parallel) homodimers that is slow on the NMR timescale. Naively, we might expect four sets of peaks in the case of both structural asymmetry and slow conformational exchange (Supplementary Fig. 1). However, as has been previously noted, there is a pseudo-twofold symmetry axis in EmrE, perpendicular to the membrane normal, such that reorientation of the helices in monomers A and B could effectively convert the structure of monomer A into that of monomer B and vice versa[18] (Fig. 1b). This unique model results in inward- and outward-facing states that are identical except for their orientation. Such a mechanism is only possible in an antiparallel homodimer, and no previous experimental data have directly tested this hypothesis.

EmrE is functional in isotropic bicelles

We first verified that EmrE is properly folded and functional in bicelles, because it has not been studied in this environment previously. To prepare bicelles, we purified EmrE in detergent, reconstituted it into liposomes and then added DHPC to form bicelles (see Methods). Extensive studies have confirmed that EmrE is a properly folded, functional dimer in dodecylmaltoside (DDM), and when reconstituted into liposomes[4,6,13,28–31]. The TROSY NMR spectra (Fig. 1c) of TPP^+-bound EmrE in dilauroylphosphatidylcholine/

[1]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, USA. [2]Department of Physics and Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. [3]Department of Biochemistry and Howard Hughes Medical Institute, Brandeis University, Waltham, Massachusetts 02454, USA.
*These authors contributed equally to this work.

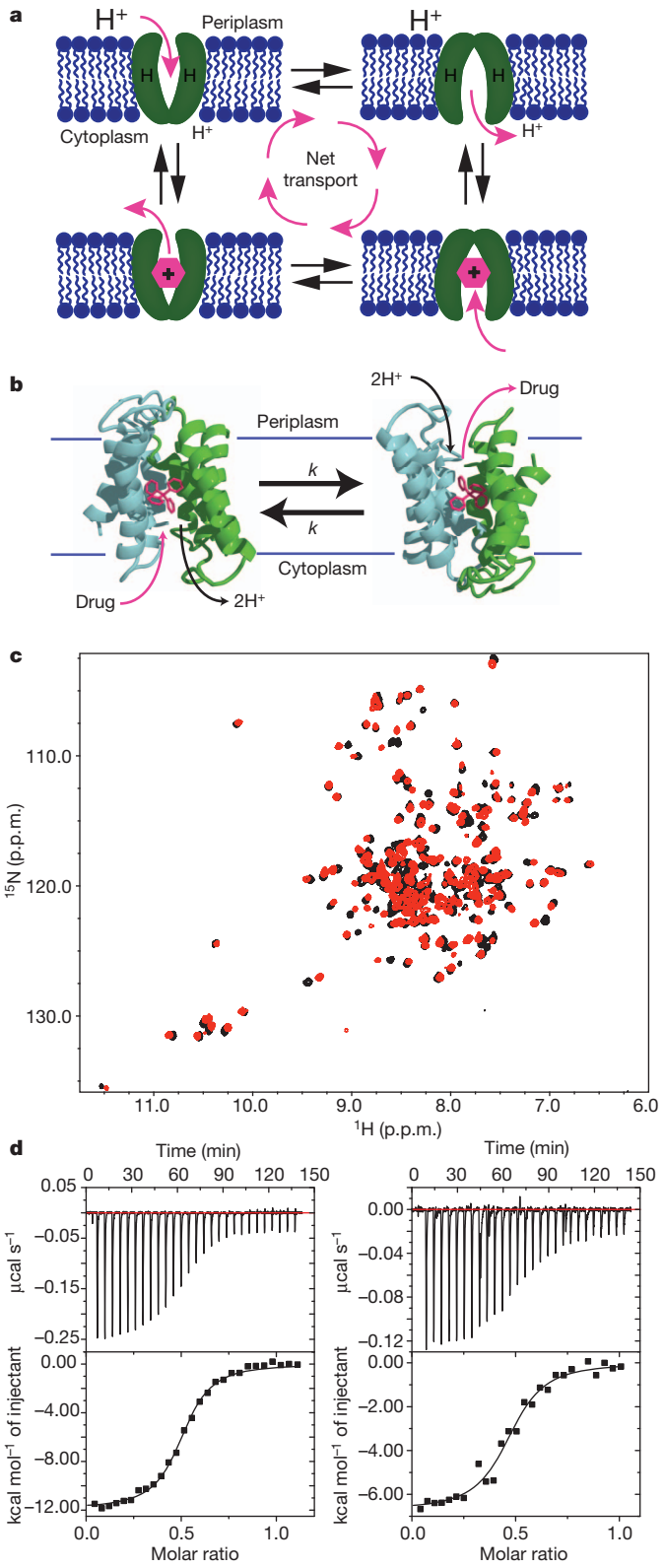


Figure 1 | Conformational interconversion and symmetry in the single-site alternating access model of EmrE transport. **a**, Each state in the transport cycle is only open to one side of the membrane, and the two states only interconvert when substrate (two protons or one polyaromatic cation) is bound. **b**, Conformational exchange as proposed for antiparallel, asymmetric EmrE[18]: the two monomers exchange conformations and the two states have identical structures. **c**, ^{1}H–^{15}N TROSY spectra of TPP^{+}-bound ^{2}H/^{15}N-EmrE at pH 7, 45 °C, are nearly identical for DMPC/DHPC (black) or DLPC/DHPC (red) bicelles and are well resolved with twice as many peaks as expected for a monomer. **d**, EmrE is functional in isotropic bicelles. The affinity of EmrE for TPP^{+} is nearly identical in DMPC/DHPC isotropic bicelles (right) and DDM (left), as measured by ITC at pH 7, 45 °C.

dihexanoylphosphatidylcholine (DLPC/DHPC) or dimyristoylphosphatidylcholine/dihexanoylphosphatidylcholine (DMPC/DHPC) isotropic bicelles are nearly identical and characteristic of a folded protein. Spectra collected on TPP^{+}-bound EmrE solubilized in DDM micelles required longer acquisition time due to poor signal to noise, a pattern previously noted for other integral membrane proteins in DDM[32]. Nevertheless, the TROSY spectra (Supplementary Fig. 2) are remarkably similar in DDM and bicelles, and the key features of good chemical shift dispersion and peak doubling are preserved, confirming that the overall structure of EmrE is the same in both environments.

It is not possible to measure transport activity of solubilized protein, so substrate affinity is the best proxy for function. In a good environment EmrE is dimeric and binds TPP^{+} tightly[4,19,29,30,33] with the affinity weakening from dissociation constant $K_d \approx 2$ nM at pH 8.5 to 50 nM at pH 7 (DDM micelles, 4 °C)[19,20]. The pH-dependent TPP^{+} affinity is consistent with substrate competition between TPP^{+} and protons for the single binding site. In a poor environment, EmrE is monomeric and TPP^{+} binding weakens by three to four orders of magnitude[33,34]. Using isothermal titration calorimetry we confirmed that EmrE in DDM micelles at 4 °C binds TPP^{+} tightly, as previously reported[4,19,29,30,33] (Supplementary Fig. 3). At 45 °C, the temperature used for the NMR experiments, EmrE has the same affinity for TPP^{+} in DDM ($K_d = 120 \pm 12$ nM) or isotropic bicelles (170 ± 70 nM) at pH 7 (Fig. 1d and Supplementary Table 1). These values fall well within the variation observed for EmrE reconstituted into liposomes with different lipid composition[33]. The binding stoichiometry is 0.5 (one TPP^{+} per dimer), as expected, indicating that all the EmrE is functional in isotropic bicelles.

TPP^{+}-bound EmrE exchanges between two conformations

With the suitability of isotropic bicelles established, we turned to the key question: does TPP^{+}-bound EmrE interconvert between two states as predicted by the single-site alternating access model? To test this hypothesis we performed a TROSY-selected ZZ-exchange NMR experiment[35], modified to measure the dynamics of the protein without interference from the high lipid concentration in our samples (see Supplementary Information). In these experiments, the ^{15}N chemical shift of each amide is recorded, followed by a mixing period, and then the ^{1}H chemical shift is recorded. In the absence of conformational exchange, the spectrum observed will be identical to the TROSY spectrum. If conformational exchange between states A and B does occur during the mixing time, cross-peaks will appear in the ZZ-exchange spectrum at $^{15}N_A/^{1}H_B$ and $^{15}N_B/^{1}H_A$, forming a 'box' connecting the TROSY peaks corresponding to the chemical shifts of a single amide in state A ($^{15}N_A/^{1}H_A$) and state B ($^{15}N_B/^{1}H_B$) (Fig. 2a). It is immediately apparent that there is widespread conformational exchange with nearly every resolved TROSY peak assigned to exchange pairs. Analysing the peak volumes of the well-resolved exchange pairs reveals equal populations for both states (Supplementary Fig. 4), consistent with a concerted global exchange process.

The kinetics of conformational exchange can be directly determined by varying the mixing time and analysing cross-peak build-up and auto-peak decay as a function of time in the TROSY-selected ZZ-exchange experiment. Cross-peaks grow in and auto-peaks decay owing to exchange between states A and B, and all peaks decay because of the intrinsic relaxation (R^*) of the $S_Z I_\beta$ spin state active during the mixing time (see Supplementary Information)[35,36] according to

$$\frac{d}{dt}\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} -R^{*A} - k & k \\ k & -R^{*B} - k \end{bmatrix}\begin{bmatrix} A \\ B \end{bmatrix} \quad (1)$$

Because the populations are the same, the forward and reverse exchange rate constants are identical (k). Inspection of the data shows

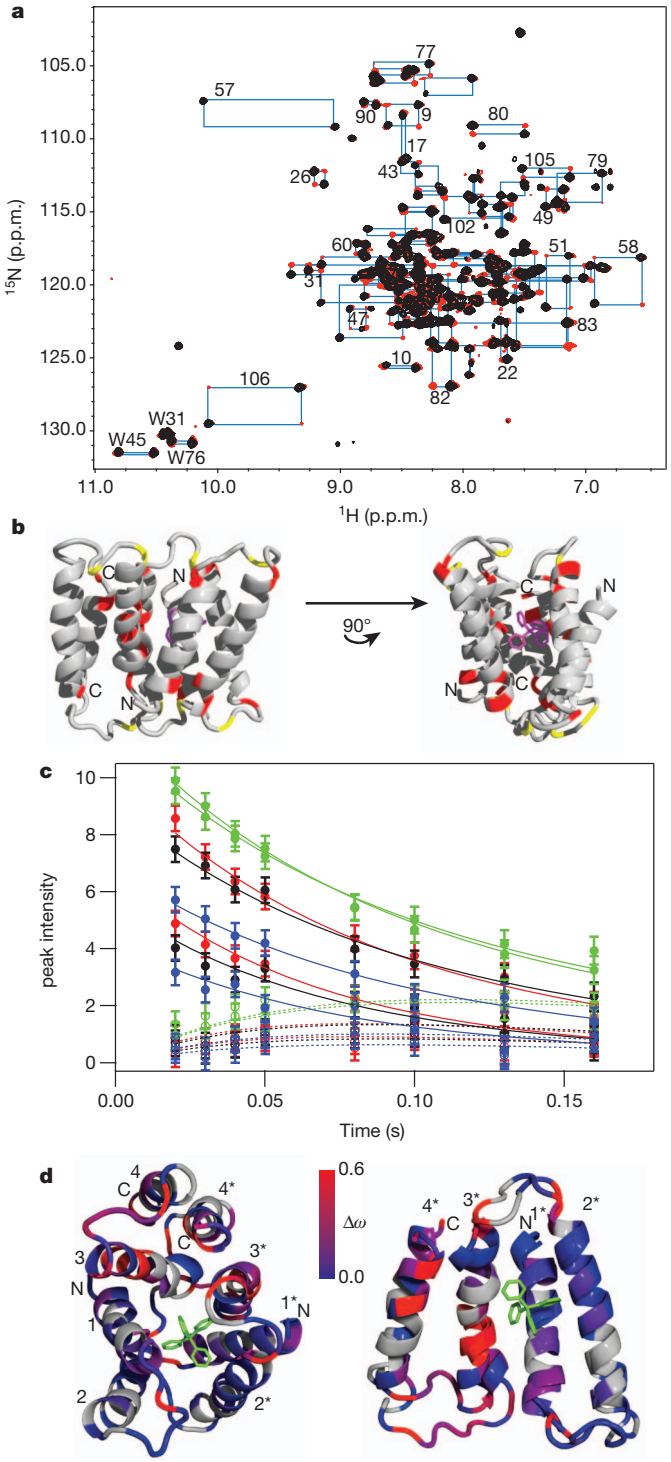


Figure 2 | TPP$^+$-bound EmrE interconverts between two conformations. **a**, Overlay of TROSY ZZ-exchange (red, 100 ms mixing time) and ^{1}H–^{15}N TROSY heteronuclear single-quantum coherence (HSQC) spectra (black) of TPP$^+$-bound ^{2}H/^{15}N-EmrE in isotropic bicelles. Cross-peaks demonstrate conformational exchange and blue boxes connect peaks corresponding to a single amide. All residues experience exchange. Assignments are shown for several well-resolved residues. **b**, Residues used for quantitative analysis are coloured based on whether the intrinsic relaxation rate is the same (red) or different (yellow) in the two states. **c**, Best fit of ZZ-exchange auto- (solid circles) and cross-peak (open circles) intensities as a function of mixing time (Supplementary equation (3)) yields a single global conformational exchange rate of $k = 4.9 \pm 0.5\ s^{-1}$. Each residue is a different colour (additional data in Supplementary Fig. 5). Error bars are estimated from the noise in each spectrum. **d**, Chemical shift difference ($\Delta\omega$) between states A and B (Supplementary equation (4)) are plotted onto the antiparallel dimer model (left) or an overlay of the two monomers with transmembrane helices 1–3 aligned to compare the two monomers within the antiparallel dimer (right).

quite clearly that the common assumption of equal intrinsic relaxation in both states ($R^{*A} = R^{*B}$) is not true for many residues (Fig. 2c and Supplementary Figs 5, 7 and 8). The Palmer group has developed an alternative method for analysing data that eliminates any effects of different R^* (ref. 37). The composite ratio of peak intensities (Ξ) calculated according to their method has a simple quadratic dependence on the mixing time (t), and in the case of equal populations $\Xi = k^2t^2$ (see Supplementary Information). For EmrE, analysis of Ξ shows that all residues collapse onto a single curve (Supplementary Fig. 6), confirming that this is a global conformational exchange process with a single timescale. Fitting the individual peak intensities as a function of time requires a more complex equation (Supplementary equation (3)), but results in the same global rate constant (Fig. 2c, Supplementary Fig. 5). This global conformational exchange rate is also detected in analysis of the tryptophan side chain dynamics (Supplementary Information and Supplementary Fig. 9).

These NMR dynamics experiments directly detect global conformational exchange between two states with equal populations and a rate constant of $k_{\text{open-in to open-out}} = k_{\text{open-out to open-in}} = k = 4.8 \pm 0.5\ s^{-1}$ for TPP$^+$-bound EmrE in isotropic bicelles at 45 °C. The atomic resolution provided by NMR demonstrates that the entire protein is involved in the conformational exchange process, as shown by the complete peak doubling. Importantly, the well-resolved residues used for quantitative analysis of the dynamics are distributed across the protein (Fig. 2b), indicating that the whole protein undergoes exchange on the same timescale. Second, residues with different intrinsic relaxation rates, R^*, map to the loop regions that alternate between a tightly packed protein environment and a loosely packed solvent-exposed environment as EmrE interconverts between inward- and outward-facing states, providing an explanation for the differential R^* (Supplementary Information). These results demonstrate the power of NMR methods to characterize quantitatively the kinetics (rate constants) and thermodynamics (populations) of global conformational exchange in a substrate-bound transporter, the key step for substrate transport across the membrane.

A previous study[20] monitoring intrinsic tryptophan fluorescence to measure substrate binding to EmrE detected an additional slow process with a rate constant of $1.5\ s^{-1}$ that was attributed to a conformational change after TPP$^+$ binding. Considering that the experiments were performed in different environments and at different temperatures, there is good agreement between the measured rates. This suggests both techniques are observing the same process and provides additional evidence that conformational exchange on this timescale in TPP$^+$-bound EmrE is an intrinsic property of the protein.

However, the topology of the exchanging EmrE is also needed to understand the molecular mechanism. Two possibilities are consistent with the NMR data (Supplementary Fig. 1): (1) parallel, symmetric EmrE dimers interconverting between inward- and outward-facing states (AA to BB exchange) and (2) the unique model of asymmetric antiparallel EmrE dimers interconverting between two states that are identical but open to opposite sides of the membrane (AB to BA exchange). With equal populations (1) requires two distinct EmrE conformations that happen to have exactly equal free energies, whereas (2) entails inherently equal populations because each dimer consists of one monomer in each conformation. To distinguish these two possibilities experimentally, we turned to Förster resonance energy transfer (FRET) experiments.

Antiparallel topology within EmrE dimers

There is still significant controversy about the topology of EmrE dimers. Several accessibility studies have shown that EmrE exists in both orientations in its native *E. coli* membrane[17,38]. Most recently,

Schuldiner and colleagues[17] have demonstrated equal populations of both monomer orientations in the native inner membrane of *E. coli* based on the accessibility of single cysteine residues to the outside of the cell. However, these studies do not directly address the relative orientation of monomers within the dimer. Structural studies consistently show an asymmetric antiparallel dimer[4,6,26,27,39,40]. Cross-linking of single-cysteine mutants and tandem genetic fusions have been used to enforce a parallel topology between monomers and these constructs are capable of transport[11,17]. However, an antiparallel genetic fusion is also capable of transport[17]. Charge bias mutants and fusion of fluorescent reporters provide more conflicting results, most probably because the topology is exquisitely sensitive to sequence alteration[16,17].

We performed bulk and single-molecule FRET and cross-linking experiments to determine the relative topology of EmrE monomers within a dimer. All the single-cysteine mutants used in these experiments are known to be functional[13,38,41]. In addition, two-dimensional ^{1}H–^{15}N TROSY spectra of TPP^{+}-bound mutants confirm that these single-cysteine mutants have the same structure as wild type (Supplementary Fig. 10). For bulk FRET measurements, we reconstituted single-cysteine EmrE mutants into liposomes to label specifically a single face of the protein (see Methods). To test for antiparallel topology, we labelled with donor and acceptor from opposite sides of the membrane (Fig. 3a). Significant acceptor fluorescence was observed upon donor excitation (Fig. 3c, red line), indicating FRET occurs and antiparallel dimers are present. To test for parallel topology, EmrE was labelled with both dyes from only one side of the membrane (Fig. 3b). Only minimal FRET was observed (Fig. 3c, blue line), which could be due to dye leakage into liposomes during labelling or a small population of higher-order oligomers (Supplementary Information and Supplementary Fig. 13). No FRET is observed in the control experiment (Fig. 3c, black line), demonstrating that monomers do not swap between dimers under these conditions.

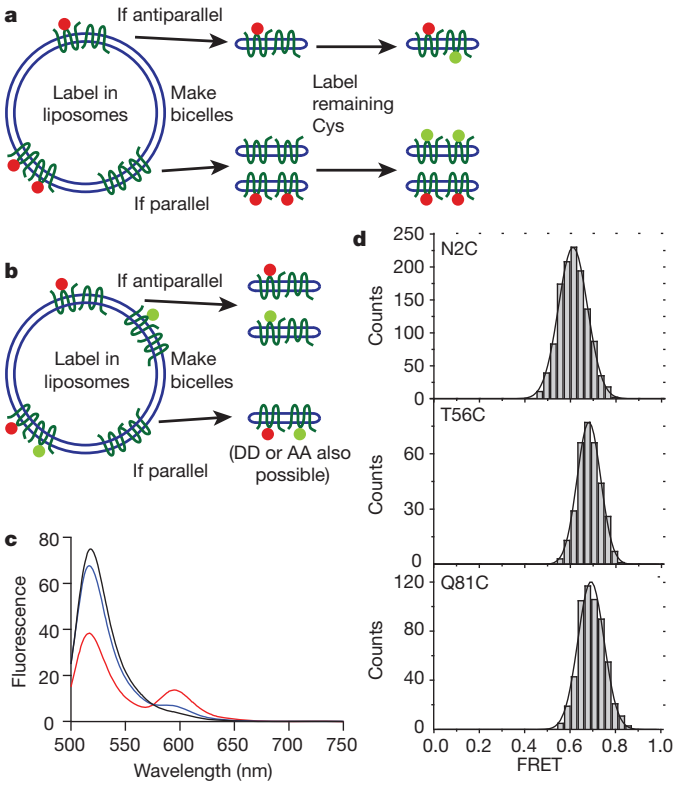


Figure 3 | EmrE is an antiparallel homodimer. T56C-EmrE was labelled in liposomes with two different schemes. **a**, Labelling with a single dye on each side of the membrane will produce FRET only if EmrE is antiparallel. **b**, Labelling with both donor and acceptor on the same side of the membrane will produce FRET only if EmrE is parallel. **c**, Observation of both donor (Alexa Fluor 488) and acceptor (Alexa Fluor 568) fluorescence upon donor excitation in sample **a** indicates FRET and antiparallel topology (red line). In contrast, only minimal FRET was observed in sample **b** (blue line). The control (black line, see Methods) excludes monomer swapping and direct excitation of acceptor. All fluorescence spectra are normalized to total donor fluorescence in SDS, which monomerizes EmrE[49]. **d**, For single-molecule FRET experiments, N2C-, T56C- or Q81C-EmrE was labelled with Cy3/Cy5 in micelles and then reconstituted into bicelles. Particles containing one donor and one acceptor were selected based on photobleaching events. A single, narrow FRET distribution is observed in each case, with a FRET efficiency consistent with an antiparallel topology.

To simplify labelling and avoid leakage issues, we used single-molecule FRET[42]. Single-cysteine EmrE was stoichiometrically labelled with Cy3/Cy5 in micelles at one of three different positions, N2C, T56C or Q81C (Supplementary Fig. 11), followed by incorporation into bicelles. These bicelles contained 0.1% biotinylated lipid to tether specifically Cy3/Cy5-EmrE-containing bicelles to neutravidine molecules on the polymer-coated slide surface (Supplementary Fig. 11). The single-molecule FRET experiments were performed with a wide-field total internal reflection fluorescence microscope set up[43]. Observation of photobleaching events during single molecule time traces allows dimers to be selected that contain one donor and one acceptor fluorophore. A single FRET distribution is observed for single-site labelled EmrE (Fig. 3d). The FRET efficiency of 0.6–0.7, depending on labelling site, corresponds to a distance of approximately 50–55 Å, consistent with transmembrane labelling. Based on the dimensions of the cryo-electron microscopy structure, donor and acceptor would not be more than about 35 Å apart, corresponding to a FRET efficiency greater than 0.9, if they were on the same side of the membrane as required for a parallel topology with this labelling scheme. These results confirm the antiparallel arrangement of monomers within an EmrE dimer.

To test this model further, we cross-linked S107C-EmrE with the heterobifunctional cross-linker sulpho-(*N*-(γ-maleimidobutyryloxy) succinimide) (s-GMBS). Nearly complete cross-linking is observed (Supplementary Fig. 14), demonstrating that residues K22 and S107C are in close proximity in the dimer. Because these positions are on opposite sides of the membrane in each monomer, heterobifunctional cross-linking is only possible with an antiparallel topology (Supplementary Information and Supplementary Fig. 14). Together with the NMR results, this means that TPP^{+}-bound EmrE must exist in an antiparallel asymmetric dimer that interconverts between two identical, oppositely oriented states (AB–BA dimer exchange) when bound to TPP^{+} (Supplementary Fig. 1).

According to this model, no conformational exchange should be observed in the single molecule FRET time traces, because the donor–acceptor distance will be identical in both states. Inspection of the time traces confirms this prediction (Supplementary Fig. 12), even though the rate of conformational exchange determined by NMR is within the range observable in these experiments. Our NMR spectra are very clean, with the double set of peaks fully accounted for by a single population of antiparallel EmrE. We do not observe additional minor peaks (Supplementary Fig. 1), so parallel EmrE is not present in our samples in measurable amounts (less than 5%).

The single-site alternating access model predicts that the two lowest energy states in the TPP^{+}-bound conformation are the inward- and outward-facing states. If EmrE is open to one side of the membrane in our NMR samples, then we should observe different water accessibility for residues at the loops and ends of helices upon exchange from state A to state B because these regions pack together to close off the active site of EmrE at one end and open up to allow access to the active site at the other. To test this, we recorded ^{1}H–^{15}N TROSY spectra with increasing concentrations of a soluble, chelated gadolinium compound. This paramagnetic compound causes line broadening of nearby amides. The results (Fig. 4a) show that some amides are similarly affected in both states whereas others are differentially broadened. The amides with differential accessibility to water map exactly to the regions expected for interconversion of an antiparallel dimer between 'symmetric' states (Fig. 1b and Supplementary Fig. 1), confirming that the conformation

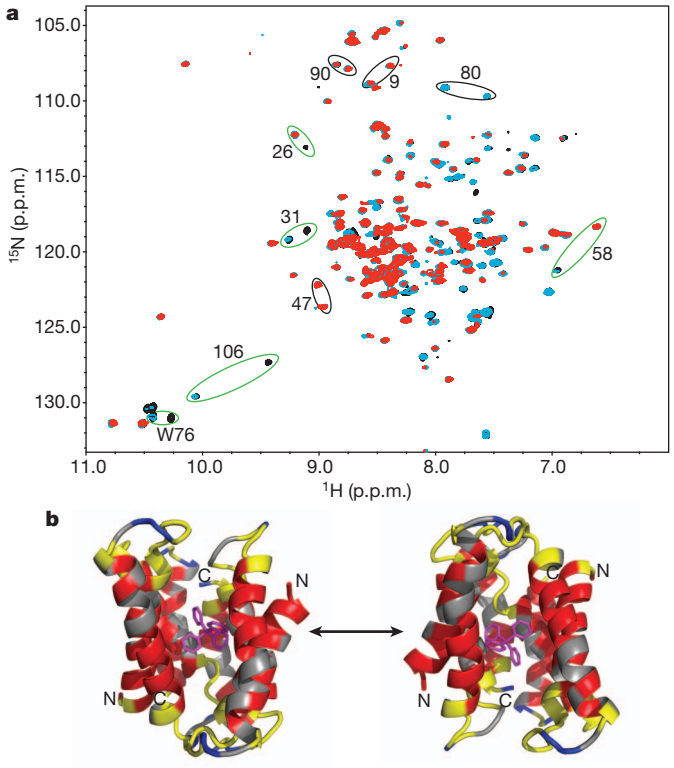


Figure 4 | EmrE has asymmetric water accessibility. **a**, ^{1}H–^{15}N TROSY spectra of TPP^{+}-bound EmrE in bicelles in the absence (black) and presence of 1 mM (blue) and 5 mM (red) paramagnetic gadobenate dimeglumine. Residues with different PRE effects (green circles) or the same PRE effect (black circles) in the two states are highlighted. **b**, Residues with equal accessibility (dark blue) or protection (red) from water in both states are plotted on the structure. Grey residues are not assigned in both states. Residues with differential accessibility to water (yellow) highlight the pore and loop regions, exactly as expected for inward- to outward-facing conformational exchange. This is consistent with an antiparallel dimer open only to one side of the membrane, as in the crystal structure (Protein Data Bank accession number 3B5D).

we observe is open to one side of the membrane and is consistent with the crystal structure (Fig. 4b).

The ZZ-exchange data assign pairs of peaks as exchange partners, representing the chemical shift of a single residue in each state. Chemical shift is quite sensitive to local structural changes. In the conformational exchange model proposed for EmrE, the two monomers swap conformations during the transition from inward- to outward-facing states (AB–BA). Thus, overlay of the two monomers (Fig. 2d) highlights regions where local structure changes upon conformational exchange in this model. The overlay reveals kinking of transmembrane helix 3 and movement of transmembrane helix 4 relative to the substrate-binding domain as the largest structural changes. This is exactly where the largest chemical shift differences between the two states (Supplementary equation (4)) are observed. The transmembrane helices lining the binding pocket move as a generally rigid body, with significant chemical shift differences only in the regions that close off or open up access to the transport pore (Fig. 2d). These chemical shift differences are entirely consistent with the structural differences between the two monomers, supporting the AB–BA exchange model. The two sets of NMR peaks correspond to the two distinct monomer conformations within an exchanging antiparallel asymmetric dimer.

Antiparallel EmrE is competent for transport

We have used solution NMR spectroscopy to observe directly conformational exchange between inward- and outward-facing states of the small multidrug resistance transporter, EmrE. This is the key step in the transport cycle that moves the substrate across the membrane from one aqueous compartment to another. Measurement of bulk transport is not possible in a solubilized system. However, solution NMR provides a powerful tool to follow directly the protein conformational changes that effectively 'transport' TPP^{+}, and we have quantitatively measured this exchange process in TPP^{+}-bound EmrE.

Our FRET experiments show that the monomers within each dimer are antiparallel. Thus, antiparallel substrate-bound EmrE is able to undergo the key conformational exchange step in the single-site alternating access model of antiport. Together with our chemical shift mapping and water accessibility data, this suggests that the low-resolution crystal and cryo-electron microscopy structures are essentially correct, and we are working to refine and improve the resolution of this structure using NMR restraints.

Our results provide experimental evidence for the unusual model proposed by Fleishman *et al.* that the inward- and outward-facing states are identical[18] (Supplementary Fig. 1). This model reconciles the asymmetric antiparallel structural data[4,6,26,27,39,40] with biochemical studies indicating functional symmetry of active site residues[3,44,45] and single distances measured for most residues by electron paramagnetic resonance (EPR)[13]. These results demonstrate that equal insertion of EmrE in both orientations in the *E. coli* inner membrane[17] is functionally relevant and they enhance the importance of EmrE as a potential model for evolution of dual topology membrane proteins[9].

METHODS SUMMARY

EmrE was expressed and purified using a 6× His-tag, which was then removed with thrombin. Purification was performed in decylmaltoside or DDM. EmrE was reconstituted into DMPC or DLPC liposomes using standard methods and then formed into isotropic bicelles by addition of DHPC and several freeze–thaw cycles. Protein concentration was determined using absorbance at 280 nm, with an extinction coefficient determined by amino-acid analysis. Isothermal titration calorimetry (ITC) was performed by titrating 54 μM TPP^{+} stock solutions into 10 μM EmrE, with matching concentrations of detergent or lipid in both solutions.

Bulk FRET labelling was performed in liposomes to label residues separately on either side of the membrane. An 'antiparallel' sample was labelled with donor and acceptor on opposite sides, and a 'parallel' sample was labelled with donor and acceptor on the same side. Donor-only and acceptor-only controls were labelled with dye only on the exterior of the liposome, reconstituted into bicelles and then mixed. Single-molecule FRET samples were labelled in micelles and experiments were performed using a wide-field total internal reflection fluorescence microscope set up[43].

NMR experiments were performed using a 700 MHz Varian NMR spectrometer or 800 MHz Bruker spectrometer equipped with a cryoprobe. All NMR samples contained 0.5–1.0 mM ^{2}H/^{15}N-EmrE in buffer conditions of 2 mM TPP^{+}, 20 mM NaCl, 20 mM potassium phosphate, 2 mM TCEP, pH 7.0, 45 °C. The membrane mimetic in each sample (DDM micelles or isotropic bicelles) was as listed. The TROSY-selected ZZ-exchange experiment[35] was modified to include a lipid 'flipback' pulse. Data were processed and analysed with NMRPipe[46], NMRView[47], Sparky[48] and IgorPro (Wavemetrics). All EmrE structure figures were created in PyMOL using Protein Data Bank 3B5D with the backbone rebuilt to render the cartoons. Full-page versions of the spectra in the main figures are included in the Supplementary Information.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 January; accepted 7 November 2011.
Published online 18 December 2011.

1. Arkin, I. T., Russ, W. P., Lebendiker, M. & Schuldiner, S. Determining the secondary structure and orientation of EmrE, a multi-drug transporter, indicates a transmembrane four-helix bundle. *Biochemistry* **35,** 7233–7238 (1996).
2. Yerushalmi, H., Lebendiker, M. & Schuldiner, S. EmrE, an *Escherichia coli* 12-kDa multidrug transporter, exchanges toxic cations and H^{+} and is soluble in organic solvents. *J. Biol. Chem.* **270,** 6856–6863 (1995).
3. Schuldiner, S. EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim. Biophys. Acta* **1794,** 748–762 (2009).
4. Tate, C. G., Ubarretxena-Belandia, I. & Baldwin, J. M. Conformational changes in the multidrug transporter EmrE associated with substrate binding. *J. Mol. Biol.* **332,** 229–242 (2003).
5. Butler, P. J., Ubarretxena-Belandia, I., Warne, T. & Tate, C. G. The *Escherichia coli* multidrug transporter EmrE is a dimer in the detergent-solubilised state. *J. Mol. Biol.* **340,** 797–808 (2004).

6. Ubarretxena-Belandia, I. & Tate, C. G. New insights into the structure and oligomeric state of the bacterial multidrug transporter EmrE: an unusual asymmetric homo-dimer. *FEBS Lett.* **564,** 234–238 (2004).
7. Chen, Y. J. *et al.* X-ray structure of EmrE supports dual topology model. *Proc. Natl Acad. Sci. USA* **104,** 18999–19004 (2007).
8. Ninio, S., Elbaz, Y. & Schuldiner, S. The membrane topology of EmrE – a small multidrug transporter from *Escherichia coli*. *FEBS Lett.* **562,** 193–196 (2004).
9. Rapp, M., Seppälä, S., Granseth, E. & von Heijne, G. Emulating membrane protein evolution by rational design. *Science* **315,** 1282–1284 (2007).
10. McHaourab, H. S., Mishra, S., Koteiche, H. A. & Amadi, S. H. Role of sequence bias in the topology of the multidrug transporter EmrE. *Biochemistry* **47,** 7980–7982 (2008).
11. Steiner-Mordoch, S. *et al.* Parallel topology of genetically fused EmrE homodimers. *EMBO J.* **27,** 17–26 (2008).
12. Korkhov, V. M. & Tate, C. G. An emerging consensus for the structure of EmrE. *Acta Crystallogr. D* **65,** 186–192 (2009).
13. Amadi, S. T., Koteiche, H. A., Mishra, S. & McHaourab, H. S. Structure, dynamics and substrate-induced conformational changes of the multidrug transporter EmrE in liposomes. *J. Biol. Chem.* **285,** 26710–26718 (2010).
14. Soskine, M., Mark, S., Tayer, N., Mizrachi, R. & Schuldiner, S. On parallel and antiparallel topology of a homodimeric multidrug transporter. *J. Biol. Chem.* **281,** 36205–36212 (2006).
15. Charalambous, K., Miller, D., Curnow, P. & Booth, P. J. Lipid bilayer composition influences small multidrug transporters. *BMC Biochem.* **9,** 31 (2008).
16. Seppala, S., Slusky, J., Lloris-Garcera, P., Rapp, M. & Von Heijne, G. Control of membrane protein topology by a single C-terminal residue. *Science* **328,** 1698–1700 (2010).
17. Nasie, I., Steiner-Mordoch, S., Gold, A. & Schuldiner, S. Topologically random insertion of EmrE supports a pathway for evolution of inverted repeats in ion-coupled transporters. *J. Biol. Chem.* **285,** 15234–15244 (2010).
18. Fleishman, S. J. *et al.* Quasi-symmetry in the cryo-EM structure of EmrE provides the key to modeling its transmembrane domain. *J. Mol. Biol.* **364,** 54–67 (2006).
19. Yerushalmi, H. & Schuldiner, S. A model for coupling of H^+ and substrate fluxes based on 'time-sharing' of a common binding site. *Biochemistry* **39,** 14711–14719 (2000).
20. Adam, Y., Tayer, N., Rotem, D., Schreiber, G. & Schuldiner, S. The fast release of sticky protons: kinetics of substrate binding and proton release in a multidrug transporter. *Proc. Natl Acad. Sci. USA* **104,** 17989–17994 (2007).
21. Jardetzky, O. Simple allosteric model for membrane pumps. *Nature* **211,** 969–970 (1966).
22. West, I. C. Ligand conduction and the gated-pore mechanism of transmembrane transport. *Biochim. Biophys. Acta* **1331,** 213–234 (1997).
23. Glover, K. J. *et al.* Structural evaluation of phospholipid bicelles for solution-state studies of membrane-associated biomolecules. *Biophys. J.* **81,** 2163–2171 (2001).
24. Whiles, J., Ceems, R. & Vold, R. Bicelles in structure–function studies of membrane-associated proteins. *Bioorg. Chem.* **30,** 431–442 (2002).
25. Raschle, T., Hiller, S., Etzkorn, M. & Wagner, G. Nonmicellar systems for solution NMR spectroscopy of membrane proteins. *Curr. Opin. Struct. Biol.* **20,** 471–479 (2010).
26. Lehner, I. *et al.* The key residue for substrate transport (Glu14) in the EmrE dimer is asymmetric. *J. Biol. Chem.* **283,** 3281–3288 (2008).
27. Ubarretxena-Belandia, I., Baldwin, J. M., Schuldiner, S. & Tate, C. G. Three-dimensional structure of the bacterial multidrug transporter EmrE shows it is an asymmetric homodimer. *EMBO J.* **22,** 6175–6181 (2003).
28. Rotem, D. & Schuldiner, S. EmrE, a multidrug transporter from *Escherichia coli*, transports monovalent and divalent substrates with the same stoichiometry. *J. Biol. Chem.* **279,** 48787–48793 (2004).
29. Soskine, M., Adam, Y. & Schuldiner, S. Direct evidence for substrate-induced proton release in detergent-solubilized EmrE, a multidrug transporter. *J. Biol. Chem.* **279,** 9951–9955 (2004).
30. Yerushalmi, H., Mordoch, S. S. & Schuldiner, S. A single carboxyl mutant of the multidrug transporter EmrE is fully functional. *J. Biol. Chem.* **276,** 12744–12748 (2001).
31. Muth, T. R. & Schuldiner, S. A membrane-embedded glutamate is required for ligand binding to the multidrug transporter EmrE. *EMBO J.* **19,** 234–240 (2000).
32. Sanders, C. R. & Sönnichsen, F. Solution NMR of membrane proteins: practice and challenges. *Magn. Reson. Chem* **44** (*NMR of Proteins in Solution* special issue) S24–S40 (2006).
33. Miller, D. *et al. In vitro* unfolding and refolding of the small multidrug transporter EmrE. *J. Mol. Biol.* **393,** 815–832 (2009).
34. Sikora, C. W. & Turner, R. J. Investigation of ligand binding to the multidrug resistance protein EmrE by isothermal titration calorimetry. *Biophys. J.* **88,** 475–482 (2005).
35. Li, Y. & Palmer, A. G. TROSY-selected ZZ-exchange experiment for characterizing slow chemical exchange in large proteins. *J. Biomol. Nucl. Magn. Reson.* **45,** 357–360 (2009).
36. Farrow, N. A., Zhang, O., Forman-Kay, J. D. & Kay, L. E. A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudinal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. Nucl. Magn. Reson.* **4,** 727–734 (1994).
37. Miloushev, V. Z. *et al.* Dynamic properties of a type II cadherin adhesive domain: implications for the mechanism of strand-swapping of classical cadherins. *Structure* **16,** 1195–1205 (2008).
38. Nara, T. *et al.* Anti-parallel membrane topology of a homo-dimeric multidrug transporter, EmrE. *J. Biochem.* **142,** 621–625 (2007).
39. Korkhov, V. M. & Tate, C. G. Electron crystallography reveals plasticity within the drug binding site of the small multidrug transporter EmrE. *J. Mol. Biol.* **377,** 1094–1103 (2008).
40. Agarwal, V., Fink, U., Schuldiner, S. & Reif, B. MAS solid-state NMR studies on the multidrug transporter EmrE. *Biochim. Biophys. Acta* **1768,** 3036–3043 (2007).
41. Mordoch, S. S., Granot, D., Lebendiker, M. & Schuldiner, S. Scanning cysteine accessibility of EmrE, an H+-coupled multidrug transporter from *Escherichia coli*, reveals a hydrophobic pathway for solutes. *J. Biol. Chem.* **274,** 19480–19486 (1999).
42. Ha, T. *et al.* Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl Acad. Sci. USA* **93,** 6264–6268 (1996).
43. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nature Methods* **5,** 507–516 (2008).
44. Sharoni, M., Steiner-Mordoch, S. & Schuldiner, S. Exploring the binding domain of EmrE, the smallest multidrug transporter. *J. Biol. Chem.* **280,** 32849–32855 (2005).
45. Weinglass, A. B. *et al.* Exploring the role of a unique carboxyl residue in EmrE by mass spectrometry. *J. Biol. Chem.* **280,** 7487–7492 (2005).
46. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. Nucl. Magn. Reson.* **6,** 277–293 (1995).
47. Johnson, B. A. & Blevins, R. A. NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. Nucl. Magn. Reson.* **4,** 603–614 (1994).
48. Goddard, T. D. & Kneller, D. G. Sparky 3. University of California, San Francisco.
49. Rotem, D., Sal-man, N. & Schuldiner, S. *In vitro* monomer swapping in EmrE, a multidrug transporter from *Escherichia coli*, reveals that the oligomer is the functional unit. *J. Biol. Chem.* **276,** 48243–48249 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Y. Liu for assistance replicating the ITC data. We thank J. Villali for assistance growing isotopically labelled EmrE. We are grateful to Y. Li and A. Palmer for providing pulse programs and G. Chang for providing the EmrE expression plasmid. This work was supported by the National Institutes of Health (1R01GM095839) and the Searle Scholars Program (K.H.W.), the US Department of Energy, Office of Basic Energy Sciences (D.K.), the Howard Hughes Medical Institute (D.K. and T.H.), and an NSF graduate research fellowship to E.M. (DGE-1143954).

Author Contributions E.M. and K.H.W. optimized EmrE sample preparation, performed ITC experiments and recorded two-dimensional NMR spectra under different conditions. E.M. and G.D. performed the bulk FRET experiments. S.D. and R.V. performed the single molecule FRET experiments with guidance from T.H. G.D. and M.C. optimized the modified ZZ-exchange NMR experiment. M.C. collected the ZZ-exchange data and K.H.W. analysed it. G.D. performed the paramagnetic relaxation enhancement NMR experiments. A.B., E.M., G.D. and K.H.W. contributed to assignments. D.K. mentored initial project development. K.H.W. conceived the project and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to K.H.W. (khenzler@wustl.edu).

METHODS

EmrE expression and purification. EmrE was expressed using a pET15b plasmid provided by G. Chang[7]. This vector produces EmrE with an amino (N)-terminal 6× His tag that can be removed by cleavage with thrombin to leave only two extra N-terminal residues (GS). BL21(DE3) cells transformed with this vector were grown in M9 minimal media. EmrE was induced with 0.33 mM IPTG at an absorbance $A_{600\,nm}$ of 0.7–0.8 at 17 °C. Cells were collected after 14–20 h. $^{1}H/^{15}N$-labelled EmrE was produced in the same way substituting 1 g $^{15}NH_4Cl$. $^{2}H/^{15}N$-labelled EmrE was produced by growing cells in $^{2}H/^{15}N$ M9 (1 g $^{15}NH_4Cl$, 2 g glucose, 12.8 g $Na_2HPO_4 \cdot 7H_2O$, 3.0 g KH_2PO_4, 0.5 g NaCl, 2 ml 1 M $MgSO_4$ in D_2O, 100 µl 1 M $CaCl_2$in D_2O, 100 mg ampicillin, one generic multivitamin, 0.5 g $^{2}H/^{15}N$ Isogro (Sigma) per litre).

Cell pellets were re-suspended in lysis buffer (100 mM NaCl, 2.5 mM $MgSO_4$, 20 mM tris pH 7.5, 250 mM sucrose, 5 mM β-mercaptoethanol, 1 mg ml^{-1} lysozyme, DNase, 1 µg ml^{-1} pepstatin, 10 µM leupeptin and 100 µM PMSF) and lysed by sonication. The membrane fraction was separated by a high-speed spin (30,000*g* for 1 h), re-suspended in the same buffer and solubilized with 40 mM decylmaltoside (Anatrace) at room temperature for 2 h. After a second high-speed spin, the supernatant was applied to Ni-NTA His•Bind beads (Novagen) prewashed with buffer A (10 mM decylmaltoside, 10 mM KCl, 90 mM NaCl, 20 mM tris, pH 7.8, 5 mM β-mercaptoethanol) and allowed to bind for 15 min at room temperature. The beads were washed with 10 bed volumes of buffer A, followed by 10 bed volumes of buffer B (buffer A plus 5 mM imidazole). EmrE was eluted with five bed volumes of elution buffer (buffer A plus 400 mM imidazole). The salt concentration was increased to 200 mM and thrombin was added to cleave the His-tag overnight at room temperature. Samples were then concentrated and 0.5 ml aliquots loaded onto a Superdex 200 column pre-equilibrated in NMR buffer (20 mM potassium phosphate, 20 mM NaCl, pH 7.0) with 10 mM decylmaltoside. The protein eluted near 14–15 ml with a symmetric peak. Fractions containing EmrE were combined and reconstituted into bicelles. For samples in DDM (Anatrace), DDM was substituted for decylmaltoside throughout the protocol and the fast protein liquid chromatography fractions were combined and concentrated to the desired final protein/detergent concentration.

Preparation of isotropic bicelles. First, the amount of EmrE in the combined fast protein liquid chromatography fractions was determined as described below. Long-chain lipid (DLPC (1,2-dilauroyl-*sn*-glycero-3-phosphocholine) or DMPC (1,2-dimyristoyl-*sn*-glycero-3-phosphocholine), Avanti Polar Lipids) was hydrated in NMR buffer at 20 mg ml^{-1}. At least 200:1 molar ratio of lipid:EmrE dimer was used for all samples. The lipids were bath sonicated for 10 min and 50 µl of 10% octyl glucoside (Anatrace) was added per millitre of solution. After 20 min, they were mixed with the fast protein liquid chromatography fractions containing EmrE and incubated for 30 min. Three aliquots of 30 mg BioBeads (BioRad) per milligram of total detergent were used to remove the detergent. After removal of the BioBeads, the vesicles were collected by ultracentrifugation at 50,000*g* for 1 h at 20 °C. The supernatant was removed and the pellet was re-suspended in NMR buffer containing DHPC (1,2-dihexanoyl-*sn*-glycero-3-phosphocholine, Avanti Polar Lipids) to break the liposomes up into bicelles. The DHPC concentration was calculated to produce a 1:3 ratio of long-chain lipid:DHPC, assuming 85% recovery of long-chain lipid. Four freeze–thaw cycles were used to produce uniform bicelles; samples were stored at −80 °C until use.

EmrE concentration determination. EmrE concentration was determined using absorbance at 280 nm. The extinction coefficient (38,370 l mol^{-1} cm^{-1}) was calibrated using amino-acid analysis of three samples of EmrE each in DDM and decylmaltoside, and was found to be the same for EmrE in bicelles.

ITC. ITC experiments were performed using a VP-ITC titration microcalorimeter (MicroCal) by titrating TPP^+ (50–80 µM) into EmrE (9–13 µM) in isotropic bicelles ($q = 0.33$, DMPC/DHPC) or 5 mM analytical grade DDM. Both the TPP^+ and EmrE solution were in NMR buffer and had matching detergent or lipid concentrations. Matching bicelle stocks were produced by acquiring proton NMR spectra of all samples (empty bicelle blank, TPP^+ stock, EmrE stock), integrating the DMPC and DHPC terminal methyl peaks, and ensuring that the lipid ratio and peak volumes matched. Isotropic bicelle samples had a total lipid concentration of 30–50 mM and $q_{effective} = 0.33$[23]. The TPP^+ concentration in the final stock solution was determined spectrophotometrically ($\varepsilon = 4{,}400$ l mol^{-1} cm^{-1} at 269 nm, 3,750 l mol^{-1} cm^{-1} at 276 nm). Heats of dilution were determined from reference titrations of the same TPP^+ stock into empty micelles or bicelles. Data were fitted to a model of the ligand TPP^+ (X) binding to n independent and identical sites on the macromolecule EmrE (M) to determine the association constant (K), enthalpy of binding (ΔH) and binding stoichiometry (n), using equation (1):

$$Q_i^{tot} = V_o \Delta H \bullet M_i^{tot} \frac{nKx}{1+Kx} \qquad (2)$$

where x is the free ligand concentration, M_i^{tot} is the total macromolecule concentration, Q_i^{tot} is total heat after the ith injection and V_o is the cell volume. The data were fitted with a nonlinear least-squares approach using the ITC Data Analysis in Origin software supplied with the calorimeter (OriginLab).

Error was determined from standard deviation between replicate experiments. The K values from each replicate were averaged, and then the average value was converted to the dissociation constant, K_d.

Bulk FRET sample preparation and measurement. All FRET and cross-linking experiments use single-cysteine mutants of EmrE: the three native cysteines are mutated to serine and a single cysteine is introduced at the desired location. T56C-EmrE was reduced with DTT and then reconstituted into DMPC liposomes with a molar ratio at least 300:1 lipid:EmrE monomer. After ultracentrifugation to collect the liposomes, they were re-suspended in deoxygenated NMR buffer with 2 mM cysteine and extruded through 400 µm filters to produce unilamellar vesicles loaded with cysteine. The sample was passed over a G25 Sephadex column to remove free cysteine from the exterior of the liposomes. The tight-binding substrate TPP^+ was maintained at saturating concentrations throughout the preparation to stabilize EmrE dimers and prevent monomer swapping.

To test for antiparallel topology, the first dye-maleimide was added to the exterior of the liposomes at 5× molar ratio relative to EmrE monomer. The reaction was allowed to proceed for 30 min, then quenched by addition of 20-fold excess of β-mercaptoethanol. Free dye was removed by collecting the liposomes by ultracentrifugation, re-suspending in fresh buffer and repeating the ultracentrifugation. The second dye-maleimide was added along with octyl glucoside to disrupt the liposomes. The reaction was allowed to proceed for 1 h and then quenched as before. Free dye and detergent were removed by passing the sample over a second G25 Sephadex column. DHPC was then added to the liposome suspension to form bicelles. Alexa Fluor 488 was used as the donor and Alexa Fluor 568 as the acceptor for bulk FRET experiments.

To test for parallel topology, samples were produced in a similar manner, but labelled with mixed donor and acceptor only from the exterior of the liposome. After the labelling reaction, remaining free dye was quenched and removed, then DHPC added to form bicelles. This should produce EmrE with only one face labelled by fluorescent dye, and any cysteine that faces the interior of the liposome should remain unlabelled. Two additional control samples were independently labelled with either donor only or acceptor only from the exterior of the liposome and then mixed.

Fluoresence measurements were made using a PTI spectrofluorimeter (Photon Technology International) using FeliX fluorescence analysis software version 1.42b (Photon Technology International). Labelled T56C-EmrE samples were diluted into isotropic bicelles or 5% SDS containing 2 mM TPP^+. The donor, Alexa Fluor 488 was excited at a wavelength of 488 nm and emission spectra were collected scanning from 500 to 750 nm. The acceptor, Alexa Fluor 568, was excited at a wavelength of 568 nm and emission spectra were collected scanning from 580 to 750 nm.

Single-molecule FRET experiments. Three different single-cysteine mutants, N2C, Q81C and T56C, were labelled with Cy3-maleimide and Cy5-maleimide for single-molecule FRET experiments. Labelling was performed in the same manner as bulk FRET samples, or by labelling EmrE in detergent micelles using an equimolar mixture of donor and acceptor before reconstitution into isotropic bicelles as previously described. The final bicelles used for single-molecule FRET experiments contained 0.1% biotinyl-DPPE for immobilizing the samples.

Single-molecule experiments were performed on a wide-field total internal reflection fluorescence microscope set up[43]. Biotinylated bicelles containing Cy3 and Cy5 labelled EmrE were specifically immobilized on a polymer-coated quartz surface. Then free bicelles were flushed out of the chamber and molecules were imaged in the imaging buffer consisting of 3 mM Trolox and the oxygen scavenger system (0.8% dextrose, 0.1 mg ml^{-1} glucose oxidase, 0.02 mg ml^{-1} catalase) in NMR buffer (2 mM TPP^+, 20 mM NaCl, 20 mM potassium phosphate, pH 7.0). A 532 green laser (Coherent) was used for Cy3 excitation and the sample was imaged by a charge-coupled-device camera (iXon DV 887-BI; Andor Technology). Homemade IDL and C++ programs were used to record and analyse the movies. FRET efficiency was calculated from $I_A/(I_D + I_A)$, where I_D and I_A are the donor (Cy3) and acceptor (Cy5) fluorescent intensities respectively.

For each sample, several minute-long movies were collected (imaging area 70 µm× 35 µm) at 100 ms time resolution. Donor and acceptor intensity time traces were corrected for the background and smoothed using four-point adjacent-averaging. FRET efficiencies from molecules that showed single Cy3 and Cy5 photobleaching steps were chosen to build the histograms.

Cross-linking of EmrE. *o*-PDM (*N*,*N*′-(*o*-phenylene)dimaleimide) and s-GMBS were used to cross-link S107C EmrE to test for parallel (*o*-PDM) or antiparallel (s-GMBS) topology. These experiments were performed with 90 µM EmrE,

20 mM potassium phosphate, 20 mM sodium chloride, 1 mM TCEP, pH 7, and cross-linking for 20 min at 37 °C followed by quenching with β-mercaptoethanol at 20× the cross-linker concentration. Cross-linking in detergent was performed in 10 mM decylmaltoside; cross-linking in lipid was performed in DLPC liposomes at the specified protein:lipid ratio. Addition of SDS monomerizes EmrE and provides a control. S107/K22R serves as a control to determine whether the lysine side chain or N-terminal amine participates in the s-GMBS cross-linking reaction.

NMR sample preparation and data acquisition. All NMR samples were 0.5–1.0 mM $^{2}H/^{15}N$-EmrE, and contained excess (2 mM) TPP^{+} to saturate the protein with substrate. The sample in DDM (Supplementary Fig. 1) had 118 mM DDM. All other NMR samples were isotropic bicelles as described above, with at least 100:1 long-chain lipid:EmrE molar ratio, total lipid concentrations of 300–400 mM and $q \approx 0.33$. The q value was confirmed for each sample by integrating the DMPC (or DLPC) and DHPC methyl resonances. All NMR samples were prepared in 20 mM potassium phosphate, 20 mM NaCl buffer, pH 7.0, and contained 0.05% NaN_3, 2 mM TCEP and 10% D_2O.

Two-dimensional TROSY spectra and the ^{15}N-separated nuclear Overhauser enhancement spectroscopy (NOESY)–HSQC and rotating frame nuclear Overhauser spectroscopy (ROESY)–HSQC spectra were acquired on a 700 MHz Varian spectrometer equipped with a room-temperature probe using standard pulse sequences with gradient coherence selection. The TROSY-selected ZZ-exchange experiment[35] was modified to include a lipid 'flipback' pulse and was acquired on a 800 MHz Bruker spectrometer equipped with a cryoprobe. ZZ-exchange spectra were acquired with mixing times of 20, 30, 40, 50, 80, 100, 130, 160 ms with 128 scans per increment and 128 complex points in the indirect dimension. Eighty per cent of the backbone resonances were assigned using a non-standard protocol combining standard triple resonance experiments (TROSY-HNCA, TROSY-HNCO, TROSY-HN(CO)CA) with amino-acid-specific labelling and ZZ-exchange data. NMR data were processed and analysed with NMRPipe[46], NMRView[47], Sparky[48] and IgorPro (Wavemetrics). All EmrE structure figures were created in PyMOL using Protein Data Bank 3B5D with the backbone rebuilt to render the cartoons. Full-page versions of the spectra in the main figures are included in the Supplementary Information.

doi:10.1038/nature10652

A gas cloud on its way towards the supermassive black hole at the Galactic Centre

S. Gillessen[1], R. Genzel[1,2], T. K. Fritz[1], E. Quataert[3], C. Alig[4], A. Burkert[4,1], J. Cuadra[5], F. Eisenhauer[1], O. Pfuhl[1], K. Dodds-Eden[1], C. F. Gammie[6] & T. Ott[1]

Measurements of stellar orbits[1–3] provide compelling evidence[4,5] that the compact radio source Sagittarius A* at the Galactic Centre is a black hole four million times the mass of the Sun. With the exception of modest X-ray and infrared flares[6,7], Sgr A* is surprisingly faint, suggesting that the accretion rate and radiation efficiency near the event horizon are currently very low[3,8]. Here we report the presence of a dense gas cloud approximately three times the mass of Earth that is falling into the accretion zone of Sgr A*. Our observations tightly constrain the cloud's orbit to be highly eccentric, with an innermost radius of approach of only ~3,100 times the event horizon that will be reached in 2013. Over the past three years the cloud has begun to disrupt, probably mainly through tidal shearing arising from the black hole's gravitational force. The cloud's dynamic evolution and radiation in the next few years will probe the properties of the accretion flow and the feeding processes of the supermassive black hole. The kilo-electronvolt X-ray emission of Sgr A* may brighten significantly when the cloud reaches pericentre. There may also be a giant radiation flare several years from now if the cloud breaks up and its fragments feed gas into the central accretion zone.

As part of our NACO[9] and SINFONI[10,11] Very Large Telescope (VLT) observation programmes studying the stellar orbits around the Galactic Centre supermassive black hole, Sgr A*, we have discovered an object moving at about 1,700 km s^{-1} along a trajectory almost straight towards Sgr A* (Fig. 1). The object has a remarkably low temperature (about 550 K; Supplementary Fig. 2) and a luminosity about five times that of the Sun, unlike any star we have so far seen near Sgr A*. It is also seen in the spectroscopic data as a redshifted emission component in the Brγ and Brδ hydrogen lines, and the 2.058 μm He I lines, with the same proper motion as the L′-band object. Its three-dimensional velocity increased from 1,200 km s^{-1} in 2004 to 2,350 km s^{-1} in 2011. The Brγ emission is elongated along its direction of motion with a spatially resolved velocity gradient (Fig. 2). Together, these findings show that the object is a dusty, ionized gas cloud.

The extinction of the ionized gas is typical for the central parsec (Supplementary Information section 1) and its intrinsic Brγ luminosity is 1.66 ($\pm$0.25) $\times$ 10^{-3} times that of the Sun. For case B recombination the implied electron density is $n_c = 2.6 \times 10^5 f_V^{-1/2} R_c^{-3/2} T_e^{0.54}$ cm^{-3}, for an effective cloud radius of $R_c \approx 15$ milli-arcseconds (mas), volume filling factor f_V ($\leq$1) and an assumed electron temperature T_e in units of 10^4 K, a value typical for the temperatures measured in the central parsec[12]. The cloud mass is $M_c = 1.7 \times 10^{28} f_V^{1/2} R_c^{3/2} T_e^{0.54}$ g or about $3 f_V^{1/2}$ Earth masses. It may plausibly be photo-ionized by the ultraviolet radiation field from nearby massive hot stars, as we infer from a comparison of the recombination rate with the number of impinging Lyman continuum photons[3,13]. This conclusion is supported by the He I/Brγ line flux ratio of approximately 0.7, which is similar to the values found in the photo-ionized gas in the central parsec (0.35–0.7). If so, the requirement of complete photo-ionization sets a lower limit to f_V of $10^{-1 \pm 0.5}$ for the extreme case that the cloud is a thin sheet.

The combined astrometric and radial velocity data tightly constrain the cloud's motion. It is on a highly eccentric ($e = 0.94$) Keplerian orbit bound to the black hole (Fig. 1, Table 1 and Supplementary Information section 2). The pericentre radius is a mere 36 light hours (3,100 Schwarzschild radii, R_S), which the cloud will reach in summer 2013. Only the two stars S2 (r_{peri} = 17 light hours) and S14 (r_{peri} = 11 light hours) have come closer to the black hole[2,3] since our monitoring started in 1992. Although the cloud's gas density may be only modestly greater than that of other ionized gas clouds in the central parsec—$n_e \approx (0.1-2) \times 10^5$ cm^{-3}; refs 12 and 14—it has a specific angular momentum about 50 times smaller[12].

For the nominal properties of the X-ray detected accretion flow onto the black hole[15,16] the cloud should stay close to Keplerian motion all the way to the pericentre (Supplementary Information sections 3 and 4). Its density is currently about $300 f_V^{-1/2}$ times greater than that of the surrounding hot gas in the accretion flow[15]: extrapolating to the pericentre, its density contrast will then still be about $60 f_V^{-1/2}$. Similarly, the cloud's ram pressure by far exceeds that of the hot gas throughout the orbit. In contrast, the thermal pressure ratio will quickly decrease from unity at apocentre and the hot gas is expected to drive a shock slowly compressing the cloud. Whereas the external pressure compresses the cloud from all directions, the black hole's tidal forces shear the cloud along the direction of its motion, because the Roche density for self-gravitational stabilization exceeds the cloud density by nine orders of magnitude[3]. In addition, the ram pressure compresses the cloud parallel to its motion. The interaction between the fast-moving cloud and the surrounding hot gas should also lead to shredding and disruption, owing to the Kelvin–Helmholtz and Rayleigh–Taylor instabilities[17–20]. Rayleigh–Taylor instabilities at the leading edge should in fact break up the cloud within the next few years if it started as a spheroidal, thick blob (Supplementary Information section 3). A thin, dense sheet would by now already have fragmented and disintegrated, suggesting that f_V is of the order of unity.

We are witnessing the cloud's disruption happening in our spectroscopic data (Fig. 2). The intrinsic velocity width more than tripled over the last eight years, and we see between 2008 and 2011 a growing velocity gradient along the orbital direction. Test particle calculations implementing only the black hole's force show that an initially spherical gas cloud placed on the orbit (Table 1) is stretched along the orbit and compressed perpendicular to it, with increasing velocity widths and velocity gradients reasonably matching our observations (Fig. 3 and Supplementary Fig. 4). There is also a tail of gas with lower surface brightness on approximately the same orbit as the cloud, which cannot be due to tidal disruption alone. It may be stripped gas, or lower-density, lower-filling-factor gas on the same orbit. The latter explanation is more plausible given that the integrated Brγ and L′-band

[1]Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstrasse 1, D-85748 Garching, Germany. [2]Department of Physics, Le Conte Hall, University of California, Berkeley, California 94720, USA. [3]Department of Astronomy, University of California, Berkeley, California 94720, USA. [4]Universitätssternwarte der Ludwig-Maximilians-Universität, Scheinerstrasse 1, D-81679 München, Germany. [5]Departamento de Astronomía y Astrofísica, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile. [6]Center for Theoretical Astrophysics, Astronomy and Physics Departments, University of Illinois at Urbana-Champaign, 1002 West Green Street, Urbana, Illinois 61801, USA.

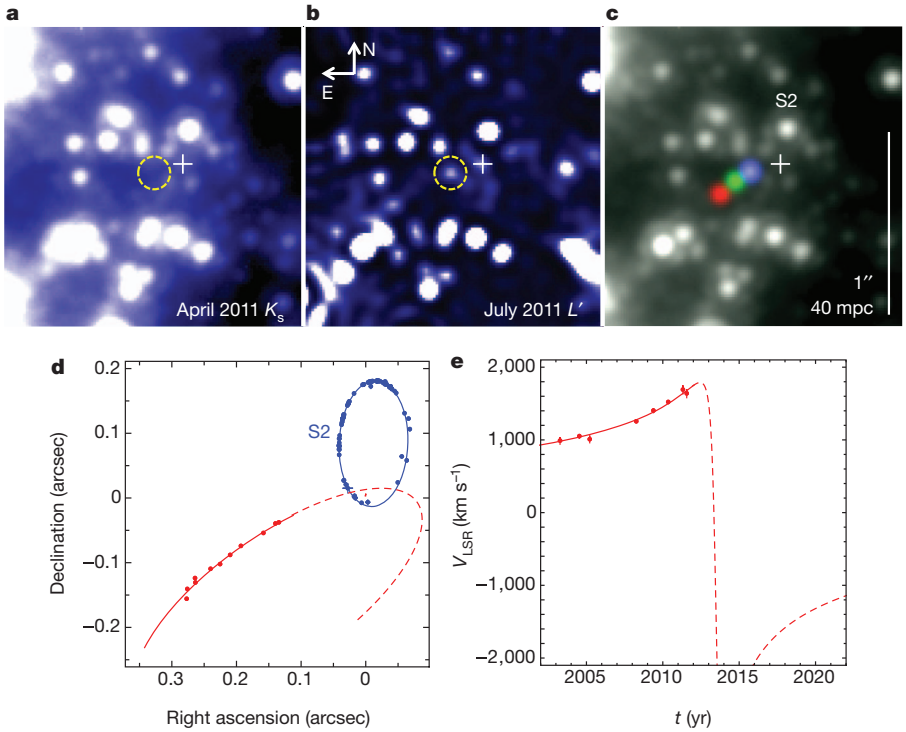


Figure 1 | Infalling dust/gas cloud in the Galactic Centre. **a**, **b**, NACO[9] adaptive optics VLT images showing that the cloud (dashed circle) is detected in the L′-band (3.76 μm, deconvolved and smoothed with a 2-pixel Gaussian kernel) but not in the K_s-band (2.16 μm, undeconvolved), indicating that it is not a star but a dusty cloud with a temperature of around 550 K (Supplementary Fig. 2). The cloud is also detected in the M-band (4.7 μm) but not seen in the H-band (1.65 μm). North is up, East is left. The white cross marks the position of Sgr A*. **c**, The proper motion derived from the L-band data is about 42 mas yr^{-1}, or 1,670 km s^{-1} (in 2011), from the southeast towards the position of Sgr A* (red for epoch 2004.5, green for 2008.3 and blue for 2011.3, overlaid on a 2011 K_s-band image). The cloud is also detected in deep spectroscopy with the adaptive-optics-assisted integral field unit SINFONI[10,11] in the H I $n = 7$–4 Brγ recombination line at 2.1661 μm and in He I at 2.058 μm, with a radial velocity of 1,250 km s^{-1} (in 2008) and 1,650 km s^{-1} (in 2011). **d**, **e**, The combination of the astrometric data in L′ and Brγ and the radial velocity (V_{LSR}) data in Brγ tightly constrains the orbit of the cloud (error bars are 1σ measurement errors). The cloud is on a highly eccentric, bound orbit ($e = 0.94$), with a pericentre radius and epoch of 36 light hours (3,100R_S) and 2013.5 (Table 1). The blue ellipse shows the orbit of the star S2 for comparison[1–3]. For further details see the Supplementary Information.

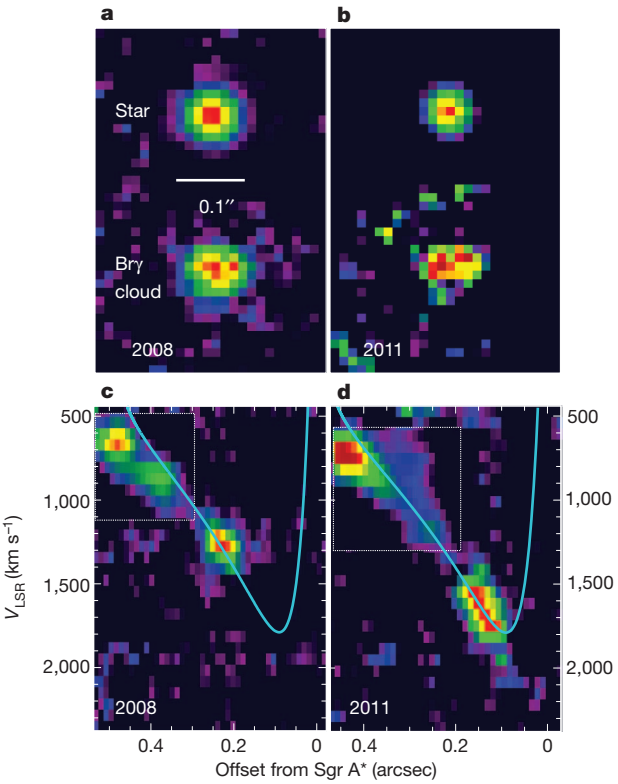


Figure 2 | The velocity shear in the gas cloud. The left column shows data from 2008.3, the right from 2011.3. Panels **a** and **b** show integrated Brγ maps of the cloud, in comparison to the point spread function from stellar images shown above. The inferred intrinsic East–West half-width at half-maximum source radii are $R_c = 21 \pm 5$ mas in 2008 and 19 ± 8 mas in 2011 (approximately along the direction of orbital motion), after removal of the instrumental broadening estimated from the stellar images above. A similar spatial extent is found from the spatial separation between the red- and blue-shifted emission of the cloud ($R_c = 23 \pm 5$ mas). The minor-axis radius of the cloud is only marginally resolved or unresolved (radius less than 12 mas). We adopt $R_c = 15$ mas as the 'effective' circular radius, by combining the results in the two directions. Panels **c** and **d** are position-velocity maps, obtained with SINFONI on the VLT, of the cloud's Brγ emission. The slit is oriented approximately along the long axis of the cloud and the projected orbital direction and has a width of 62 mas for the bright 'head' of the emission. For the lower-surface-brightness 'tail' of emission (in the enclosed white dotted regions) we smoothed the data with 50 mas and 138 km s^{-1} and used a slit width of 0.11″. The gas in the tail is spread over around 200 mas downstream of the cloud. The trailing emission appears to be connected by a smooth velocity gradient (of about 2 km s^{-1} mas^{-1}), and the velocity field in the tail approximately follows the best-fit orbit of the head (cyan curves; see also Table 1). An increasing velocity gradient has formed in the head between 2008 (2.1 km s^{-1} mas^{-1}) and 2011 (4.6 km s^{-1} mas^{-1}). As a result of this velocity gradient, the intrinsic integrated full-width at half-maximum (FWHM) velocity width of the cloud increased from 89 (±30) km s^{-1} in 2003 and 117 (±25) km s^{-1} in 2004, to 210 (±24) km s^{-1} in 2008, and 350 (±40) km s^{-1} in 2011.

Table 1 | Orbit parameters of the infalling cloud

Parameters of Keplerian orbit around the $4.31 \times 10^6 M_\odot$ black hole at $R_0 = 8.33$ kpc	Best-fitting value
Semi-major axis, a	521 ± 28 mas
Eccentricity, e	0.9384 ± 0.0066
Inclination of ascending node, i	106.55 ± 0.88 deg
Position angle of ascending node, Ω	101.5 ± 1.1 deg
Longitude of pericentre, ω	109.59 ± 0.78 deg
Time of pericentre, t_{peri}	2013.51 ± 0.035
Pericentre distance from black hole, r_{peri}	$4.0 \pm 0.3 \times 10^{15}$cm $= 3{,}140 R_S$
Orbital period, t_o	137 ± 11 years

luminosities did not drop by more than 30% between 2004 and 2011, and the integrated Brγ flux of the tail is comparable to that of the cloud.

The disruption and energy deposition processes in the next years until and after the pericentre passage are powerful probes of the physical conditions in the accretion zone (Supplementary section 3). We expect that the interaction between hot gas and cloud will drive a strong shock into the cloud. Given the densities of cloud and hot gas, the cloud as a whole should remain at low temperature until just before it reaches the pericentre. Near the pericentre the post-shock temperature of the cloud may increase rapidly to $T_{\mathrm{postshock,c}} \approx 6-10 \times 10^6$ K, resulting in X-ray emission. We estimate the observable 2−8-keV luminosity to be less than 10^{34} erg s^{-1} there, somewhat larger than the current 'quiescent' X-ray luminosity[6,15,21] of Sgr A* (10^{33} erg s^{-1}). Rayleigh–Taylor instabilities may by then have broken up the cloud into several fragments, in which case the emission may be variable throughout this period. Our predictions depend sensitively on the density and disruption state of the cloud, as well as on the radial dependencies of the hot gas properties, none of which we can fully quantify. The steeper the radial profiles are and the higher the value of f_V, the more X-ray emission will occur. Shallower profiles and a low value of f_V could shift the emission into the un-observable soft X-ray and ultraviolet bands. Together the evolution of the 2–8-keV and Brγ luminosities, as well as the Brγ velocity distribution, will strongly constrain the thermal states and interaction of the cloud and the ambient hot gas in the at present unprobed regime of $10^3 R_S - 10^4 R_S$, when compared with test particle calculations and more detailed numerical simulations (Fig. 3 and Supplementary Fig. 4).

The radiated energy estimated above is less than 1% of the total kinetic energy of the cloud, about $10^{45.4}$ erg. As the tidally disrupted filamentary cloud passes near pericentre some fraction of the gas may well collide with itself, dissipate and circularize[22]. This is probable because of the large velocity dispersion of the cloud, its size comparable to the impact parameter and because the Rayleigh–Taylor and Kelvin–Helmholtz timescales are similar to the orbital timescale. Because the mass of the cloud is larger than the mass of hot gas within the central $3{,}100 R_S$ or so (approximately $10^{27.3}$ g; ref. 15), it is plausible that then the accretion near the event horizon will be temporarily dominated by accretion of the cloud. This could in principle release up to around 10^{48} erg over the next decade, although the radiative efficiency of the inflow at these accretion rates is of the order of 1–10% (refs 23 and 24). Observations of the emission across the electromagnetic spectrum during this post-circularization phase will provide stringent constraints on the physics of black-hole accretion with unusually good knowledge of the mass available.

What was the origin of the low-angular-momentum cloud? Its orbital angular momentum vector is within 15° of the so-called 'clockwise' disk of young, massive O and Wolf–Rayet stars at radii of about 1″ to 10″ from Sgr A* (refs 3 and 25). Several of these stars have powerful winds. One star, IRS16SW, about 1.4″ southeast of Sgr A*, is a massive Wolf–Rayet contact binary[26]. Colliding winds in the stellar disk, and especially

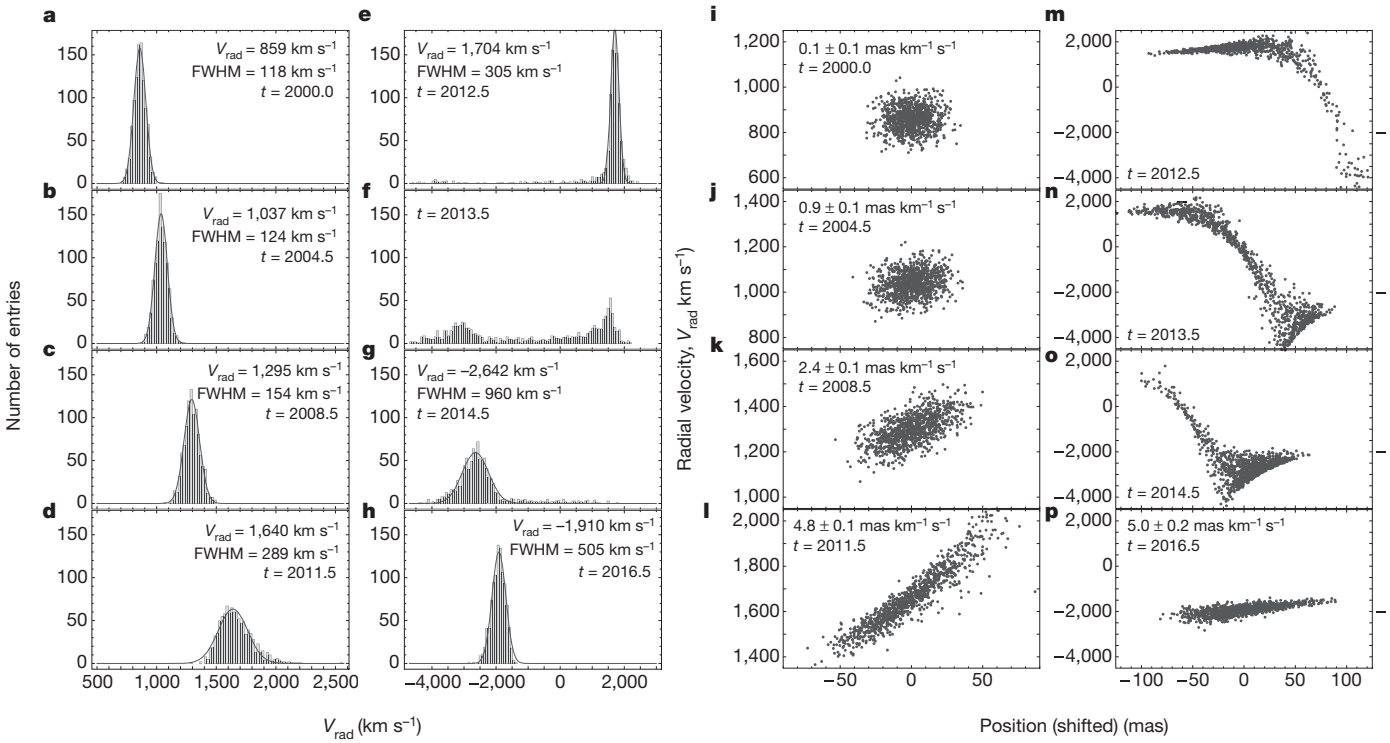


Figure 3 | Test particle simulation of the orbital tidal disruption. An initially Gaussian cloud of initial FWHM diameter 25 mas and FWHM velocity width 120 km s^{-1} is placed on the orbit in Table 1. **a–h**, As described in Supplementary Information section 4 (see also Supplementary Fig. 4) these panels show the evolution of the cloud integrated velocity width for eight epochs. The mean velocity and FWHM are given for those epochs at which the distribution is unimodal (that is, not **f**). **i–p**, These panels show the evolution of the velocity change as a function of position (measured in milli-arcseconds) along the orbital direction, purely on the basis of the tidal disruption of the cloud by the gravitational force of the supermassive black hole. This toy model is a good description of the velocity (and spatial) data between 2004 and 2011, thus allowing plausible forward projections until pericentre passage. Beyond that, the test particle approach will probably fail owing to the hydrodynamic effects, which will then most probably dominate further evolution. The cloud cannot be described by a simple velocity gradient between 2012 and 2015, and hence we omit the corresponding value in panels **m**, **n** and **o**.

in binaries, may create low-angular-momentum gas that then falls deep into the potential of the supermassive black hole[27,28].

Received 25 August; accepted 17 October 2011.
Published online 14 December 2011.

1. Ghez, A. M. *et al.* Measuring distance and properties of the Milky Way's central supermassive black hole with stellar orbits. *Astrophys. J.* **689,** 1044–1062 (2008).
2. Gillessen, S. *et al.* Monitoring stellar orbits around the massive black hole in the Galactic Center. *Astrophys. J.* **692,** 1075–1109 (2009).
3. Genzel, R., Eisenhauer, F. & Gillessen, S. The Galactic Center massive black hole and nuclear star cluster. *Rev. Mod. Phys.* **82,** 3121–3195 (2010).
4. Reid, M. J., Menten, K. M., Trippe, S., Ott, T. & Genzel, R. The position of Sagittarius A*. III. Motion of the stellar cusp. *Astrophys. J.* **659,** 378–388 (2007).
5. Doeleman, S. S. *et al.* Event-horizon-scale structure in the supermassive black hole candidate at the Galactic Centre. *Nature* **455,** 78–80 (2008).
6. Baganoff, F. K. *et al.* Rapid X-ray flaring from the direction of the supermassive black hole at the Galactic Centre. *Nature* **413,** 45–48 (2001).
7. Genzel, R. *et al.* Near-infrared flares from accreting gas around the supermassive black hole at the Galactic Centre. *Nature* **425,** 934–937 (2003).
8. Marrone, D. P., Moran, J. M., Zhao, J.-H. & Rao, R. An unambiguous detection of Faraday rotation in Sagittarius A*. *Astrophys. J.* **654,** L57–L60 (2007).
9. Lenzen, R., Hofmann, R., Bizenberger, P. & Tusche, A. CONICA: the high-resolution near-infrared camera for the ESO VLT. *Proc. SPIE* (ed. Fowler, A. M.) **3354,** 606–614 (1998).
10. Eisenhauer, F. *et al.* SINFONI—Integral field spectroscopy at 50 milli-arcsecond resolution with the ESO VLT. *Proc. SPIE* (eds Iye, M. & Moorwood, A. F. M.) **4841,** 1548–1561 (2003).
11. Bonnet, H. *et al.* Implementation of MACAO for SINFONI at the Cassegrain focus of VLT, in NGS and LGS modes. *Proc. SPIE* (ed. Wizinowich, P.) **4839,** 329–343 (2003).
12. Zhao, J. H., Morris, M. M., Goss, W. M. & An, T. Dynamics of ionized gas at the Galactic Center: Very Large Array observations of the three-dimensional velocity field and location of the ionized streams in Sagittarius A West. *Astrophys. J.* **699,** 186–214 (2009).
13. Martins, F. *et al.* Stellar and wind properties of massive stars in the central parsec of the Galaxy. *Astron. Astrophys.* **468,** 233–254 (2007).
14. Scoville, N. Z., Stolovy, S. R., Rieke, M., Christopher, M. & Yusef-Zadeh, F. Hubble Space Telescope Paα and 1.9 micron imaging of Sagittarius A West. *Astrophys. J.* **594,** 294–311 (2003).
15. Xu, Y.-D., Narayan, R., Quataert, E., Yuan, F. & Baganoff, F. K. Thermal X-ray iron line emission from the Galactic Center black hole Sagittarius A*. *Astrophys. J.* **640,** 319–326 (2006).
16. Yuan, F., Quataert, E. & Narajan, R. Nonthermal electrons in radiatively inefficient accretion flow models of Sagittarius A*. *Astrophys. J.* **598,** 301–312 (2003).
17. Klein, R. I., McKee, C. F. & Colella, P. On the hydrodynamic interaction of shock waves with interstellar clouds. 1: Nonradiative shocks in small clouds. *Astrophys. J.* **420,** 213–236 (1994).
18. Chandrasekhar, S. *Hydrodynamic and Hydromagnetic Stability* 428–514 (Dover Publications, 1961).
19. Murray, S. D., &. Lin, D. N. C. Energy dissipation in multiphase infalling clouds in galaxy halos. *Astrophys. J.* **615,** 586–594 (2004).
20. Cooper, J. L., Bicknell, G. V., Sutherland, R. S. & Bland-Hawthorn, J. Starburst-driven galactic winds: filament formation and emission processes. *Astrophys. J.* **703,** 330–347 (2009).
21. Baganoff, F. K. *et al.* Chandra X-ray spectroscopic imaging of Sgr A* and the central parsec of the Galaxy. *Astrophys. J.* **591,** 891–915 (2003).
22. Sanders, R. H. The circumnuclear material in the Galactic Centre—a clue to the accretion process. *Mon. Not. R. Astron. Soc.* **294,** 35–46 (1998).
23. Sharma, P. Quataert, E. Hammett, G. W. & Stone, J. M. Electron heating in hot accretion flows. *Astrophys. J.* **667,** 714–723 (2007).
24. Blandford, R. D. & Begelman, M. C. On the fate of gas accreting at a low rate on to a black hole. *Mon. Not. R. Astron. Soc.* **303,** L1–L5 (1999).
25. Bartko, H. *et al.* Evidence for warped disks of young stars in the Galactic Center. *Astrophys. J.* **697,** 1741–1763 (2009).
26. Martins, F. *et al.* GCIRS 16SW: a massive eclipsing binary in the Galactic Center. *Astrophys. J.* **649,** L103–L106 (2006).
27. Ozernoy, L. M., Genzel, R. & Usov, V. V. Colliding winds in the stellar core at the Galactic Centre: some implications. *Mon. Not. R. Astron. Soc.* **288,** 237–244 (1997).
28. Cuadra, J., Nayakshin, S., Springel, V. & Di Matteo, T. Accretion of cool stellar winds on to Sgr A*: another puzzle of the Galactic Centre? *Mon. Not. R. Astron. Soc.* **260,** L55–L59 (2005).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This paper is based on observations at the Very Large Telescope (VLT) of the European Observatory (ESO) in Chile. We thank C. McKee and R. Klein for discussions on the cloud destruction process. J.C. acknowledges support from FONDAP, FONDECYT, Basal and VRI-PUC. A.B. acknowledges the support of the excellence cluster 'Origin and Structure of the Universe'.

Author Contributions S.G. collected and analysed the data and discovered the orbit of the gas cloud. R.G. and S.G. wrote the paper. T.K.F. detected the high proper motion and extracted the astrometric positions and the photometry. R.G., A.B. and E.Q. derived the cloud's properties, its evolution and the estimate of the X-ray luminosity. R.G., E.Q., A.B. and C.F.G. contributed to the analytical estimates. C.A. and J.C. set up numerical simulations to check the analysis. F.E., O.P. and K.D.-E. helped in the data analysis and interpretation. T.O. provided valuable software tools.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.G. (ste@mpe.mpg.de) or R.G. (genzel@mpe.mpg.de).

doi:10.1038/nature10612

Fast core rotation in red-giant stars as revealed by gravity-dominated mixed modes

Paul G. Beck[1], Josefina Montalban[2], Thomas Kallinger[1,3], Joris De Ridder[1], Conny Aerts[1,4], Rafael A. García[5], Saskia Hekker[6,7], Marc-Antoine Dupret[2], Benoit Mosser[8], Patrick Eggenberger[9], Dennis Stello[10], Yvonne Elsworth[7], Søren Frandsen[11], Fabien Carrier[1], Michel Hillen[1], Michael Gruberbauer[12], Jørgen Christensen-Dalsgaard[11], Andrea Miglio[7], Marica Valentini[2], Timothy R. Bedding[10], Hans Kjeldsen[11], Forrest R. Girouard[13], Jennifer R. Hall[13] & Khadeejah A. Ibrahim[13]

When the core hydrogen is exhausted during stellar evolution, the central region of a star contracts and the outer envelope expands and cools, giving rise to a red giant. Convection takes place over much of the star's radius. Conservation of angular momentum requires that the cores of these stars rotate faster than their envelopes; indirect evidence supports this[1,2]. Information about the angular-momentum distribution is inaccessible to direct observations, but it can be extracted from the effect of rotation on oscillation modes that probe the stellar interior. Here we report an increasing rotation rate from the surface of the star to the stellar core in the interiors of red giants, obtained using the rotational frequency splitting of recently detected 'mixed modes'[3,4]. By comparison with theoretical stellar models, we conclude that the core must rotate at least ten times faster than the surface. This observational result confirms the theoretical prediction of a steep gradient in the rotation profile towards the deep stellar interior[1,5,6].

The asteroseismic approach to studying stellar interiors exploits information from oscillation modes of different radial order n and angular degree l, which propagate in cavities extending at different depths[7]. Stellar rotation lifts the degeneracy of non-radial modes, producing a multiplet of $(2l+1)$ frequency peaks in the power spectrum for each mode. The frequency separation between two mode components of a multiplet is related to the angular velocity and to the properties of the mode in its propagation region. More information on the exploitation of rotational splitting of modes may be found in the Supplementary Information. An important new tool comes from mixed modes that were recently identified in red giants[3,4]. Stochastically excited solar-like oscillations in evolved G and K giant stars[8] have been well studied in terms of theory[9–12], and the main results are consistent with recent observations from space-based photometry[13,14]. Whereas pressure modes are completely trapped in the outer acoustic cavity, mixed modes also probe the central regions and carry additional information from the core region, which is probed by gravity modes. Mixed dipole modes ($l=1$) appear in the Fourier power spectrum as dense clusters of modes around those that are best trapped in the acoustic cavity. These clusters, the components of which contain varying amounts of influence from pressure and gravity modes, are referred to as 'dipole forests'.

We present the Fourier spectra of the brightness variations of stars KIC 8366239 (Fig. 1a), KIC 5356201 (Supplementary Fig. 3a) and KIC 12008916 (Supplementary Fig. 5a), derived from observations with the Kepler spacecraft. The three spectra show split modes, the spherical degree of which we identify as $l=1$. These detected multiplets cannot have been caused by finite mode lifetime effects from mode damping, because that would not lead to a consistent multiplet appearance over several orders such as that shown in Fig. 1. The spacings in period between the multiplet components (Supplementary Fig. 7) are too small to be attributable to consecutive unsplit mixed modes[4] and do not follow the characteristic frequency pattern of unsplit mixed modes[3]. Finally, the projected surface velocity, $v \sin i$, obtained from ground-based spectroscopy (Table 1), is consistent with the rotational velocity measured from the frequency splitting of the mixed mode that predominantly probes the outer layers. We are thus left with rotation as the only cause of the detected splittings.

The observed rotational splitting is not constant for consecutive dipole modes, even within a given dipole forest (Fig. 1b and Supplementary Figs 3b and 5b). The lowest splitting is generally present for the mode at the centre of the dipole forest, which is the mode with the largest amplitude in the outer layers. Splitting increases for modes with a larger gravity component, towards the wings of the dipole mode forest. For KIC 8366239, we find that the average splitting of modes in the wings of the dipole forests is 1.5 times larger than the mean splitting of the centre modes of the dipole forests.

We compared the observations (Fig. 1b) with theoretical predictions for a model representative of KIC 8366239, as defined in the Supplementary Information. The effect of rotation on the oscillation frequencies can be estimated in terms of a weight function, called a rotational kernel (K_{nl}). From the kernels, it is shown that at least 60% of the frequency splitting for the $l=1$ mixed modes with a dominant gravity component is produced in the central region of the star (Fig. 2). This substantial core contribution to mixed modes enables us to investigate the rotational properties of the core region, which was hitherto not possible for the Sun, owing to a lack of observed modes that probe the core region (within a radius $r < 0.2\,R_{\odot}$; ref. 15). The solar rotational profile is known in great detail for only those regions probed by pressure-dominated modes[16–18]. In contrast to these modes in the wings of the dipole forest, only 30% of the splitting of the centre mode originates from the central region of the star, whereas the outer third by mass of the star contributes 50% of the frequency splitting. By comparing the rotational velocity derived from the splitting of such pressure-dominated modes with the projected surface velocity from spectroscopy, we find that the asteroseismic value is systematically larger. This offset cannot be explained by inclination of the rotation axis towards the observer alone (Supplementary Tables 1 and 2), but originates from the contribution of the fast-rotating core (Fig. 2). Furthermore, internal non-rigid rotation leads to a larger splitting for modes in the wings of the dipole forest than for centre modes.

[1]Instituut voor Sterrenkunde, Katholieke Universiteit Leuven, 3001 Leuven, Belgium. [2]Institut d'Astrophysique et de Géophysique de l'Université de Liège, 4000 Liège, Belgium. [3]Institut für Astronomie der Universität Wien, Türkenschanzstraße 17, 1180 Wien, Austria. [4]Afdeling Sterrenkunde, Institute for Mathematics Astrophysics and Particle Physics (IMAPP), Radboud University Nijmegen, 6500GL Nijmegen, The Netherlands. [5]Laboratoire Astrophysique, Instrumentation et Modélisation (AIM), CEA/DSM—CNRS—Université Paris Diderot; Institut de Recherche sur les lois Fondamentales de l'Univers/Service d'Astrophysique (IRFU/Sap), Centre de Saclay, 91191 Gif-sur-Yvette Cedex, France. [6]Astronomical Institute 'Anton Pannekoek', University of Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands. [7]School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK. [8]Laboratoire d'études spatiales et d'instrumentation (LESIA), CNRS, Université Pierre et Marie Curie, Université Denis Diderot, Observatoire de Paris, 92195 Meudon Cedex, France. [9]Observatoire de Genève, Université de Genève, 51 Ch. des Maillettes, 1290 Sauverny, Switzerland. [10]Sydney Institute for Astronomy (SIfA), School of Physics, University of Sydney 2006, Australia. [11]Department of Physics and Astronomy, Aarhus University, DK-8000 Aarhus C, Denmark. [12]Department of Astronomy and Physics, Saint Marys University, Halifax, NS B3H 3C3, Canada. [13]Orbital Sciences Corporation/NASA Ames Research Center, Moffett Field, 94035 California, USA.

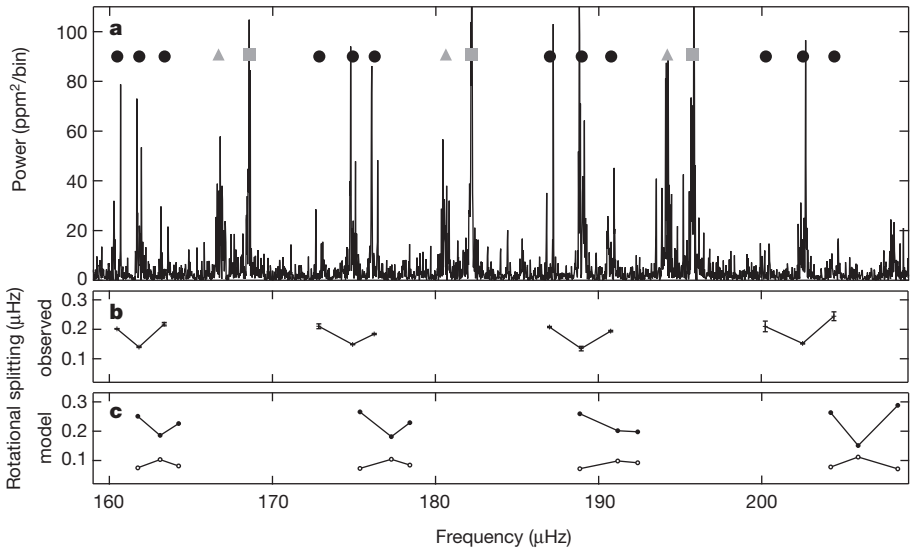


Figure 1 | Oscillation spectrum of KIC 8366239. **a**, Grey squares, radial modes; grey triangles, $l = 2$ modes; black circles, $l = 1$ rotational multiplets. A zoom on the region of 185–195 μHz and the analysis of the comb-like structure of the oscillation spectrum in an échelle diagram are shown in the online material (Supplementary Figs 1 and 2). The spectral window of the Fourier analysis can be found in the online material (Supplementary Fig. 11). The *y*-axis indicates the flux variation power in parts-per-million squared for each frequency bin. **b**, The observed rotational splitting for individual $l = 1$ modes. Error bars, standard deviation of the measured rotational splitting of dipole modes. Similar analyses of the stars KIC 5356201 (Supplementary Figs 3 and 4) and KIC 12008916 (Supplementary Figs 5 and 6) are discussed in the online material. **c**, Theoretically predicted rotational splitting assuming two different rotation laws. The values are calculated for a representative model of KIC 8366239 as defined in the Supplementary Information. Solid circles, splitting for non-rigid rotation for the case of a core rotation ten times faster than the surface rotation of 2.5 km s^{-1} resembles the observations qualitatively well. Open circles, theoretical splittings for rigid rotation and an equatorial surface rotation velocity of 3 km s^{-1} show a trend opposite to the observed one, with the largest splitting in the centre of the dipole forest and lower splitting in gravity-dominated modes. In the case of rigid rotation, the variable splitting is governed purely by the variation of the Ledoux[24] constant across the dipole forest (Supplementary Fig. 8). Because the representative model (**c**) has not been corrected for surface effects, there is a slight offset to the observations (**b**).

For a model rotating rigidly, the reverse behaviour is expected (Fig. 1c). Because the prediction for the rigidly rotating model is incompatible with the observed trends of the splittings (Fig. 1b) but can be well reproduced qualitatively under the assumption of non-rigid rotation (Fig. 1c), we conclude that the three stars investigated here (see Table 1, last column) rotate non-rigidly, with the central region rotating much faster than the surface.

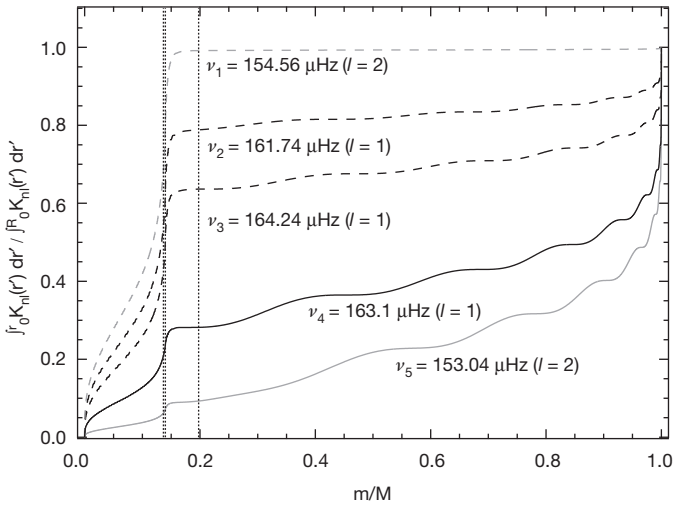


Figure 2 | Contributions to the total rotational splitting. Partial integrals of normalized rotation kernels, illustrating the contribution from different regions to the rotational splitting for pressure-dominated modes (ν_4, ν_5; solid lines) and gravity-dominated modes (ν_1, ν_2, ν_3; dashed lines) of degree $l = 1$ and $l = 2$, as a function of the stellar mass-fraction, m/M. The kernels have been calculated for modes from a representative model of KIC 8366239, as defined in the Supplementary Information, with oscillation frequencies given. Vertical dotted lines, left to right: boundary of the helium core, the hydrogen-burning shell and the bottom of the convective envelope.

The above interpretation is consistent with the correlation between the mode lifetime and the corresponding rotational splitting that has been observed in our data. Mixed modes in the wings of the dipole forest are predicted to have large amplitudes in the central regions of the star and, therefore, larger values of inertia and lifetime. These modes have narrower mode profiles in the frequency spectrum than have the centre modes that are predominantly trapped in the outer cavity[11]. This behaviour of the mode profiles (Supplementary Fig. 1)

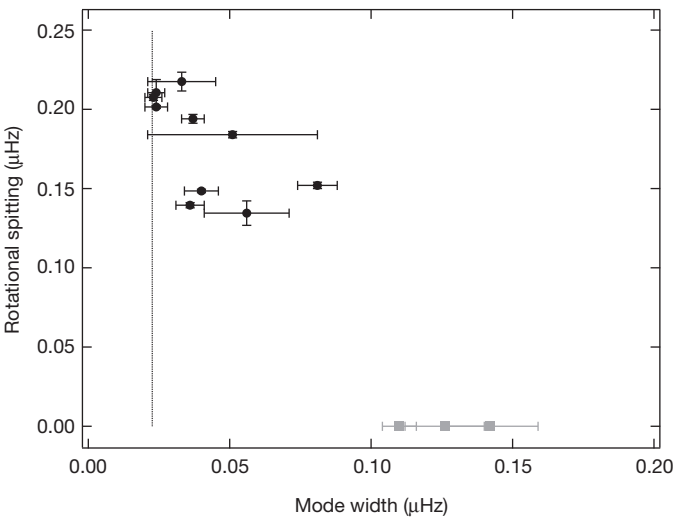


Figure 3 | Rotational splitting versus mode linewidth for KIC 8366239. The lifetime is inversely proportional to the mode linewidth. Black circles, $l = 1$ modes. Grey squares, linewidth of the pure acoustic radial modes ($l = 0$), for comparison. Dotted vertical line, formal frequency resolution. Error bars, standard deviation of the measured rotational splitting and mode width of dipole modes. Similar diagrams for the two other stars from Table 1 are shown in Supplementary Figs 9 and 10, respectively.

Table 1 | Observational parameters of three stars, showing rotational splitting.

KIC	ν_{max} (μHz)	R ($R_\odot$)	ΔP_{obs} (s)	T_{eff} (K)	$\nu \sin i$ (km s^{-1})	Asteros. rot. vel. (km s^{-1})	$l = 1$ min rot. split. (μHz)	Averaged split. ratio for dipole modes
8366239	182 ± 1	5.30 ± 0.08	56 ± 11	4980 ± 120	<1	3.9	0.135 ± 0.008	1.5
5356201	209.7 ± 0.7	4.47 ± 0.03	50 ± 10	4840 ± 90	2.4	3.8	0.154 ± 0.003	1.7
12008916	159.9 ± 0.6	5.18 ± 0.05	52 ± 7	4830 ± 100	-	5.7	0.200 ± 0.001	1.8

We present three cases of firm detections of rotational splitting in red giants from around 500 day-long time series of photometric data obtained with the Kepler satellite in long-cadence mode (about 30-min time sampling)[23]. KIC, stars' identifiers in the Kepler Input Catalog. ν_{max}, frequency of maximum oscillation power. The stellar radius R and the effective temperature T_{eff} were computed according to ref. 14. The mean period spacing ΔP_{obs} and its standard deviation were derived from the central multiplet components of the dipole modes. These three stars are classified as low-luminosity red-giant stars[22]. The observationally derived mean period spacing of the mixed modes indicates that they are in an early phase of red-giant evolution[4], in which they are burning hydrogen in a shell around the helium core. The projected surface velocity, $\nu \sin i$, was determined from ground-based spectra. The asteroseismic rotation velocity was computed from the minimum dipole splitting and assumed rigid rotation. Although this is the most pressure-dominated mode found in the spectrum, it contains a contribution from the core. The inferred rotation velocity from this mode is therefore slightly higher than the surface rotation velocity. The observed ratio of the average rotational splitting found in the wings of the dipole forest to the mean splitting of the forest centre is given in the last column. Extensive lists of further asteroseismic parameters are given in Supplementary Table 1 and 2.

was recently confirmed by observations[3,4]. We observe that short-lived modes (with broader profiles) exhibit smaller rotational splitting in KIC 8366239 (Fig. 3) and the two other stars from Table 1 (Supplementary Figs 9 and 10). With increasing lifetime (dominant gravity component) we see a substantial increase in the size of the rotational splitting of the modes. The frequencies of the narrowest modes are mainly affected by the rotation in the central regions of the star. Taking a representative model for KIC 8366239, and assuming that the convective envelope rotates rigidly and that the radiative interior rotates rigidly but faster than the convective envelope, we find that the observed splitting ratio of 1.5 (Table 1) is obtained from a core rotating at least ten times faster than the surface.

So far, our understanding of the evolution of the angular momentum as a function of the evolutionary stage is still poor. Before this study, it was possible to deliver evidence of non-rigid rotation in only two core-hydrogen-burning massive stars, for which the seismic analysis revealed a near-core rotation rate about three to five times faster than the envelope rotation rate[19,20]. Claims of non-rigid rotation in white dwarfs have also been discussed[21]. The low-luminosity red giants[22] discussed here are in the early phase of hydrogen-shell burning, and have a core-to-envelope rotation rate more than twice as high as those of the massive stars[19,20]. When these stars enter the subsequent phase of core-helium burning, we believe that the core will undergo a slight expansion and the envelope will shrink, leading to a surface rotation on the horizontal branch larger than that on the red-giant branch[1]. The basic assumption of conservation of angular momentum predicts a rotational gradient between surface and core that is less steep than the one detected. Exploiting rotational splitting of mixed modes in a large sample of red giants in various evolutionary stages will provide an excellent tool to inspect how the internal angular-momentum distribution evolves with time towards the end of the life of the star.

Received 29 July; accepted 29 September 2011.
Published online 7 December 2011.

1. Sills, A. & Pinsonneault, M. H. Rotation of horizontal-branch stars in globular clusters. *Astrophys. J.* **540,** 489–503 (2000).
2. Sweigart, A. V. & Mengel, J. G. Meridional circulation and CNO anomalies in red giant stars. *Astrophys. J.* **229,** 624–641 (1979).
3. Beck, P. G. *et al.* Kepler detected gravity-mode period spacings in a red giant star. *Science* **332,** 205 (2011).
4. Bedding, T. R. *et al.* Gravity modes as a way to distinguish between hydrogen- and helium-burning red giant stars. *Nature* **471,** 608–611 (2011).
5. Zahn, J.-P. Circulation and turbulence in rotating stars. *Astron. Astrophys.* **265,** 115–132 (1992).
6. Eggenberger, P. *et al.* Effects of rotation on the evolution and asteroseismic properties of red giants. *Astron. Astrophys.* **509,** A72 (2010).
7. Aerts, C., Christensen-Dalsgaard, J. & Kurtz, D. W. *Asteroseismology* Ch. 3 (Springer-Verlag, 2010).
8. De Ridder, J. *et al.* Non-radial oscillation modes with long lifetimes in giant stars. *Nature* **459,** 398–400 (2009).
9. Dziembowski, W. A. *et al.* Oscillations of α UMa and other red giants. *Mon. Not. R. Astron. Soc.* **328,** 601–610 (2001).
10. Christensen-Dalsgaard, J. Physics of solar-like oscillations. *Sol. Phys.* **220,** 137–168 (2004).
11. Dupret, M.-A. *et al.* Theoretical amplitudes and lifetimes of non-radial solar-like oscillations in red giants. *Astron. Astrophys.* **506,** 57–67 (2009).
12. Montalbán, J., Miglio, A., Noels, A., Scuflaire, R. & Ventura, P. Seismic diagnostics of red giants: first comparison with stellar models. *Astrophys. J.* **721,** L182–L188 (2010).
13. Mosser, B. *et al.* The universal red-giant oscillation pattern. An automated determination with CoRoT data. *Astron. Astrophys.* **525,** L9 (2011).
14. Kallinger, T. *et al.* Asteroseismology of red giants from the first four months of *Kepler* data: fundamental stellar parameters. *Astron. Astrophys.* **522,** A1 (2010).
15. García, R. A. *et al.* Tracking solar gravity modes: the dynamics of the solar core. *Science* **316,** 1591–1593 (2007).
16. Elsworth, Y. *et al.* Slow rotation of the Sun's interior. *Nature* **376,** 669–672 (1995).
17. Chaplin, W. J. *et al.* Rotation of the solar core from BiSON and LOWL frequency observations. *Mon. Not. R. Astron. Soc.* **308,** 405–414 (1999).
18. Thompson, M. J. *et al.* The internal rotation of the Sun. *Annu. Rev. Astron. Astrophys.* **41,** 599–643 (2003).
19. Aerts, C. *et al.* Asteroseismology of HD 129929: core overshooting and nonrigid rotation. *Science* **300,** 1926–1928 (2003).
20. Pamyatnykh, A. A. *et al.* Asteroseismology of the β Cephei star ν Eridani: interpretation and applications of the oscillation spectrum. *Mon. Not. R. Astron. Soc.* **350,** 1022–1028 (2004).
21. Kawaler, S. D., Sekii, T. & Gough, D. Prospects for measuring differential rotation in white dwarfs through asteroseismology. *Astrophys. J.* **516,** 349–365 (1999).
22. Bedding, T. R. *et al.* Solar-like oscillations in low-luminosity red giants: first results from *Kepler*. *Astrophys. J. Lett.* **713,** L176–L181 (2010).
23. García, R. A. *et al.* Preparation of *Kepler* light curves for asteroseismic analyses. *Mon. Not. R. Astron. Soc.* **414,** L6–L10 (2011).
24. Ledoux, P. The nonradial oscillations of gaseous stars and the problem of beta Canis Majoris. *Astrophys. J.* **114,** 373–384 (1951).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We acknowledge the work of the team behind Kepler. Funding for the Kepler Mission is provided by NASA's Science Mission Directorate. P.G.B. and C.A. were supported by the European Community's Seventh Framework Programme (ERC grant PROSPERITY); J.D.R. and T.K. were supported by the Fund for Scientific Research, Flanders. S.H. was supported by the Netherlands Organisation for Scientific Research. J.M. and M.V. were supported by the Belgian Science Policy Office. The work is partly based on observations with the High Efficiency and Resolution Mercator Echelle Spectrograph at the Mercator Telescope, which is operated at La Palma in Spain by the Flemish Community.

Author Contributions P.G.B., T.K., J.D.R., C.A., R.A.G., S.H., B.M., Y.E., S.F., F.C. and M.G. measured the mode parameters, and derived and interpreted the rotational splitting and period spacings. J.M., M.-A.D., P.E., J.C.-D. and A.M. calculated stellar models and provided theoretical interpretation of the rotational splitting. M.H. and M.V. observed and analysed the spectra. J.D.R., S.H., S.F., Y.E., D.S., T.R.B., H.K., F.R.G., J.R.H. and K.A.I. contributed to the coordination of the project, including the acquisition and distribution of the data. C.A. defined and supervised the research. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.G.B. (paul.beck@ster.kuleuven.be).

doi:10.1038/nature10666

Polar methane accumulation and rainstorms on Titan from simulations of the methane cycle

T. Schneider[1], S. D. B. Graves[1], E. L. Schaller[2] & M. E. Brown[1]

Titan has a methane cycle akin to Earth's water cycle. It has lakes in polar regions[1,2], preferentially in the north[3]; dry low latitudes with fluvial features[4,5] and occasional rainstorms[6,7]; and tropospheric clouds mainly (so far) in southern middle latitudes and polar regions[8–15]. Previous models have explained the low-latitude dryness as a result of atmospheric methane transport into middle and high latitudes[16]. Hitherto, no model has explained why lakes are found only in polar regions and preferentially in the north; how low-latitude rainstorms arise; or why clouds cluster in southern middle and high latitudes. Here we report simulations with a three-dimensional atmospheric model coupled to a dynamic surface reservoir of methane. We find that methane is cold-trapped and accumulates in polar regions, preferentially in the north because the northern summer, at aphelion, is longer and has greater net precipitation than the southern summer. The net precipitation in polar regions is balanced in the annual mean by slow along-surface methane transport towards mid-latitudes, and subsequent evaporation. In low latitudes, rare but intense storms occur around the equinoxes, producing enough precipitation to carve surface features. Tropospheric clouds form primarily in middle and high latitudes of the summer hemisphere, which until recently has been the southern hemisphere. We predict that in the northern polar region, prominent clouds will form within about two (Earth) years and lake levels will rise over the next fifteen years.

Explanations for Titan's tropospheric clouds range from control by local topography and cryovolcanism[11,12] to control by a seasonally varying global Hadley circulation with methane condensation in its ascending branch[8,17]. General circulation models (GCMs) have suggested that clouds either form primarily in mid-latitudes and near the poles in both hemispheres[18]—which is inconsistent with the observed hemispheric differences[14]—or that they form where insolation is at a maximum[17], which is likewise not fully consistent with newer observations[14,15]. Similarly, explanations for the polar hydrocarbon lakes range from control by local topography and a subsurface methane table[2] to control by evaporation and precipitation, which depend on the atmospheric circulation[16,19]. GCMs have suggested that surface methane accumulates near the poles, but also show it in mid-latitudes[16], where no lakes have been observed, and they have failed to reproduce the observed hemispheric asymmetry. Additionally, no model is fully consistent with the cloud distribution or has shown low-latitude precipitation that is intense enough to carve fluvial features. Indeed, it has been suggested that the atmosphere is too stable for rainstorms to occur in low latitudes[20]. Yet intense and, in at least one case, apparently precipitating storms have been observed[6,7].

To investigate the extent to which major features of Titan's climate and methane cycle can be explained by large-scale processes, we use a GCM that includes an atmospheric model with a methane cycle and surface reservoir. The atmospheric model is three-dimensional (3D), in contrast to previous two-dimensional (2D) models[16–18], because the intermittency of clouds and features such as equatorial super-rotation[21] (impossible in axisymmetric circulations[22]) point to the importance of 3D dynamics. The surface reservoir gains or loses methane according to the local rates of precipitation (P) and evaporation (E), with horizontal diffusion as a simple representation of slow surface flows from moister to drier regions. We show zonal and temporal averages from a simulation in a statistically steady state, which does not depend on initial conditions except for the (conserved) total methane amount present in the atmosphere-surface system (here, the equivalent of 12 m liquid methane in the global mean if all methane were condensed at the surface). See Supplementary Information for details.

In the statistically steady state, our model shows an annual- and global-mean methane concentration in the atmosphere corresponding to 7 m liquid methane—possibly higher than, but broadly consistent with, observations[23,24]. The remainder is at the surface. Methane has accumulated near the poles (Fig. 1a). It is transported there from spring into summer by a global Hadley circulation (Fig. 1b), with ascent (Fig. 2) and a precipitation maximum (Fig. 1c) over the summer pole. Some of the methane accumulating near the summer pole flows along the surface (on Titan it may also flow beneath the surface[2]) towards mid-latitudes. It evaporates and is transported back towards the opposite pole when the circulation reverses, around equinox. Polar regions lose methane from late summer to winter, with zonal-mean net evaporation rates ($E - P$) reaching around 0.2 m yr^{-1} (1 yr referring to 1 Earth year) in the southern summer (Fig. 1b). This is consistent with observations: the zonal-mean evaporative loss rate is of similar magnitude to (but smaller than) the recently observed local drop in south-polar lake levels[25]. We predict that the north-polar region will gain methane for roughly the next 15 yr, with zonal-mean net precipitation rates ($P - E$) reaching about 1.4 m yr^{-1} around the northern summer solstice (NSS; Fig. 1b). This should lead to an observable rise in lake levels.

Methane is cold-trapped at the poles. Along with annual-mean insolation, annual-mean evaporation is at its lowest near the poles. (Evaporation is the dominant loss term in the surface energy balance and scales with insolation.) Surface temperatures decrease from low latitudes towards the poles throughout the year (Fig. 1d), because almost all solar energy absorbed at the poles is used to evaporate methane. At the same time, precipitation is highest near the summer pole (Fig. 1c) because the insolation maximum destabilizes the atmosphere with respect to moist convection. This can be seen from the moist static energy (MSE), which is a more direct measure than temperature for the energetic effect of solar radiation on the atmosphere. This is because if, as on Titan and in our GCM, the surface heat capacity is negligible[26], only net radiation at the top of the atmosphere drives the vertically integrated MSE balance; similarly, insolation variations are the dominant driver of variations in the MSE balance integrated over the planetary boundary layer. Indeed, along with insolation, near-surface MSE is highest near the summer poles (Fig. 1e). This implies a propensity for deep convection, because the slow rotation and large thermal inertia of Titan's atmosphere constrain horizontal and temporal variations in temperature and MSE above the boundary layer, keeping them weak[27], and the vertical MSE stratification controls convective stability (MSE increasing with altitude indicates stability; ref. 26). Lower latitudes, by contrast, do not favour deep

[1]California Institute of Technology, Pasadena, California 91125, USA. [2]NASA Dryden Aircraft Operations Facility, National Suborbital Education and Research Center, Palmdale, California 93550, USA.

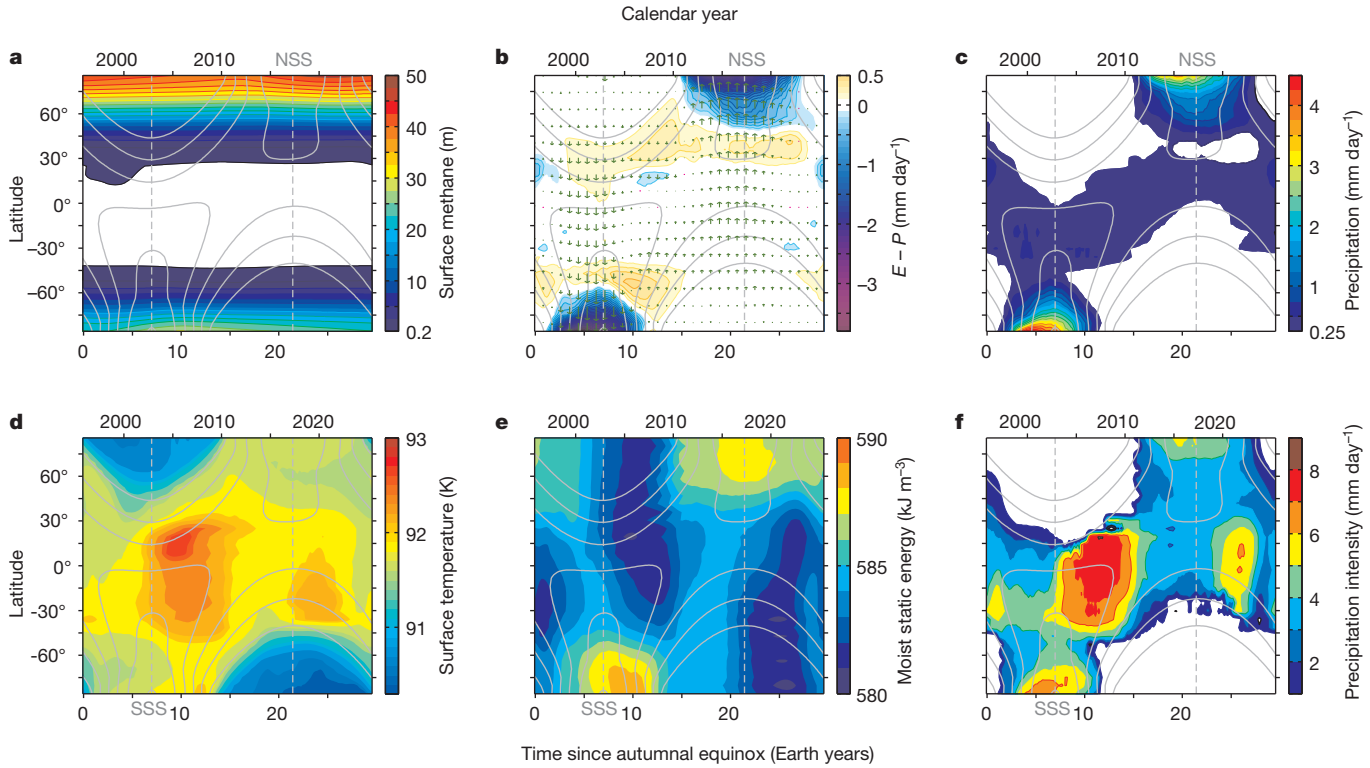


Figure 1 | Annual cycle of zonal-mean climate statistics in Titan GCM. The lower time axes start at the autumnal equinox (corresponding to November 1995); the upper axes indicate corresponding calendar years. Solid grey lines are contours of absorbed solar radiation at the surface (contour interval 0.25 W m^{-2}), and dashed grey lines mark the northern summer solstice (NSS) and southern summer solstice (SSS). **a**, Methane surface-reservoir depth (colour scale truncated at 20 cm). **b**, Net evaporation ($E - P$) at the surface (colours, contour interval 0.1 mm day^{-1}, with 1 day = 86,400 s) and column-integrated meridional methane flux (arrows, with the longest corresponding to a flux of 5.4 kg m^{-1} s^{-1}). The methane flux is largely accomplished by the mean meridional (Hadley) circulation; eddy fluxes are weaker by a factor of 3 or larger, and are strongest in the middle and low latitudes. **c**, Precipitation rate. **d**, Surface temperature. The surface temperatures are roughly consistent with observations[30], with similar equator–pole temperature contrasts and with a winter pole that is around 1 K colder than the summer pole. (Generally, tropospheric temperatures in the simulation are within 1 K of Titan observations; see Supplementary Fig. 2.) **e**, MSE per unit volume averaged between surface and 2-km altitude. MSE is the sum of gravitational potential energy and moist enthalpy, including the contribution of the latent heat of methane vapour[26]. **f**, Precipitation intensity (precipitation rate when it rains). The precipitation intensity is the mean precipitation conditional on the precipitation rate being non-zero (exceeding 10^{-3} mm day^{-1}). All fields are averages over longitude and time (25 Titan yr) in the statistically steady state of a simulation with a total of 12 m liquid methane in the atmosphere–surface system. The statistically steady state was reached over a long (135 Titan yr) spin-up period. Although the GCM climate is statistically zonally symmetric, climate statistics such as the methane surface-reservoir depth exhibit instantaneous zonal asymmetries ('lakes'), which can persist for several Titan years.

convection, even though they are warmer than the summer poles, because their maximum insolation is weaker, so their maximum near-surface MSE is also lower (Fig. 1e). The result is the drying of lower latitudes and the accumulation of methane at the poles, similar to the results of previous models but without methane accumulation in mid-latitudes[16] (see Supplementary Information for possible reasons for this difference).

In the model's low latitudes, rainstorms are rare but intense: that is, although low latitudes are dry and have smaller mean precipitation rates than the poles (Fig. 1c), they have the largest precipitation intensity (Fig. 1f). Precipitation rates in the more intense storms are greater than 10 mm per (Earth) day, similar to rough estimates of the rates needed to carve the observed fluvial surface features[4,28]. The precipitation intensity is largest before and around the equinoxes, when the reversal of the Hadley circulation (see Fig. 1b) is associated with dynamic instabilities. These instabilities generate waves strong enough to trigger deep convection and intense rain if they advect relatively moist air from higher latitudes over the warm low latitudes. This is in contrast to 2D and local models that suggest that intense precipitation does not occur in low latitudes[16–18,20]; however, it is consistent with the observations of intense and apparently precipitating low-latitude storms on Titan[6,7].

In the GCM as on Titan[3], there is more surface methane in the northern polar regions than in the southern (Fig. 1a). This asymmetry is a consistent feature, irrespective of initial conditions (for example, whether methane is initially in the atmosphere or at the surface). It is caused by Saturn's orbital eccentricity, which leads to a northern summer (currently around aphelion) that is longer and cooler than the southern, and therefore allows more methane to be cold-trapped. Although the maximum rate of net precipitation (Fig. 1b) is greatest in the warmer southern summer, polar net precipitation integrated over a Titan year (about 30 Earth years) is greatest in the north, primarily because its rainy season is longer. For example, averaged over the polar caps bounded by the polar circles (63.3° N/S), the period during which absorbed insolation exceeds 0.5 W m^{-2} is 14% longer in the north than in the south. Annually integrated net precipitation is 0.86 m (38%) greater in the north; 0.76 m of this is attributable to increased precipitation, and 0.10 m to decreased evaporation. (The contribution of evaporation to the hemispheric asymmetry is small because it scales with insolation, which, in the annual mean, is equal in the north and south; this contradicts a previous hypothesis that the hemispheric asymmetry is due to evaporation differences[3].) In the model's annual mean, the excess net precipitation in polar regions is balanced by along-surface methane transport towards mid-latitudes and subsequent evaporation, so surface methane extends farther towards the equator in the north than in the south (Fig. 1a). The same is true on Titan[3]. Thus, surface or subsurface transport of methane is essential for maintaining a statistically steady state with non-zero net precipitation in polar regions and asymmetries between the hemispheres; we expect that such transport occurs on Titan. One reason that previous models[16] did not produce hemispheric asymmetries is that they were lacking a representation of this transport.

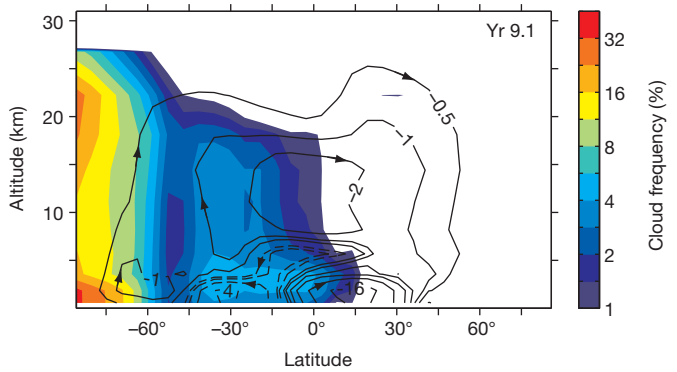


Figure 2 | Relationship between tropospheric cloudiness and atmospheric circulation in southern summer in Titan GCM. Colours show the tropospheric methane-cloud frequency and contours show the stream function of the mean meridional mass circulation, both at 9.1 yr past autumnal equinox (corresponding to January 2005, the time of the Huygens landing on Titan). The contouring is logarithmic, with factors of $\sqrt{2}$ and 2 between adjacent contour levels for cloud frequency and stream function, respectively. Solid stream-function contours for clockwise rotation, dashed contours for anticlockwise rotation (contour levels at $\pm 2^{-1}$, 1, 2, 2^2, $\ldots \times 10^9\,\mathrm{kg\,s^{-1}}$). The cloud frequency is the relative frequency of phase changes of methane on the grid scale and in the convection scheme of the GCM. Moist convection and maxima in cloud frequency above the boundary layer occur above the polar near-surface MSE maximum and above relatively high surface temperatures in mid-latitudes (here, at around 30° S) in the summer hemisphere. (Moist convection rarely occurs above the local surface-temperature maximum in the winter hemisphere (here, at 15°–30° N), because this lies under the descending branch of the Hadley circulation, where the free troposphere is relatively dry.)

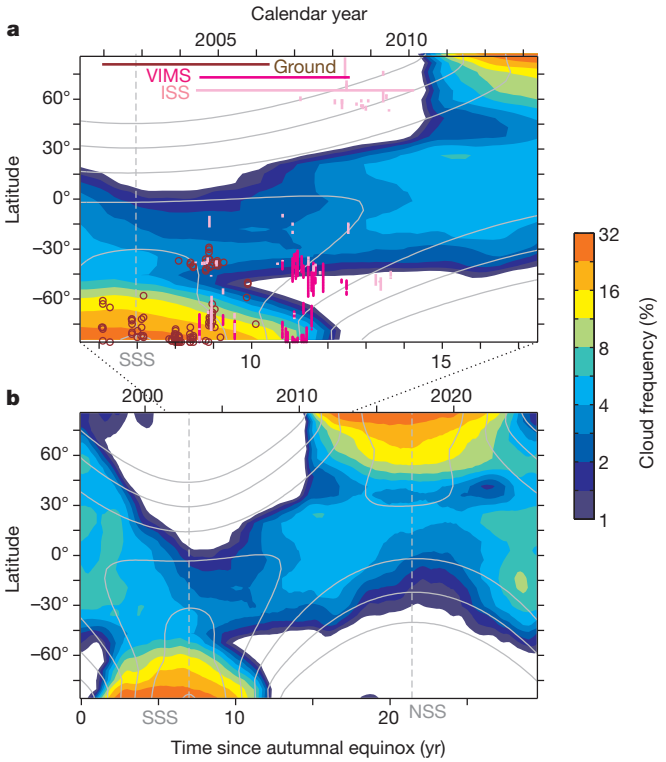


Figure 3 | Annual cycle of tropospheric methane-cloud frequency. **a**, Cloud frequency in GCM and cloud observations, focusing on the time period for which detailed observations are available or will be available soon. The GCM cloud frequency (colours) is a mass-weighted vertical average of the cloud frequency (Fig. 2) in the troposphere between the surface and 32-km altitude. Brown circles, ground-based cloud observations[13]; magenta lines, Cassini VIMS observations[14]; light pink lines, Cassini ISS observations[15]. The bars across the top indicate the time period over which the observations were made. (However, the coverage of VIMS and ISS observations is not continuous, so a lack of VIMS and ISS cloud observations in the periods indicated by the bars does not necessarily mean clouds were in fact absent.) Solid grey lines, contours of absorbed solar radiation at the surface, as in Fig. 1. **b**, As in **a**, except over a complete Titan year, beginning at autumnal equinox, and excluding observations.

Our GCM also reproduces the observed tropospheric methane-cloud distribution (Fig. 3). For the period with detailed observations (2001 onwards), the GCM reproduces the observed prevalence of clouds in the southern hemisphere mid-latitudes and polar region; it also reproduces the decreasing frequency of south-polar clouds since 2005 (Fig. 3a)[6,13–15]. For 2001 onwards, the GCM indicates a lack of clouds in the northern hemisphere, consistent with observations of Titan made with ground-based telescopes[6,13] and the Cassini Visual and Infrared Mapping Spectrometer (VIMS; ref. 14). Observations with the Cassini Imaging Science Subsystem (ISS; ref. 15) indicate more frequent northern-hemisphere clouds (Fig. 3a), but these seem to be lake-effect clouds: they are associated with stationary zonal inhomogeneities in topography and the lake distribution[29] that are not captured in the GCM. The relative frequency of clouds in the GCM fits observations better than previous 2D models[16–18]. A 3D model that resolves waves and instabilities in the atmosphere is essential for reproducing the non-sinusoidal seasonal variations of cloudiness.

Deep convective clouds in the GCM form not only above the polar near-surface MSE maximum, but also in the summer-hemisphere mid-latitudes (Fig. 2). There they form above relatively high surface temperatures (Fig. 1d), which destabilize the boundary layer with respect to dry convection. This occasionally leads to moist convection and mean ascent in the free troposphere (above shallow boundary-layer circulation cells), resulting in a secondary cloud-frequency maximum above the boundary layer (Fig. 2). The surface reservoir underneath these clouds is depleted (less than 7 cm depth in the mean, see Fig. 1a), consistent with the observed lack of lakes in mid-latitudes[2]. We predict that, with the reversal of the Hadley circulation in spring, north-polar clouds will emerge within about 2 yr, earlier than suggested by other models[17], and should be clearly observable for roughly 10 yr; around NSS, prominent cloudiness may extend into mid-latitudes (Fig. 3b). The validity of our predictions of seasonal changes will soon be testable as Titan's northern-hemisphere spring proceeds into summer and new observations become available.

Received 4 August; accepted 20 October 2011.

1. Stofan, E. *et al.* The lakes of Titan. *Nature* **445,** 61–64 (2007).
2. Hayes, A. *et al.* Hydrocarbon lakes on Titan: distribution and interaction with a porous regolith. *Geophys. Res. Lett.* **35,** L09204 (2008).
3. Aharonson, O. *et al.* An asymmetric distribution of lakes on Titan as a possible consequence of orbital forcing. *Nature Geosci.* **2,** 851–854 (2009).
4. Lebreton, J. *et al.* An overview of the descent and landing of the Huygens probe on Titan. *Nature* **438,** 758–764 (2005).
5. Lorenz, R. *et al.* The sand seas of Titan: Cassini RADAR observations of longitudinal dunes. *Science* **312,** 724–727 (2006).
6. Schaller, E. L., Roe, H. G., Schneider, T. & Brown, M. E. Storms in the tropics of Titan. *Nature* **460,** 873–875 (2009).
7. Turtle, E. P. *et al.* Rapid and extensive surface changes near Titan's equator: evidence of April showers. *Science* **331,** 1414–1417 (2011).
8. Brown, M. E., Bouchez, A. H. & Griffith, C. A. Direct detection of variable tropospheric clouds near Titan's south pole. *Nature* **420,** 795–797 (2002).
9. Porco, C. C. *et al.* Imaging of Titan from the Cassini spacecraft. *Nature* **434,** 159–168 (2005).
10. Griffith, C. A. *et al.* The evolution of Titan's mid-latitude clouds. *Science* **310,** 474–477 (2005).
11. Roe, H. G., Bouchez, A. H., Trujillo, C. A., Schaller, E. L. & Brown, M. E. Discovery of temperate latitude clouds on Titan. *Astrophys. J.* **618,** L49–L52 (2005).
12. Roe, H. G., Brown, M. E., Schaller, E. L., Bouchez, A. H. & Trujillo, C. A. Geographic control of Titan's mid-latitude clouds. *Science* **310,** 477–479 (2005).
13. Schaller, E. L., Brown, M. E., Roe, H. G., Bouchez, A. H. & Trujillo, C. A. Dissipation of Titan's south polar clouds. *Icarus* **184,** 517–523 (2006).
14. Brown, M. E., Roberts, J. E. & Schaller, E. L. Clouds on Titan during the Cassini prime mission: A complete analysis of the VIMS data. *Icarus* **205,** 571–580 (2010).
15. Turtle, E. P. *et al.* Seasonal changes in Titan's meteorology. *Geophys. Res. Lett.* **38,** L03203 (2011).
16. Mitchell, J. The drying of Titan's dunes: Titan's methane hydrology and its impact on atmospheric circulation. *J. Geophys. Res.* **113,** E08015 (2008).

17. Mitchell, J. L., Pierrehumbert, R. T., Frierson, D. M. W. & Caballero, R. The dynamics behind Titan's methane clouds. *Proc. Natl Acad. Sci. USA* **103,** 18421–18426 (2006).
18. Rannou, P., Montmessin, F., Hourdin, F. & Lebonnois, S. The latitudinal distribution of clouds on Titan. *Science* **311,** 201–205 (2006).
19. Mitri, G., Showman, A. P., Lunine, J. I. & Lorenz, R. D. Hydrocarbon lakes on Titan. *Icarus* **186,** 385–394 (2007).
20. Griffith, C. A., McKay, C. P. & Ferri, F. Titan's tropical storms in an evolving atmosphere. *Astrophys. J.* **687,** L41–L44 (2008).
21. Flasar, F. M., Baines, K. H., Bird, M. K., Tokano, T. & West, R. A. Atmospheric dynamics and meteorology. In Brown, R. H., Lebreton, J. P. & Waite, J. H. (eds) *Titan from Cassini-Huygens* Chap. 13, 323–352 (Springer, 2009).
22. Schneider, T. The general circulation of the atmosphere. *Annu. Rev. Earth Planet. Sci.* **34,** 655–688 (2006).
23. Penteado, P. F., Griffith, C. A., Greathouse, T. K. & de Bergh, C. Measurements of CH_3D and CH_4 in Titan from infrared spectroscopy. *Astrophys. J.* **629,** L53–L56 (2005).
24. Tokano, T. *et al.* Methane drizzle on Titan. *Nature* **442,** 432–435 (2006).
25. Hayes, A. *et al.* Transient surface liquid in Titan's polar regions from Cassini. *Icarus* **211,** 655–671 (2011).
26. Neelin, J. D. & Held, I. M. Modeling tropical convergence based on the moist static energy budget. *Mon. Weath. Rev.* **115,** 3–12 (1987).
27. Charney, J. G. A note on large-scale motions in the tropics. *J. Atmos. Sci.* **20,** 607–609 (1963).
28. Perron, J. T. *et al.* Valley formation and methane precipitation rates on Titan. *J. Geophys. Res.* **111,** E11001 (2006).
29. Brown, M. E. *et al.* Discovery of lake-effect clouds on Titan. *Geophys. Res. Lett.* **36,** L01103 (2009).
30. Jennings, D. E. *et al.* Titan's surface brightness temperatures. *Astrophys. J.* **691,** L103–L105 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful for support by a NASA Earth and Space Science Fellowship and a David and Lucile Packard Fellowship. We thank I. Eisenman for code for the insolation calculations, and O. Aharonson, A. Hayes and A. Soto for comments on a draft. The simulations were done on the California Institute of Technology's Division of Geological and Planetary Sciences Dell cluster.

Author Contributions T.S. and M.E.B. conceived the study; T.S., S.D.B.G. and E.L.S. developed the GCM; E.L.S. and M.E.B. provided data; and T.S. and S.D.B.G. wrote the paper, with contributions and comments from all authors.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to T.S. (tapio@caltech.edu).

doi:10.1038/nature10695

Demonstration of temporal cloaking

Moti Fridman[1], Alessandro Farsi[1], Yoshitomo Okawachi[1] & Alexander L. Gaeta[1]

Recent research has uncovered a remarkable ability to manipulate and control electromagnetic fields to produce effects such as perfect imaging and spatial cloaking[1,2]. To achieve spatial cloaking, the index of refraction is manipulated to flow light from a probe around an object in such a way that a 'hole' in space is created, and the object remains hidden[3–14]. Alternatively, it may be desirable to cloak the occurrence of an event over a finite time period, and the idea of temporal cloaking has been proposed in which the dispersion of the material is manipulated in time, producing a 'time hole' in the probe beam to hide the occurrence of the event from the observer[15]. This approach is based on accelerating the front part of a probe light beam and slowing down its rear part to create a well controlled temporal gap—inside which an event occurs—such that the probe beam is not modified in any way by the event. The probe beam is then restored to its original form by the reverse manipulation of the dispersion. Here we present an experimental demonstration of temporal cloaking in an optical fibre-based system by applying concepts from the space–time duality between diffraction and dispersive broadening[16]. We characterize the performance of our temporal cloak by detecting the spectral modification of a probe beam due to an optical interaction and show that the amplitude of the event (at the picosecond timescale) is reduced by more than an order of magnitude when the cloak is turned on. These results are a significant step towards the development of full spatio-temporal cloaking.

The detection of an object or an event is often performed by measuring a change in the properties of a light probe that interacts with the object or with elements participating in the event. The idea of spatial cloaking consists of the probe light being bent in a precise fashion to prevent it from being scattered by the object, which thus remains hidden from an observer. This has been done typically through use of exotic materials, such as ones with a negative index of refraction or through sophisticated manipulation of the refractive index[7–9]. Analogously, it could be possible to cloak an event in the time domain from an observer by manipulating the dispersion of a material such that a temporal gap is created in the probe beam: any event that occurs within this gap does not modify the temporal/spectral properties of the probe beam and thus remains undetected[15]. This requires rapid changes in the dispersion and a recently proposed approach[15] involves the use of optical fibres that are pumped to high power levels to produce large changes in the intensity-dependent refractive index. However, at such powers, other optical processes such as stimulated Raman and Brillouin scattering could limit the ability to achieve cloaking. We propose an alternative means of creating the conditions that allow for temporal cloaking in which we apply concepts from the space–time duality associated with diffraction and dispersion[17].

The space–time duality represents the analogy between diffraction and dispersion that arises from the mathematical equivalence between the equations describing the diffraction of a beam of light and the one-dimensional temporal propagation of a pulse through a dispersive medium[17,18]. Similar to a spatial lens that imparts a quadratic phase in space, a time-lens can be implemented that produces a quadratic phase shift in time[19–21]. This time-lens can, for example, magnify[22] or compress[23] signals in time and obeys an equivalent of the lens law. Time-lenses can be created with an electro-optic modulator[18] or via a parametric nonlinear optical process such as four-wave mixing with a chirped pump wave[19,20].

To produce the time gap necessary to achieve temporal cloaking, we create a temporal element that imparts a suitable nonlinear frequency chirp on a probe beam over a time window in which the gap will be produced. After propagating through a dispersive element, the different frequency components of the probe created by the nonlinear chirp are advanced and retarded as required to produce the desired temporal gap. We create such a temporal element by implementing a split time-lens (STL) which is composed of two half time-lenses connected at the tips. This temporal element is created by the nonlinear process of Bragg scattering via four-wave mixing with two pump waves in which the wavelength of one pump wave is constant in time and the other pump wave is pulsed and has a linear frequency chirp. To produce the appropriate chirp on the pump wave that will lead to a temporal gap, the red and blue spectral components of the pump pulse are separated and delayed with respect to each other. It is this step that splits the time lens to create the STL, and further details of how this is accomplished and modelled are provided in the Supplementary Methods.

A schematic of the full temporal cloaking device is presented in Fig. 1. A continuous-wave probe beam, which could be used to detect an event, is incident from the left. After passing through the STL the light propagates through a dispersive element (for example, a single-mode fibre) that translates the red and blue parts of the chirped probe to the edges of the time window, creating a temporal gap in the probe beam so that any event that might lead to a temporal or spectral change during this window will have no effect, and the event will remain undetected. Finally, a medium with a dispersion opposite in sign to the previous element (for example, a dispersive-compensating optical fibre), together with a second, similar STL, closes the gap so that neither the occurrence of the event nor the presence of the time-lenses are detected. (We note that the spatial equivalent to this second negative dispersive element would be one that produces negative diffraction[17] and thus would require a metamaterial with a negative index of refraction.)

The actual experimental configuration for temporal cloaking is presented in Fig. 2. We numerically simulated the probe beam as it passes

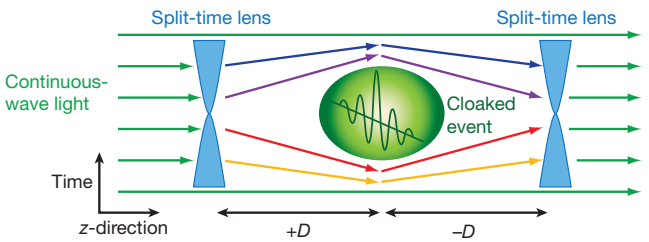


Figure 1 | Schematics of the temporal cloak using a pair of STLs. The STLs are used to create a temporal 'hole' in a probe beam such that any temporal or spectral changes caused by an event within this hole do not occur. The figure is oriented such that the probe light is described by horizontal lines, and lines at different orientations represent different wavelengths. *D* denotes the magnitude of the total negative or positive group-velocity dispersion.

[1]School of Applied and Engineering Physics, Cornell University, Ithaca, New York 14853, USA.

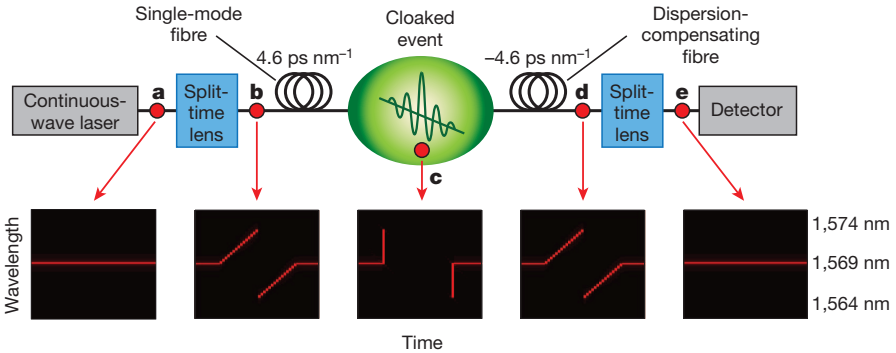


Figure 2 | Experimental configuration for cloaking an event in time. Top, schematics of the experimental configuration. Bottom, the wavelength of the probe beam as a function of time from numerical simulations. **a**, Before the first STL; **b**, after the first STL; **c**, after propagating through the single-mode fibre; **d**, before the second STL; **e**, after the second STL. When both STLs are in operation the event is temporally cloaked. Both STLs are identical, with the same *f*-number and are described in the Supplementary Information.

through the system by solving the nonlinear Schrödinger equation. The resulting wavelength of the probe beam as a function of time before and after the STL is shown in Fig. 2a and b, respectively. The light then propagates through a dispersive element consisting of a single-mode fibre such that the shorter (or longer) wavelengths propagate faster (or slower) than the initial probe beam wavelength. The wavelength distribution as a function of time shown in Fig. 2b becomes that shown in Fig. 2c, in which a temporal gap opens at the focal point of the STL. The gap is synchronized such that the event occurs within this gap and therefore is not sensed by the probe. The probe then propagates through a dispersion-compensating fibre, and the temporal gap is closed as shown in Fig. 2d. Finally, a second STL restores the probe light back to its initial wavelength, shown in Fig. 2e, so that the probe beam is restored to its initial state, and both the event and the presence of the time-lenses are undetected. We note that both after the time-lenses and after the event, we remove the pump waves from the system with wavelength division multiplexers.

Figure 3 shows experimental data illustrating how a temporal gap is created in a probe beam at 1,542 nm. Using a tuneable narrow-band filter followed by a fast (30 GHz) photo-detector connected to a sampling scope, we measured the exact time for each wavelength to obtain a spectrogram of the probe beam, after the first STL (blue dots), after the single-mode fibre at the focal length of the STL (red circle), and after the dispersion-compensating fibre before the second STL (green asterisks). After the first STL the probe acquires a frequency chirp, illustrated in Fig. 2b. After the single-mode fibre the higher wavelengths are delayed while the lower wavelengths are advanced so that a temporal gap opens from −25 ps to 25 ps according to Fig. 2c. This gap is closed after the dispersion-compensating fibre, and the frequency chirp returns to that shown in Fig. 2d. Before the first STL and after the second STL we detect only the probe at its original wavelength. The inset shows the power of the probe beam as a function of time when the gap is opened (red) and after the gap is closed (green), which provides additional evidence that during the temporal gap the power of the probe beam is significantly lower and that after closing the gap, no appreciable disturbance is observed.

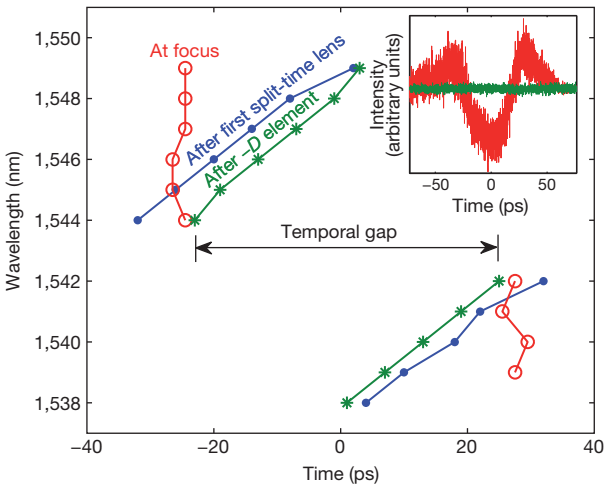


Figure 3 | Temporal gap in a probe beam. Measured wavelength distribution as a function of time at three locations in the system. After the first STL, the wavelengths are distributed according to Fig. 2b (blue dots). At the focal point, after the positive dispersion element, a temporal gap is opened according to Fig. 2c (red circles). After the negative dispersion element and before the second STL, the gap is closed and the wavelength distribution is returned to its initial state as described by Fig. 2b (green asterisks). Inset, power of the signal as a function of time when the gap is opened (red) and after the gap is closed (green).

To demonstrate the temporal cloaking capability of this system, we create an event that results in the generation of new frequencies due to the presence of the probe beam. It consists of a nonlinear interaction of a short pump pulse with a probe beam via four-wave mixing with a repetition rate of 41 kHz. When the cloak is off, the probe beam at 1,569 nm interacts in a highly nonlinear fibre with a short (5 ps) pump at 1,554 nm such that a frequency component is generated at 1,539 nm every 24 μs. Thus, the signature of the event is the detection of the 1,539 nm signal, and when the cloaking is turned off, it is clearly observed, as shown in Fig. 4a by the dashed (blue) curve. However, when the cloaking is turned on, the amplitude of the detected signal is reduced below the detection noise level, as shown in Fig. 4a (solid red curve) for several interaction events. The spectrum of the output signal in the regime near the idler wavelength is presented in Fig. 4b when the event is uncloaked (blue), when the event is cloaked (red), and without the event (green). To investigate the cloaking efficiency further, in Fig. 4c we present the signal without the event (green) and with the event when the cloaking is off (solid blue curve) and on (red dots) using a highly sensitive detector with a 300-MHz bandwidth. When the cloaking is on, the detected amplitude of the event is more than ten times lower than when the event is uncloaked.

Finally, we investigate the efficiency of the cloak as a function of the pump power of the STLs by measuring the amplitude of the detected event as a function of the pump power used by both STLs. The pump power of the STLs governs the amount of light in the probe beam that is shifted in frequency, so as the pump power increases, less light remains in the probe beam during the temporal gap when the event occurs (see Fig. 5). When the average pump power is 17 mW, the amplitude of the detected signal is 6.1 mV (also shown in the upper inset). As the pump power is increased, the detected amplitude decreases until it reaches the noise level at 2.3 mV (shown in the lower inset), when the pump power of the STL is 37 mW. Increasing the pump power of the STLs further increases the amplitude due to higher pump noise. Using numerical simulations of the propagation of the

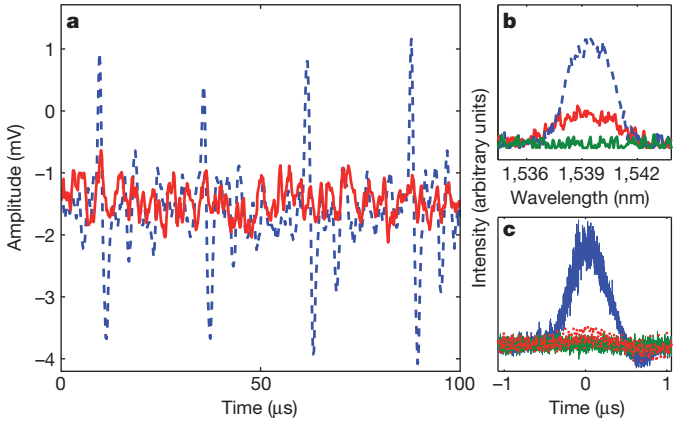


Figure 4 | Experimental results of the temporal cloaking. Experimental results showing the detection of signal, indicating that the probe beam has undergone an interaction (that is, an event) with a short pump pulse. **a**, The events occur every 24 μs when the cloaking is turned off (dashed blue curve) and on (solid red curve). When the cloaking is turned on, the amplitude of the signal probing the event is below the detection noise level, indicating that the event has been hidden. **b**, Spectra of the output signal around 1,539 nm as generated from the four-wave-mixing event when the event is uncloaked (blue), when the event is cloaked (red) and without the event (green). **c**, Temporal measurement of an uncloaked event (blue), cloaked event (red), and without an event (green). The amplitude of the event is more than ten times smaller when the cloaking is on, and the probe beam is nearly identical to that without an event.

signal through our cloaking system, we can predict the output signal amplitude as a function of pump power for our system. The results are presented as the solid blue curve in Fig. 5 and show good agreement with the experimental results. A detailed description of the numerical model is presented in the Supplementary Information.

The temporal gap can be readily widened by increasing the dispersive broadening of the pump and the dispersion between the STLs. However, as the dispersion is increased, effects due to third-order dispersion arise, which will prevent the gap from closing completely unless the third-order dispersion is compensated. In our experiment, the spectral and temporal width of the pump pulses before chirping are 9 nm and 0.4 ps, respectively, which results in third-order dispersion limiting the width of the temporal gap to 110 ps, as long as the pump power is increased to efficiently deplete the probe beam completely within the gap. Nevertheless, given that the third-order dispersion is proportional to $\Delta\lambda^3$ whereas the amount of linear frequency chirp is proportional to $\Delta\lambda^2$ (ref. 16), it is possible to increase the temporal gap by resorting to narrower pump pulses and introducing more dispersion. The limitation in this case is stimulated Brillouin scattering, which limits the length of the single-mode fibre to 50 km and the temporal gap to a width of a few nanoseconds.

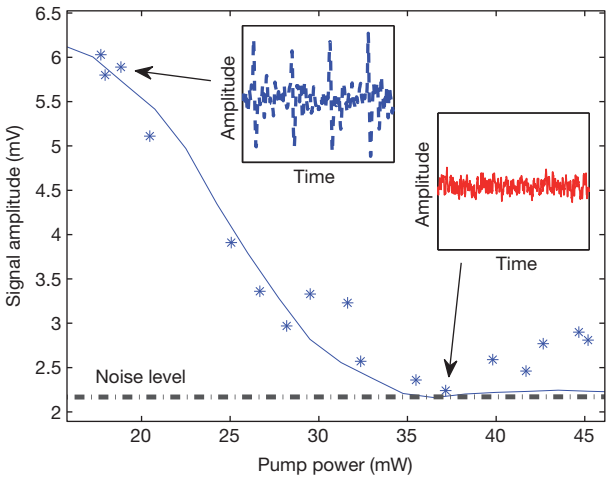


Figure 5 | Intermediate cloaking. Amplitude of the detected event as a function of the pump power of the STLs. Asterisks show experimental results and the solid curve shows the numerical simulation. The insets show the experimental results of the signal as a function of time when the cloaking is off (upper, blue) and on (lower, red).

Our results could also find application beyond full spatial temporal cloaking. Although spatial cloaking is still limited in angular acceptance, wavelength, polarization and efficiency, we have shown that we can deplete a probe wave and restore it to its original state after the event. This suggests it is possible to use similar technology for routeing different optical data streams coming simultaneously into a single optical data processing unit by opening a gap in one of the data streams and restoring it to a different time so that it will arrive after the other stream has been processed.

In summary, we have presented the first experimental demonstration of temporal cloaking that successfully hides an event from a probe beam in the time domain. Our scheme is based on the space–time duality, using a pair of STLs. The STLs use parametric four-wave mixing such that dispersion manipulation is highly efficient, and can readily be adapted for other wavelengths in the electromagnetic spectrum. Our results represent a significant step towards obtaining a complete spatio-temporal cloaking device.

METHODS

The first STL needs to focus light to the edges of the time window rather than to the centre so that after propagating through a medium with normal group-velocity dispersion, part of the probe light is delayed in time while the other part is advanced. This is realized by imposing a parabolic phase shift with a discontinuity in the middle. To obtain such a STL, we use Bragg scattering via four-wave mixing process with two pump waves. This parametric process shifts the frequency of the probe wave by the frequency difference between the two pumps. One pump is linearly chirped, so the resulting phase shift will be parabolic. Our phase-matching conditions mean that the resulting chirp on the probe wave is opposite to that of the pump wave. With such a four-wave mixing process it is possible to impose phase shifts as large as 4,000π on the probe wave.

We use a tuneable continuous-wave laser (SANTAC TSL-210) to generate the probe wave. To generate the pump waves, we use a femtosecond laser based on an erbium-doped fibre ring laser with a carbon nanotube saturable absorber. Both for the STLs and for the four-wave mixing process used to create the event we use a highly nonlinear single-mode fibre from Optical Fiber Solutions in Denmark.

Received 13 June; accepted 1 November 2011.

1. Leonhardt, U. Optical conformal mapping. *Science* **312,** 1777–1780 (2006).
2. Pendry, J. B., Schurig, D. & Smith, D. R. Controlling electromagnetic fields. *Science* **312,** 1780–1782 (2006).
3. Leonhardt, U. & Tyc, T. Broadband invisibility by non-Euclidean cloaking. *Science* **323,** 110–112 (2009).
4. Cai, W. Chettiar, U. K., Kildishev, A. V. & Shalaev, V.M. Optical cloaking with metamaterials. *Nature* **1,** 224–227 (2007).
5. Cummer, S. A. *et al.* Scattering theory derivation of a 3d acoustic cloaking shell. *Phys. Rev. Lett.* **100,** 024301 (2008).
6. Lai, Y. Chen, H. Zhang, Z. Q. & Chan, C. T. Complementary media invisibility cloak that cloaks objects at a distance outside the cloaking shell. *Phys. Rev. Lett.* **102,** 093901 (2009).
7. Gabrielli, L. H. Cardenas, J. Poitras, C. B. & Lipson, M. Silicon nanostructure cloak operating at optical frequencies. *Nature Photon.* **3,** 461–463 (2009).
8. Li Valentine, J., Zentgraf, L., Bartal, T. G., &. Zhang, X. An optical cloak made of dielectrics. *Nature Mater.* **8,** 568–571 (2009).
9. Li, J. & Pendry, J. B. Hiding under the carpet: a new strategy for cloaking. *Phys. Rev. Lett.* **101,** 203901 (2008).
10. Miller, D. A. B. On perfect cloaking. *Opt. Express* **14,** 12457–12466 (2006).
11. Weder, R. A. A rigorous analysis of high-order electromagnetic invisibility cloaks. *J. Phys. A* **41,** 065207 (2008).
12. Greenleaf, A. Lassas, M. & Uhlmann, G. Anisotropic conductivities that cannot be detected by EIT. *Physiol. Meas.* **24,** 413, doi:10.1088/0967-3334/24/2/353 (2003).
13. Schurig, D. *et al.* Metamaterial electromagnetic cloak at microwave frequencies. *Science* **314,** 977–980 (2006).
14. Chen, H. & Wu, B. I. Zhang, B. & Kong, J. A. Electromagnetic wave interactions with a metamaterial cloak. *Phys. Rev. Lett.* **99,** 063903 (2007).
15. McCall, M. W. Favaro, A. Kinsler, P. & Boardman, A. A spacetime cloak, or a history editor. *J. Opt.* **13,** 024003 (2011).

16. Agrawal, G. P. *Nonlinear Fiber Optics* 4th edn (Academic Press, 2007).
17. Kolner, B. H. Space-time duality and the theory of temporal imaging. *IEEE J. Quantum Electron.* **30,** 1951–1963 (1994).
18. Kolner, B. H. & Nazarathy, M. Temporal imaging with a time lens. *Opt. Lett.* **14,** 630–632 (1989).
19. Bennett, C. V. & Kolner, B. H. Principles of parametric temporal imaging. I. System configurations. *IEEE J. Quantum Electron.* **36,** 430–437 (2000).
20. Bennett, C. V. & Kolner, B. H. Principles of parametric temporal imaging. II. System performance. *IEEE J. Quantum Electron.* **36,** 649–655 (2000).
21. Salem, R. *et al.* Optical time lens based on four-wave mixing on a silicon chip. *Opt. Lett.* **33,** 1047–1049 (2008).
22. Foster, M. A. *et al.* Silicon-chip-based ultrafast optical oscilloscope. *Nature* **456,** 81–84 (2008).
23. Foster, M. A. *et al.* Ultrafast waveform compression using a time-domain telescope. *Nature Photon.* **3,** 581–585 (2009).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank D. J. Gauthier for his comments. This work was supported by the Defence Advanced Research Project Agency and by the Center for Nanoscale Systems, supported by the National Science Foundation for Science, Technology, and Innovation (NYSTAR).

Author Contributions M.F., A.F. and Y.O. performed the experiments and the numerical simulations. A.L.G. supervised the project.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to A.L.G. (alg3@cornell.edu).

doi:10.1038/nature10705

Changing Arctic Ocean freshwater pathways

James Morison[1], Ron Kwok[2], Cecilia Peralta-Ferriz[1], Matt Alkire[1], Ignatius Rigor[1], Roger Andersen[1] & Mike Steele[1]

Freshening in the Canada basin of the Arctic Ocean began in the 1990s[1,2] and continued[3] to at least the end of 2008. By then, the Arctic Ocean might have gained four times as much fresh water as comprised the Great Salinity Anomaly[4,5] of the 1970s, raising the spectre of slowing global ocean circulation[6]. Freshening has been attributed to increased sea ice melting[1] and contributions from runoff[7], but a leading explanation has been a strengthening of the Beaufort High—a characteristic peak in sea level atmospheric pressure[2,8]—which tends to accelerate an anticyclonic (clockwise) wind pattern causing convergence of fresh surface water. Limited observations have made this explanation difficult to verify, and observations of increasing freshwater content under a weakened Beaufort High suggest that other factors[2] must be affecting freshwater content. Here we use observations to show that during a time of record reductions in ice extent from 2005 to 2008, the dominant freshwater content changes were an increase in the Canada basin balanced by a decrease in the Eurasian basin. Observations are drawn from satellite data (sea surface height and ocean-bottom pressure) and *in situ* data. The freshwater changes were due to a cyclonic (anticlockwise) shift in the ocean pathway of Eurasian runoff forced by strengthening of the west-to-east Northern Hemisphere atmospheric circulation characterized by an increased Arctic Oscillation[9] index. Our results confirm that runoff is an important influence on the Arctic Ocean and establish that the spatial and temporal manifestations of the runoff pathways are modulated by the Arctic Oscillation, rather than the strength of the wind-driven Beaufort Gyre circulation.

A comparison between the results of large-scale trans-Arctic hydrographic sections in 1993 (ref. 10) and 1994 (ref. 11) and data from climatology[12] revealed a large-scale cyclonic shift in the boundary between Atlantic-derived and Pacific-derived water masses across the Arctic deep basins. This cyclonic shift was related to an increase in the cyclonic atmospheric circulation of the Northern Hemisphere associated with low Arctic sea level atmospheric pressure and characterized by an increased AO index[13] (the AO is the strength of the Northern Hemisphere Annular Mode; Supplementary Information 2).

Arctic regional indices have also been proposed to characterize Arctic Ocean change, including the doming of the sea surface characterized by the sea surface height gradient in a wind-forced model of the Arctic Ocean[6]. The doming is related to the strength of the Beaufort High and has been linked to changes in Arctic Ocean freshwater content[2,8,14], because anticyclonic wind stress drives convergence of Ekman transport in the ocean surface boundary layer, thickening the fresh surface layer and increasing the doming.

In contrast, we use a new combination of satellite altimetry and gravity, along with traditional hydrography, to show that from 2005 to 2008, an increased AO index caused the circulation to become more cyclonic and the Eurasian river runoff to be increasingly diverted eastward to the Canada basin at the expense of the Eurasian basin, with a nearly negligible increase in average Arctic Ocean fresh water.

The recent Canada basin freshening[2,4] is illustrated by *in situ* salinity observations taken in 2008 (Fig. 1) at ocean depths of 50 to 60 m that are 1–3 p.s.u. (practical salinity units, essentially equivalent to parts per thousand by weight) lower than values from pre-1990s climatology[12], a difference that is about five times the climatological root-mean-square interannual variability (Supplementary Fig. 4b). Geostrophic water velocities at 50–60 m computed from density-determined dynamic heights[15] relative to the depth where ocean pressure equals 500 dbar (1 dbar corresponds to about 1 m water equivalent pressure) for 2008 (Fig. 1) show the anticyclonic Beaufort Gyre current pattern as an intense southern core with westward motion along the Alaskan coast, but with broader eastward return flows farther north.

In contrast to the Canada basin, the Makarov basin upper-ocean salinities in 2008 are 1–2 p.s.u. greater than values from climatology[12] (Fig. 1). The corresponding trough in dynamic heights forces the geostrophic upper-ocean currents to sweep cyclonically around the southeastern part of the Makarov basin and across the Chukchi borderland (see location in Fig. 1) into the Canada basin's anticyclonic gyre.

Hydrochemistry sampling in 2008 indicates that, relative to a reference salinity of 34.87 p.s.u., Pacific water and Eurasian runoff provide the dominant fractions of freshwater in the upper 200 m of the Beaufort Sea[16]. The sea ice melt fraction is comparatively small and almost always negative, indicating the dominance of sea ice production and export over melt. We have compared spring 2008 (ref. 16) and summer 2003 (ref. 17) hydrochemistry data, and adjusting for seasonal differences[2] of the order of 1 m (Supplementary Information 3 and

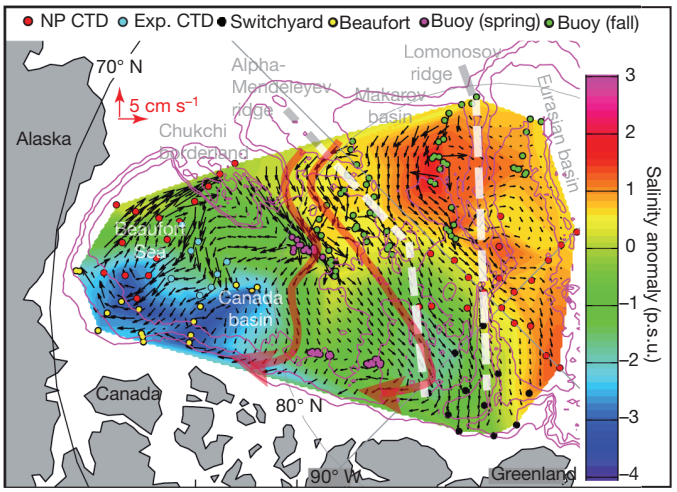


Figure 1 | 2008 Arctic Ocean salinity anomaly and geostrophic velocity at 50–60 m depth. The salinity anomaly (colour shading) is relative to the pre-1990 winter climatology[12] given in practical salinity units. The velocities (vectors) are derived from dynamic heights relative to the 500-dbar pressure surface[15]. Red arrows highlight S-shaped pathways from the Russian shelves into the Canada basin anticyclonic gyre. Dashed lines highlight the Alpha-Mendeleyer and Lomonosor ridges. 'NP CTD', 'Exp. CTD' and 'Switchyard' are the hydrographic stations done in spring 2008, 'Beaufort' represents the hydrographic stations done in summer 2007, and 'Buoy (spring)' and 'Buoy (fall)' indicate hydrographic data[15] from Ice Tethered Profilers gathered in autumn 2007 and spring 2008.

[1]Polar Science Center, Applied Physics Laboratory, University of Washington, 1013 Northeast 40th Street, Seattle, Washington 98105, USA. [2]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91109, USA.

Supplementary Figs 5 and 6), we find that the change in average freshwater inventories in the top 195 m of the Beaufort Sea are 0.6 m of sea ice melt, −0.8 m of Pacific water and 3.6 m of Eurasian runoff, with a total increase of 3.4 m. Only the increases in Eurasian runoff and the total inventories are substantially greater than the standard errors (s.e.) of these averages, owing to the variability among stations (s.e. value is 0.93 m for sea ice melt, 0.3 m for Pacific water and 0.99 m for Eurasian runoff, with an s.e. value for the total of 1.12 m; see Supplementary Information 3). From 2003 to 2008, the dominant source of freshening is increased Eurasian runoff, especially in the depth range of 50 to 115 m (Supplementary Information 3 and Supplementary Fig. 5), consistent with transport by geostrophic currents extending down into the upper halocline that come cyclonically around the Makarov basin into the Canada basin (Fig. 1).

Combined with spot verification by hydrography[15,16], dynamic ocean topography (DOT, deviation of the sea surface from the geoid) from the Ice Cloud and Land Elevation Satellite (ICESat) laser altimeter[15] and ocean bottom pressure (OBP) from the Gravity Recovery and Climate Experiment Gravity (GRACE) satellites (Supplementary Information 4) provide the spatial and temporal coverage needed to understand the 2005–2008 Arctic Ocean changes. OBP is the sum of DOT and the steric pressure anomaly (SPA, due to changes in water density). Comparisons among satellite-derived DOT and OBP and *in situ* observations of SPA have been done for temperate oceans[18,19], but this is the first such effort for the Arctic Ocean. Models indicate that at shorter than seasonal timescales, Arctic OBP variations are barotropic (Supplementary Information 4) and reflect DOT variations[20,21]. At interannual and longer timescales the deep ocean response is baroclinic, OBP variations are smaller than DOT variations, and the SPA is comparable to −DOT (Supplementary Fig. 7). Comparisons of ICESat DOT with hydrography[15] (Supplementary Information 4) and with GRACE OBP confirm this for the Arctic Ocean; multiyear variations in DOT, with a fractional correction by OBP, yield the multiyear variations in SPA.

DOT from 2005 to 2008 increased at ~5–8 cm yr^{-1} (uncertainty is standard deviation (s.d.) of 0.9 cm yr^{-1}; Supplementary Information 4) in the Canada basin (Fig. 2a), resulting in anticyclonic spin-up of the surface velocity. However, DOT decreased by 3−4 cm yr^{-1} in a trough aligned with the Eurasian continental shelf-break (seaward edge of the shallow continental shelf), resulting in increasing DOT-gradient-driven eastward surface velocities and transport of Eurasian river water along the Russian shelf. Variations in the pattern include cyclonic cells that allow for the runoff-rich coastal water to be carried across the shelf and into the eastern Makarov basin and Chukchi borderland regions. Furthermore, the increase in DOT towards the Russian coast could be driving a seaward secondary flow in the bottom boundary layer over the shelf that injects runoff-enriched water into the upper halocline of the central basin, where we find increases in the Eurasian runoff.

ICESat DOT rates of change agree with those inferred from the difference between GRACE OBP and the SPA from repeat hydrographic stations (Fig. 2a, correlation 0.84, for number of samples $N = 19$; 99% confidence limits: 0.5 < correlation coefficient < 0.95); the s.d. of DOT relative to OBP minus DOT, 1.2 cm yr^{-1}, is also applicable to SPA rate of change from the difference between GRACE OBP and ICES at DOT rates of change (Supplementary Fig. 10). Taking the difference of the DOT and OBP rates of change (Supplementary Information 4, OBP 4-year, 42-sample, rate-of-change uncertainty= ±0.37 cm yr^{-1}; ref. 22) yields estimates of SPA changes over the whole Arctic Ocean (Fig. 2b)—to our knowledge the first such estimates. The increasing SPA (3–5 cm yr^{-1}) in the Eurasian basin and along the Russian shelf-break balances the decreasing (−4 to −6 cm yr^{-1}) SPA in the Canada basin associated with declining salinity.

SPA is negatively related to freshwater content in cold polar oceans[23]. The correlation of SPA and freshwater content calculated directly for the Beaufort Sea (Supplementary Fig. 12) suggests that

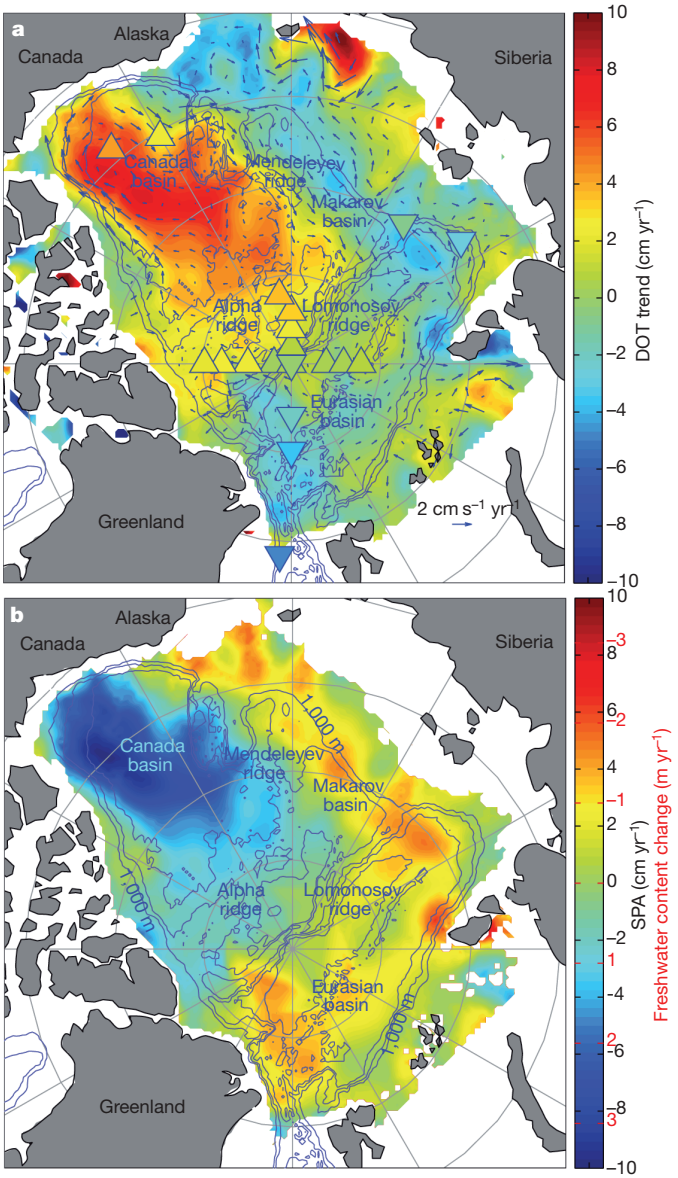


Figure 2 | Rates of change between 2005 and 2008 of DOT, SPA and freshwater content. **a**, DOT rate of change from ICESat altimetry (s.d. 0.9 cm yr^{-1}, Supplementary Information 4). Arrows show rate of change in near-surface velocity driven by DOT. The DOT rate of change equal to GRACE OBP rate of change minus SPA rate of change from hydrography are shown as colour-coded triangles (s.d. of difference is 1.2 cm yr^{-1}, Supplementary Fig. 10). **b**, SPA rate of change (±1.2 cm yr^{-1}) equal to the OBP rate of change (Supplementary Fig. 9) minus the DOT rate of change (Fig. 2a) for water depths over 50 m (Supplementary Information 5). The scale labelled in red indicates the change in freshwater content, −35.6 × SPA (±0.42 m yr^{-1}, Supplementary Information 6).

freshwater content is approximately −35.6 × SPA (Fig. 2b). Freshwater content changes are dominated by strong increases in the Canada basin balanced by decreases in the Eurasian basin and along the Russian shelf-break, reflecting change in Eurasian runoff pathways.

The area-averaged 2005–2008 freshwater rate of change (Fig. 2b) of 0.04 m yr^{-1} is almost insignificant (s.e. 0.034 m yr^{-1}, Supplementary Information 6). If we consider the deep basin only (water depths exceeding 500 m), the average is greater at 0.18 m yr^{-1} (s.e. 0.039 m yr^{-1}, Supplementary Information 6), because the 500-m depth contour runs near the middle of the freshwater minimum along the Russian shelf-break (Fig. 2b). This is essentially equal to the rate of change estimated for the deep basin over the previous decade[3]. The

rate of change of deep-basin freshwater volume and its difference from the larger area average are comparable to recent rates of change in ice volume[24] and variations in liquid freshwater exports[25] (Supplementary Information 7), illustrating the importance of observations with sufficiently broad spatial coverage (Supplementary Information 9).

Maps of 2006–2008 sea level pressure (SLP), DOT and SPA anomalies (Fig. 3) are consistent with variations in the AO. Just as highs in SLP force convergence of near-surface Ekman transport, doming of DOT and deepening of isohaline surfaces, SLP lows cause divergence, development of a trough in DOT and shoaling of isohalines. Whereas the Beaufort High dominates the mean SLP pattern, the AO manifests itself over the Arctic Ocean as a trough of low pressure extending from the Greenland–Norwegian seas into the Eurasian and Makarov basins (Supplementary Fig. 2). When the winter AO index increased in 2007, the SLP anomaly decreased over the Eurasian and Makarov basins, reflecting the AO pattern (Fig. 3a). The trough in SLP anomaly forces a trough in DOT aligned with the Russian shelf-break (Fig. 3b and Fig. 2a). The trough pattern includes upwelling of isohaline surfaces under the centre of the trough, as indicated by increased SPA (Fig. 3c, Fig. 2b). It also includes increased DOT and downwelling of isohalines across the Russian shelf. Raised DOT towards the coast moves fresher runoff-rich water eastward. The average upper-ocean circulation patterns for 2004–2005 (Supplementary Fig. 13e) and 2007–2009 (Supplementary Fig. 13f) confirm in absolute terms the increased cyclonic circulation on the Russian side of the Arctic Ocean (Figs 2a and 3b).

Our observations suggest idealized modes of Arctic Ocean circulation (Fig. 4). In the low-AO-index mode (Fig. 4a), an expanded high in SLP drives an anticyclonic surface circulation over most of the basin. The SLP pattern is similar to the mean SLP pattern (Supplementary Fig. 2a) but strengthened and expanded westward. Eurasian runoff leaves the Arctic directly across the Eurasian basin. In the high-AO-index pattern (Fig. 4b), cyclonic motion occurs on the Russian side of the Arctic Ocean, and the anticyclonic cell shifts to the southeast in the Canada basin. Eurasian runoff is diverted eastward and off the East Siberian shelf into the Canada basin circulation, where it can increase freshwater content through Ekman transport at the surface and by geostrophic currents at depth. The dipole character of the cyclonic mode and its connection to the AO cannot be captured by the doming index[6] because the doming criterion considers only a single DOT feature.

The 2005–2008 high-AO-index shift exemplifies a change in Arctic Ocean characteristics that began in 1989 and largely characterized the next 20 years (Supplementary Information 2). Then, as in 2005–2008, the AO index increased relative to its pre-1989 average, and the transpolar drift of sea ice and surface water shifted cyclonically[10] (Supplementary Fig. 13a–f). Salinity increased in the Makarov and Eurasian basins[10,26,27] and decreased in the Beaufort Sea[1,2] (Supplementary Fig. 4a) owing to an increase in the fraction of runoff[7], specifically caused by a diversion of Eurasian runoff to the east[26,27]. An important difference is that, although the 2005–2008 Canada basin circulation was increasingly anticyclonic, it became less anticyclonic and doming decreased in the early 1990s[2] (Supplementary Figs 13a–f). Clearly, increased doming was not the cause of the 1990s Beaufort Sea freshening.

The climate implications of cyclonic AO-induced shifts in freshwater pathways include increasing deep thermohaline convection in

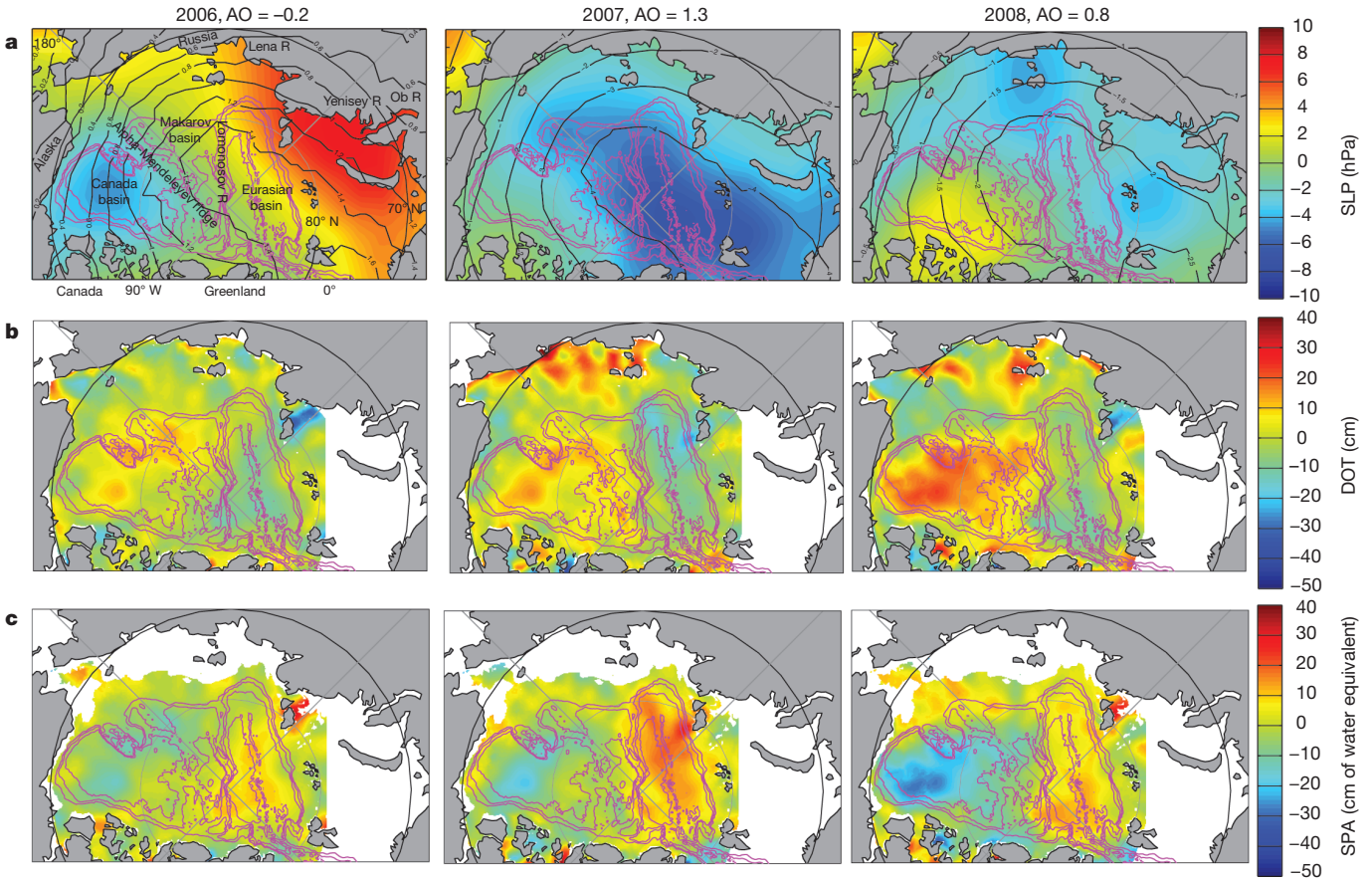


Figure 3 | 2006–2008 anomalies relative to 2004–2005 averages of SLP, DOT and SPA. a, Winter sea level atmospheric pressure from the International Arctic Buoy Program (http://iabp.apl.washington.edu/data_slp.html). Black contours are the mean SLP anomaly relative to 2004–2005 plus the AO contribution. **b**, DOT (February–March) (in cm). **c**, SPA, equal to the February–March-average OBP minus DOT. The SLP is the winter (previous November–April) average. The text AO value is the winter (November–April) average anomaly relative to the 1950–1989 average winter AO.

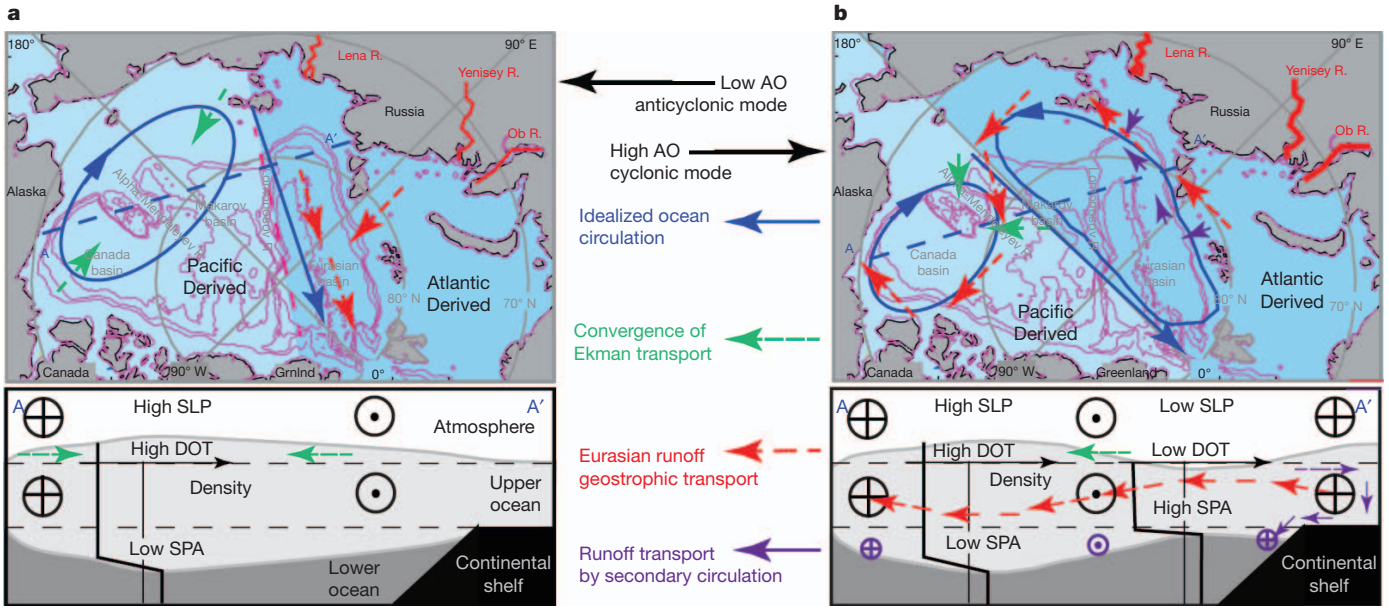


Figure 4 | Schematic views of the idealized Arctic Ocean circulation patterns under low and high AO anomalies. At the top are plan views and at the bottom are section views for the anticyclonic pattern for low AO index (**a**) and the cyclonic pattern for high AO index (**b**). The blue arrows indicate the prevailing surface geostrophic ocean circulation. The Ob, Yenisey and Lena rivers are the dominant sources of runoff to the Arctic Ocean. The red, green and purple arrows show the Eurasian runoff freshwater paths as indicated.

the Greenland Sea at the expense of the Labrador Sea, and enhancing sea ice melt by weakening the cold halocline layer of the Eurasian basin[27]. Climate models suggest an increasing AO with greenhouse warming[28], but the atmospheric models used typically produce low-pressure anomalies centred over the central Arctic Ocean[28–30] or show a negative AO bias[30] (Supplementary Information 1). Climate models need to capture the asymmetric effect of the AO on SLP in the Arctic to predict the role of the Arctic Ocean in our changing climate.

METHODS SUMMARY

The 2005–2008 repeat hydrographic stations (Figs 2 and 3) are from the North Pole Environmental Observatory (NPEO, http://psc.apl.washington.edu/northpole/) and Lincoln Sea Switchyard (http://psc.apl.washington.edu/switchyard/index.html) airborne surveys, plus the Beaufort Gyre Exploration Project (http://www.whoi.edu/beaufortgyre/) ship-borne surveys. Data from near the Laptev Sea shelf-break are from the Nansen and Amundsen Basin Observational System (http://nabos.iarc.uaf.edu/) and Fram Strait data comes from the Norwegian Polar Institute (http://www.npolar.no/en/).

The spring 2008 NPEO programme (Fig. 1) included conductivity–temperature–depth and hydrochemistry (Supplementary Information-3) stations in the North Pole and Beaufort Sea regions[4,16] augmented by Ice Tethered Profile buoy conductivity–temperature–depth data (http://www.whoi.edu/science/PO/arcticgroup/projects/ipworkshop.html). For Fig. 1, salinity anomalies and dynamic heights relative to 500 dbar at each station are linearly interpolated, and the gradients in dynamic height determine geostrophic currents[15].

We use GRACE monthly fields of OBP from the University of Texas Center for Space Research release 4 (dpc201012), from August 2002 to December 2009 (http://grace.jpl.nasa.gov/data) processed from spherical harmonic gravity coefficients by the Center for Space Research following ref. 31. The values represent anomalies relative to the mean from January 2003 to December 2007. We use data filtered with a 300-km half-amplitude radius Gaussian smoother. The GRACE Arctic OBP has been validated with *in situ* pressure at the North Pole[22].

We derive DOT (here filtered with a 100-km radius Gaussian smoother) as the difference between ICESat laser altimeter measurements of sea surface height in open water leads relative to the WGS84 ellipsoid and the EGM2008 geoid[15]. These data are available for download at http://rkwok.jpl.nasa.gov/icesat/data_topography.html. The DOT measured with ICESat in February–March 2008 are well correlated (correlation coefficient, $r_{corr} = 0.92$, $N = 176$, 99% confidence limits: $0.884 < r_{corr} < 0.945$) with dynamic height relative to 500 dbar calculated from 2008 hydrographic data (Supplementary Fig. 8). The resulting geostrophic surface velocities show the same features derived from the dynamic heights relative to 500 dbar (Fig. 1)[15].

Received 9 February; accepted 9 November 2011.

1. McPhee, M. G., Stanton, T. P., Morison, J. H. & Martinson, D. G. Freshening of the upper ocean in the Arctic: is perennial sea ice disappearing? *Geophys. Res. Lett.* **25,** 1729–1732 (1998).
2. Proshutinsky, A. *et al.* Beaufort Gyre freshwater reservoir: state and variability from observations. *J. Geophys. Res.* **114,** C00A10, http://dx.doi.org/10.1029/2008JC005104 (2009).
3. Rabe, B. *et al.* An assessment of Arctic Ocean freshwater content changes from the 1990s to the 2006–2008 period. *Deep Sea Res. I* **58,** 173–185 (2011).
4. McPhee, M. G., Proshutinsky, A., Morison, J., Steele, M. & Alkire, M. Rapid change in freshwater content of the Arctic Ocean. *Geophys. Res. Lett.* **36,** L10602, http://dx.doi.org/10.1029/2009GL037525 (2009).
5. Dickson, R. R., Meincke, J., Malmberg, S. A. & Lee, A. J. The 'Great Salinity Anomaly' in the northern North Atlantic 1968–1982. *Prog. Oceanogr.* **20,** 103–151 (1988).
6. Proshutinsky, A. Y. & Johnson, M. A. Two circulation regimes of the wind-driven Arctic Ocean. *J. Geophys. Res.* **102,** 12493–12514 (1997).
7. Macdonald, R. W., Carmack, E. C., McLaughlin, F. A., Falkner, K. K. & Swift, J. H. Connections among ice, runoff and atmospheric forcing in the Beaufort Gyre. *Geophys. Res. Lett.* **26,** 2223–2226 (1999).
8. Proshutinsky, A., Bourke, R. H. & McLaughlin, F. A. The role of the Beaufort Gyre in Arctic climate variability: seasonal to decadal climate scales. *Geophys. Res. Lett.* **29,** 2100, http://dx.doi.org/10.1029/2002GL015847 (2002).
9. Thompson, D. W. J. & Wallace, J. M. The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophys. Res. Lett.* **25,** 1297–1300 (1998).
10. Morison, J., Steele, M. & Andersen, R. Hydrography of the upper Arctic Ocean measured from the nuclear submarine USS Pargo. *Deep Sea Res. I* **45,** 15–38 (1998).
11. Carmack, E. C. *et al.* Changes in temperature and tracer distributions within the Arctic Ocean: results from the 1994 Arctic Ocean Section. *Deep Sea Res. II* **44,** 1487–1502 (1997).
12. Environmental Working Group (EWG). *Joint U.S.-Russian Atlas of the Arctic Ocean, Oceanography Atlas for the Winter Period* (National Ocean Data Center (NODC), 1997).
13. Morison, J., Aagaard, K. & Steele, M. Recent environmental changes in the Arctic. *Arctic* **53,** 359–371 (2000).
14. Lique, C. *et al.* Evolution of the Arctic Ocean Salinity, 2007–08: contrast between the Canadian and the Eurasian Basins. *J. Clim.* **24,** 1705–1717 (2011).
15. Kwok, R. & Morison, J. Dynamic topography of the ice-covered Arctic Ocean from ICESat. *Geophys. Res. Lett.* **38,** L02501, http://dx.doi.org/10.1029/2010GL046063 (2011).
16. Alkire, M. B. *et al.* Sensor-based profiles of the NO parameter in the central Arctic and southern Canada Basin: new insights regarding the cold halocline. *Deep Sea Res. I* **57,** 1432–1443 (2010).
17. Yamamoto-Kawai, M., McLaughlin, F. A., Carmack, E. C., Nishino, S. & Shimada, K. Freshwater budget of the Canada Basin, Arctic Ocean, from salinity, $\delta^{18}O$, and nutrients. *J. Geophys. Res.* **113,** C01007, http://dx.doi.org/10.1029/2006JC003858 (2008).

18. Chambers, D. P. Observing seasonal steric sea level variations with GRACE and satellite altimetry. *J. Geophys. Res.* **111,** C03010, http://dx.doi.org/10.1029/2005JC002914 (2006).
19. Willis, J. K., Chambers, D. P. & Nerem, R. S. Assessing the globally averaged sea level budget on seasonal to interannual timescales. *J. Geophys. Res.* **113,** C06015, http://dx.doi.org/10.1029/2007JC004517 (2008).
20. Vinogradova, N., Ponte, R. M. & Stammer, D. Relation between sea level and bottom pressure and the vertical dependence of oceanic variability. *Geophys. Res. Lett.* **113,** L03608, http://dx.doi.org/10.1029/2006GL028588 (2007).
21. Bingham, R. J. & Hughes, C. W. The relationship between sea-level and bottom pressure variability in an eddy permitting ocean model. *Geophys. Res. Lett.* **35,** L03602, http://dx.doi.org/10.1029/2007GL032662 (2008).
22. Morison, J., Wahr, J., Kwok, R. & Peralta-Ferriz, C. Recent trends in Arctic Ocean mass distribution revealed by GRACE. *Geophys. Res. Lett.* **34,** L07602, http://dx.doi.org/10.1029/2006GL029016 (2007).
23. Steele, M. & Ermold, W. Steric sea level change in the Northern Seas. *J. Clim.* **20,** 403–417 (2007).
24. Kwok, R. *et al.* Thinning and volume loss of the Arctic Ocean sea ice cover: 2003–2008. *J. Geophys. Res.* **114,** C07005, http://dx.doi.org/10.1029/2009JC005312 (2009).
25. Serreze, M. C. *et al.* The large-scale freshwater cycle of the Arctic. *J. Geophys. Res.* **111,** C11010, http://dx.doi.org/10.1029/2005JC003424 (2006).
26. Ekwurzel, B., Schlosser, P., Mortlock, R. A., Fairbanks, R. G. & Swift, J. H. River runoff, sea ice meltwater, and Pacific water distribution and mean residence times in the Arctic Ocean. *J. Geophys. Res.* **106,** 9075–9092 (2001).
27. Steele, M. & Boyd, T. Retreat of the cold halocline layer in the Arctic Ocean. *J. Geophys. Res.* **103,** 10419–10435 (1998).
28. Shindell, D. T., Miller, R. L., Schmidt, G. A. & Pandolfo, L. Simulation of recent northern winter climate trends by greenhouse-gas forcing. *Nature* **399,** 452–455 (1999).
29. Koldunov, N. V., Stammer, D. & Marotzke, J. Present-day Arctic sea ice variability in the coupled ECHAM5/MPI-OM model. *J. Clim.* **23,** 2520–2543 (2010).
30. Walsh, J. E., Chapman, W. L., Romanovsky, V., Christensen, J. H. & Stendel, M. Global Climate Model performance over Alaska and Greenland. *J. Clim.* **21,** 6156–6174 (2008).
31. Chambers, D. P. Evaluation of new GRACE time-variable gravity data over the ocean. *Geophys. Res. Lett.* **33,** L17603 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported chiefly by NSF grants OPP 0352754, ARC-0634226, ARC-0856330 and NASA grant NNX08AH62G. R.K. was supported at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. GRACE ocean data were processed by D. P. Chambers, supported by the NASA MEASURES Program. We thank the NASA ICESat and GRACE programmes, K. Falkner, R. Collier, M. McPhee, W. Ermold, L. de Steur, A. Proshutinsky and the Beaufort Gyre Exploration Project, J. Toole and R. Krishfield and the Ice Tethered Profiler project at WHOI, and W. Smethie of the Switchyard project for the observations that made this work possible.

Author Contributions The main idea was developed by J.M. and R.K. J.M. wrote most of the text and with R.A. and C.P.-F. drew most of the figures. R.K. developed the DOT records. The SLP and OBP anomaly plots were originally developed by C.P.-F. The 2008 hydrography observations were made by J.M., M.A., R.A. and M.S. The AO spatial pattern data, figures and insight were provided by I.R. The hydrographic data processing was done by R.A. and the chemistry analysis was done by M.A. Switchyard data and freshwater insight was provided by M.S. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.M. (morison@apl.washington.edu).

doi:10.1038/nature10709

Reversal of cocaine-evoked synaptic potentiation resets drug-induced adaptive behaviour

Vincent Pascoli[1], Marc Turiault[1] & Christian Lüscher[1,2]

Drug-evoked synaptic plasticity is observed at many synapses and may underlie behavioural adaptations in addiction[1]. Mechanistic investigations start with the identification of the molecular drug targets. Cocaine, for example, exerts its reinforcing[2] and early neuroadaptive effects[3] by inhibiting the dopamine transporter, thus causing a strong increase in mesolimbic dopamine. Among the many signalling pathways subsequently engaged, phosphorylation of the extracellular signal-regulated kinase (ERK) in the nucleus accumbens[4] is of particular interest because it has been implicated in NMDA-receptor and type 1 dopamine (D1)-receptor-dependent synaptic potentiation[5] as well as in several behavioural adaptations[6–8]. A causal link between drug-evoked plasticity at identified synapses and behavioural adaptations, however, is missing, and the benefits of restoring baseline transmission have yet to be demonstrated. Here we find that cocaine potentiates excitatory transmission in D1-receptor-expressing medium-sized spiny neurons (D1R-MSNs) in mice via ERK signalling with a time course that parallels locomotor sensitization. Depotentiation of cortical nucleus accumbens inputs by optogenetic stimulation *in vivo* efficiently restored normal transmission and abolished cocaine-induced locomotor sensitization. These findings establish synaptic potentiation selectively in D1R-MSNs as a mechanism underlying a core component of addiction, probably by creating an imbalance between distinct populations of MSNs in the nucleus accumbens. Our data also provide proof of principle that reversal of cocaine-evoked synaptic plasticity can treat behavioural alterations caused by addictive drugs and may inspire novel therapeutic approaches involving deep brain stimulation or transcranial magnetic stimulation.

We first tested whether cocaine treatment interfered with activity-dependent long-term potentiation (LTP) in the nucleus accumbens. When excitatory afferents onto MSNs were challenged with a high-frequency stimulation (HFS) train, LTP of the excitatory postsynaptic currents (EPSCs) was observed (Fig. 1a). The magnitude of the LTP was halved in brain slices from mice that had received a single injection of cocaine 7 days before the recording. If the cocaine was injected a month before measuring synaptic plasticity *ex vivo*, the difference was no longer present (Fig. 1b). MSNs of the nucleus accumbens fall into two classes of about equal proportions defined by the type of dopamine receptor expressed, with a small fraction of neurons (6–17%) that express both receptors[9]. A possible explanation for the partial change in LTP magnitude is therefore that cocaine exposure abolishes plasticity selectively in one class. To test this, we attempted LTP induction *ex vivo* after cocaine exposure in bacterial artificial chromosome (BAC) transgenic mice expressing enhanced green fluorescent protein (eGFP) either in D1 receptor (D1R)- or D2R-MSNs. We identified D1R-MSNs by a crossover strategy in which we recorded from green cells in *drd1a*-eGFP mice and non-green cells in *drd2*-eGFP mice (and vice versa for D2R-MSNs). Because the two approaches to identify the cell type yielded very similar results, we pooled the data (Fig. 1c, d). The main finding of this first experiment was that HFS, which reliably induced LTP in both types of MSNs after saline injection, became inefficient after cocaine treatment in D1R-MSNs (Supplementary Fig. 1a–d). The D2R overexpression recently reported in *drd2*-eGFP mice[10] interferes with neither the reported synaptic effects nor with the acute locomotor response to cocaine (Supplementary Fig. 2). After a cocaine injection, we were unable to induce HFS LTP regardless of whether we recorded from MSNs in the shell or the core of the nucleus accumbens (Supplementary Fig. 3), but because MSNs in the shell are the immediate targets of the dopamine neurons of the medial ventral tegmental area, which undergo the most significant changes in response to a single cocaine injection[11], we focused on nucleus accumbens shell neurons in the present study. Without distinguishing between D1R- and D2R-MSNs, several reports have already suggested that drug-evoked synaptic plasticity in the nucleus accumbens may underlie drug-related behavioural adaptations (reviewed in ref. 12). For example, when 5–7 daily cocaine injections were followed by a 10–21-day withdrawal, an overall increase of the AMPA/NMDA receptor (AMPAR/NMDAR) ratio[13] or GluA1/2 surface expression[14,15] was observed, and both observations were reversed by a challenge injection of cocaine.

The failure of HFS to induce LTP selectively in D1R-MSNs after cocaine treatment may be due either to the impairment of LTP induction or the occlusion of LTP expression. To distinguish between the two scenarios, we recorded miniature EPSCs (mEPSCs) in both cell types (Fig. 1e and Supplementary Fig. 4a) and observed a significant increase in amplitude along with a modest change in frequency of unitary events in D1R-MSNs (whereas these parameters remained unchanged in D2R-MSNs). Given that the paired pulse ratio also remained unchanged (Fig. 1f and Supplementary Fig. 4b), a postsynaptic mechanism underlying the increase of transmission and hence an occlusion scenario is the most likely explanation.

We characterized the induction criteria for HFS LTP in MSNs *in vitro* and found that it depends on NMDAR activation (Fig. 2a). Applying the MEK (MAP ERK kinase) inhibitor U0126 for the duration of the induction protocol also led to a complete block (Fig. 2b), indicating that activation of the ERK pathway is essential for this form of LTP, akin to NMDA-dependent LTP in the hippocampus and the dorsal striatum[16,17], the latter also depending on D1Rs[18].

Given that our results indicate an occlusion of LTP, we hypothesized that cocaine could drive synaptic potentiation selectively in D1R-MSNs via ERK activity. Indeed, a sharp increase in phosphorylated ERK is detected soon after cocaine exposure in D1R-MSNs, but not in D2R-MSNs[9]. In the nucleus ERK modulates gene expression, whereas in dendrites ERK is probably involved in the regulation of activity-dependent spine dynamics, synaptic glutamate receptor insertion and local dendritic protein synthesis[19]. To provide *in vivo* evidence for ERK dependence of cocaine-evoked synaptic plasticity, we treated mice with SL327, a blood–brain barrier penetrant ERK pathway inhibitor, before the saline or cocaine injection. We found that this manipulation rescued HFS LTP in D1R-MSNs one week later (Fig. 2c, d) without modification of the acute locomotor response to cocaine

[1]Department of Basic Neurosciences, Medical Faculty, University of Geneva, CH-1211 Geneva, Switzerland. [2]Clinic of Neurology, Department of Clinical Neurosciences, Geneva University Hospital, CH-1211 Geneva, Switzerland.

(Supplementary Fig. 5). A crossover control design with the two mouse lines again showed no difference between the green cells of one line and the non-green cells of the other (Supplementary Fig. 6).

Because inhibition of the ERK pathway blocks locomotor sensitization to cocaine[20,21], and we found LTP in D1R-MSNs to be dependent on ERK, we reasoned that cocaine-evoked potentiation might be a cellular correlate of the behavioural adaptation. Moreover, ERK activation is correlated both with AMPAR expression at the cell surface throughout the nucleus accumbens and with locomotor sensitization[14], but the synapses involved have not been identified, and no causal link has been established.

We then confirmed that a single injection of cocaine was sufficient to cause locomotor sensitization to a second injection of the same dose[22]. This was the case when challenged a week but not a month after the initial cocaine injection (Supplementary Fig. 7). The behavioural alteration therefore followed a time course similar to the cocaine-evoked synaptic potentiation (Fig. 1a, b). If cocaine-evoked potentiation is causally involved in locomotor sensitization, then depotentiating these synapses may reverse the behavioural alterations. To test this prediction experimentally, we injected channelrhodopsin (ChR2)-expressing adeno-associated virus (AAV) into the infralimbic cortex and implanted light guides into the nucleus accumbens to be able to selectively activate *in vivo* the terminals of this major excitatory input. Histological verification of the injection site did confirm robust expression in the infralimbic cortex, along with sparser expression in the prelimbic cortex (Fig. 3a and Supplementary Fig. 8). To validate further this approach, we recorded photocurrents from infected cortical neurons and prepared slices of the nucleus accumbens shell in which wide-field light exposure led to robust AMPAR-mediated EPSCs (Fig. 3b). We next applied light pulses at 1 Hz for 10 min (an established long-term depression (LTD) protocol to reduce synaptic transmission, see Methods). We found that this protocol strongly depressed transmission in nucleus accumbens slices from both saline and cocaine-treated mice (Fig. 3c). Interestingly, the magnitude of the depression was significantly larger in the latter, in line with an efficient depotentiation added to the LTD. As the LTD/depotentiation was NMDAR dependent and occurred without change of the paired pulse ratio, it is probably mediated by a postsynaptic expression mechanism and therefore constitutes an actual reversal of cocaine-evoked potentiation (Supplementary Fig. 9).

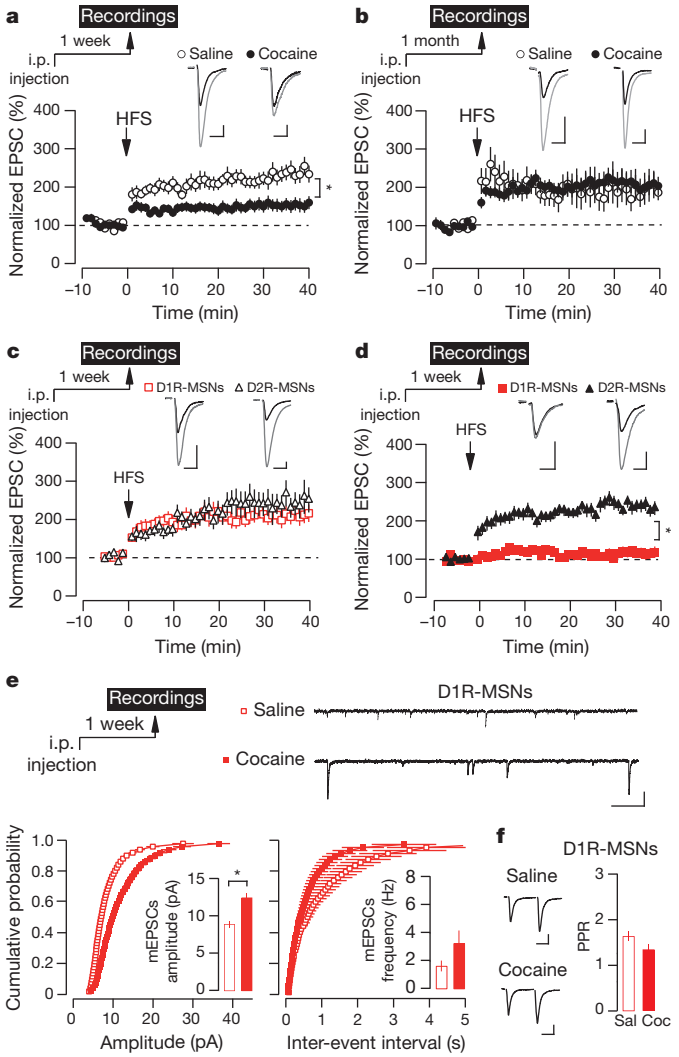


Figure 1 | Cocaine disrupts HFS-induced LTP in D1R-MSNs of the nucleus accumbens by potentiation of excitatory afferents. a–d, Graphs show normalized EPSCs as a function of time and the overlay of averaged (20 trials) traces of AMPAR EPSCs before (black line) and after (grey line) HFS. Symbols represent average of 6 trials. **a**, One week after cocaine injection, LTP was halved (224 ± 17.9% to 149 ± 14.7%, Student's *t*-test $t_{43} = 3.06$). **b**, After 1 month, no difference between treatments was detected (204 ± 26.1% to 206 ± 21.3%, $t_{21} = 0.07$). **c**, Using *drd1a*- and *drd2*-eGFP mice, HFS reliably induced LTP in D1R-MSNs and D2R-MSNs after saline injection (206 ± 16.0% and 229 ± 24.1%, respectively, $t_{27} = 0.78$). **d**, Cocaine abolished HFS LTP in D1R-MSNs but not in D2R-MSNs (115 ± 8.5% and 239 ± 12.7%, respectively, $t_{39} = 7.98$). **e**, Sample traces, cumulative probability and mean values of amplitude and frequency for mEPSCs recorded from D1R-MSNs one week after injection of saline or cocaine. Cocaine significantly increases mean amplitude of mEPSCs in D1R-MSNs ($t_{14} = 4.68$; Kolmogorov–Smirnov test: $P < 0.01$) whereas the frequency was unaffected ($t_{14} = 1.635$; Kolmogorov–Smirnov: $P > 0.05$). **f**, Paired-pulse ratio (PPR, 50 ms inter-stimulus interval) measured in D1R-MSNs was not different a week after saline (Sal) or cocaine (Coc) injection (1.7 ± 0.11 and 1.4 ± 0.1, respectively, $t_{34} = 1.73$). Scale bars: 20 ms and 20 pA. $n = 8–30$, $*P < 0.05$. Error bars show s.e.m.

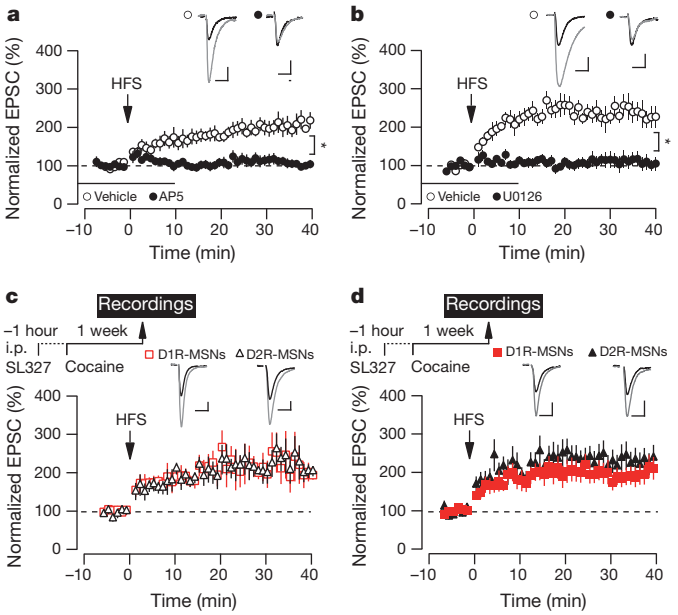


Figure 2 | HFS LTP and cocaine-evoked potentiation both depend on ERK activation. a, HFS LTP was blocked by an NMDAR antagonist (AP5, 100 μM; control, 209 ± 15% versus AP5, 93 ± 9%, $t_{14} = 6.78$). **b**, Bath application of the MEK inhibitor (U0126, 5 μM, 15 min before until 10 min after HFS protocol) also blocked LTP (control, 234 ± 17.8% versus U0126, 109 ± 16.2%, $t_{13} = 5.047$). **c**, Intraperitoneal administration of the MEK inhibitor (SL237, 40 mg kg^{-1}) in *drd1a*-eGFP mice or *drd2*-eGFP mice 1 h before saline did not modify the magnitude of HFS LTP in D1R- or D2R-MSNs when assessed one week later (224 ± 31% and 221 ± 27%, respectively, $t_{10} = 0.09$). **d**, SL237, administered 1 h before cocaine restored HFS LTP in D1R-MSNs (200 ± 22% versus 232 ± 27% for D2R-MSNs, $t_{22} = 0.93$). Scale bars: 10 ms and 20 pA. $n = 6–13$, $*P < 0.05$. Error bars show s.e.m.

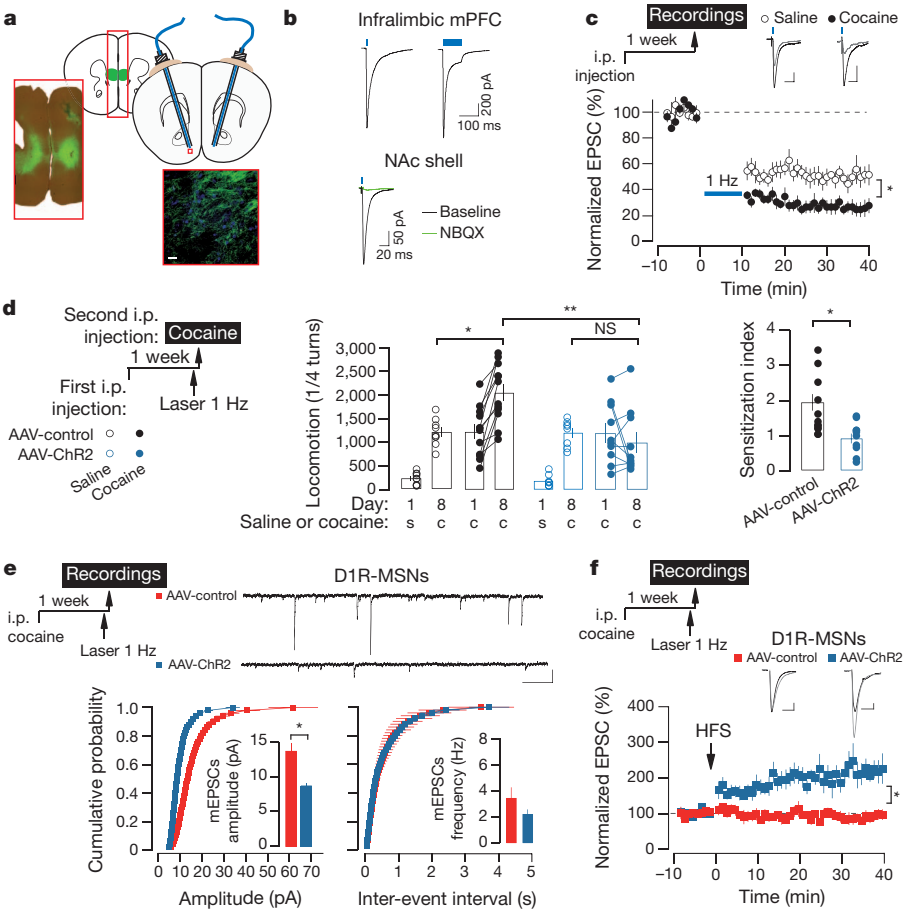


Figure 3 | Reversal of cocaine-evoked potentiation abolishes locomotor sensitization to cocaine. **a**, Schematic illustration of the site infected with AAV coding for eGFP-ChR2 and of the bilateral cannula implantation with optic fibres inserted in the nucleus accumbens shell. Insets show AAV-eGFP-ChR2 expression in the infralimbic cortex and high-resolution image of axonal inputs onto nucleus accumbens shell neurons (blue stains nuclei, ×60 magnification). **b**, Averaged current traces in response to light pulses (470 nm) of 4 ms and 100 ms in infected cells in slices of the infralimbic medial prefrontal cortex (mPFC; top). In nucleus accumbens (NAc) slices, light pulses (4 ms) evoked AMPAR EPSCs (blocked by NBQX 20 μM, green trace, bottom). **c**, Low-frequency stimulation (light pulses of 4 ms at 1 Hz for 10 min) induced depression of AMPAR EPSCs evoked with light in nucleus accumbens slices from mice injected with saline or cocaine one week previously (48 ± 5.7% and 27 ± 4.5%, respectively, $t_{14} = 2.851$). $n = 8$, $^*P < 0.05$. **d**, Quarter turns done by mice in the circular corridor for 60 min after injection represent locomotion (scatter plots of individual score and bars of mean ± s.e.m. are shown). Multiple-way repeated-measures analysis of variance for matching data yielded: interaction between day, virus and treatment $F_{(1,35)} = 13.41$, $P < 0.001$; effect of day $F_{(1,35)} = 78.16$, $P < 0.001$; effect of virus $F_{(1,35)} = 17.78$, $P < 0.001$; effect of treatment $F_{(1,35)} = 88.65$, $P < 0.001$. Wilcoxon or *t*-test: $^*P < 0.001$ for cocaine versus saline pre-treatment on day 8 in AAV-control and $P = 0.203$ cocaine versus saline pre-treatment on day 8 in AAV-ChR2 infected mice; $^{**}P = 0.0015$ for locomotor response to cocaine on day 8 for AAV-ChR2 versus AAV-control infected mice). No effect of light protocol on locomotor response to cocaine on day 8 when mice receive saline on day 1 was detected ($P > 0.05$ AAV-ChR2 versus AAV control). The sensitization index (see Methods) was significantly reduced in AAV-ChR2 infected mice treated with the light stimulation protocol (4 ms pulse, 1 Hz, 10 min) 45 min before the cocaine injection on day 8 ($t_{20} = 3.47$). $n = 8$–12, $^*P < 0.05$. NS, not significant. **e**, Acute slices of cocaine- and laser-treated AAV-ChR2 or AAV-control infected mice were prepared on day 8 (45 min after light stimulation). Sample traces, cumulative probability and mean values of amplitude and frequency for mEPSCs recorded from D1R-MSNs are shown. mEPSC amplitude was decreased in AAV-ChR2 compared to AAV-control infected mice ($t_{17} = 5.08$; Kolmogorov–Smirnov: $P < 0.001$). mEPSC frequency was unaltered ($t_{17} = 1{,}45$; Kolmogorov–Smirnov: $P > 0.05$). Scale bars: 200 ms and 20 pA. $n = 9$–10, $^*P < 0.05$. **f**, The light stimulation protocol restored HFS LTP in D1R-MSNs from AAV-ChR2 but not AAV-control infected mice (191.0 ± 21.2% and 93.0 ± 9.1%, respectively, $t_{12} = 4.26$). Scale bars: 10 ms and 20 pA. $n = 7$, $^*P < 0.05$. Error bars show s.e.m.

We next applied the above-validated protocol *in vivo* with the goal of establishing a causal link between cocaine-evoked plasticity and behavioural sensitization. We placed light guides into the ventral striatum, past the nucleus accumbens core, thus preferentially aiming at the principal cortical input onto MSNs of the nucleus accumbens shell, that is, axons that have their origin in the infralimbic cortex (Supplementary Fig. 8). When freely moving mice were treated with the optogenetic depotentiation protocol 45 min before the injection of the cocaine challenge at day 8, locomotor sensitization was completely erased. In control experiments, light stimulation did not affect the locomotor response at day 8 when the first injection was saline instead of cocaine or when a control virus was used (Fig. 3d). To ensure that the light stimulation restored normal transmission, we recorded *ex vivo* mEPSCs and found that an effective light treatment significantly reduced their mean amplitude in D1R-MSNs (Fig. 3e) to a level comparable to baseline transmission (Fig. 1e) without any effect on paired pulse ratio (Supplementary Fig. 10). In contrast, the mEPSC amplitude and frequency were not modified in D2R-MSNs after cocaine (Supplementary Fig. 11). Lastly, after cocaine treatment and optogenetic depotentiation the HFS protocol resulted in LTP *ex vivo* (Fig. 3f), confirming the restoration of baseline transmission in D1R-MSNs.

Behavioural sensitization typically refers to the observation of increased locomotor responses with repeated injections of cocaine. However, sensitization becomes apparent already after the second injection and is best observed about a week later[22] (Fig. 3). Although certainly not sufficient to induce addiction, such early forms of drug-induced adaptations are considered permissive building blocks for more definite behavioural alterations. If this is the case, the synaptic

potentiation onto D1R-MSNs should still be observed after chronic cocaine exposure, for example after 5 daily injections followed by a 10-day withdrawal. To confirm this hypothesis we recorded mEPSCs at the end of this protocol, as well as 1 h and 24 h after a challenge injection (Supplementary Fig. 12). This experiment confirmed that excitatory synapses onto D1R-MSNs were selectively potentiated at the end of the withdrawal period.

We then tested whether, after this chronic exposure to cocaine, light stimulation was still effective in reversing locomotor sensitization. After the 5 days of injection we imposed 10 days of withdrawal, and treated the mice with light stimulation 45 min before injecting a challenge dose. This completely reversed locomotor sensitization (Fig. 4a). When a second challenge injection was administered 24 h later, sensitized responses were again only observed in control mice. As a control, we did not observe an acute effect on locomotor behaviour during light stimulation (Supplementary Fig. 14).

Lastly, to estimate how long the effect of the light stimulation lasted, we tested for behavioural sensitization 5 days after the intervention (Fig. 4b). No significant locomotor sensitization was observed even when the mice were challenged at this extended time after light stimulation.

Our results identify NMDAR- and ERK-dependent LTP in D1R-MSNs of the nucleus accumbens as a form of synaptic plasticity required for locomotor sensitization to cocaine. ERK activation probably constitutes a general feature of addictive drugs, because in all brain regions that receive dopamine inputs tetrahydrocannabinol, amphetamines, morphine and nicotine also activate ERK signalling[4]. Through cocaine-driven ERK phosphorylation, potentiation is induced selectively in D1R-MSNs, which leads to the occlusion of HFS-driven LTP. These findings are in line with observations that ERK activation may control AMPAR trafficking directly, an effect that may also be maintained over days through activation of ERK nuclear targets leading to gene regulation[23].

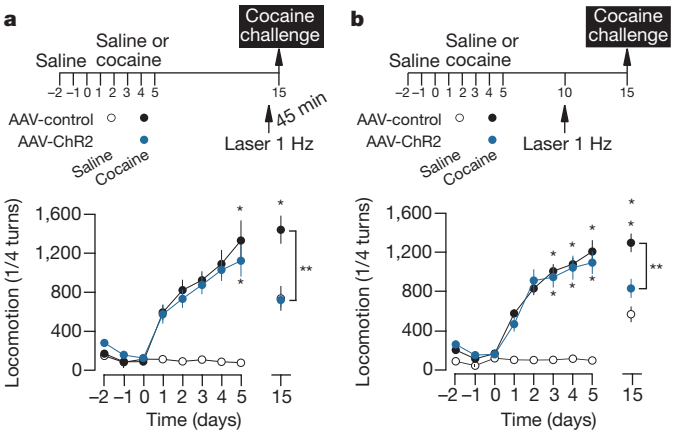


Figure 4 | Optogenetic depotentiation resets behavioural sensitization induced by chronic cocaine injections. **a**, After 5 daily cocaine injections a robust locomotor sensitization is measured in AAV-control and AAV-ChR2 infected mice. The light stimulation protocol (45 min before a challenge injection of cocaine on day 15) abolished the locomotor sensitized response in AAV-ChR2 infected mice (one-way analysis of variance: effect of treatment $F_{(2,41)} = 88.65$, $P < 0.001$. Post-hoc comparison by Bonferroni test yielded: $*P < 0.001$ for cocaine versus saline pre-treatment in AAV control; $**P < 0.001$ for AAV-ChR2 versus AAV-control infected mice, pre-treated with cocaine ($n = 13$–15). **b**, Light stimulation protocol applied on day 10 still reduced the locomotor sensitized response to a cocaine challenge injection on day 15 (one-way analysis of variance: effect of treatment $F_{(2,45)} = 15.31$, $P < 0.001$. Post-hoc comparison by Bonferroni test yielded: $*P < 0.001$ for cocaine versus saline pre-treatment in AAV-control infected mice; $**P = 0.003$ for AAV-ChR2 versus AAV-control infected mice, pre-treated with cocaine. No sensitization in AAV-ChR2 ($P = 0.27$, compared to saline pre-treatment; $n = 11$–21). Error bars show s.e.m.

We have explored and gained insight into the molecular mechanisms of synaptic adaptations to develop a strategy for reversal of cocaine-evoked potentiation with the goal of normalizing behaviour. We chose an optogenetic depotientiaton of inputs from the infralimbic cortex to the nucleus accumbens shell because of the strong anatomical connection and the functional implication of this projection in cocaine-seeking behaviour[24,25]. It is an appealing idea that behavioural adaptation, which closely reflects the potentiation of excitatory transmission onto D1R-MSNs, is due to an imbalance of the two classes of MSNs[26].

Several studies have already reported that pharmacological and molecular manipulations of key players in synaptic plasticity in the nucleus accumbens can affect adaptive behaviours associated with addictive drug exposure. For example, the inhibition of calcium-permeable AMPARs, a hallmark of late-stage cocaine-evoked synaptic plasticity in the nucleus accumbens[27] and the viral expression of a peptide that impairs GluA1 trafficking[28], reduce cue-induced cocaine seeking and cocaine-primed reinstatement, respectively.

We provide proof of principle that optogenetic manipulations can be used to reverse cocaine-evoked synaptic plasticity and thus abolish locomotor sensitization. Although light stimulation fully resets locomotor behaviour, sensitization begins to reappear after a few days, suggesting that several treatment sessions may be required to obtain long-lasting effects. This is not surprising, as chronic cocaine exposure also induces a number of additional adaptive changes including structural remodelling (for example, increase in spines[29]) and alterations of gene expression[23].

Sensitization to cocaine-associated stimuli has been linked with incentive saliency[30] and may explain the exceptionally strong motivation of addicts to obtain the drug. With chronic use, early adaptive changes such as those described here may build up to enhance craving during cocaine withdrawal[30]. Successful interventions that reverse these changes in animal models could inspire novel treatments for human addiction, a disease with a high social burden. Indeed, novel protocols of deep brain stimulation or transcranial magnetic stimulation may induce forms of synaptic plasticity that reverse drug-evoked adaptations, thus curbing the risk of relapse.

METHODS SUMMARY

All experiments were reviewed by the institutional ethics committee and approved by the relevant authorities of the Canton of Geneva. C57BL/6 or heterozygous BAC transgenic female and male mice, in which eGFP expression was driven by either D1R (*drd1a*-eGFP) or D2R (*drd2*-eGFP) gene regulatory elements were injected intraperitoneally with saline or cocaine. We then prepared brain slices for electrophysiological recordings as previously described[3] or placed the mice in an apparatus to quantify locomotor behaviour. Locomotor sensitization or synaptic plasticity was monitored at the various time points after the cocaine injection. A two-injection or five-injection protocol to induce locomotor sensitization was used as previously described[22]. Standard surgical procedures[3] were used to infect mice with ChR2-AAV or a control AAV (0.5 μl) in the infralimbic medial prefrontal cortex while the light guides were aimed at both nucleus accumbens (shell). In depotentiation experiments, a 473-nm solid-state laser was used to carry out the *in vivo* light stimulation protocol in awake mice (600 pulses of 4 ms duration at 1 Hz, 10–20 mW), 45 min or 5 days before behavioural testing or *ex vivo* electrophysiology recordings, respectively.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 10 June; accepted 10 November 2011.
Published online 7 December 2011.

1. Lüscher, C. & Malenka, R. C. Drug-evoked synaptic plasticity in addiction: from molecular changes to circuit remodeling. *Neuron* **69,** 650–663 (2011).
2. Chen, R. *et al.* Abolished cocaine reward in mice with a cocaine-insensitive dopamine transporter. *Proc. Natl Acad. Sci. USA* **103,** 9333–9338 (2006).
3. Brown, M. T. *et al.* Drug-driven AMPA receptor redistribution mimicked by selective dopamine neuron stimulation. *PLoS ONE* **5,** e15870 (2010).
4. Valjent, E., Pages, C., Herve, D., Girault, J. A. & Caboche, J. Addictive and non-addictive drugs induce distinct and specific patterns of ERK activation in mouse brain. *Eur. J. Neurosci.* **19,** 1826–1836 (2004).

5. Pascoli, V. *et al.* Cyclic adenosine monophosphate-independent tyrosine phosphorylation of NR2B mediates cocaine-induced extracellular signal-regulated kinase activation. *Biol. Psychiatry* **69,** 218–227 (2011).
6. Lu, L., Koya, E., Zhai, H., Hope, B. T. & Shaham, Y. Role of ERK in cocaine addiction. *Trends Neurosci.* **29,** 695–703 (2006).
7. Girault, J. A., Valjent, E., Caboche, J. & Herve, D. ERK2: a logical AND gate critical for drug-induced plasticity? *Curr. Opin. Pharmacol.* **7,** 77–85 (2007).
8. Thomas, M. J., Kalivas, P. W. & Shaham, Y. Neuroplasticity in the mesolimbic dopamine system and cocaine addiction. *Br. J. Pharmacol.* **154,** 327–342 (2008).
9. Bertran-Gonzalez, J. *et al.* Opposing patterns of signaling activation in dopamine D_1 and D_2 receptor-expressing striatal neurons in response to cocaine and haloperidol. *J. Neurosci.* **28,** 5671–5685 (2008).
10. Kramer, P. F. *et al.* Dopamine D2 receptor overexpression alters behavior and physiology in *Drd2-EGFP* mice. *J. Neurosci.* **31,** 126–132 (2011).
11. Lammel, S., Ion, D. I., Roeper, J. & Malenka, R. C. Projection-specific modulation of dopamine neuron synapses by aversive and rewarding stimuli. *Neuron* **70,** 855–862 (2011).
12. Wolf, M. E. The Bermuda Triangle of cocaine-induced neuroadaptations. *Trends Neurosci.* **33,** 391–398 (2010).
13. Kourrich, S., Rothwell, P. E., Klug, J. R. & Thomas, M. J. Cocaine experience controls bidirectional synaptic plasticity in the nucleus accumbens. *J. Neurosci.* **27,** 7921–7928 (2007).
14. Boudreau, A. C. & Wolf, M. E. Behavioral sensitization to cocaine is associated with increased AMPA receptor surface expression in the nucleus accumbens. *J. Neurosci.* **25,** 9144–9151 (2005).
15. Boudreau, A. C., Reimers, J. M., Milovanovic, M. & Wolf, M. E. Cell surface AMPA receptors in the rat nucleus accumbens increase during cocaine withdrawal but internalize after cocaine challenge in association with altered activation of mitogen-activated protein kinases. *J. Neurosci.* **27,** 10621–10635 (2007).
16. English, J. D. & Sweatt, J. D. A requirement for the mitogen-activated protein kinase cascade in hippocampal long term potentiation. *J. Biol. Chem.* **272,** 19103–19106 (1997).
17. Xie, G. Q. *et al.* Ethanol attenuates the HFS-induced, ERK-mediated LTP in a dose-dependent manner in rat striatum. *Alcohol. Clin. Exp. Res.* **33,** 121–128 (2009).
18. Shen, W., Flajolet, M., Greengard, P. & Surmeier, D. J. Dichotomous dopaminergic control of striatal synaptic plasticity. *Science* **321,** 848–851 (2008).
19. Zhu, J. J., Qin, Y., Zhao, M., Van Aelst, L. & Malinow, R. Ras and Rap control AMPA receptor trafficking during synaptic plasticity. *Cell* **110,** 443–455 (2002).
20. Valjent, E. *et al.* Regulation of a protein phosphatase cascade allows convergent dopamine and glutamate signals to activate ERK in the striatum. *Proc. Natl Acad. Sci. USA* **102,** 491–496 (2005).
21. Pierce, R. C., Pierce-Bancroft, A. F. & Prasad, B. M. Neurotrophin-3 contributes to the initiation of behavioral sensitization to cocaine by activating the Ras/mitogen-activated protein kinase signal transduction cascade. *J. Neurosci.* **19,** 8685–8695 (1999).
22. Valjent, E. *et al.* Mechanisms of locomotor sensitization to drugs of abuse in a two-injection protocol. *Neuropsychopharmacology* **35,** 401–415 (2010).
23. Brami-Cherrier, K., Roze, E., Girault, J. A., Betuing, S. & Caboche, J. Role of the ERK/MSK1 signalling pathway in chromatin remodelling and brain responses to drugs of abuse. *J. Neurochem.* **108,** 1323–1335 (2009).
24. Sesack, S. R., Deutch, A. Y., Roth, R. H. & Bunney, B. S. Topographical organization of the efferent projections of the medial prefrontal cortex in the rat: an anterograde tract-tracing study with *Phaseolus vulgaris* leucoagglutinin. *J. Comp. Neurol.* **290,** 213–242 (1989).
25. LaLumiere, R. T., Niehoff, K. E. & Kalivas, P. W. The infralimbic cortex regulates the consolidation of extinction after cocaine self-administration. *Learn. Mem.* **17,** 168–175 (2010).
26. Lobo, M. K. *et al.* Cell type-specific loss of BDNF signaling mimics optogenetic control of cocaine reward. *Science* **330,** 385–390 (2010).
27. Conrad, K. L. *et al.* Formation of accumbens GluR2-lacking AMPA receptors mediates incubation of cocaine craving. *Nature* **454,** 118–121 (2008).
28. Anderson, S. M. *et al.* CaMKII: a biochemical bridge linking accumbens dopamine and glutamate systems in cocaine seeking. *Nature Neurosci.* **11,** 344–353 (2008).
29. Robinson, T. E. & Kolb, B. Structural plasticity associated with exposure to drugs of abuse. *Neuropharmacology* **47** (suppl. 1), 33–46 (2004).
30. Vanderschuren, L. J. & Pierce, R. C. Sensitization processes in drug addiction. *Curr. Top. Behav. Neurosci.* **3,** 179–195 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. Kalivas and the members of the Lüscher laboratory for suggestions on the manuscript. This work is supported by the Swiss National Science Foundation (C.L.) and ''Synapsy'', a National Competence Center in Research (NCCR) of the Swiss Confederation on the synaptic basis of mental disorders.

Author Contributions V.P. carried out all electrophysiology experiments and was helped by M.T. with the behavioural experiments. C.L. designed the study and wrote the manuscript together with V.P. and M.T.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.L. (christian.luscher@unige.ch).

METHODS

Mice. C57BL/6 or heterozygous BAC transgenic mice, in which eGFP expression was driven by either D1R (*drd1a*-eGFP) or D2R (*drd2*-eGFP) gene regulatory elements were backcrossed[31] in C57Bl/6 mice for three to four generations, were used. Mice were housed in groups of 3–4 except for those implanted with guide cannulae, in which case animals were housed separately. All animals were kept in a temperature- and hygrometry-controlled environment with a 12 h light/12 h dark cycle. Mice were injected intraperitoneally with 20 mg kg^{-1} cocaine, 40 mg kg^{-1} SL327 (dissolved in 25% DMSO) or 0.9% saline (injection volume 10 ml kg^{-1}). Immediately after injection, mice were placed in the locomotor recording apparatus for 1 h. All procedures were approved by the Institutional Animal Care and Use Committee of the University of Geneva.

Locomotor sensitization. Locomotor activity was measured as the number of quarter turns entirely crossed by a mouse in a circular corridor. Locomotor chamber apparatus was placed under a video tracking system (Any-maze, Stoelting) and measurements were made automatically by the software. After 3 days of habituation to the test apparatus, mice underwent the experimental procedure, which consisted of two sessions of 60 min separated by 1 week (or 1 month), called day 1 and day 8 (or day 30). During the day 1 session, mice received saline or cocaine and were placed immediately in the corridor for 60 min. One week or 1 month later (day 8 or day 30), a second session was performed during which all mice were injected with cocaine before being placed in the circular corridor for 60 min. The light stimulation protocol (600 pulses of 4 ms at 1 Hz) was done 45 min before the second cocaine injection. To compare the effects of various times after the first injections or various virus infections, locomotor activity in response to the second cocaine injection was normalized to the mean locomotor activity of saline-pre-treated mice and the sensitization index was calculated by dividing the normalized locomotor response to the second injection by the normalized response to the first injection. Locomotor sensitization was also evaluated during challenge sessions that followed a chronic treatment (5 days of cocaine 15 mg kg^{-1}, 10 days withdrawal). The light stimulation protocol was done 45 min or 5 days before the challenge injection of cocaine.

Virus stereotaxic injection of ChR2-AAV or control AAV. AAV1 viruses produced at the University of North Carolina (Vector Core Facility) were injected into the infralimbic cortex of 15–20 g wild-type or BAC transgenic mice. Anaesthesia was induced and maintained with isoflurane (Baxter AG). The animal was placed in a stereotaxic frame (Angle One) and craniotomies were performed using stereotaxic coordinates (anterio-posterior, +1.9; medio-lateral, ±0.3; dorso-ventral, 2.4–2.6). Injections of AAV1 viruses (0.5 μl) were carried out using graduated pipettes (Drummond Scientific Company), broken back to a tip diameter of 10–15 μm, at a rate of ~0.05 μl min^{-1}. In all experiments the viruses were allowed a minimum of 3 weeks to incubate before any other procedures were carried out. As a control, some mice were injected with an AAV containing only GFP.

Cannula implantation. Following anaesthesia and craniotomy over the nucleus accumbens, two holes were drilled around the craniotomy and screws were placed in the holes. Two weeks after viral injections, guide cannulae (Plastics One) were lowered slowly into position using stereotaxic coordinates (bilaterally anterio-posterior, +1.5; medio-lateral, ±1.6; dorso-ventral 4.1; 15° angle) and cemented in place using dental cement (Lang Dental MFG Company) to encase the base of the guide cannulae and the screws. Once the cement had dried, a dummy cannula (Plastics One) was placed inside each guide cannula to prevent infection.

Slice electrophysiology. Coronal 200–250-μm slices of mouse forebrain were prepared in cooled artificial cerebrospinal fluid (ACSF) containing (in mM): NaCl 119, KCl 2.5, MgCl 1.3, $CaCl_2$ 2.5, Na_2HPO_4 1.0, $NaHCO_3$ 26.2 and glucose 11, bubbled with 95% O_2 and 5% CO_2. Slices were kept at 32–34 °C in a recording chamber superfused with 2.5 ml min^{-1} ACSF. Visualized whole-cell voltage-clamp recording techniques were used to measure holding and synaptic responses of MSNs of the nucleus accumbens shell, identified in some experiments by the presence of the GFP of BAC transgenic mice by using a fluorescent microscope (Olympus BX50WI, fluorescent light U-RFL-T). The holding potential was −70 mV, and the access resistance was monitored by a hyperpolarizing step of −14 mV with each sweep, every 10 s. Experiments were discarded if the access resistance varied by more than 20%. Synaptic currents were evoked by stimuli (50–100 μs) at 0.1 Hz through bipolar stainless steel electrodes placed at the cortex–nucleus-accumbens border. The internal solution contained (in mM): 140 k-gluconate, 5 KCl, 130 CsCl, 10 HEPES, 0.2 EGTA, 2 $MgCl_2$, 4 Na_2ATP, 0.3 Na_3GTP and 10 sodium creatine-phosphate. Currents were amplified (Multiclamp 700B, Axon Instruments), filtered at 5 kHz and digitized at 20 kHz (National Instruments Board PCI-MIO-16E4, Igor, WaveMetrics). The liquid junction potential was small (−3 mV), and therefore traces were not corrected. All experiments were carried out in the presence of picrotoxin (100 μM). LTP was induced by using the following HFS protocol: 100 pulses at 100 Hz repeated 4 times at 0.1 Hz paired with depolarization at 0 mV[32,33].

Miniature EPSCs were recorded in the presence of tetrodotoxin (0.5 μM). The frequency, amplitudes and kinetic properties of these currents were then analysed using the Mini Analysis software package (v.4.3, Synaptosoft). Cocaine or light illumination of ChR2-induced changes in cumulative miniature EPSC amplitude and inter-event interval distribution were analysed for statistical significance using the nonparametric two-sample Kolmogorov–Smirnov test (KyPlot) with a conservative critical probability level of $P < 0.05$.

Paired pulse ratio (PPR) was calculated by dividing the second evoked EPSC by the first with a 50-ms interval in between.

AMPAR EPSCs evoked with ChR2 stimulation by 4-ms light pulses (LED, Thorlabs) were recorded in the same conditions as electrically evoked synaptic currents. Low-frequency stimulation (LFS; 1 Hz for 10 min) was applied with light pulses and the magnitude of LTD was determined by comparing average EPSCs that were recorded 20–30 min after induction to EPSCs recorded immediately before induction.

***In vivo* stimulation of infralimbic cortex projections in the nucleus accumbens shell.** Virus injected and cannulated animals were allowed a minimum of 1 week to recover and to express the virus. 473-nm solid-state lasers (GMP, CH) were used to carry out the *in vivo* stimulation protocol in awake mice. A fibre optic (Thorlabs) was customized to enable the mouse to move freely during stimulation. Briefly, the plastic cap of a dummy cannula (Plastics One) was hollowed out and a hole of sufficient diameter for the fibre optic to pass through made in the top. This was threaded onto the fibre optic, one end of which was stripped to leave a 200 μm external diameter. The fibre was then lowered into the guide cannula on the mouse and the hollowed-out dummy cannula cap screwed onto the guide cannula. An FC/PC rotative fibre-optic rotary joint (Doric lenses) was used to release torsion in the fibre caused by the animal's rotation. The fibre was connected to the laser, which delivered 4 ms pulses at 1 Hz for 10 min, an established LTD protocol at excitatory synapses in the nucleus accumbens[34]. All stimulations were carried out in the mouse home cage (except in the experiment shown in Supplementary Fig. 10, in which stimulation was performed during locomotor recordings in the circular corridor) 45 min or 5 days before behavioural testing or *ex vivo* electrophysiology recordings.

31. Mameli, M. *et al.* Cocaine-evoked synaptic plasticity: persistence in the VTA triggers adaptations in the nucleus accumbens. *Nature Neurosci.* **12,** 1036–1041 (2009).
32. Gong, S. *et al.* Targeting Cre recombinase to specific neuron populations with bacterial artificial chromosome constructs. *J. Neurosci.* **27,** 9817–9823 (2007).
33. Kombian, S. B. & Malenka, R. C. Simultaneous LTP of non-NMDA- and LTD of NMDA-receptor-mediated responses in the nucleus accumbens. *Nature* **368,** 242–246 (1994).
34. Pennartz, C. M., Ameerun, R. F., Groenewegen, H. J. & Lopes da Silva, F. H. Synaptic plasticity in an *in vitro* slice preparation of the rat nucleus accumbens. *Eur. J. Neurosci.* **5,** 107–117 (1993).

doi:10.1038/nature10715

Modulation of TRPA1 thermal sensitivity enables sensory discrimination in *Drosophila*

Kyeongjin Kang[1]*, Vincent C. Panzano[1]*, Elaine C. Chang[1], Lina Ni[1], Alexandra M. Dainis[1], Adam M. Jenkins[2], Kimberly Regna[2], Marc A. T. Muskavitch[2,3,4] & Paul A. Garrity[1]

Discriminating among sensory stimuli is critical for animal survival. This discrimination is particularly essential when evaluating whether a stimulus is noxious or innocuous. From insects to humans, transient receptor potential (TRP) channels are key transducers of thermal, chemical and other sensory cues[1,2]. Many TRPs are multimodal receptors that respond to diverse stimuli[1–3], but how animals distinguish sensory inputs activating the same TRP is largely unknown. Here we determine how stimuli activating *Drosophila* TRPA1 are discriminated. Although *Drosophila* TRPA1 responds to both noxious chemicals[4] and innocuous warming[5], we find that TRPA1-expressing chemosensory neurons respond to chemicals but not warmth, a specificity conferred by a chemosensory-specific TRPA1 isoform with reduced thermosensitivity compared to the previously described isoform. At the molecular level, this reduction results from a unique region that robustly reduces the channel's thermosensitivity. Cell-type segregation of TRPA1 activity is critical: when the thermosensory isoform is expressed in chemosensors, flies respond to innocuous warming with regurgitation, a nocifensive response. TRPA1 isoform diversity is conserved in malaria mosquitoes, indicating that similar mechanisms may allow discrimination of host-derived warmth—an attractant—from chemical repellents. These findings indicate that reducing thermosensitivity can be critical for TRP channel functional diversification, facilitating their use in contexts in which thermal sensitivity can be maladaptive.

Highly temperature-responsive TRP cation channels, thermoTRPs, mediate thermosensation from insects to mammals[1,2] and are important for human pain and inflammation[6]. Like mammalian thermoTRPs, *Drosophila melanogaster* TRPA1 is both a thermal and chemical sensor, responding to innocuous warmth (above ~25–27 °C)[5,7] and noxious chemicals[4]. TRPA1 acts in thermosensors within the brain to modulate thermal preference over 18–32 °C (ref. 5), innocuous temperatures compatible with fly survival[8], and in gustatory chemosensors to inhibit ingestion of electrophiles[4], reactive chemicals like allyl isothiocyanate (AITC, found in wasabi) and *N*-methylmaleimide (NMM) that rapidly incapacitate flies (Supplementary Fig. 1). The responsiveness of TRPA1 to both innocuous and noxious stimuli raises the question of how these stimuli are distinguished to elicit distinct behavioural responses. Mammals face similar issues; for example, TRPM8 transduces both innocuous and noxious cold[1–3].

We previously reported TRPA1-expressing chemosensors in the labral sense organ[4]; using improved immunostaining conditions, we now also detect specific TRPA1 protein expression in labellar chemosensors (Fig. 1a, b). Extracellular tip recording[9] indicated these neurons were TRPA1-dependent chemosensors; they responded to the electrophile NMM with robust spiking in wild type but not *TrpA1* mutants

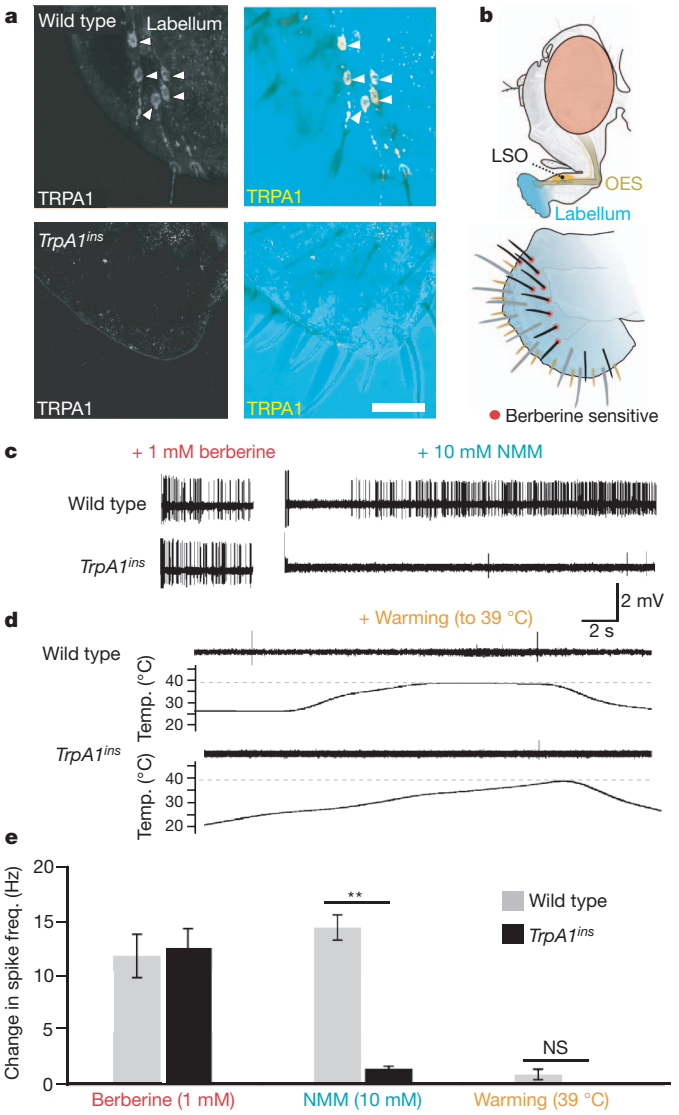


Figure 1 | TRPA1-dependent gustatory neurons do not respond to heat. **a**, TRPA1 immunostaining of wild-type (top) and *TrpA1*^{ins} (bottom) labella. Right, differential interference contrast overlay of labellar structures. Arrowheads, chemosensor cell bodies. Scale bar, 20 μm. **b**, Top, *Drosophila* gustatory organs: LSO, labral sense organ; OES, oesophagus. Bottom, labellar bristles. Brown, s-type; grey, L-type; black, i-type; berberine-sensitive bristles were targeted for electrophysiology. **c**, **d**, Bristle responses to: berberine (1 mM) and NMM (10 mM) (**c**); warming (**d**). Temp., temperature. **e**, Average spike rate after subtracting electrolyte-only baseline. freq., frequency. $^{**}P < 0.01$; NS, not significant ($P > 0.05$), *t*-test. All data are mean ± standard error of the mean (s.e.m.). Warming reached an average maximum temperature of 39.0 ± 0.6 °C (mean ± standard deviation).

[1]National Center for Behavioral Genomics and Volen Center for Complex Systems, Department of Biology, Brandeis University, Waltham, Massachusetts 02454, USA. [2]Biology Department, Boston College, Newton, Massachusetts 02467, USA. [3]Broad Institute, Cambridge, Massachusetts 02142, USA. [4]Harvard School of Public Health, Boston, Massachusetts 02115, USA.
*These authors contributed equally to this work.

(Fig. 1c and Supplementary Fig. 2). The mutant defect was electrophile specific, as *TrpA1* mutants responded like wild type to berberine chloride (Fig. 1c), a bitter compound that also activates these neurons[10]. In contrast, warming to ~39 °C, from innocuously warm to the noxious range, elicited no spiking in these cells (Fig. 1d, e). This is notable as the effectiveness of TRPA1 in conferring warmth sensitivity has led to its use as a thermogenetic tool[5,11]. Thus, despite the known sensitivity of TRPA1 to both temperature and chemicals, these chemosensors are warmth insensitive.

In addition to the previously characterized transcript, *TrpA1(B)*, a transcript with an alternative 5′ end, *TrpA1(A)*, has been annotated[12] (Fig. 2a). These transcripts encode protein isoforms with distinct amino termini, but the same ankyrin and transmembrane domains (Fig. 2b). Polymerase chain reaction with reverse transcription (RT–PCR) demonstrated differential expression: *TrpA1(A)* was expressed in the proboscis, which houses the TRPA1-expressing chemosensors, whereas *TrpA1(B)* predominated elsewhere in the head, where TRPA1-expressing thermosensors are located (Fig. 2c).

Examined in *Xenopus* oocytes, TRPA1(A) was much less thermosensitive than TRPA1(B), as reflected in its temperature coefficient (Q_{10}), the fold change in current per 10 °C change[1,2]. Arrhenius plot analysis[13] yielded a Q_{10} of ~9 for TRPA1(A) versus ~116 for TRPA1(B) (Fig. 2d–f). In addition, whereas TRPA1(B) was essentially inactive at low temperatures, TRPA1(A)-dependent currents were observed ≤15 °C, further reducing the temperature-dependent activity differential of TRPA1(A) (Supplementary Fig. 3). The maximum heat-activated current for TRPA1(A) was also significantly lower (Supplementary Fig. 4). Lastly, the transition (or threshold) temperature for increased temperature responsiveness was 29.7 ± 0.3 °C for TRPA1(A) versus 27.8 ± 0.4 °C for TRPA1(B) ($P < 0.01$, *t*-test). As the innocuous warm temperature range in *Drosophila* is of particular behavioural relevance, the Q_{10} from 27–37 °C (below the ~38 °C nociceptive threshold in *Drosophila*[14]) was also calculated, yielding 6.2 ± 0.5 for TRPA1(A) and 90 ± 8 for TRPA1(B) (Fig. 2f). Other properties were largely unaffected; both channels responded robustly to electrophiles and had similar voltage sensitivities (Fig. 2g, h). TRPA1(A) and TRPA1(B) had similar maximum current amplitudes at 300 μM NMM, with half-maximum effective concentration (EC_{50}) values of 176 ± 12 and 128 ± 9 μM, respectively (Fig. 2i).

The reduced thermosensitivity of TRPA1(A) could account for the chemosensors' warmth insensitivity. But although TRPA1(A) is less temperature sensitive than TRPA1(B), its Q_{10} resembles several TRPs implicated in warmth sensitivity[15–17]. To assess whether TRPA1(A) could confer warmth sensitivity upon *Drosophila* chemosensors, each isoform was used to rescue a *TrpA1* mutant. We previously demonstrated that expressing TRPA1(B) in TRPA1-dependent chemosensors using Gr66a-Gal4 rescues the *TrpA1* mutant behavioural defect[4]. Using electrophysiology, we found both isoforms restored NMM responsiveness (Fig. 3a and Supplementary Fig. 5a, b), but only TRPA1(B) conferred warmth sensitivity (Fig. 3b, c). These differences did not require properties unique to TRPA1-dependent chemosensors. Each isoform was expressed ectopically in sweet-responsive chemosensors using Gr5a-Gal4 (ref. 18). Both isoforms conferred electrophile sensitivity upon these normally electrophile-insensitive neurons, but only TRPA1(B) conferred thermosensitivity (Fig. 3d–f and Supplementary Fig. 4c, d). The inability of TRPA1(A) to confer warmth sensitivity on fly chemosensors emphasizes that although a Q_{10} above 5 makes TRPA1(A) more thermally sensitive than most ion channels, *in vivo* testing is important in evaluating whether a channel is sufficiently thermosensitive to make a specific neuron warmth responsive.

These data support a model in which the specificity of TRPA1-expressing gustatory neurons for chemicals is established by their selective expression of TRPA1(A), an isoform unable to confer warmth sensitivity. In contrast, the chemical sensitivity of TRPA1(B) should render TRPA1-dependent thermosensors sensitive to reactive chemicals. However, the location of TRPA1-dependent anterior cell thermosensors

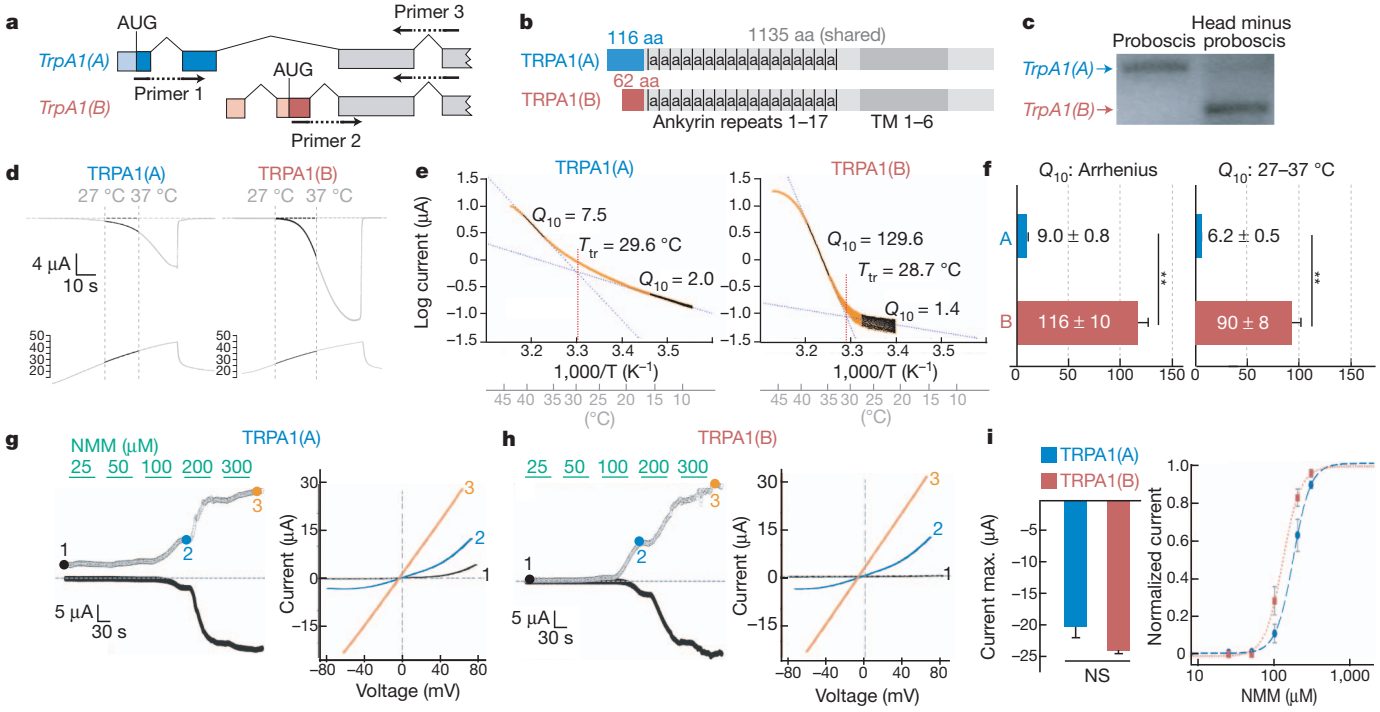


Figure 2 | TRPA1 isoform diversity yields tissue-specific channels with different thermal sensitivities. **a**, *TrpA1* gene structure and primer locations. **b**, TRPA1 proteins. Dark red and blue boxes denote isoform-specific sequences. Dark grey, transmembrane (TM) region. a, ankyrin repeat; aa, amino acids. **c**, RT–PCR analysis of dissected tissue. **d**, **e**, TRPA1(A)- and TRPA1(B)-dependent currents (**d**) and Arrhenius plots (**e**) in oocytes. T_{tr}, transition temperature. **f**, Q_{10} values from Arrhenius plot (left) or 27–37 °C (right). **g**, **h**, Left panels, NMM responsiveness of TRPA1(A) (**g**) and TRPA1(B) (**h**). Right panels, *I–V* relationships at points marked at left. **i**, Mean amplitudes at 300 μM NMM (left) and NMM dose–response (right). All data, mean ± s.e.m. $^{**}P < 0.01$; NS, not significant, *t*-test.

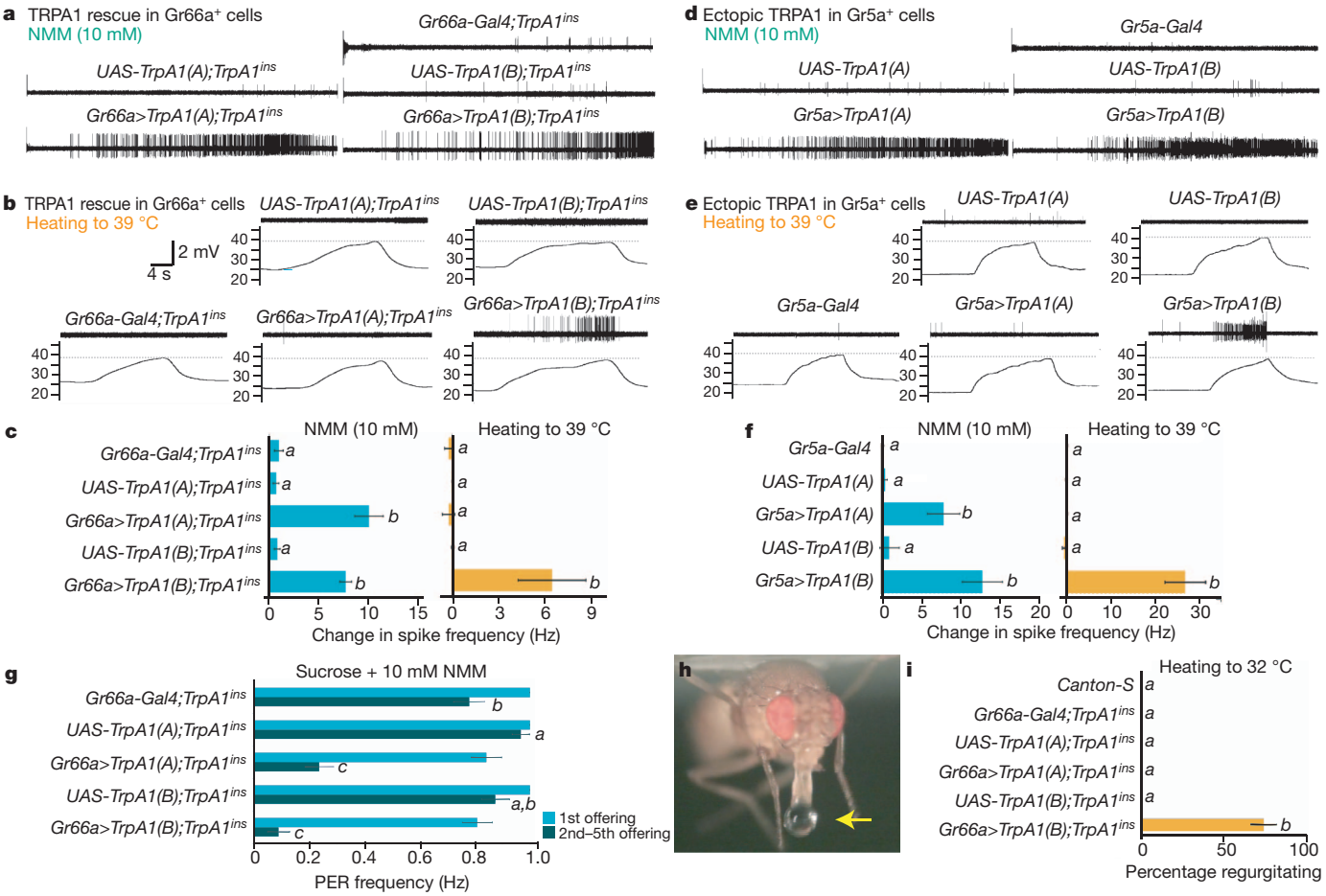


Figure 3 | TRPA1 isoform diversity determines sensory specificity of gustatory neurons. **a–c**, *TrpA1* mutant, berberine-sensitive i-type bristles expressing different TRPA1 isoforms. **a**, **b**, Electrophysiological responses to NMM (**a**) and warming (**b**). **c**, Quantification. **d–f**, L-type bristles expressing TRPA1 isoforms. Responses to NMM (**d**) and warming (**e**). **f**, Quantification. **g**, Rescue of *TrpA1* mutant behavioural response to NMM-containing food. PER, proboscis extension response. **h**, Warmth-induced regurgitation in *TrpA1* mutant rescued with TRPA1(B). **i**, Regurgitation upon warming from room temperature (~23 °C) to 32 °C. In **c**, **f**, **g** and **i**, statistically distinct groups are marked by *a*, *b* and *c*. (Tukey honestly significant difference (HSD) test, $\alpha = 0.01$). Data are mean ± s.e.m.

inside the head[5] should minimize exposure to environmental irritants. Interestingly, multiple TRPV1 and TRPM1 isoforms are present in humans and other mammals[2,17,19,20], suggesting the potential generality of isoform diversity in modulating TRP functions.

The behavioural significance of discriminating noxious from innocuous TRPA1 activators was examined by testing gustatory responses of *TrpA1* mutants rescued by chemosensor expression of each isoform. *TrpA1* mutants exhibit decreased avoidance of reactive-electrophile-containing food[4]. Each isoform rescued this behaviour (Fig. 3g). However, TRPA1(B) also triggered a nocifensive response to innocuous warming. When allowed to ingest water to satiation and warmed to ~32 °C, neither wild-type nor TRPA1(A) rescue animals showed detectable gustatory responses (Fig. 3h, i). However, warming TRPA1(B) rescue flies caused ~75% to regurgitate (Fig. 3h, i and Supplementary Movie 1). Thus, substituting TRPA1(B) for TRPA1(A) in chemosensors disrupts discrimination of noxious from innocuous stimuli and demonstrates the negative behavioural consequence of misregulated thermosensitivity.

To probe how the alternative N termini in TRPA1 confer distinct properties, conserved residues within these regions were mutated (Fig. 4a). Mutating either a cysteine (C105) or two basic residues (R113, R116) in TRPA1(A) markedly increased temperature responsiveness (Fig. 4a–c and Supplementary Figs 6 and 7). Whereas the Q_{10} of wild-type TRPA1(A) was <10, the TRPA1(A) mutants exhibited Q_{10} values of >50 (Fig. 4b, c), greater than the reported Q_{10} values of canonical thermoTRPs like TRPM8 (~24)[21] and TRPV1 (~40)[22]. In addition, the TRPA1(A) mutants conducted little current below the threshold, increasing the temperature-dependent activity differential (Fig. 4c and Supplementary Fig. 8). The enhanced sensitivities of the mutants seemed to be temperature specific, as NMM sensitivity was not increased (Supplementary Fig. 8). These data indicate that TRPA1(A) retains all the requirements for robust thermosensation, but contains a modulatory region preventing those elements from exerting their full effect.

For TRPA1(B), mutating either a conserved tryptophan or two basic residues in the N terminus yielded channels retaining robust thermosensitivity ($Q_{10} > 50$; Fig. 4b and Supplementary Fig. 6). The thresholds of the TRPA1(A) and TRPA1(B) mutants were all ~30–34 °C, within the innocuous warm range but above the ~28 °C value for wild-type TRPA1(B) (Fig. 4b, c). Thus, although TRPA1(B)-specific sequences are unnecessary for robust responsiveness to innocuous warming, they may tune channel threshold within this range.

In insect disease vectors, TRPA1 orthologues have been implicated in detecting both warmth and chemical repellents[4,23,24], cues with opposing effects on host seeking. We found the malaria mosquito *Anopheles gambiae* also contains TRPA1(A) and TRPA1(B) isoforms of differing thermosensitivity (Fig. 4d–h). In oocytes, the Q_{10} of *A. gambiae* TRPA1(A) was ~4 versus a Q_{10} of ~200 for *A. gambiae* TRPA1(B); from 27–37 °C, the Q_{10} of *A. gambiae* TRPA1(A) was ~2 versus ~60 for *A. gambiae* TRPA1(B) (Fig. 4h). *A. gambiae* TRPA1(A) yielded lower maximum heat-induced current than *A. gambiae* TRPA1(B) (Supplementary Fig. 4) and had a higher threshold

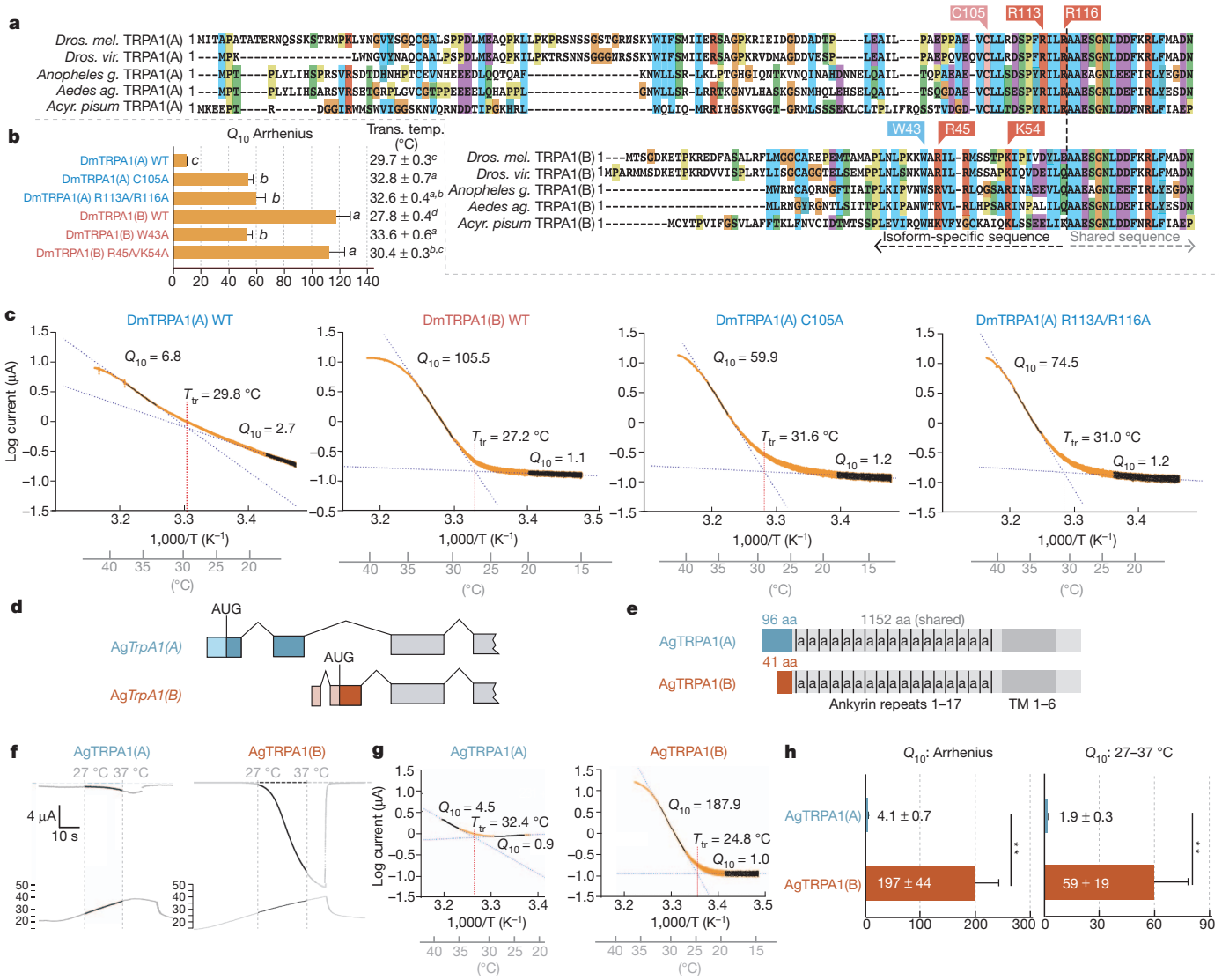


Figure 4 | Regulation of insect TRPA1 thermosensitivity by alternative N termini. **a**, TRPA1 sequence alignments. *Acyr. pisum*, *Acyrthosiphon pisum*; *Aedes ag.*, *A. aegypti*; *Anopheles g.*, *A. gambiae*; *Dros. mel.*, *D. melanogaster*; *Dros. vir.*, *D. virilis*. **b**, Q_{10} values and transition temperatures (Trans. temp.) of wild-type (WT) and mutant TRPA1. Dm, *D. melanogaster*. *a*, *b* and *c* denote statistically distinct groups (Tukey HSD, $\alpha = 0.02$). **c**, Arrhenius plots of indicated channels. **d**, *A. gambiae* (Ag)*TrpA1* gene structure. **e**, AgTRPA1 isoforms. a, ankyrin repeat; aa, amino acids. Blue and dark red indicate isoform-specific amino acids. Dark grey indicates transmembrane (TM) region. **f–h**, Temperature sensitivity of AgTRPA1(A) and AgTRPA1(B). **f**, **g**, Traces (**f**) and Arrhenius plots (**g**) of temperature-dependent current recordings at −60 mV in *Xenopus* oocytes. **h**, Q_{10} values from Arrhenius plot (left) or 27–37 °C (right) (**$P < 0.01$, *t*-test). Error bars, ± s.e.m.

(34.2 ± 1.8 °C versus 25.2 ± 0.9 °C, $P < 0.01$). *A. gambiae* TRPA1(A) also exhibited significant conductance below threshold (Fig. 4f, g). Both channels responded to electrophiles (Supplementary Fig. 9). TRPA1(A) and TRPA1(B) are conserved in other haematophagous insects including *Aedes aegypti* and *Culex quinquefasciatus* mosquitoes and *Pediculus humanus corporis* lice (Fig. 4a and Supplementary Fig. 10), which transmit dengue, West Nile fever and typhus, respectively. The functional diversity of TRPA1 provides a potential explanation for how insect vectors discriminate noxious chemicals from host-derived warmth, indicating that TRPA1 presents two distinct molecular targets for disrupting pest behaviour.

TRPA1-based electrophile detection appears to have emerged ≥500 million years ago in a common vertebrate/invertebrate ancestor[4]. However, the larger TRPA family extends to choanoflagellates, separated from animals ≥600 million years[4]. As divergent TRPA clades contain highly temperature-sensitive channels[1], thermosensitivity may be ancestral. In this scenario, the specialization of TRPA1 for noxious chemical detection would necessitate reducing thermosensitivity, consistent with the effect of the N terminus in TRPA1(A). The ability of N-terminal variation to sculpt channel properties is intriguing as the N terminus is the most divergent region of TRPA1 within insects and from insects to mammals[5].

TRPs are a large family of channels, with 27 human and 13 *Drosophila* members, which vary greatly in thermosensitivity and function[2]. Considerable diversity is evident even among closely related TRPs. In mammals, for example, TRPM8 (Q_{10} ~24; ref. 21) mediates thermosensation[3], whereas the less thermosensitive TRPM4 and TRPM5 (Q_{10} ~8.5–10; ref. 15) mediate insulin secretion[25] and TRPM7 (with no reported thermal sensitivity) is implicated in ion homeostasis[2]. The mechanisms underlying such diversification are unclear. Whereas studies of thermal sensing by TRPs have focused on identifying regions promoting thermosensitivity[1,26–29], our work indicates that regions reducing thermosensitivity are also critical. Here we find that selectively reducing the thermosensitivity of

Drosophila TRPA1 facilitates its use in a context in which thermosensitivity is undesirable. Similar mechanisms could mediate functional diversification not only among isoforms of a single TRP, but also contribute to the remarkable functional diversification observed between different TRP family channels.

METHODS SUMMARY

Fly strains and immunohistochemistry. *UAS-TrpA1(B)* and *Gr66a-Gal4* transgenic strains and the *TrpA1*^{ins} mutant have been described[4]. The *UAS-TrpA1(A)* transgene was amplified from fly complementary DNA with an isoform-specific primer (5′-TATAAAGCTTAAGCCACCATGATTACAGCTCCGGCCACGGCCA-3′) and a reverse primer (5′-GAGACTCGAGCTACATGCTCTTATTGAAGCTCAGGGCG-3′). As detailed in Methods, the *UAS-TrpA1(A)* transgene was inserted in the same genomic location used for the *UAS-TrpA1(B)* transgene to control for transgene position effects. Anti-TRPA1 immunohistochemistry was as described[4], except that the secondary antibody was incubated for 3 days.

Behaviour. The proboscis extension assay was conducted as previously described[4], with seven flies per experiment, three experiments per genotype. For heat-sensitive regurgitation, >20 flies per genotype (2–3 days old) were starved overnight with water, then glued to glass slides. After 2–3 h recovery, flies were satiated with water. Only flies drinking longer than 5 s were tested. Drinking times did not significantly differ between wild-type and rescue flies (E.C.C. and P.A.G., unpublished data), consistent with similar ingestion behaviours. Flies were heated with a radiant heater at 800 W (H-4438, Optimus) and temperature was monitored by adjacent thermocouple microprobe (IT-23, Physitemp Instruments) wrapped in fly cuticle.

Physiology. Oocyte physiology was performed as described[4], with additional details provided in Methods. Extracellular recordings of gustatory neurons were obtained by tip recording[9], as detailed in Methods.

Molecular biology. cDNA was prepared from dissected *D. melanogaster* tissue (RETROscript, Ambion) and subjected to PCR with three primers, two forward isoform-specific primers and a reverse common primer detailed in Methods. Each primer was designed to straddle a splice junction, minimizing amplification of contaminating genomic DNA. In all cases, similar results were obtained from four independent tissue preparations.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 September; accepted 16 November 2011.
Published online 4 December 2011.

1. Dhaka, A., Viswanath, V. & Patapoutian, A. TRP ion channels and temperature sensation. *Annu. Rev. Neurosci.* **29,** 135–161 (2006).
2. Wu, L. J., Sweet, T. B. & Clapham, D. E. International Union of Basic and Clinical Pharmacology. LXXVI. Current progress in the mammalian TRP ion channel family. *Pharmacol. Rev.* **62,** 381–404 (2010).
3. Daniels, R. L. & McKemy, D. D. Mice left out in the cold: commentary on the phenotype of TRPM8-nulls. *Mol. Pain* **3,** 23 (2007).
4. Kang, K. *et al.* Analysis of *Drosophila* TRPA1 reveals an ancient origin for human chemical nociception. *Nature* **464,** 597–600 (2010).
5. Hamada, F. N. *et al.* An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* **454,** 217–220 (2008).
6. Patapoutian, A., Tate, S. & Woolf, C. J. Transient receptor potential channels: targeting pain at the source. *Nature Rev. Drug Discov.* **8,** 55–68 (2009).
7. Viswanath, V. *et al.* Opposite thermosensor in fruitfly and mouse. *Nature* **423,** 822–823 (2003).
8. Cohet, Y. Epigenetic influences on the lifespan of the *Drosophila*: existence of an optimal growth temperature for adult longevity. *Exp. Gerontol.* **10,** 181–184 (1975).
9. Hodgson, E. S., Lettvin, J. Y. & Roeder, K. D. Physiology of a primary chemoreceptor unit. *Science* **122,** 417–418 (1955).
10. Weiss, L. A., Dahanukar, A., Kwon, J. Y., Banerjee, D. & Carlson, J. R. The molecular and cellular basis of bitter taste in *Drosophila*. *Neuron* **69,** 258–272 (2011).
11. von Philipsborn, A. C. *et al.* Neuronal control of *Drosophila* courtship song. *Neuron* **69,** 509–522 (2011).
12. Graveley, B. R. *et al.* The developmental transcriptome of *Drosophila melanogaster*. *Nature* **471,** 473–479 (2011).
13. Vyklicky, L. *et al.* Temperature coefficient of membrane currents induced by noxious heat in sensory neurones in the rat. *J. Physiol. (Lond.)* **517,** 181–192 (1999).
14. Tracey, W. D. Jr, Wilson, R. I., Laurent, G. & Benzer, S. *painless*, a *Drosophila* gene essential for nociception. *Cell* **113,** 261–273 (2003).
15. Talavera, K. *et al.* Heat activation of TRPM5 underlies thermal sensitivity of sweet taste. *Nature* **438,** 1022–1025 (2005).
16. Gracheva, E. O. *et al.* Molecular basis of infrared detection by snakes. *Nature* **464,** 1006–1011 (2010).
17. Gracheva, E. O. *et al.* Ganglion-specific splicing of TRPV1 underlies infrared sensation in vampire bats. *Nature* **476,** 88–91 (2011).
18. Marella, S. *et al.* Imaging taste responses in the fly brain reveals a functional map of taste category and behavior. *Neuron* **49,** 285–295 (2006).
19. Lu, G., Henderson, D., Liu, L., Reinhart, P. H. & Simon, S. A. TRPV1b, a functional human vanilloid receptor splice variant. *Mol. Pharmacol.* **67,** 1119–1127 (2005).
20. Vos, M. H. *et al.* TRPV1b overexpression negatively regulates TRPV1 responsiveness to capsaicin, heat and low pH in HEK293 cells. *J. Neurochem.* **99,** 1088–1102 (2006).
21. Brauchi, S., Orio, P. & Latorre, R. Clues to understanding cold sensation: thermodynamics and electrophysiological analysis of the cold receptor TRPM8. *Proc. Natl Acad. Sci. USA* **101,** 15494–15499 (2004).
22. Liu, B., Hui, K. & Qin, F. Thermodynamics of heat activation of single capsaicin ion channels VR1. *Biophys. J.* **85,** 2988–3006 (2003).
23. Wang, G. *et al. Anopheles gambiae* TRPA1 is a heat-activated channel expressed in thermosensitive sensilla of female antennae. *Eur. J. Neurosci.* **30,** 967–974 (2009).
24. Maekawa, E. *et al.* The role of proboscis of the malaria vector mosquito *Anopheles stephensi* in host-seeking behavior. *Parasit. Vectors* **4,** 10 (2011).
25. Uchida, K. & Tominaga, M. The role of thermosensitive TRP (transient receptor potential) channels in insulin secretion. *Endocr. J.* doi:10.1507/endocrj.EJ11-0130 (2011).
26. Grandl, J. *et al.* Pore region of TRPV3 ion channel is specifically required for heat activation. *Nature Neurosci.* **11,** 1007–1013 (2008).
27. Grandl, J. *et al.* Temperature-induced opening of TRPV1 ion channel is stabilized by the pore domain. *Nature Neurosci.* **13,** 708–714 (2010).
28. Yang, F., Cui, Y., Wang, K. & Zheng, J. Thermosensitive TRP channel pore turret is part of the temperature activation pathway. *Proc. Natl Acad. Sci. USA* **107,** 7083–7088 (2011).
29. Yao, J., Liu, B. & Qin, F. Modular thermal sensors in temperature-gated transient receptor potential (TRP) channels. *Proc. Natl Acad. Sci. USA* **108,** 11109–11114 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the P.A.G. laboratory, H. Garrity, L. Griffith, L. Huang, J. Rodriguez and M. Rosbash for helpful comments, and F. Marion-Poll and A. Dahanukar for guidance with tip-recording. This work was supported by grants from the National Science Foundation (IOS-1025307), National Institute of Mental Health (EUREKA R01 MH094721) and National Institute of Neurological Disorders and Stroke (NINDS) (PO1 NS044232) to P.A.G., a National Research Service Award from NINDS to V.C.P. (F31 NS071897-02) and the Boston College DeLuca Professorship to M.A.T.M.

Author Contributions: K.K., V.C.P. and P.A.G. designed experiments. K.K. performed molecular biology, genetics and oocyte physiology. V.C.P. performed genetics and sensory neuron electrophysiology. E.C.C. performed genetics and behavioural experiments. A.M.D. performed behavioural experiments. L.N. performed immunohistochemistry. A.M.J., K.R. and M.A.T.M. grew and harvested mosquitoes. P.A.G. performed bioinformatics. K.K., V.C.P. and P.A.G. wrote the paper.

Author Information TRPA1(A) sequence is deposited in GenBank under accession number JQ015263. 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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to P.A.G. (pgarrity@brandeis.edu).

METHODS

Fly strains and immunohistochemistry. The *UAS-TrpA1(A)* transgene was inserted into the genome by site-specific transgenesis[30] at the same landing site as *UAS-TrpA1(B)*, attp16 (ref. 31).

Behavioural analysis. Chemicals used in incapacitation assays were sucrose (Calbiochem LC8510), sorbitol (Sigma S-1876), ficoll (Sigma F-4375), agarose (Invitrogen 15510-027), caffeine (Sigma C0750), NMM (Sigma 389412), isopropanol (100%, J. T. Baker 9083-03), ethanol (100%, Decon Lab 2716) and allyl isothiocyanate (95%, Sigma 377430).

Characterization of TRPA1 isoforms in *Xenopus* oocytes. TRPA1 currents were recorded as described[4,5]. To evaluate temperature sensitivities, oocytes were perfused in the recording buffer (96 mM NaCl, 1 mM $MgCl_2$, 4 mM KCl, and 5 mM HEPES, pH 7.6), the temperature of which was increased ~0.5 °C s^{-1} from 10 to 45 °C by SC-20 in-line heater/cooler (Warner Instruments) with a CL100 bipolar temperature controller (Warner Instruments). Temperature-evoked current was recorded at −60 mV. From the recorded current, Q_{10} was calculated as described[13,16]. Arrhenius $Q_{10} = 10^{(10(-S_{\mathrm{arrhe}})/(T_1 T_2))}$, where S_{arrhe} is the slope of linear phase of an Arrhenius plot between absolute temperatures, T_1 and T_2. Transition temperature was assessed as the temperature at which the least-squares fit lines from the two linear phases intersect[13,16]. Q_{10} from 27–37 °C was calculated from currents at temperatures of interest using the equation, $Q_{10} = (I_2/I_1)^{10/(T_2 - T_1)}$, where $I_1$and I_2 are currents observed at temperatures of T_1 and T_2, respectively. Q_{10} determinations were validated by using *Crotalus atrox* TRPA1 (ref. 16) as a control with known Q_{10} (K.K. and P.A.G., unpublished data). To assess sensitivity to NMM, voltage across the membrane was initially held at −80 mV, and a 300-ms voltage ramp (−80 mV to 80 mV) per second was applied. The oocytes were perfused for 1 min with the recording buffer containing indicated concentrations of NMM with 30-s washes between NMM applications. Current amplitudes at −80 mV after application of each NMM concentration were fitted to the Hill equation through Sigmaplot 10. The first coding exon of *A. gambiae TrpA1(B)* was chemically synthesized (Genscript).

Gustatory neuron electrophysiology. Extracellular recordings of gustatory neurons were obtained using the tip-recording method[9]. Adult female flies, aged 1–4 days, were prepared by inserting a glass reference electrode containing *Drosophila* Ringer's solution into the thorax and advancing the electrode through the head to the labellum. A glass recording electrode with an ~15-μm opening was used to apply tastants to individual sensilla. Raw signals were amplified using a TasteProbe preamplifier (Syntech) and were digitized and analysed using a PowerLab data-acquisition system with LabChart software (ADInstruments). Amplified signals were digitized at a rate of 20 kb s^{-1} and filtered using a 100–3,000 Hz band-pass filter before analysis. Individual action potentials were sorted using a visually adjusted threshold and average spike rate was calculated beginning 200 ms after electrode contact. Recording times varied by experiment: berberine chloride and sucrose positive controls, 5 s; electrolyte only, 20 s; NMM on i-type bristles, 60 s; NMM on L-type bristles, 120 s; heat ramps, >60 s. For heat-ramp experiments, recordings were performed using electrolyte only as tastant. After ~30 s of recording to determine baseline activity, heat was applied manually to the fly using a radiant heater (PRESTO HeatDish, National Presto Industries). Application of heat was maintained for ~10–30 s and the distance between the heat source and the preparation was reduced to obtain a temperature of ≥39 °C. Bristle temperature was estimated using thermocouple microprobe (IT-23, Physitemp Instruments) wrapped in fly cuticle. All tastants were dissolved in 30 mM tricholine citrate as the electrolyte to inhibit the activity of the water cell in L-type bristles[32]. Tastants were stored at −20 °C and aliquots maintained at 4 °C for up to 1 week. For all experiments, a positive control was used to confirm the viability of the target bristle. For i-type bristles, 1 mM berberine chloride was used as control. For L-type bristles, 30 mM sucrose was used. Individual tastant presentations were separated by a minimum delay of 60 s. At least two animals and six bristles were examined for each condition.

Molecular biology. Primers for RT–PCR reactions: *D. melanogaster TrpA1(A)* forward, 5′-GCCGGAACAGCAAGTATT3-3′; *D. melanogaster TrpA1(B)* forward, 5′-GTGGACTATCTGGAGGCG-3′; *D. melanogaster TrpA1* common reverse, 5′-TATCCT TCGCATTAAAGTCGC-3′.

Mutagenesis of *Drosophila* TRPA1 was performed as described[4]. Briefly, for a desired mutation, each of two mutually complementary mutant primers was paired for PCR with a primer (outer primer) that anneals outside of either SalI or HpaI restriction recognition site. The two resulting PCR fragments that overlap only in the region of the two mutant primers were combined and served as template for the next PCR reaction that contained only outer primers. The second PCR product was digested by SalI and HpaI, and subsequently replaced the corresponding wild type region of TRPA1 cDNA. The fragment between the two restriction sites was sequenced. Sequences were aligned using MUSCLE 3.7[33].

30. Groth, A. C., Fish, M., Nusse, R. & Calos, M. P. Construction of transgenic *Drosophila* by using the site-specific integrase from phage ϕC31. *Genetics* **166,** 1775–1782 (2004).
31. Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. & Perrimon, N. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature Genet.* **40,** 476–483 (2008).
32. Wieczorek, H. & Wolff, G. The labellar sugar receptor of *Drosophila*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **164,** 825–834 (1989).
33. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32,** 1792–1797 (2004).

doi:10.1038/nature10660

Antibody-based protection against HIV infection by vectored immunoprophylaxis

Alejandro B. Balazs[1], Joyce Chen[1], Christin M. Hong[1], Dinesh S. Rao[2], Lili Yang[1] & David Baltimore[1]

Despite tremendous efforts, development of an effective vaccine against human immunodeficiency virus (HIV) has proved an elusive goal. Recently, however, numerous antibodies have been identified that are capable of neutralizing most circulating HIV strains[1–5]. These antibodies all exhibit an unusually high level of somatic mutation[6], presumably owing to extensive affinity maturation over the course of continuous exposure to an evolving antigen[7]. Although substantial effort has focused on the design of immunogens capable of eliciting antibodies *de novo* that would target similar epitopes[8–10], it remains uncertain whether a conventional vaccine will be able to elicit analogues of the existing broadly neutralizing antibodies. As an alternative to immunization, vector-mediated gene transfer could be used to engineer secretion of the existing broadly neutralizing antibodies into the circulation. Here we describe a practical implementation of this approach, which we call vectored immunoprophylaxis (VIP), which in mice induces lifelong expression of these monoclonal antibodies at high concentrations from a single intramuscular injection. This is achieved using a specialized adeno-associated virus vector optimized for the production of full-length antibody from muscle tissue. We show that humanized mice receiving VIP appear to be fully protected from HIV infection, even when challenged intravenously with very high doses of replication-competent virus. Our results suggest that successful translation of this approach to humans may produce effective prophylaxis against HIV.

Previous efforts to engineer humoral immunity using adeno-associated virus (AAV)-based vectors resulted in modest antibody production[11], which was subsequently improved through the use of alternative capsids[12] and self-complementary AAV (scAAV) vectors[13] that increase expression at the expense of carrying capacity. Recently, scAAV vectors were used to direct expression of simian immunodeficiency virus (SIV)-neutralizing immunoadhesins consisting of small, artificially fused antibody fragments[14]. However, the efficacy of this prophylaxis was limited by an endogenous immune response directed against the immunoadhesin proteins. To ask whether newer capsid serotypes and vector configurations might support long-lived expression of full-length human antibodies from muscle, we produced AAV vectors with the capsid from serotype 8 (ref. 15) that expressed either luciferase or 4E10 HIV neutralizing antibody driven from cytomegalovirus (CMV) promoters and administered them through a single injection of the gastrocnemius muscle (Fig. 1a). Within one week of vector administration, either luciferase or antibody gene expression was detectable (Supplementary Fig. 1a, left and right respectively). Expression continued to rise, achieving maximum levels after 12–16 weeks and then decreasing two- to threefold before stabilizing for the duration of the 64-week study. Given the long-lived nature of this expression, it seemed possible that these vectors could be used to engineer lifelong humoral immunity provided by full-length, fully human antibodies. Hence, we performed a systematic process of vector and transgene optimization to improve the expression characteristics of this system (Supplementary Information). The heavy- and light-chain variable regions of the HIV-neutralizing b12 antibody were cloned into the vector, and AAV stock was produced for intramuscular administration of 1×10^{11} genome copies into the gastrocnemius muscle of two immunodeficient and two immunocompetent mouse strains: NOD/SCID/γc (NSG), Rag2/γc (RAG), C57BL/6 (B6) and

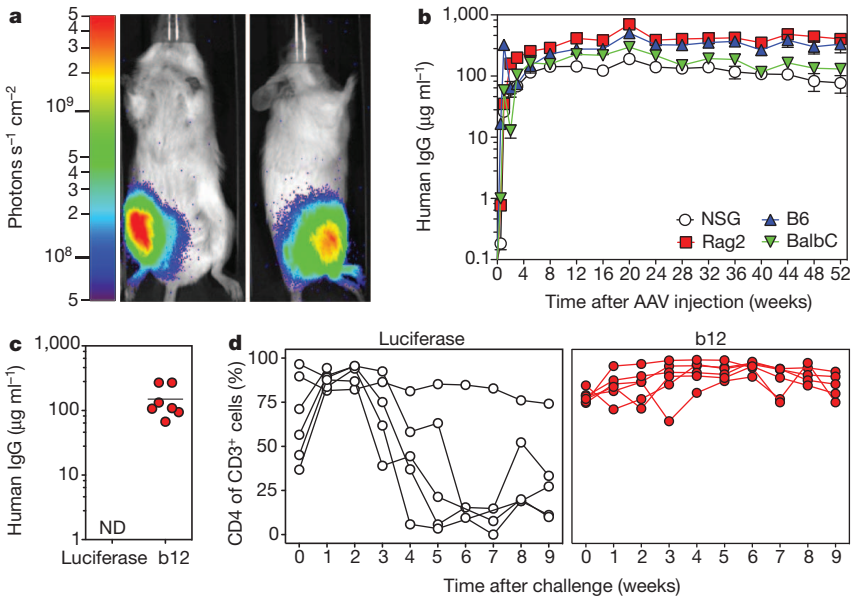


Figure 1 | VIP protects against HIV-mediated CD4 cell depletion in humanized mice. **a**, Xenogen imaging of a representative Rag2/γc mouse 15 weeks after intramuscular injection of 1×10^{10} genome copies of AAV2/8 expressing luciferase. **b**, Quantification of human IgG by ELISA after intramuscular injection of 1×10^{11} genome copies of the optimized expression vector producing b12-IgG in either immunodeficient NOD/SCID/γc (NSG) and Rag2/γc (Rag2) or immunocompetent C57BL/6 (B6) and Balb/C mice (plot shows mean and standard error, $n = 4$). **c**, Concentration of human IgG in circulation as measured by ELISA on serum samples taken 6 weeks after intramuscular injection of vector expressing either luciferase or b12-IgG (ND, not detected). **d**, Depletion of CD4 T cells in humanized mice after intraperitoneal challenge with 10 ng p24 NL4-3 into animals that received AAV2/8 vectors expressing luciferase (left) or b12-IgG1 (right) 6 weeks earlier ($n = 6$).

[1]Division of Biology, California Institute of Technology, 1200 East California Boulevard, Pasadena, California 91125, USA. [2]Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California at Los Angeles, 10833 Le Conte Avenue, Los Angeles California 90095, USA.

Balb/C. Mice produced the encoded antibody at serum concentrations that were 100-fold higher than the levels achieved with the non-optimized vector, and this level of expression persisted for at least 52 weeks (Fig. 1b compared with Supplementary Fig. 1a, right). In agreement with previous studies of AAV-induced tolerance in mice[16], we detected very limited mouse antibodies raised against human b12-IgG in B6 mice, whereas Balb/C animals generated detectable mouse antibodies against the transgene (data not shown) that did not appear to impact human IgG levels.

To test the ability of VIP to protect mice from challenge *in vivo*, we adapted a previously described humanized mouse model[17] that exhibits CD4 cell depletion following challenge with replication-competent HIV (Supplementary Fig. 5). We administered vector expressing either luciferase or b12 antibody to NSG mice, producing stable serum b12 antibody concentrations of approximately 100 μg ml^{-1} within 6 weeks (Fig. 1c). These mice were adoptively populated with expanded human peripheral blood mononuclear cells (huPBMCs), which engrafted over a period of 2 weeks. Mice were then challenged by intraperitoneal injection of the NL4-3 strain of HIV. After HIV challenge, most mice expressing luciferase showed dramatic loss of CD4 cells whereas mice expressing b12 antibody showed no CD4 cell depletion (Fig. 1d).

To compare the protective abilities of the historically available broadly neutralizing antibodies, vectors expressing b12, 2G12, 4E10 and 2F5 were produced and administered to NSG mice. Seven weeks after administration, NSG mice produced 20–250 μg ml^{-1} of the indicated antibodies (Fig. 2a). Interestingly, *in vivo* serum concentrations of 4E10 and 2F5 were somewhat lower than those of b12 and 2G12, despite comparable expression *in vitro* (Supplementary Fig. 3b), perhaps resulting from the previously described self-reactivity of these clones[18]. Transduced mice were adoptively populated with huPBMCs, challenged by intravenous injection with HIV and sampled weekly to quantify CD4 cell depletion over time (Fig. 2b). Animals expressing b12 were completely protected from infection, whereas those expressing 2G12, 4E10 and 2F5 were partly protected. Groups demonstrating partial protection consisted of animals with delayed CD4 cell depletion as well as animals that maintained high CD4 cell levels throughout the course of the experiment. Interestingly, mice expressing 250 μg ml^{-1} of the 2G12 antibody were only partly protected, despite antibody levels being over 300-fold higher than previously established half-maximum inhibitory concentration (IC_{50}) values for this antibody-strain combination *in vitro*[19]. Eight weeks after challenge, mice were killed and paraffin-embedded spleen sections underwent immunohistochemical staining for the HIV-expressed p24 antigen to quantify the extent of infection (Fig. 2c). Remarkably, mice expressing b12 had no detectable p24-expressing cells, whereas those expressing other antibodies exhibited significant positive staining (Fig. 2d).

To determine the robustness of protection mediated by VIP, a large cohort of mice expressing b12 antibody were adoptively populated with huPBMCs. Before challenge, all mice expressed high levels of human IgG, presumably owing to engrafted human B-cells (Supplementary Fig. 6a), but only those receiving the b12-expressing vector produced IgG specific for gp120, which reached 100 μg ml^{-1} (Supplementary Fig. 6b). Mock-infected mice expressing either luciferase or b12 demonstrated consistent high-level CD4 cell engraftment throughout the course of the experiment, showing that transgene toxicity was not contributing to CD4 cell loss (Fig. 3). In contrast, mice expressing luciferase that received 1 ng of HIV experienced rapid and extensive CD4 cell depletion. At higher doses, infection in luciferase-expressing mice became more consistent and resulted in depletion of CD4 cells below the level of detection in some cases (25, 125 ng doses). Remarkably, all mice expressing b12 demonstrated protection from CD4 cell loss, despite receiving HIV doses over 100-fold higher than necessary to deplete seven out of eight control animals (Fig. 3).

As newer anti-HIV antibodies have become available, we have compared the relative efficacy of b12 to VRC01 antibody. VRC01 neutralizes over 90% of circulating HIV strains *in vitro*[1], making it an excellent candidate for human trials. We administered decreasing doses of vector expressing either b12 or VRC01 to NSG mice and monitored expression of the antibodies over time. For both antibodies, we observed clear dose-dependent expression at all time points analysed (Supplementary Fig. 7a and Fig. 4, top). Mice expressing luciferase or antibodies at various levels were adoptively populated with huPBMCs. Just before challenge, a gp120-specific enzyme-linked immunosorbent assay (ELISA) confirmed the effective antibody concentration in each group (Supplementary Fig. 7b and Fig. 4, middle). After intravenous challenge with 10 ng of HIV, CD4 cells were monitored to determine the impact of antibody concentration. An average

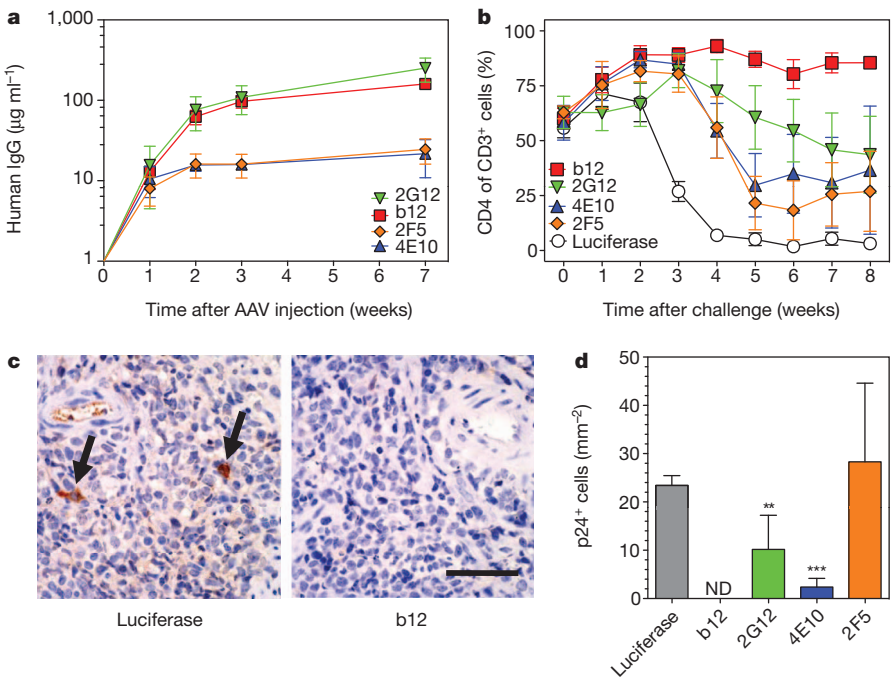


Figure 2 | Comparison of protection mediated by various broadly neutralizing HIV antibodies. **a**, Concentration of antibody in circulation as measured by total human IgG ELISA on serum samples taken after intramuscular injection of vectors expressing four broadly neutralizing HIV antibodies ($n = 8$). **b**, Comparison of the relative effectiveness of four broadly neutralizing HIV antibodies in protecting huPBMC-NSG humanized mice against CD4 cell depletion after intravenous HIV challenge with 5 ng p24 NL4-3 ($n = 8$). **c**, HIV p24 detection by immunohistochemical staining of sections taken from spleens 8 weeks after challenge. Arrows indicate cells scored as positive for p24 expression. Scale bar, 40 μm. **d**, Quantification of immunohistochemical staining of spleen denoting the relative frequency of p24-expressing cells in spleens of infected animals. ND, not detected. Asterisks indicate outcomes significantly different from luciferase control mice versus mice expressing antibodies by two-tailed *t*-test ($n = 4$–6) $^{**}P < 0.01$, $^{***}P < 0.0001$. **a**, **b**, Show mean and s.e.m.; **d** shows mean and s.d.

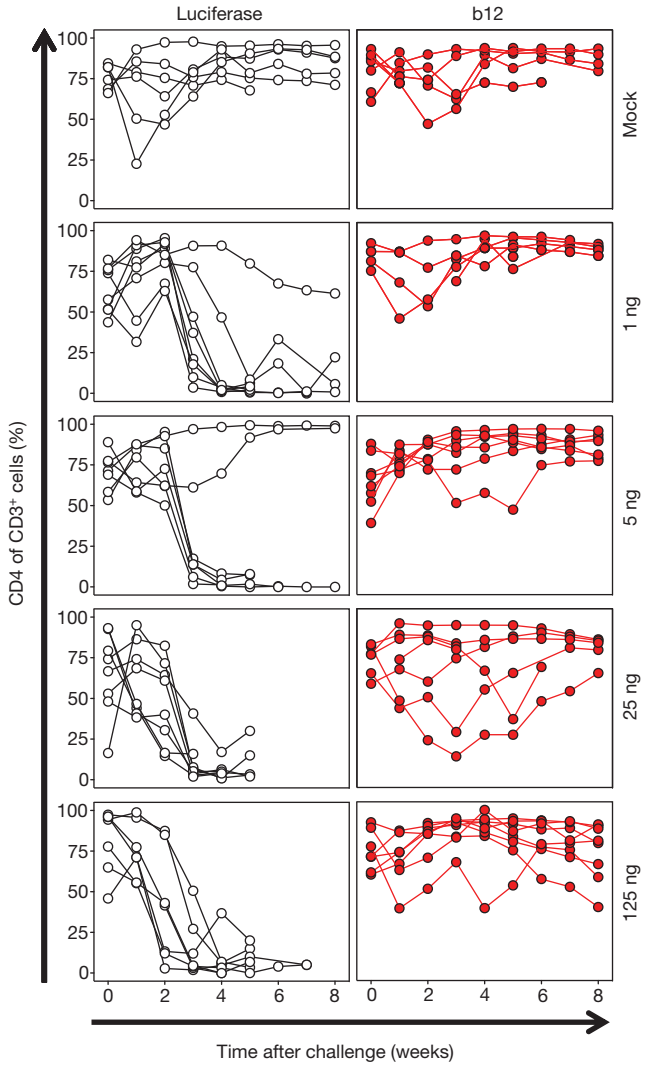


Figure 3 | Robustness of CD4 cell protection mediated by b12 antibody. CD4 cell depletion in huPBMC-NSG humanized mice as a result of intravenous challenge with the dose of NL4-3 indicated on the far right. Mice expressing luciferase (left plots) were susceptible to CD4 cell loss, whereas those expressing b12 (right plots) demonstrated protection from HIV at all doses ($n = 8$).

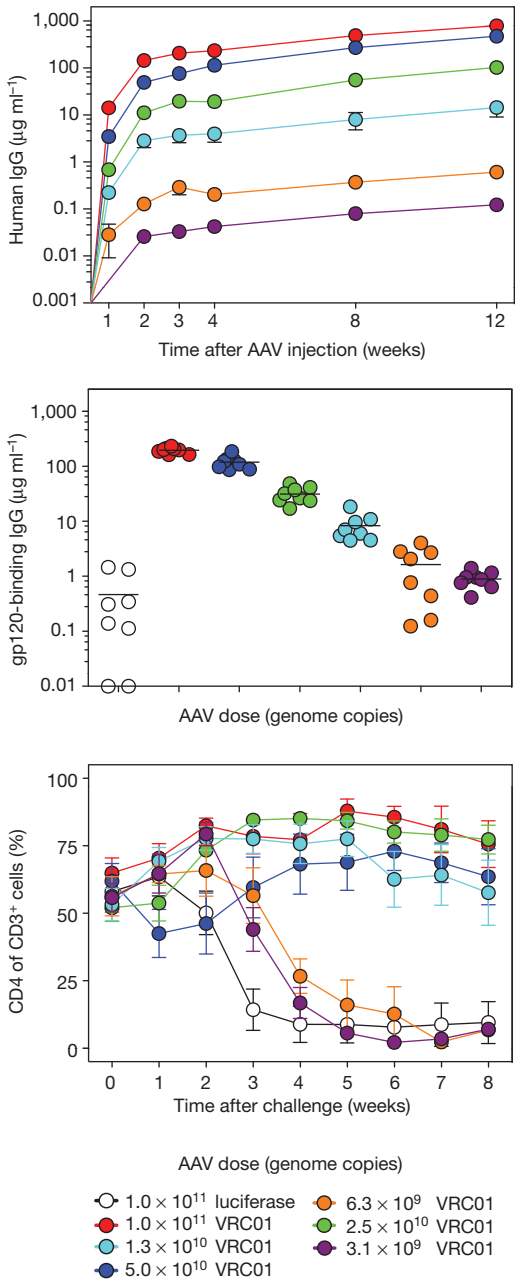


Figure 4 | Determination of the minimum protective dose of VRC01 *in vivo*. Top, VRC01 expression over time as a function of dose as determined by total human IgG ELISA on serum samples taken after AAV administration ($n = 8$). Mice receiving luciferase-expressing vector exhibited no detectable human antibodies ($n = 12$). Middle, concentration of VRC01 in serum 1 day before challenge, 3 weeks after adoptive transfer of human PBMCs and 15 weeks after intramuscular administration of the indicated dose of AAV, as determined by a gp120-specific ELISA to measure the fraction of antibodies capable of binding HIV ($n = 8$–12). Bottom, CD4 cell depletion in huPBMC-NSG humanized mice as a result of intravenous challenge with 10 ng of NL4-3 into animals expressing a range of VRC01, demonstrating the minimum dose of antibody necessary to protect against infection. Top and bottom plots show mean and standard error, middle plot shows individual animals and mean ($n = 8$–12).

b12 concentration of 34 μg ml^{-1} and VRC01 concentration of 8.3 μg ml^{-1} protected mice from infection (Supplementary Fig. 7c and Fig. 4, bottom). Groups expressing lower concentrations of b12 and VRC01 were only partly protected, with several animals showing no detectable loss of CD4 cells and others exhibiting delayed CD4 cell depletion.

Here we demonstrate that broadly neutralizing human antibodies expressed by VIP are capable of protecting animals from even high-dose HIV exposure *in vivo*. Human-to-human HIV transmission rates vary with behaviour but do not generally exceed one per hundred heterosexual exposures[20], and recent studies have demonstrated that infections are generally initiated by a single founder strain[21]. Humanized mice with b12 serum concentrations of 100 μg ml^{-1} were resistant to HIV infection at challenge doses 100-fold higher than necessary to infect the vast majority of animals, suggesting that this level of protection may far exceed what would be necessary to provide protection against HIV infection in humans. In contrast to previous approaches, VIP produces full-length antibodies that are identical in sequence to those produced by the immune system[22]. Recent experiments have suggested that full-length antibody structures possess superior *in vitro* neutralization activity compared with modified architectures such as immunoadhesins[23]. Use of such naturally occurring antibody architectures should also reduce immune responses against the transgene, which were previously shown to reduce the effectiveness of prophylaxis against SIV[14].

Our results demonstrate that VIP administration results in long-lived production of human antibodies at super-prophylactic levels in immunocompetent animals. Clinical trials involving AAV have demonstrated remarkable success when targeting immunoprivileged sites such as retinal tissue[24], but transduction of liver resulted in an

adaptive immune response against vector capsid[25]. Studies in non-human primates have shown that the elicitation of capsid-specific cytotoxic T-lymphocytes is limited to AAV capsids that exhibit heparin-binding activity[26]. Interestingly, serotypes lacking heparin-binding activity, including AAV8, did not induce CTL responses, suggesting that AAV8-based vectors, like the one we have used, may circumvent previously observed immunological barriers to long-term transduction. Additionally, in contrast to liver transduction, administration of AAV by intramuscular injection has been shown to result in very long-lived, albeit low-level, factor IX expression[27], suggesting that the route of administration can significantly impact the duration of expression. Although the expression level attainable in humans remains to be determined, it is worth noting that the significantly longer serum half-life of a human antibody in humans[28], as opposed to mice[29], may result in higher steady-state levels than those observed in the present study. Regardless of this, our results suggest that even if VIP administration in humans results in serum antibody concentrations 100-fold lower than those observed in mice, it may still confer protection against HIV infection.

Given the urgency that exists in combating the ongoing global HIV pandemic and the incremental progress towards a vaccine, novel modalities of prophylaxis must be explored towards solving this global health crisis. Our work demonstrates the feasibility of directly translating the existing repertoire of broadly neutralizing antibodies into functional immunoprophylaxis with robust protective abilities *in vivo*. As more potent broadly neutralizing HIV antibodies continue to be isolated[30], VIP can deliver these in concert with existing antibodies to provide increased potency, broader coverage and greater resistance to escape mutations. This approach may find broad use in the rapid development of effective prophylaxis against any existing or future infectious disease for which broadly neutralizing antibodies can be isolated. Beyond infectious diseases, VIP can be applied to therapeutic regimens in which continuous production of monoclonal antibodies *in vivo* is desirable. Given the level of protection that VIP has demonstrated *in vivo*, we believe that highly effective prophylaxis through expression of existing monoclonal antibodies against HIV in humans is achievable.

METHODS SUMMARY

AAV2/8 was produced by transient transfection and purification from culture supernatant by PEG precipitation and caesium chloride ultracentrifugation. Virus was quantified by quantitative PCR (qPCR) against CMV sequences and functionally validated *in vitro* to confirm gene expression before use *in vivo*. Mice were given single injections with purified vector in the gastrocnemius muscle. Antibody concentration in the serum was determined using an ELISA specific for either total human IgG or human IgG against HIV-gp120. Humanized mice expressing antibodies were produced by adoptive transfer of expanded huPBMCs into mice previously transduced with AAV vectors. HIV challenge was performed by intraperitoneal or intravenous injection, and blood was sampled weekly to determine the ratio of CD4 to CD8 cells by flow cytometry.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 25 July; accepted 21 October 2011.
Published online 30 November 2011.

1. Wu, X. *et al.* Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science* **329,** 856–861 (2010).
2. Walker, L. M. *et al.* Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. *Science* **326,** 285–289 (2009).
3. Scheid, J. F. *et al.* Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. *Science* **333,** 1633–1637 (2011).
4. Wu, X. *et al.* Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. *Science* **333,** 1593–1602 (2011).
5. Walker, L. M. *et al.* Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477,** 466–470 (2011).
6. Zhou, T. *et al.* Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. *Science* **329,** 811–817 (2010).
7. Scheid, J. F. *et al.* Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. *Nature* **458,** 636–640 (2009).
8. Kwong, P. D. & Wilson, I. A. HIV-1 and influenza antibodies: seeing antigens in new ways. *Nature Immunol.* **10,** 573–578 (2009).
9. Burton, D. R. *et al.* HIV vaccine design and the neutralizing antibody problem. *Nature Immunol.* **5,** 233–236 (2004).
10. Dormitzer, P. R., Ulmer, J. B. & Rappuoli, R. Structure-based antigen design: a strategy for next generation vaccines. *Trends Biotechnol.* **26,** 659–667 (2008).
11. Lewis, A. D., Chen, R., Montefiori, D. C., Johnson, P. R. & Clark, K. R. Generation of neutralizing activity against human immunodeficiency virus type 1 in serum by antibody gene transfer. *J. Virol.* **76,** 8769–8775 (2002).
12. Fang, J. *et al.* Stable antibody expression at therapeutic levels using the 2A peptide. *Nature Biotechnol.* **23,** 584–590 (2005).
13. McCarty, D. M. Self-complementary AAV vectors; advances and applications. *Mol. Ther.* **16,** 1648–1656 (2008).
14. Johnson, P. R. *et al.* Vector-mediated gene transfer engenders long-lived neutralizing activity and protection against SIV infection in monkeys. *Nature Med.* **15,** 901–906 (2009).
15. Gao, G. P. *et al.* Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl Acad. Sci. USA* **99,** 11854–11859 (2002).
16. Breous, E., Somanathan, S. & Wilson, J. M. BALB/c mice show impaired hepatic tolerogenic response following AAV gene transfer to the liver. *Mol. Ther.* **18,** 766–774 (2010).
17. Kumar, P. *et al.* T cell-specific siRNA delivery suppresses HIV-1 infection in humanized mice. *Cell* **134,** 577–586 (2008).
18. Haynes, B. F. *et al.* Cardiolipin polyspecific autoreactivity in two broadly neutralizing HIV-1 antibodies. *Science* **308,** 1906–1908 (2005).
19. Binley, J. M. *et al.* Comprehensive cross-clade neutralization analysis of a panel of anti-human immunodeficiency virus type 1 monoclonal antibodies. *J. Virol.* **78,** 13232–13252 (2004).
20. Wawer, M. J. *et al.* Rates of HIV-1 transmission per coital act, by stage of HIV-1 infection, in Rakai, Uganda. *J. Infect. Dis.* **191,** 1403–1409 (2005).
21. Salazar-Gonzalez, J. F. *et al.* Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing. *J. Virol.* **82,** 3952–3970 (2008).
22. Fang, J. *et al.* An antibody delivery system for regulated expression of therapeutic levels of monoclonal antibodies *in vivo*. *Mol. Ther.* **15,** 1153–1159 (2007).
23. West, A. P. Jr, Galimidi, R. P., Gnanapragasam, P. N. P. & Bjorkman, P. J. Single chain Fv-based anti-HIV proteins: potential and limitations. *J. Virol.* doi:10.1128/JVI.05848-11 (19 October 2011).
24. Maguire, A. M. *et al.* Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N. Engl. J. Med.* **358,** 2240–2248 (2008).
25. Manno, C. S. *et al.* Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nature Med.* **12,** 342–347 (2006).
26. Vandenberghe, L. H. *et al.* Heparin binding directs activation of T cells against adeno-associated virus serotype 2 capsid. *Nature Med.* **12,** 967–971 (2006).
27. Jiang, H. *et al.* Evidence of multiyear factor IX expression by AAV-mediated gene transfer to skeletal muscle in an individual with severe hemophilia B. *Mol. Ther.* **14,** 452–455 (2006).
28. Morell, A., Terry, W. D. & Waldmann, T. A. Metabolic properties of IgG subclasses in man. *J. Clin. Invest.* **49,** 673–680 (1970).
29. Petkova, S. B. *et al.* Enhanced half-life of genetically engineered human IgG1 antibodies in a humanized FcRn mouse model: potential application in humorally mediated autoimmune disease. *Int. Immunol.* **18,** 1759–1769 (2006).
30. Diskin, R. *et al.* Increasing the potency and breadth of an HIV antibody using structure-based rational design. *Science* doi:10.1126/science.1213782 (27 October 2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank J. Wilson for AAV8-related plasmids and assistance, D. Burton for b12 and 2G12 expression plasmids, G. Nabel for 4E10, 2F5 and VRC01 expression plasmids, and the Caltech Protein Expression Center for providing purified antibodies. The following reagents were obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH: pNL4-3 from M. Martin, and TZM-bl cells from J. Kappes and X. Wu. We thank J. Bloom, D. Kotton, D. Majumdar, G. Mostoslavsky, R. O'Connell and A. Sigal for comments, and other members of the Baltimore laboratory for their assistance in performing this work. This project was supported by the Bill and Melinda Gates Foundation through Grand Challenges in Global Health Initiative (awarded to D.B.) Grand Challenge grant 37866 and by the National Institutes of Health (HHSN266200500035C) through a contract from the National Institute of Allergy and Infectious Disease (NIAID) and by the Joint Center for Translational Medicine. A.B.B. is supported by amfAR postdoctoral research fellowship 107756-47-RFVA. D.S.R. is supported by career development award 1K08CA133521 from the National Institutes of Health.

Author Contributions A.B.B. and D.B. conceived the study with assistance from L.Y. A.B.B. designed the experiments. A.B.B., J.C. and C.M.H. performed experiments. A.B.B., J.C. and C.M.H. analysed the data. D.S.R. performed immunohistochemistry and analysis. A.B.B. and D.B. wrote the paper with contributions from all authors.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to D.B. (baltimo@caltech.edu).

METHODS

Construction and cloning of modular AAV transfer vectors. To construct the AAV transfer vector, oligonucleotides encoding the 145-base-pair (bp) AAV2-derived inverted terminal repeat 1 (ITR1) in the 'flip' orientation and ITR2 in the 'flop' orientation flanked by unique restriction sites were synthesized (Integrated DNA Technologies) and annealed before ligation into PBR322 plasmid vector. Subsequently, promoters, transgenes and polyadenylation signals flanked by compatible sites were amplified by PCR and cloned between the ITRs, resulting in a modular AAV transfer vector in which unique combinations of restriction sites flanked each element.

AAV virus production and purification. AAV8 was purified from culture supernatants as described[31,32] with some modifications. 293T cells were maintained in DMEM medium supplemented with 10% fetal bovine serum, 1% penicillin–streptomycin mix (Mediatech) and 1% glutamine (Mediatech) in a 5% CO_2 incubator at 37 °C. Three days before transfection, four 15 cm plates were seeded with 3.75×10^6 cells each in 25 ml media. Two hours before transfection, media was changed to 15 ml of fresh media. The AAV backbone vector was co-transfected with helper vectors pHELP (Applied Viromics) and pAAV 2/8 SEED (University of Pennsylvania Vector Core) at a ratio of 0.25:1:2 using BioT transfection reagent (Bioland Scientific). The total amount of DNA used per transfection was 80 µg. Five AAV virus collections were performed at 36, 48, 72, 96 and 120 h after transfection. For each time point, media was filtered through a 0.2 µm filter and 15 ml of fresh media was gently added to the plate. After collection, approximately 75 ml of 5× PEG solution (40% polyethylene glycol, 2.5 M NaCl) was added to the total volume of supernatant collected (~300 ml) and the virus was precipitated on ice for at least 2 h (ref. 33). Precipitated virus was pelleted at 7,277*g* for 30 min (Sorvall RC 3B Plus, H-6000A rotor) and re-suspended in 1.37 g ml^{-1} caesium chloride. Resuspended virus was split evenly into two Quick-Seal tubes (Beckman) and spun at 329,738*g* at 20 °C for 24 h (Beckman Coulter, Optima LE-80K, 70Ti rotor). Fractions of 100–200 µl were collected in a 96-well flat-bottom tissue culture plate, and a refractometer was used to quantify the refractive index of 5 µl of each fraction. Wells exhibiting refractive indexes between 1.3755 and 1.3655 were combined and diluted to a final volume of 15 ml using Test Formulation Buffer 2 (TFB2, 100 mM sodium citrate, 10 mM Tris, pH 8)[34]. Virus was loaded onto 100 kDa MWCO centrifugal filters (Millipore) and subjected to centrifugation at 500*g* at 4 °C until 1 ml retentate remained. Retained virus was then again diluted to 15 ml total volume in TFB2 and this process was repeated such that the virus was washed three times. Final retentate volume was between 500–1000 µl total, which was aliquoted and stored at −80 °C.

AAV quantification and functional validation. Purified AAV was quantified by qPCR essentially as described[35] with the following modifications. Frozen aliquots of AAV were thawed and diluted tenfold in digestion buffer containing 10 units of DNase I (Roche) and incubated at 37 °C for 30 min. DNase-digested virus was serially diluted and 5 µl of each dilution was used in a 15 µl qPCR reaction with PerfeCTa SYBR Green SuperMix, ROX (Quanta Biosciences) and primers designed against the CMV enhancer (5′ CMV: AACGCCAATAGGGACTT TCC and 3′ CMV: GGGCGTACTTGGCATATGAT) or the luciferase transgene (5′ Luc: ACGTGCAAAAGAAGCTACCG and 3′ Luc: AATGGGAAGTCACGA AGGTG). Samples were run in duplicate on an Applied Biosystems 7300 Real Time PCR System. The following cycling conditions were used: one cycle of 50 °C for 2 min, one cycle of 95 °C for 10 min, 40 cycles of 95 °C for 15 s and 60 °C for 60 s. Virus titre was determined by comparison with a standard curve generated using either a purified DNA fragment cut with XhoI/NheI from the pVIP luciferase-expressing vector or a reference standard consisting of purified AAV2/8 expressing 4E10 antibody previously titred against the DNA standard.

To validate the functional activity of each lot of the titred virus, we performed *in vitro* infection assays using 293T cells and measured the concentration of the antibody in the cell supernatant. Twenty-four hours before infection, 12-well plates were seeded with 500K cells in 1 ml of media. Two hours before infection, media was replaced with 500 µl per well of fresh media. Genome copies (10^{11}) of each virus were added to each well and allowed to infect for 6 days. Supernatants were removed and quantified for total IgG production by ELISA.

Mouse strains. Immunodeficient NOD/SCID/γc (NSG), immunocompetent C57BL/6 (B6) and Balb/C mice were obtained from the Jackson Laboratory. Immunodeficient Rag2/γc mice were obtained from A. Berns.

AAV intramuscular injection and bioluminescent imaging. Aliquots of previously titered viruses were thawed slowly on ice and diluted in TFB2 to achieve the predetermined dose in a 40 µl volume. Mice were anaesthetized by isofluorane inhalation and a single 40 µl injection was administered into the gastrocnemius muscle with a 28G insulin syringe. At various times after vector administration, mice were either bled to determine antibody concentration in serum or imaged using a Xenogen IVIS 200 Series imaging system (Caliper Lifesciences). To image, mice were anaesthetized by isofluorane inhalation and given 100 µl of 15 mg ml^{-1} D-luciferin (Gold Biotechnology) by intraperitoneal injection. Images were taken between 5 and 10 min after D-luciferin injection.

Quantification of antibody production by ELISA. For detection of total human IgG, ELISA plates were coated with 1 µg per well of goat anti-human IgG-Fc antibody (Bethyl) for 1 h. Plates were blocked with 1% BSA (KPL) in TBS for at least 2 h. Samples were incubated for 1 h at room temperature in TBST containing 1% BSA (KPL), then incubated for 30 min with HRP-conjugated goat anti-human kappa light chain antibody (Bethyl). Sample was detected with TMB Microwell Peroxidase Substrate System (KPL). A standard curve was generated using either Human Reference Serum (Lot 3, Bethyl) or purified Human IgG/Kappa (Bethyl).

For detection of gp120-binding IgG, ELISA plates were coated with 0.04–0.10 µg per well HIV-1 gp120MN protein (Protein Sciences) for 1 h. Plates were blocked with 1% BSA (KPL) in TBS for at least 2 h. Samples were incubated for 1 h at room temperature in TBST containing 1% BSA (KPL), then incubated for 30 min with HRP-conjugated goat anti-human IgG-Fc antibody (Bethyl). Sample was detected with TMB Microwell Peroxidase Substrate System (KPL). A standard curve was generated using either purified b12 or VRC01 protein as appropriate for the samples.

HIV virus production and titring. 293T cells were maintained in DMEM medium supplemented with 10% fetal bovine serum, 1% penicillin–streptomycin mix (Mediatech), 1% glutamine (Mediatech) in a 5% CO_2 incubator at 37 °C. Three days before transfection, two 15 cm plates were seeded with 3.75×10^6 cells each in 25 ml media. Two hours before transfection, media was changed to 15 ml of new media. Forty micrograms of the pNL4-3 plasmid[36] encoding an infectious molecular clone of HIV was transfected using Trans-IT reagent (Mirus) according to the manufacturer's instructions. Supernatant collections were performed at 24, 48 and 72 h after transfection and 15 ml of fresh media was gently added back to plate after each harvest. Pooled supernatants were filtered using a 0.45 µm filter to remove cell debris and aliquoted for storage at −80 °C. HIV was quantified following the manufacturer's instructions using an Alliance HIV-1 p24 antigen ELISA kit (Perkin-Elmer).

***In vitro* HIV protection assay.** *In vitro* neutralization assays in luciferase reporter cells were performed as described[37] with the following modifications. TZM-bl cells from the National Institutes of Health AIDS Research and Reference Reagent Program were maintained in DMEM medium supplemented with 10% fetal bovine serum, 1% penicillin–streptomycin mix (Mediatech), 1% glutamine (Mediatech) in a 5% CO_2 incubator at 37 °C. Before the assays, TZM-bl cells were trypsinzed, counted and re-suspended in a concentration of 10^5 cells per millilitre, in a total volume of 15 ml. Cells were mixed with 75 µg ml^{-1} DEAE-dextran and varying concentrations of each antibody as indicated and allowed to incubate on ice during the preparation of the virus. To prepare virus dilutions, stock NL4-3 was diluted to 250 ng ml^{-1} in growth media and subsequently fourfold serially diluted in the assay plate. One hundred microlitres of media containing 10,000 cells pre-incubated with antibody were added to wells containing previously diluted virus. Infection was allowed to proceed for 48 h in a 5% CO_2 incubator at 37 °C. Before reading the plate, 100 µl of BriteLite reagent (Perkin Elmer) was added to each well, and the plate was incubated for 2 min at room temperature. One hundred and twenty microlitres of each well was then transferred to an opaque plate and read by VICTOR3 (Wallac 1420 VICTOR3 plate reader, PerkinElmer).

Production of humanized mice for *in vivo* challenge. Humanized mice were produced essentially as described[17] with the following modifications. Human peripheral mononuclear blood cells (AllCells) were thawed from −80 °C, expanded in RPMI medium supplemented with 10% fetal bovine serum, 1% penicillin–streptomycin mix(Mediatech),1% glutamine (Mediatech), 50 µM β-mercaptoethanol, 10 mM HEPES(Gibco), 1× non-essential amino acids (Gibco), 1× sodium pyruvate (Gibco) and stimulated for T-cell expansion with 5 µg ml^{-1} phytohemagglutinin (Sigma) and 10 ng ml^{-1} human IL-2 (Peprotech) in a 5% CO_2 incubator at 37 °C. Cells were expanded for 7–13 days before use. For engraftment, 2 million to 4 million cells were injected intraperitoneally into NSG mice in a 300 µl volume of media.

HIV protection experiments. One day before HIV challenge, blood samples from each mouse were subjected both to ELISA for antibody quantification and flow cytometry to determine baseline CD4/CD8 ratios. The following day, mice were challenged through either intraperitoneal or intravenous injection of 100 µl containing the specified dose of HIV diluted in PBS. Infected mice were subjected to weekly blood sampling to determine the ratio of CD4 to CD8 cells in the T-lymphocyte subset by flow cytometry.

Flow cytometry. Blood samples were taken from mice by retro-orbital bleeding and were centrifuged for 5 min at 1,150*g* in a microcentrifuge to separate plasma from cell pellets. Plasma was removed and frozen for future analysis and cell pellets were re-suspended in 1.1 ml of 1× RBC lysis buffer (Biolegend) and incubated on

ice for at least 10 min to remove red blood cells. After lysis, samples were pelleted at 1,150*g* in a microcentrifuge for 5 min at room temperature, and stained with 65 µl of a cocktail containing 5 µl anti-human CD3-FITC, 5 µl anti-human CD4-PE, 5 µl anti-human CD8a-APC antibodies (Biolegend) and 50 µl of phosphate buffered saline supplemented with 2% fetal bovine serum (PBS+). Samples were washed with 1 ml PBS+ and again pelleted at 1,150*g* in a microcentrifuge for 5 min. Pelleted cells were re-suspended in 200 µl of PBS+ supplemented with 2 µg ml^{-1} propidium iodide (Invitrogen) and analysed on a FACSCalibur flow cytometer (Beckton-Dickinson). Samples were first gated by CD3 expression before determining the ratio of CD4 to CD8 cells within this subset. Samples containing fewer than 20 $CD3^+$ events were excluded from the analysis.

Histological staining for HIV p24. At the conclusion of the *in vivo* challenge experiments, spleens were removed from mice and immersed in 10% neutral buffered formalin for 24 h. After fixation, tissues were removed and placed in 70% ethanol until standard paraffin embedding and processing. Sections (4 µm thick) were then taken and immunohistochemical staining was performed for HIV-p24 detection using the Kal-1 murine monoclonal antibody and standard antigen retrieval techniques[38]. The slides were reviewed by a pathologist (D.S.R.) on an Olympus BX51 light microscope and images obtained using a SPOT Insight Digital Camera (Diagnostic Instruments).

31. Lock, M. *et al.* Rapid, simple, and versatile manufacturing of recombinant adeno-associated viral vectors at scale. *Hum. Gene Ther.* **21,** 1259–1271 (2010).
32. Ayuso, E. *et al.* High AAV vector purity results in serotype- and tissue-independent enhancement of transduction efficiency. *Gene Ther.* **17,** 503–510 (2010).
33. Matsushita, T. *et al.* Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Ther.* **5,** 938–945 (1998).
34. Wright, J. F. *et al.* Identification of factors that contribute to recombinant AAV2 particle aggregation and methods to prevent its occurrence during vector purification and formulation. *Mol. Ther.* **12,** 171–178 (2005).
35. Rohr, U. P. *et al.* Fast and reliable titration of recombinant adeno-associated virus type-2 using quantitative real-time PCR. *J. Virol. Methods* **106,** 81–88 (2002).
36. Adachi, A. *et al.* Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone. *J. Virol.* **59,** 284–291 (1986).
37. Montefiori, D. C. Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. *Curr. Protocol. Immunol.* Ch. 12.11, doi:10.1002/0471142735.im1211s64 (2005).
38. Kaluza, G. *et al.* A monoclonal antibody that recognizes a formalin-resistant epitope on the p 24 core protein of HIV-1. *Pathol. Res. Pract.* **188,** 91–96 (1992).

doi:10.1038/nature10694

Interactions between cancer stem cells and their niche govern metastatic colonization

Ilaria Malanchi[1]*, Albert Santamaria-Martínez[1]*, Evelyn Susanto[1], Hong Peng[1,2], Hans-Anton Lehr[3], Jean-Francois Delaloye[4] & Joerg Huelsken[1]

Metastatic growth in distant organs is the major cause of cancer mortality. The development of metastasis is a multistage process with several rate-limiting steps[1]. Although dissemination of tumour cells seems to be an early and frequent event[2], the successful initiation of metastatic growth, a process termed 'metastatic colonization', is inefficient for many cancer types and is accomplished only by a minority of cancer cells that reach distant sites[3,4]. Prevalent target sites are characteristic of many tumour entities[5], suggesting that inadequate support by distant tissues contributes to the inefficiency of the metastatic process. Here we show that a small population of cancer stem cells is critical for metastatic colonization, that is, the initial expansion of cancer cells at the secondary site, and that stromal niche signals are crucial to this expansion process. We find that periostin (POSTN), a component of the extracellular matrix, is expressed by fibroblasts in the normal tissue and in the stroma of the primary tumour. Infiltrating tumour cells need to induce stromal POSTN expression in the secondary target organ (in this case lung) to initiate colonization. POSTN is required to allow cancer stem cell maintenance, and blocking its function prevents metastasis. POSTN recruits Wnt ligands and thereby increases Wnt signalling in cancer stem cells. We suggest that the education of stromal cells by infiltrating tumour cells is an important step in metastatic colonization and that preventing *de novo* niche formation may be a novel strategy for the treatment of metastatic disease.

We aimed to explore limiting factors that determine metastatic success using the MMTV-PyMT mouse breast cancer model, which spontaneously metastasizes to the lungs[6]. We reasoned that the recently identified cancer stem cells (CSCs, also called tumour-initiating cells), a subset of cancer cells that allow long-term tumour growth and are thought to be responsible for remissions[7,8], might also be relevant to the development of metastatic disease (Supplementary Fig. 1). We measured the relative size of the population of CSCs from primary MMTV-PyMT tumours and their pulmonary metastases using the previously established markers CD90 and CD24, which label a subset of the $CD24^+CD29^{hi}$ or $CD24^+CD49f^{hi}$ population used earlier to isolate CSCs and normal mammary gland stem cells[9–13] (Supplementary Fig. 2). This CSC subset accounts for $3 \pm 2.1\%$ (s.d.) of all tumour cells from both primary tumours and metastases (Fig. 1a). When $CD90^+CD24^+$ CSCs or $CD90^+CD24^+$-depleted non-CSCs are separately isolated from GFP^+ tumours and directly introduced into mice through tail vein injection (GFP, green fluorescent protein), only the CSC population is able to produce lung metastases (Fig. 1b). Moreover, $CD90^+CD24^+$ cells isolated subsequently from pulmonary metastases are again the only tumour cell population that efficiently initiates secondary metastases (Fig. 1c). This is not due to differences in the extravasation capabilities of CSCs and non-CSCs (Supplementary Fig. 3).

In time course experiments, the relative size of the CSC population changes drastically during metastatic colonization. When injecting unfractionated tumour cells, the total number of tumour cells in the lung declines rapidly within the first 7 d after seeding[14] (Fig. 1d and Supplementary Fig. 4). Notably, the relative amount of CSCs transiently increases within the first and second weeks to more than 20%. This is due to selective expansion of the stem cell population as a result of their proliferation rate increasing over sixfold relative to the primary tumour in this early phase of metastatic colonization (as measured by BrdU incorporation; Fig. 1e and Supplementary Fig. 5). By contrast, non-CSCs show reduced proliferation 7 d after injection and, notably, fail to generate CSCs (Fig. 1e and Supplementary Fig. 6). Consequently, only CSCs are able to form metastatic colonies, and non-CSCs remain as single cells (Fig. 1f). Together, this demonstrates that selective expansion of CSCs is responsible for the initiation of metastasis.

The number of injected CSCs evidently exceeds the number of metastatic nodules, indicating that additional factors restrict successful metastatic colonization. Stem cells are suspected to rely on signals from their stromal environment, such as localized growth factors that can affect stem cell maintenance and proliferation[15–17]. We identified *Postn* as a stromal factor of normal stem cell niches and the metastatic niche (Supplementary Methods). *Postn* encodes for periostin, which becomes incorporated in the extracellular matrix[18] and has a role in bone, tooth and heart development and function[19–21]. Its expression is downregulated in the adult except in niches in direct contact with tissue-specific stem cells in mammary gland, bone, skin and intestine (Fig. 2a and Supplementary Figs 7 and 8). In tumours, POSTN is produced by stromal αSMA^+VIM^+ (αSMA also known as ACTA2) fibroblasts according to RNA *in situ* hybridization, immunostaining and quantitative PCR (Fig. 2b–h). Notably, POSTN expression is induced in the lung stroma by infiltrating cancer cells (Fig. 2d), but does not occur in the alveolar lung tissue of tumour-bearing but metastasis-free animals (Supplementary Fig. 8). Human breast cancer patients show induction of stromal POSTN expression in 75% of lymph node metastases (Fig. 2i, j and Supplementary Fig. 9). Whereas POSTN-deficient mice maintain normal mammary gland development (Supplementary Figs 10–12), MMTV-PyMT *Postn*$^{-/-}$ breast cancers show a dramatic decrease in the number and size of pulmonary metastases, to less than 10% of controls, despite unaltered primary tumour size and morphology ($n = 37$, $P < 0.003$; Fig. 3a–d and Supplementary Figs 13–15). Metastasis formation from POSTN-deficient tumour cells is rescued in wild-type recipients (Supplementary Figs 16 and 17), where expression of POSTN is at the same level as in controls (Supplementary Fig. 18), indicating that stromal production of POSTN determines metastatic efficiency.

In searching for tumour-derived factors that can induce stromal POSTN, *in silico* promoter analysis predicted regulation mediated

[1]Ecole Polytechnique Fédérale de Lausanne, Swiss Institute for Experimental Cancer Research and National Center of Competence in Research 'Molecular Oncology', 1015 Lausanne, Switzerland. [2]Nanfang Hospital, Department of Otorhinolaryngology, Head and Neck Surgery, Southern Medical University, Guangzhou 510515, China. [3]University Institute of Pathology, CHUV, University of Lausanne, 1011 Lausanne, Switzerland. [4]Department of Gynecology and Obstetrics, Centre Hospitalier Universitaire Vaudois, 1011 Lausanne, Switzerland.
*These authors contributed equally to this work.

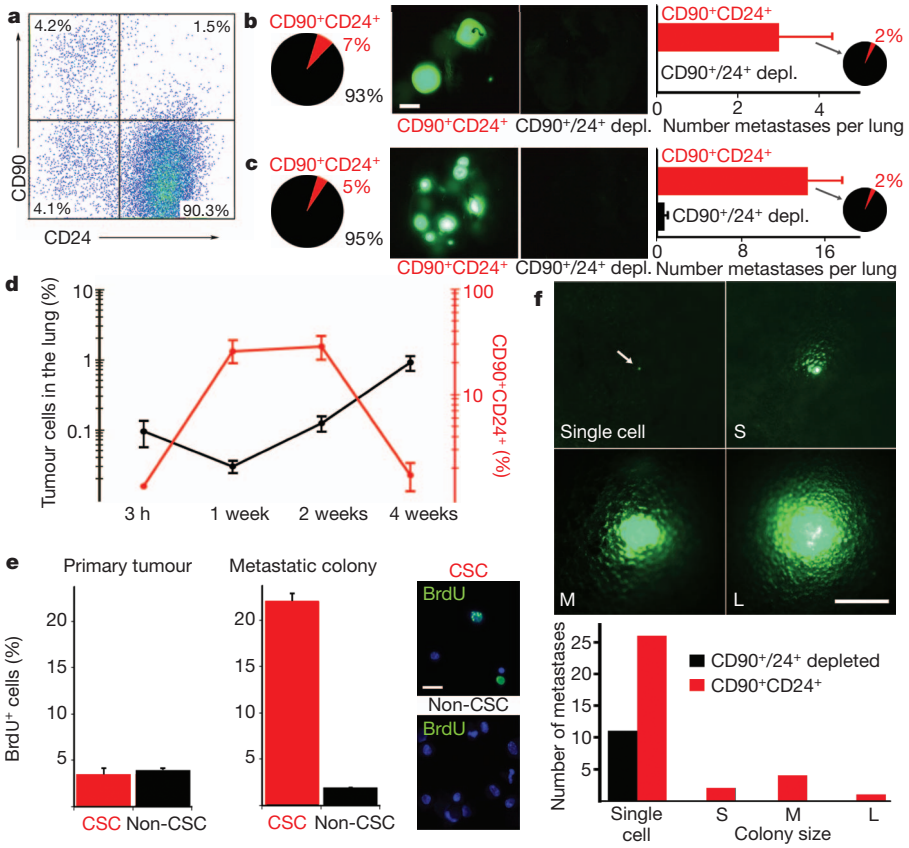


Figure 1 | Cancer stem cells initiate metastasis. (**a**) Representative density plot showing the abundance of cancer stem cells in MMTV-PyMT breast tumours defined as $CD24^+CD90^+$ after gating for viable (7-AAD$^-$) and lin$^-$ (CD31$^-$CD45$^-$TER119$^-$) cells (not shown). **b**, **c**, CSCs are the only cells to form pulmonary metastases on tail vein injection. $CD24^+CD90^+$ CSCs and non-CSC populations from GFP$^+$ tumour cells freshly isolated from primary tumours (**b**; 10^4 cells injected each) and metastases (**c**; 10^5 cells injected each) were separately injected in recipient mice. The frequency of CSCs (red; $P < 0.01$ (**b**), $P < 0.05$ (**c**)) is maintained at the metastatic site (shown as pie charts; $n = 6$ each; errors, s.e.m.). Scale bar, 1 mm. **d**, Time course experiments show selective proliferation of CSCs during metastatic colonization. GFP$^+$ tumour cells (10^6) were intravenously injected into recipient mice. At the indicated time points, the total proportion of GFP$^+$ tumour cells in the lung (black line) and the relative amount of the CSC population (red line) were evaluated by analysis using fluorescence-activated cell sorting (FACS). Note the transient increase of CSCs after one week ($n = 4$ per time point; errors, s.e.m.). **e**, **f**, CSCs are the only cells able to initiate growth in a short-term *in vivo* colonization assay. **e**, One week after tail vein injection of 10^6 GFP$^+$ tumour cells, mice were injected with BrdU. After 2 h, CSCs and non-CSCs were isolated by FACS and the frequency of proliferating cells was evaluated by BrdU staining on cytospins ($n = 12$, $P < 0.05$; errors, s.d.; representative example is shown on the right). Note the increase of proliferation in the CSC population during the early phase of lung colonization; by contrast, proliferation does not significantly differ between the two populations in the primary tumour. **f**, GFP$^+$ CSCs (2×10^5) or GFP$^+$ non-CSCs (4×10^5) were tail vein injected and analysed after two weeks. CD90$^+$/CD24$^+$-depleted non-CSCs remained as single cells whereas CSCs were able to initiate growth and form metastases of different sizes (S, 30–300 cells; M, 300–3,000 cells; L, >3,000 cells). The upper panel shows a representative example of the different metastasis colonies and the results are cumulatively quantified in the bottom panel ($n = 6$). Scale bars, 20 μm (**e**) and 250 μm (**f**).

by SMAD, NF-κB and LEF/TCF. Primary lung fibroblasts upregulate POSTN in response to TGF-β3 and TGF-β2, but are not responsive to BMP4, Wnt3A, Wnt5A or TNF-α (Fig. 3e, Supplementary Fig. 19 and data not shown). Co-culture experiments revealed that tumour cells are sufficient to trigger POSTN production (Supplementary Fig. 20). Notably, CSCs and non-CSCs both produce high levels of TGF-β3 (Supplementary Fig. 21), and blocking the action of TGF-β3 by expression of a secreted decoy receptor[22] (TGFβR2ΔTM; see Methods) in tumour cells blocks POSTN expression and prevents metastasis formation (Fig. 3f and Supplementary Fig. 22) in line with earlier results[23]. Together, these experiments demonstrate that infiltrating tumour cells need to educate the host stroma of the target organ to support metastasis initiation.

Even small metastatic colonies are strongly diminished in POSTN-deficient animals (Supplementary Fig. 23). Having shown that metastatic colonization depends on cancer stem cells, we went on to assess a potential role of POSTN in stem cell maintenance. Growth under conditions of ultralow attachment has been used to study CSCs *in vitro*[24]. Surprisingly, establishment of such cancer stem cell cultures is not possible using POSTN-deficient tumours (Fig. 3g–j; $n = 16$; compare with Supplementary Fig. 24, which shows unaltered cell survival in standard two-dimensional culture). In wild-type tumour spheres, POSTN is expressed by tumour-derived, stromal fibroblasts (Supplementary Fig. 25). Adding periostin protein to the mutant tumour cells rescues sphere formation (Fig. 3j). Conversely, a blocking antibody targeting POSTN prevents maintenance of wild-type cancer stem cells (Supplementary Fig. 30), but affects neither cellular survival nor growth *per se* (Supplementary Figs 26 and 27). Furthermore, CSCs fail to proliferate when co-cultured with POSTN-mutant, pulmonary fibroblasts (Fig. 3k). In lung metastases, we observe a preferential localization of CD90$^+$ CSCs adjacent to stromal niches (Fig. 3l), whereas lung metastases in POSTN-deficient animals show a reduction in the size of the CSC population (Fig. 3m). Thus, POSTN is an essential niche factor that supports CSC growth during metastatic colonization.

To gain insight into how POSTN is related to the mechanisms that control stem cell maintenance, we characterized the interactome of POSTN by tandem affinity purification (TAP)-tag enrichment and

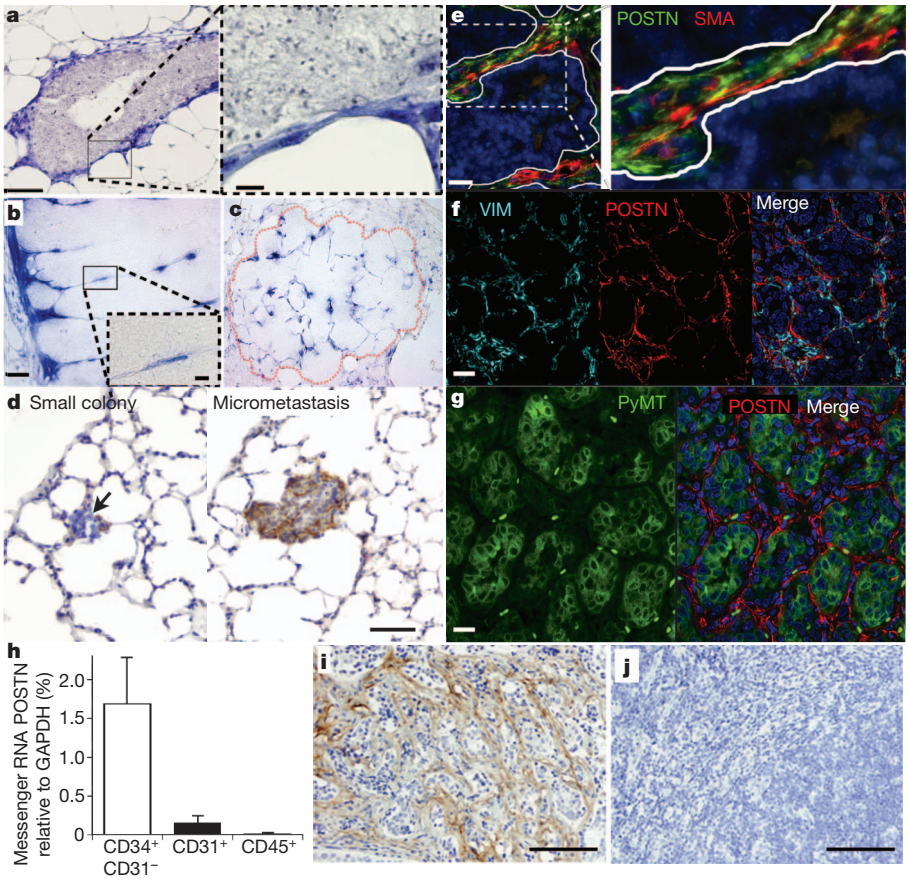


Figure 2 | POSTN is a stromal niche component that is induced on metastasis formation. **a**–**c**, Stromal POSTN expression as detected by RNA *in situ* hybridization (POSTN-expressing cells in blue) around terminal end buds, which are enriched for mammary gland stem cells during puberty (**a**), tumour cells at the primary site (**b**) and in pulmonary metastases (**c**; outlined in red). **d**, Immunohistochemistry of POSTN expression in pulmonary metastases at different stages. **e**–**g**, Immunofluorescence analysis of POSTN and fibroblast markers in pulmonary metastases. POSTN is not expressed by PyMT tumour cells (**g**), but is from αSMA$^+$ (**e**) and VIM$^+$ (**f**) fibroblasts. The white outline in **e** defines the tumour–stroma border. **h**, Quantitative PCR with reverse transcription detects POSTN expression in CD34$^+$CD31$^-$ pulmonary fibroblasts, in CD31$^+$ endothelial cells but not in CD45$^+$ immune cells that were isolated by FACS from lungs of mice with macrometastases ($n = 3$; errors, s.d.). **i**, **j**, Immunohistochemistry of POSTN expression in auxiliary lymph nodes of human breast cancer patients with metastasis-free N0 disease (**j**) or metastatic disease (**i**). A total of 17 of 23 lymph node metastases were positive for POSTN expression. Scale bars: 20 μm (**a**) and 2.5 μm (**a**, inset); 50 μm (**b**; same scale in **c**) and 10 μm (**b**, inset); 50 μm (**d**); 20 μm (**e**–**g**); 200 μm (**i**, **j**).

tandem mass spectrometry (Supplementary Methods), and this pointed to binding of POSTN to Wnt ligands[25]. Using TAP-tag pull-downs, we confirmed interaction between POSTN and Wnt1 and Wnt3A, but found that a known inhibitory ligand of the pathway, DKK1, does not associate with POSTN (Fig. 4a, b and Supplementary Fig. 28). The interaction between POSTN and Wnt ligands boosts Wnt signalling activity as measured by LRP6 phosphorylation and reporter assays using Wnt1, Wnt2 and Wnt6 (Fig. 4c, d). The Wnt pathway is known to control stem cell maintenance in a variety of tissues, including mammary gland[26] and tumours[8,27,28]. Indeed, addition of Wnt3A can rescue *in vitro* stem cell expansion in the absence of POSTN (Supplementary Figs 30 and 31). Furthermore, we find Wnt signalling activity to be concentrated in the CSC population *in vivo* as analysed by FACS (Fig. 4e, f). Here, we use a lentiviral reporter that allows Wnt-inducible GFP expression and constitutive labelling of individual reporter-containing cells by human CD2 (Supplementary Fig. 29). Wnt signalling activity in metastases is abrogated in the absence of POSTN (Fig. 4g; see Supplementary Fig. 29 for Wnt activity in tumour spheres). Similarly, we detect higher levels of the general Wnt target gene *Axin2* in metastases in wild-type hosts than in mutant hosts (Fig. 4h, i). Remarkably, metastasis of Wnt-driven breast cancer proves to be independent of stromal POSTN (Fig. 4j). Thus, POSTN acts as a niche component that can promote stem cell maintenance and metastatic colonization by augmenting Wnt signalling (Supplementary Fig. 1).

Thus, instrumental factors that contribute to the inefficiency of the metastatic process are the low abundance of CSCs, which are required to initiate growth at the secondary site, and incompatibilities with 'foreign' niches, which necessitate education of the target organ to resemble the primary microenvironment. Periostin is an essential component of this CSC-supportive niche and needs to be induced in the lung stroma by infiltrating tumour cells. Together with a recent report that implicated tumour-cell-derived tenascin C in breast cancer metastasis, these examples demonstrate an essential role for single proteins of the extracellular matrix in metastatic colonization[29,30].

Surprisingly, POSTN deficiency affects neither normal mammary gland development nor tumour formation or the relative size of the CSC population at the primary site (Supplementary Fig. 32). This may be due to a greater complexity and redundancy in the primary niche, such that the loss of a single factor can be tolerated. By contrast, loss of only one important factor, such as POSTN, prevents metastatic colonization in the presumably less supportive and less complex secondary site. Accordingly, the early phase of metastasis can be anticipated to be particularly sensitive to therapeutic intervention because the dependence of cancer cells on niche signals is probably highest in that phase. Targeting this metastatic niche promises to be less sensitive to rapid genetic changes in cancer cells and may not only prevent metastatic colonization but may also interfere with the survival of disseminated, dormant cancer cells.

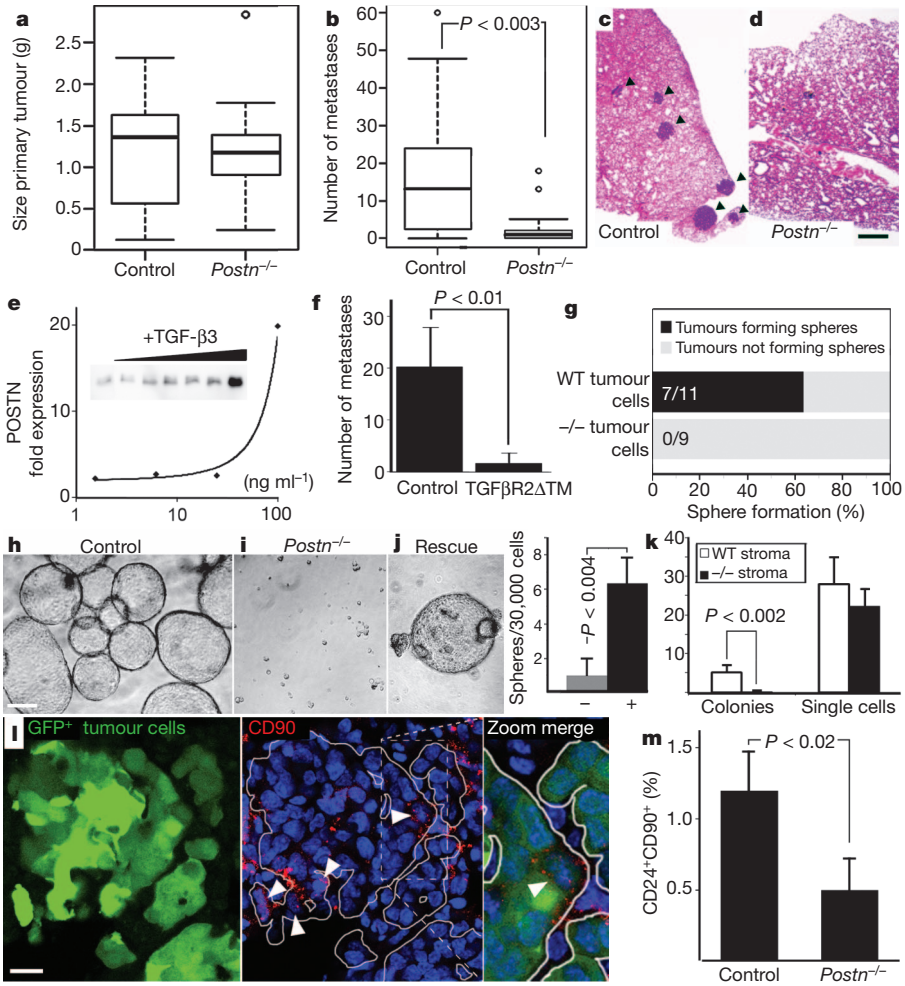


Figure 3 | POSTN is required for metastatic colonization by affecting CSC maintenance. **a**, **b**, Primary tumour size (**a**) and number of spontaneously formed macrometastases (**b**) in control and *Postn*$^{-/-}$ mice ($n = 37$). **c**, **d**, HE section of spontaneous, pulmonary metastases in control (**c**) and POSTN-mutant (**d**) animals. **e**–**f**, TGF-β3 induces POSTN expression in lung stromal cells. **e**, TGF-β3 was added to lung stromal cells to test its ability to induce POSTN expression. RNA expression by quantitative PCR and protein expression by western blot analysis (inset) were measured at the indicated concentrations. **f**, Blocking of TGF-β3 abrogates metastasis formation. Cancer cells were infected either with a control virus or with a virus expressing secreted, dominant negative TGFβR2ΔTM and separately intravenously injected into recipient mice. Metastasis formation was evaluated 6 weeks after injection ($n = 12$; errors, s.d.). **g**, Frequency of mammosphere formation from primary tumours. WT, wild type. **h**, Representative picture of control tumour mammospheres. **i**, Loss of CSCs in secondary mammosphere assays from *Postn*$^{-/-}$ tumours. **j**, Sphere formation by *Postn*$^{-/-}$ tumour cells can be rescued by addition of 50 ng ml^{-1} of periostin to primary cultures. Control tumour cells produce 13.3 ± 2.5 spheres per 30,000 cells. Quantification is shown in the histogram on the right ($n = 4$; errors, s.d.). **k**, Only stromal cells isolated from wild-type lungs are able to support CSC colony formation *in vitro*. CSCs were sorted by FACS and seeded on lung stromal cells isolated either from wild-type or *Postn*$^{-/-}$ mice. Colony formation was evaluated after two weeks ($n = 3$; errors, s.d.). **l**, Immunofluorescence analysis shows that GFP$^+$CD90$^+$ CSCs (white arrow heads) are concentrated in areas with immediate contact to the stroma. **m**, Reduced frequency of CD24$^+$CD90$^+$ CSCs in rare pulmonary metastases in POSTN-deficient mice ($n = 12$; errors, s.d.). Scale bars: 500 μm (**c**, **d**); 50 μm (**h**–**j**); 5 μm (**l**).

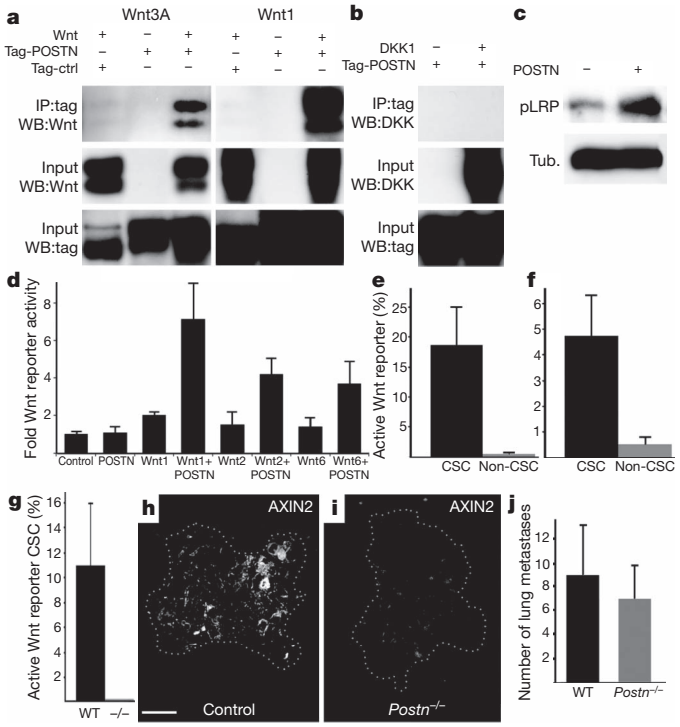


Figure 4 | POSTN promotes stem cell maintenance and metastasis formation by augmenting Wnt signalling. **a**, Interaction of TAP-tagged POSTN and Wnt3A and Wnt1 as determined by pull down and western blotting. **b**, Interaction of TAP-tagged POSTN and DKK1 as determined by pull down and western blotting. Note the absence of interaction between POSTN and DKK1. **c**, LRP6 phosphorylation (pLRP) increases on POSTN expression. **d**, Wnt reporter activity in response to Wnt1, Wnt2 or Wnt6 is stimulated by POSTN ($n = 3$; errors, s.d.). **e**, **f**, *In vivo* Wnt reporter assay. Tumour cells were infected with a lentivirus containing both an artificial Wnt-responsive promoter driving GFP expression and a constitutive promoter expressing human CD2 (Supplementary Fig. 29). Cells were analysed by FACS considering only human-CD2$^+$, reporter-containing cells. CD90$^+$CD24$^+$ CSCs and non-CSCs were scored for Wnt reporter activity as GFP$^+$. The signal concentrates in CSCs in both primary tumours (**e**) and pulmonary metastases (**f**) ($n = 3$; errors, s.e.m.). **g**, CD90$^+$CD24$^+$ wild-type CSCs containing a Wnt reporter (human CD2$^+$) show signalling activity (GFP$^+$) only when seeded in wild-type as opposed to *Postn*$^{-/-}$ lungs ($n = 3$; errors, s.e.m.). **h**, **i**, Immunofluorescence staining of AXIN2 in control (**h**) and POSTN-deficient (**i**) lung metastases (outlined in white). Note the absence of staining in POSTN$^-$ lung metastases. Scale bar, 10 μm. **j**, MMTV-Wnt1 tumour cells can form pulmonary metastases in the absence of POSTN expression in the stroma. Tumour cells were orthotopically transplanted into mammary glands of wild-type or POSTN-deficient recipients and analysed after ten weeks for spontaneous lung metastasis ($n = 5$ each; errors, s.d.).

METHODS SUMMARY

Details on the mouse models used and the generation of the *Postn* gene ablation can be found in Supplementary Information. Tumours derived from MMTV-PyMT mice were collected and cells were isolated and directly stained for FACS analysis or sorting using antibodies against CD31, CD45, TER119 (LY76), CD90, CD24 and CD49f (ITGA6). FACS-isolated tumour cells were directly transplanted into recipient mice.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 26 July 2010; accepted 1 November 2011.
Published online 7 December 2011.

1. Chambers, A. F., Groom, A. C. & MacDonald, I. C. Dissemination and growth of cancer cells in metastatic sites. *Nature Rev. Cancer* **2,** 563–572 (2002).
2. Hüsemann, Y. *et al.* Systemic spread is an early step in breast cancer. *Cancer Cell* **13,** 58–68 (2008).
3. Kouros-Mehr, H. *et al.* GATA-3 links tumor differentiation and dissemination in a luminal breast cancer model. *Cancer Cell* **13,** 141–152 (2008).
4. Nguyen, D. X., Bos, P. D. & Massague, J. Metastasis: from dissemination to organ-specific colonization. *Nature Rev. Cancer* **9,** 274–284 (2009).
5. Hess, K. R. *et al.* Metastatic patterns in adenocarcinoma. *Cancer* **106,** 1624–1633 (2006).
6. Lin, E. Y. *et al.* Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am. J. Pathol.* **163,** 2113–2126 (2003).
7. Al-Hajj, M., Wicha, M. S., Benito-Hernandez, A., Morrison, S. J. & Clarke, M. F. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl Acad. Sci. USA* **100,** 3983–3988 (2003).
8. Reya, T. & Clevers, H. Wnt signalling in stem cells and cancer. *Nature* **434,** 843–850 (2005).
9. Cho, R. W. *et al.* Isolation and molecular characterization of cancer stem cells in MMTV-Wnt-1 murine breast tumors. *Stem Cells* **26,** 364–371 (2008).
10. Shackleton, M. *et al.* Generation of a functional mammary gland from a single stem cell. *Nature* **439,** 84–88 (2006).
11. Stingl, J. *et al.* Purification and unique properties of mammary epithelial stem cells. *Nature* **439,** 993–997 (2006).
12. Liu, J. C., Deng, T., Lehal, R. S., Kim, J. & Zacksenhaus, E. Identification of tumorsphere- and tumor-initiating cells in HER2/Neu-induced mammary tumors. *Cancer Res.* **67,** 8671–8681 (2007).
13. Zhang, M. *et al.* Identification of tumor-initiating cells in a p53-null mouse model of breast cancer. *Cancer Res.* **68,** 4674–4682 (2008).
14. Podsypanina, K. *et al.* Seeding and propagation of untransformed mouse mammary cells in the lung. *Science* **321,** 1841–1844 (2008).
15. Vermeulen, L. *et al.* Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nature Cell Biol.* **12,** 468–476 (2010).
16. Visvader, J. E. & Lindeman, G. J. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nature Rev. Cancer* **8,** 755–768 (2008).
17. Psaila, B. & Lyden, D. The metastatic niche: adapting the foreign soil. *Nature Rev. Cancer* **9,** 285–293 (2009).
18. Takeshita, S., Kikuno, R., Tezuka, K. & Amann, E. Osteoblast-specific factor 2: cloning of a putative bone adhesion protein with homology with the insect protein fasciclin I. *Biochem. J.* **294,** 271–278 (1993).
19. Shimazaki, M. & Kudo, A. Impaired capsule formation of tumors in periostin-null mice. *Biochem. Biophys. Res. Commun.* **367,** 736–742 (2008).
20. Oka, T. *et al.* Genetic manipulation of periostin expression reveals a role in cardiac hypertrophy and ventricular remodeling. *Circ. Res.* **101,** 313–321 (2007).
21. Rios, H. *et al.* Periostin null mice exhibit dwarfism, incisor enamel defects, and an early-onset periodontal disease-like phenotype. *Mol. Cell. Biol.* **25,** 11131–11144 (2005).
22. Zhao, W. *et al.* Suppression of *in vivo* tumorigenicity of rat hepatoma cell line KDH-8 cells by soluble TGF-beta receptor type II. *Cancer Immunol. Immunother.* **51,** 381–388 (2002).
23. Muraoka, R. S. *et al.* Blockade of TGF-beta inhibits mammary tumor cell viability, migration, and metastases. *J. Clin. Invest.* **109,** 1551–1559 (2002).
24. Pece, S. *et al.* Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell* **140,** 62–73 (2010).
25. Milovanovic, T. *et al.* Expression of Wnt genes and frizzled 1 and 2 receptors in normal breast epithelium and infiltrating breast carcinoma. *Int. J. Oncol.* **25,** 1337–1342 (2004).
26. Zeng, Y. A. & Nusse, R. Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture. *Cell Stem Cell* **6,** 568–577 (2010).
27. Malanchi, I. *et al.* Cutaneous cancer stem cell maintenance is dependent on β-catenin signaling. *Nature* **452,** 650–653 (2008).
28. Barker, N. *et al.* Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457,** 608–611 (2009).
29. Wels, J., Kaplan, R. N., Rafii, S. & Lyden, D. Migratory neighbors and distant invaders: tumor-associated niche cells. *Genes Dev.* **22,** 559–574 (2008).
30. Oskarsson, T. *et al.* Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nature Med.* **17,** 867–874 (2011).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to U. Koch for advice on FACS, to M. Moniatte for advice on mass spectrometry and to S. Leuba for technical assistance with histology. I.M., E.S., A.S.-M. and J.H. were supported in part by the Swiss League against Cancer, the SNF, the NCCR in Molecular Oncology and the Anna Fuller Fund. J.H. holds the EPFL chair for Signal Transduction in Oncogenesis sponsored by Debiopharm.

Author Contributions I.M., A.S.-M. and J.H. designed and performed most of the experiments, analysed data and prepared the manuscript; E.S. and H.P. performed experiments; H.-A.L. performed experiments and analysed data; J.-F.D. provided clinical samples; and J.H. designed and supervised the study.

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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to J.H. (joerg.huelsken@epfl.ch).

METHODS

Mouse work. The design of the POSTN targeting vector is shown in Supplementary Fig. 10. E14.1 129/Ola ES cells were used for targeting and were injected into C57Bl6 blastocysts. The POSTN mutant allele was then backcrossed onto the FVB background for at least eight generations. The MMTV-PyMT[31], MMTV-Wnt1[32] and the ACT–GFP[33] mouse strains have been described previously. All used strains were backcrossed onto FVB for ten generations except for the Nude allele, which was used on a NMRI background. For tumour cell transplantations, *Nude* or *Rag1*$^{-/-}$ mice were used when tumour cells expressed GFP or had been modified by lentiviral vectors. In general, we did not observe major differences in these different strains with respect to cancer stem cell phenotype, tumour take rate or metastasis number, except for an overall faster growth in *Rag1*$^{-/-}$ versus *Nude* versus wild type. For tail vein injections, tumour cells were resuspended in 100 μl of PBS. For orthotopic transplantations, tumour cells were resuspended in 50 μl growth-factor-reduced matrigel (BD Biosciences) and transplanted into a small pocket within the fourth mammary fat pad. All animal procedures were performed in accordance with the Swiss legislation on animal experimentation.

Antibodies. We used TER119, CD24, CD29, CD34, CD45 and CD90.1 (eBioscience); CD31 and GFP (Invitrogen); Cytokeratin 8, POSTN, AXIN2 and Wnt1 (Abcam); POSTN (Adipogen); Cytokeratin 14 (Covance); GATA-3 and PyMT (Santa Cruz Biotechnology); αSMA (Sigma); VIM (Lifespan Biosciences); and BrdU (BD Pharmingen and ImmunologicalsDirect). The antibody against human CD2 was a kind gift by Professor Werner Held (Ludwig Center for Cancer Research, Lausanne).

Tumour and metastasis evaluation. For the quantification of spontaneous micrometastasis, lungs were completely sectioned and HE stained, and the size and the number of metastatic nodules were analysed by microscopy. Quantification of lung metastases derived from GFP^{+} tumour cells was performed by fluorescent microscopy (Leica M205 FA stereomicroscope) or by FACS analysis of cell preparations from total lungs. All statistical evaluation used a homoscedastic Student's *t*-test with a two-tailed distribution.

For expression of a secreted TGFβ decoy receptor (TGFβR2ΔTM), tumour cells were transduced with lentiviral vectors expressing the extracellular domain (amino acids 1–185) of murine TGFβR2 (NM_009371.3) under the control of a murine phosphoglycerate kinase (PGK) promoter, and metastases were scored as described above following tail vein injection of transduced tumour cells.

Tumour cell isolation. Tumours or lungs containing pulmonary metastases were minced with a razor blade and digested with a mixture of DNase and Liberase (Roche Diagnostics). On enzymatic digestion, samples were passed through a 100-μm filter and washed once in growth medium (DMEM/F12 with 2% FBS, 20 ng ml^{-1} EGF, 10 μg ml^{-1} insulin; Invitrogen) and twice in PBS. Cells were then directly used for FACS staining and analysed in a Cyan ADP flow cytometer (Beckman Coulter) or sorted with a custom-built FACSAria II (Beckton Dickinson). For tail vein injection, cells were plated overnight on collagen. In the experiments evaluating the metastasis initiation potential of CSCs versus non-CSCs, we were able to prepare and inject higher cell numbers from metastases than from primary tumours, where larger necrotic areas limit the obtainable cell amounts. In both settings, only CSCs were able to give rise to pulmonary metastases, but this experimental setting is unsuitable for a direct comparison of the metastatic potential of CSCs isolated from primary tumours with that of those isolated from metastases. In the extravasation experiments, we performed several pilot experiments perfusing the lungs before dissection to distinguish cells trapped inside capillaries from those that successfully extravasated. Perfusion was done with Ringer's buffer through the right ventricle for 5–10 min.

Tumour mammosphere cultures. Sphere cultures were established from total tumour cell preparations. After dissociation, tumour cells were plated on collagen overnight, trypsinized and plated in 100 μl of sphere media (DMEM/F12 with B27, 20 ng ml^{-1} EGF, 20 ng ml^{-1} FGF and 4 μg ml^{-1} heparin) into 96-well, low-attachment plates (Constar) at a concentration of (1–3) $\times$ 10^{4} cells per well. For the anti-POSTN blocking antibody assay, half of the culture medium was replaced by hybridoma supernatant. Secondary mammosphere culture was performed by collecting the spheres through gentle centrifugation (130*g*) followed by enzymatic (10 min in 0.05% trypsin-EDTA at 37 °C) and mechanical dissociation using a 25G needle. Immunostainings were performed as whole mounts of PFA-fixed and 0.1% Triton-X100-permeabilized spheres and were analysed by confocal microscopy.

Identification of POSTN as a niche-expressed factor. To identify niche-expressed factors, we performed microarray RNA profiling studies where we analysed laser-capture microdissected material from normal stem cells together with their adjacent stroma. This was compared with laser-capture microdissected material isolated from differentiated cells together with their adjacent stroma from the same tissue (data not shown). Skin and intestine contain stem cells in morphologically distinct structures (the bulge in the skin, the crypt in the intestine). On the basis of these profiles (accessible through GEO GSE31730 and 8818), we evaluated potential niche localization of candidate genes by RNA *in situ* hybridization and immunofluorescence analysis (Fig. 2a, Supplementary Fig. 7 and data not shown). Moreover, we profiled the candidate genes by quantitative PCR with reverse transcription and performed RNA *in situ* hybridization in other tissues, in tumours and in metastases (data not shown). This identified POSTN as a stromal marker in primary murine breast cancer and its metastases (Fig. 2).

Immunodetection and *in situ* hybridization. Immunostaining was performed on 7-μm OCT- or gelatin-embedded (to maintain GFP signals) frozen sections permeabilized with 0.1% Triton-X100 or on 5-μm paraffin sections using antigen retrieval for 20 min in boiling, 10 mM citric acid, pH 6.0. For BrdU immunostaining, an additional incubation period of 10 min in 2 N HCl was performed. For immunohistochemistry, endogenous peroxidases were blocked by incubation with 0.6% H_2O_2 in methanol. After incubation with the indicated antibodies, secondary fluorescently labelled antibodies Alexa Fluor 488, 567 and 647 (Molecular Probes, Invitrogen) or HRP-conjugated secondary antibodies were used to reveal the primary antibodies. For immunohistochemistry, Dako Envision^{+} was used together with diaminobenzidine tablets (Sigma) to detect HRP. Fluorescent images were made with an automated upright microscope system (DM5500, Leica) or an LSM700 upright or inverted confocal microscope (Zeiss). Light images were made with an AX70 widefield microscope (Olympus). Figures were generated with PHOTOSHOP (Adobe Systems) and CYTOSKETCH (CytoCode).

In situ hybridizations were performed on paraffin sections permeabilized by proteinase K digestion using DIG-labelled (Roche), antisense transcripts of the mouse POSTN complementary DNA. After hybridization, section were washed in 50% formamide, ×5 SSC (0.75 M NaCl, 85 mM sodium citrate, pH 7.0) and 0.1% Tween at 70 °C, and non-specifically bound probe was digested by RnaseA treatment, followed by extensive washing in 50% formamide, ×2 SSC and 0.1% Tween at 65 °C. The DIG label was detected by an anti-DIG Fab (Roche) coupled to alkaline phosphatase using FAST NBT/BCIP (Sigma).

Western blot. Protein was extracted with complete RIPA buffer, separated by electrophoresis, transferred to PVDF membranes, blocked with 5% dry milk or BSA and incubated overnight with primary antibodies. Immunoreactive bands were visualized using HRP-conjugated secondary antibodies (Promega) and the detection reagent SuperSignal West Cheminoluminiscent Substrate (Thermo Scientific).

Real-time RT–PCR. RNA was prepared using a miniRNA or microRNA kit (Qiagen) from FACS-isolated primary cells. Complementary DNAs were generated using oligo-T priming and quantitative PCR was performed in a LightCycler (Roche) or a StepOnePlus thermocycler (Applied Biosystems) using the Power SYBR green PCR Master Mix (Applied Biosystems) and the primers aatgctgccctggc tatatg and gcatgacccttttccttcaa (POSTN); agcgcagacacaacccatag and gggttctgccca catagtacaag (TGF-β3); and caagctcatttcctggtatgacaat and gttgggatagggcctctcttg (GAPDH).

Periostin monoclonal antibody production. In collaboration with Adipogen, we injected *Postn*$^{-/-}$ mice with purified human POSTN protein. Hybridomas were generated and tested for antibody production by ELISA. Antibodies were selected for their ability to recognize POSTN and for their ability to block tumour mammosphere formation *in vitro*.

Periostin interactome and pull-downs. Freshly isolated tumour cells were grown under mammosphere conditions in the presence of purified, polyHis-tagged and streptavidin binding peptide (SBP)-tagged POSTN protein (Supplementary Fig. 28) for 72 h. Then proteins were crosslinked by 1 mM DTSSP (Pierce) for 10 min and solubilized by DDM and SDS (Anatrace). Protein purification was achieved by two successive affinity purification steps using the SBP and the polyHis tags and elution by biotin and imidazole, respectively. After elution, the crosslink was released by DTT treatment, the resulting protein mixture was trypsin-digested and peptides were identified by one-dimensional nano-liquid chromatography tandem mass spectrometry. Among the POSTN-attached proteins, we found Wnt6 to be a potential interactor. Because these experiments involved crosslinkers, we consider these results to be preliminary unless confirmed by independent experiments. For Wnt6, an interaction with POSTN could not be validated owing to a lack of Wnt6 antibodies and a suitable expression system for Wnt6. Nevertheless, this preliminary finding prompted us to test directly for interactions of POSTN with other Wnt ligands where the necessary tools are available. For these pull-downs, cDNAs encoding Wnt1, HA-tagged Wnt3A, DKK1 and SBP/His-tagged POSTN (Supplementary Fig. 28) or SBP/His-tagged mouse immunoglobulin-G heavy chain (as a control) were transfected separately into 293T cells. Cells secreting POSTN or immunoglobulin-G and cells producing Wnt or DKK were mixed after one day and collected for pull-downs after three additional days. Without the use of crosslinkers, POSTN-interacting proteins were enriched by the SBP/His tag as described above using NP40 as detergent. Western blot was performed for identification of the pulled-down proteins.

Wnt reporter activity. HEK293T cells were transduced with a lentiviral vector containing a LEF/TCF-responsive luciferase reporter consisting of a minimal promoter and 20 copies of the LEF/TCF consensus site or a control construct in which the LEF/TCF sites had been mutated[34]. These stable Wnt reporter cell lines were then transiently transfected with a POSTN expression vector before these cells were mixed with transiently transfected HEK293T or COS7 cells (the latter for Wnt6 expression) producing Wnt1, Wnt2 or Wnt6, respectively. The media on these co-cultures of Wnt reporter and Wnt producer cells were replaced after 24 h to remove soluble Wnt ligands. After 48 h of co-culture, luciferase activity in cell extracts was measured with a kit from Promega.

For *in vivo* reporter assays, we generated a lentiviral vector containing the same transcription factor sites and minimal promoter as described above driving GFP expression. In addition, this vector contained a ubiquitous PGK promoter driving expression of a truncated human CD2 cDNA that allows identification of reporter-containing cells by FACS (Supplementary Fig. 29). Lentiviruses were generated in 293T cells using third-generation lentivirus packaging vectors, and virus particles were concentrated by ultracentrifugation. Freshly isolated tumour cells were infected with lentiviruses overnight, enriched by anti-CD2 MACS (Miltenyi) for cells carrying the reporter construct and orthotopically transplanted for tumour formation and spontaneous metastasis. Wnt signalling activity in cells of primary tumours or spontaneous pulmonary metastases was quantified by FACS analysis by measuring GFP^+ cells after gating for human-$CD2^+$ cells. Similarly, human-$CD2^+$ tumour cells were used for tumour mammosphere cultures to measure Wnt pathway activity in the presence or absence of the POSTN blocking antibody.

Isolation of lung stromal cells. Lungs were minced with a razor blade and digested with a mixture of DNase and Liberase (Roche Diagnostics). On enzymatic digestion, samples were passed through a 100-μm filter and washed once in growth medium (DMEM/F12 with 2% FBS, 20 ng ml^{-1} EGF and 10 μ ml^{-1} insulin; Invitrogen) and twice in PBS. Cells were plated onto plates pretreated with collagen and allowed to attach for three days. The medium was then changed to serum-free medium and tumour cells or soluble factors were added to test POSTN induction. POSTN expression was evaluated by western blot, PCR with reverse transcription or immunofluorescence analysis.

31. Guy, C. T., Cardiff, R. D. & Muller, W. J. Induction of mammary tumours by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol. Cell. Biol.* **12,** 954–961 (1992).
32. Tsukamoto, A. S., Grosschedl, R., Guzman, R. C., Parslow, T. & Varmus, H. E. Expression of the int-1 gene in transgenic mice is associated with mammary gland hyperplasia and adenocarcinomas in male and female mice. *Cell* **55,** 619–625 (1988).
33. Okabe, M., Ikawa, M., Kominami, K., Nakanishi, T. & Nishimune, Y. 'Green mice' as a source of ubiquitous green cells. *FEBS Lett.* **407,** 313–319 (1997).
34. Jeannet, G. *et al.* Long-term, multilineage hematopoiesis occurs in the combined absence of β-catenin and γ-catenin. *Blood* **111,** 142–149 (2008).

doi:10.1038/nature10688

FBXO11 targets BCL6 for degradation and is inactivated in diffuse large B-cell lymphomas

Shanshan Duan[1,2]*, Lukas Cermak[1,2]*, Julia K. Pagan[1,2], Mario Rossi[1], Cinzia Martinengo[3], Paola Francia di Celle[4], Bjoern Chapuy[5], Margaret Shipp[5], Roberto Chiarle[3] & Michele Pagano[1,2]

BCL6 is the product of a proto-oncogene implicated in the pathogenesis of human B-cell lymphomas[1,2]. By binding specific DNA sequences, BCL6 controls the transcription of a variety of genes involved in B-cell development, differentiation and activation. BCL6 is overexpressed in the majority of patients with aggressive diffuse large B-cell lymphoma (DLBCL), the most common lymphoma in adulthood, and transgenic mice constitutively expressing BCL6 in B cells develop DLBCLs similar to the human disease[3,4]. In many DLBCL patients, BCL6 overexpression is achieved through translocation (~40%) or hypermutation of its promoter (~15%). However, many other DLBCLs overexpress BCL6 through an unknown mechanism. Here we show that BCL6 is targeted for ubiquitylation and proteasomal degradation by a SKP1–CUL1–F-box protein (SCF) ubiquitin ligase complex that contains the orphan F-box protein FBXO11 (refs 5, 6). The gene encoding FBXO11 was found to be deleted or mutated in multiple DLBCL cell lines, and this inactivation of *FBXO11* correlated with increased levels and stability of BCL6. Similarly, *FBXO11* was either deleted or mutated in primary DLBCLs. Notably, tumour-derived FBXO11 mutants displayed an impaired ability to induce BCL6 degradation. Reconstitution of FBXO11 expression in *FBXO11*-deleted DLBCL cells promoted BCL6 ubiquitylation and degradation, inhibited cell proliferation, and induced cell death. *FBXO11*-deleted DLBCL cells generated tumours in immunodeficient mice, and the tumorigenicity was suppressed by FBXO11 reconstitution. We reveal a molecular mechanism controlling BCL6 stability and propose that mutations and deletions in *FBXO11* contribute to lymphomagenesis through BCL6 stabilization. The deletions/mutations found in DLBCLs are largely monoallelic, indicating that *FBXO11* is a haplo-insufficient tumour suppressor gene.

Because the degradation of BCL6 has been reported to be phosphorylation-dependent[7,8], we examined the hypothesis that BCL6 is degraded by an SCF ubiquitin ligase complex, as most SCF ligases target phosphorylated substrates[5]. The expression of a dominant-negative CUL1 mutant (CUL1(1–242) or CUL1(1–385)) results in the accumulation of SCF substrates[9–11]. Western blotting of extracts from Ramos cells (a Burkitt's lymphoma cell line) infected with a retrovirus expressing CUL1(1–385) revealed that BCL6 level was increased compared to control cells, indicating that BCL6 is an SCF substrate (Supplementary Fig. 1a). Similarly, silencing RBX1, another SCF subunit, induced the accumulation of BCL6 (Supplementary Fig. 1b). Therefore, we investigated the ability of BCL6 to be recruited to the SCF via a panel of F-box proteins. Screening of a library of F-box proteins (ten shown) revealed that BCL6 specifically bound FBXO11 (Fig. 1a and Supplementary Fig. 2a), and this binding was confirmed with endogenous proteins in both Ramos cells (Fig. 1b) and primary B cells from mice (Supplementary Fig. 2b). Moreover, BCL6 was found to interact with endogenous SKP1 and neddylated CUL1, the form of CUL1 that

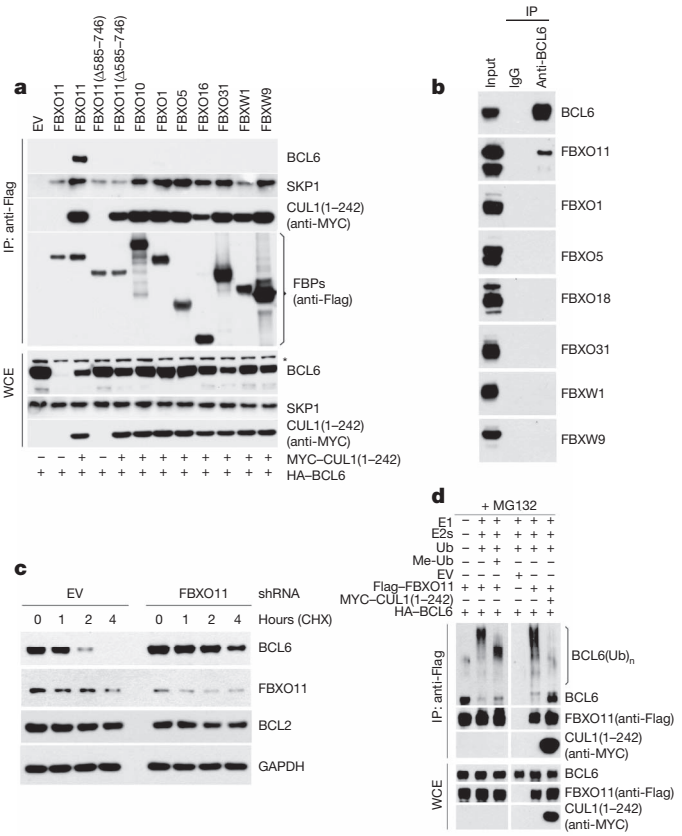


Figure 1 | FBXO11 controls the ubiquitylation and degradation of BCL6. **a**, HEK-293T cells were transfected with HA-tagged BCL6 and the indicated Flag-tagged F-box proteins (FBPs) or an empty vector (EV). Where indicated, MYC-tagged CUL1(1–242) was also transfected. Twenty-four hours after transfection, cells were harvested and lysed. Whole-cell extracts (WCE) were subjected to immunoprecipitation (IP) with anti-Flag resin and immunoblotting as indicated. The asterisk denotes a nonspecific band present in the anti-BCL6 blot. **b**, Ramos cells were treated with MG132 during the last 5 h before lysis. Lysates were immunoprecipitated with either an antibody against BCL6 or a nonspecific IgG and analysed by immunoblotting as indicated. **c**, SU-DHL6 cells were infected with viruses expressing two different *FBXO11* shRNAs (in combination) or a control shRNA, selected, and treated with cycloheximide for the indicated times. Protein extracts were immunoblotted for the indicated proteins. **d**, HEK-293T cells were transfected with HA-tagged BCL6, Flag-tagged FBXO11, MYC-tagged CUL1(1–242), and/or an empty vector (EV) as indicated. After immunopurification with anti-Flag resin, *in vitro* ubiquitylation of BCL6 was performed in the presence or absence of E1, E2s and ubiquitin (Ub). Where indicated, methylated ubiquitin (Me-Ub) was added. Samples were analysed by immunoblotting with an anti-BCL6 antibody. The bracket on the right marks a ladder of bands >86 kDa corresponding to ubiquitylated BCL6. Immunoblots of whole-cell extracts are shown at the bottom.

[1]Department of Pathology, NYU Cancer Institute, New York University School of Medicine, 522 First Avenue, SRB 1107, New York, New York 10016, USA. [2]Howard Hughes Medical Institute, New York University School of Medicine, 522 First Avenue, SRB 1107, New York, New York 10016, USA. [3]Department of Biomedical Sciences and Human Oncology, CERMS, University of Torino, 10126 Torino, Italy. [4]San Giovanni Battista Hospital, Via Santena 7, 10126 Torino, Italy. [5]Medical Oncology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, Massachusetts 02115, USA.
*These authors contributed equally to this work.

preferentially binds SCF substrates[12] (Supplementary Fig. 2a). We also found that BCL6 and FBXO11 co-localized in the nucleus where they showed overlapping, punctate staining throughout the nucleoplasm (Supplementary Fig. 3). Moreover, expression of FBXO11 resulted in a marked reduction of BCL6 levels (Supplementary Fig. 4). This reduction in protein level was due to enhanced proteolysis, as shown by the decrease in BCL6 half-life (Supplementary Fig. 4a, b) and the rescue of BCL6 levels by either CUL1(1–242) (compare lanes 2 and 3 in whole-cell extract of Fig. 1a) or the addition of MG132, a proteasome inhibitor (Supplementary Fig. 3). In contrast, none of the other 20 F-box proteins tested promoted a reduction in BCL6 levels or stability (Supplementary Fig. 4 and data not shown).

To test further whether FBXO11 regulates the degradation of BCL6, we used two short hairpin RNA (shRNA) constructs to reduce FBXO11 expression in Ramos and SU-DHL6 cells. Depletion of FBXO11 by both shRNAs induced an increase in the steady-state levels and stability of BCL6 (Fig. 1c and Supplementary Fig. 5a, b). Similar effects were observed when the cellular expression of FBXO11 was silenced using three additional shRNAs or four different short interfering RNAs (siRNAs; Supplementary Fig. 5c–e).

We also observed that treatment of cells expressing FBXO11 with MG132 induced the appearance of high molecular mass species of BCL6 (Supplementary Fig. 6). These high molecular mass species are strongly induced by the expression of FBXO11 and are probably ubiquitylated forms of BCL6 as they became more evident in the presence of overexpressed ubiquitin. Moreover, we reconstituted the ubiquitylation of BCL6 *in vitro*. Immunopurified FBXO11 promoted the *in vitro* ubiquitylation of BCL6 only when the E1 and E2 enzymes were present in the reaction, and the ubiquitylation was inhibited by the presence of CUL1(1–242) (Fig. 1d). Methylated ubiquitin inhibited the formation of the highest molecular mass forms of BCL6, demonstrating that the high molecular mass forms of BCL6 are indeed polyubiquitylated. These results support the hypothesis that FBXO11 directly controls the ubiquitin-mediated degradation of BCL6.

Notably, FBXO11 silencing induced BCL6 stabilization in the absence of ligand-mediated receptor activation (Fig. 1c and Supplementary Fig. 5), which has been reported to accelerate BCL6 degradation by stimulating its ERK-dependent phosphorylation[7]. Accordingly, the binding of FBXO11 to BCL6 was unaffected in Ramos cells incubated with anti-IgM antibodies (a treatment that mimics B-cell antigen-receptor signalling and activates ERK2 (ref. 7)) both in the presence or absence of PD98059, a MEK inhibitor (Supplementary Fig. 7a). Similarly, when immunopurified BCL6 was phosphorylated *in vitro* by ERK2 or treated with λ-phosphatase, its binding to FBXO11 was unaffected (Supplementary Fig. 7b, c). Finally, a BCL6 mutant lacking the residues phosphorylated by ERK1/2 still bound FBXO11 efficiently (Supplementary Fig. 7d). Together, these results indicate that the FBXO11-dependent degradation of BCL6 is ERK-independent.

Together, the results in Fig. 1 and Supplementary Figs 1–6 demonstrate that FBXO11 mediates the ubiquitylation and degradation of BCL6. As the levels of BCL6 are elevated in B-cell malignancies and increased BCL6 expression has a crucial role in the pathogenesis of DLBCL, we determined whether the loss of FBXO11 might account for the elevated levels of BCL6. We examined the expression of FBXO11 and BCL6 in 22 B-cell lymphoma cell lines, including 16 DLBCL cell lines. We found that FBXO11 was absent or expressed at low levels in three DLBCL cell lines (FARAGE, MHHPREB1 and HT), and the expression was inversely correlated with BCL6 expression (Supplementary Fig. 8a). Using quantitative real-time polymerase chain reaction (qRT–PCR) on genomic DNA, we found that the *FBXO11* gene was homozygously deleted in FARAGE and MHHPREB1 cells and hemizygously deleted in HT cells (Supplementary Fig. 8b, c). We then measured the stability of BCL6 in the 16 DLBCL cell lines by treating cells with either cycloheximide, which blocks protein synthesis, or etoposide, which enhances BCL6 degradation[8]. In some samples, UO126, a MEK inhibitor, was also added. First, we observed that treatment of cells with UO126 did not stabilize BCL6 (compare the 4-h time point with cycloheximide in the presence or absence of UO126 in Fig. 2), further supporting the conclusion that the FBXO11-mediated degradation of BCL6 is ERK-independent. Importantly, we found that in FARAGE, MHHPREB1 and HT cells, BCL6 was not only expressed at high levels, but it was also more stable than in other cell lines (Fig. 2).

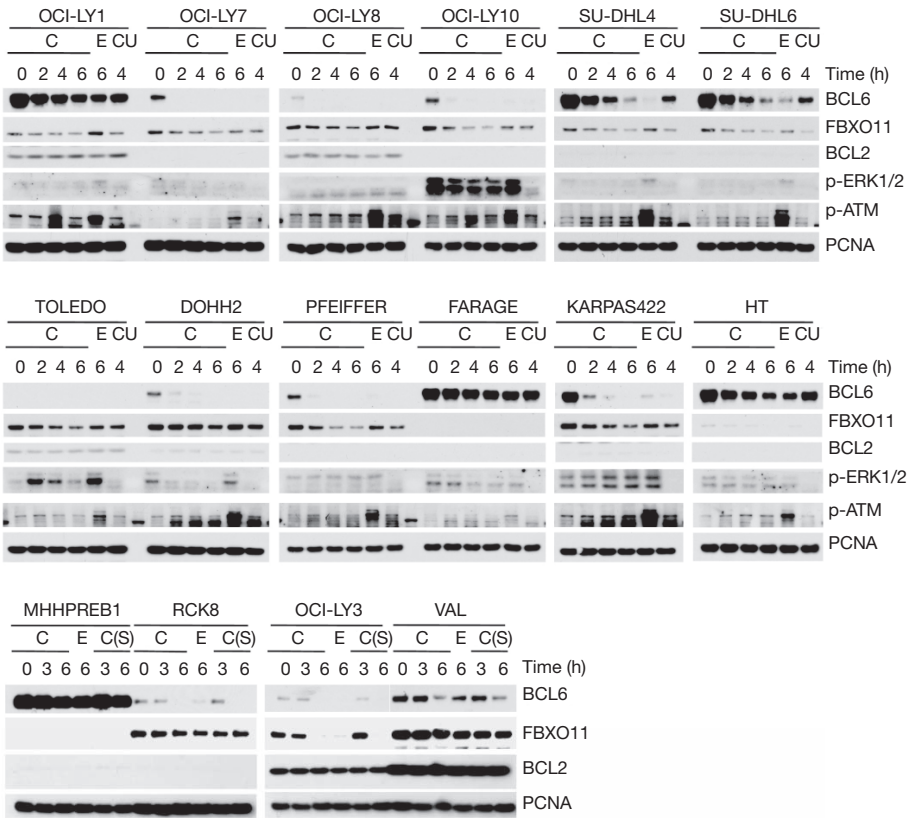


Figure 2 | DLBCL cell lines with *FBXO11* deletions or an *FBXO11* mutation display increased levels and stability of BCL6. Sixteen DLBCL cell lines were incubated with either cycloheximide (C), etoposide (E), or cycloheximide and UO126 (CU) for the indicated hours before immunoblotting for the indicated proteins. Where indicated, the intra S-phase (S) checkpoint was activated by incubating cells with thymidine for 20 h prior to treatment with cycloheximide.

Notably, the OCI-LY1 cell line also showed increased BCL6 stability, and sequencing of the *FBXO11* locus in OCI-LY1 cells revealed a point mutation in one of the CASH domains (Supplementary Fig. 9), the presumptive substrate recognition domains of FBXO11 (ref. 13).

We then used high-density, single nucleotide polymorphism (HD-SNP) genotyping in a cohort of 27 DLBCL cell lines and confirmed the presence of *FBXO11* deletions (including a focal deletion) in 14.8% of the samples (Supplementary Fig. 10). Public databases also indicate that the *FBXO11* locus is deleted in B-cell malignancies. For example, the GSK cancer cell line genomic profiling database shows that deletions in the *FBXO11* locus are present in six of 77 lymphoid cell lines, whereas only three of 255 non-lymphoid cell lines display *FBXO11* deletions (Supplementary Fig. 11). (Of these six lines, three are DLBCLs and one is a Burkitt's lymphoma.) Additionally, gene copy analyses in two different studies show *FBXO11* deletions in 6 of 69 primary DLBCLs (Supplementary Fig. 12) (refs 14, 15).

To investigate *FBXO11* gene inactivation in tumour samples further, we screened for mutations of *FBXO11* in 100 primary DLBCL samples (Supplementary Table 1). After filtering for known polymorphisms and synonymous mutations, potential inactivating mutations were further analysed and confirmed. This strategy identified a total of six sequence changes in four tumours, of which two were in the same codon (Supplementary Fig. 9). None of these four samples displayed either *BCL6* translocation or mutation of its promoter (Supplementary Table 2), and the analysis of paired normal DNA demonstrated the somatic origin of these events. Five of the six identified mutations are in the CASH domains, and one of these five changes creates a frame shift, resulting in chain termination and elimination of all three CASH domains. The fact that the point mutations are in highly conserved residues (Supplementary Fig. 9c) suggests that these FBXO11 mutants may be non-functional. Our analyses show that *FBXO11* is deleted (in most cases hemizygously) in 14.8% of DLBCL cell lines ($n = 27$) and

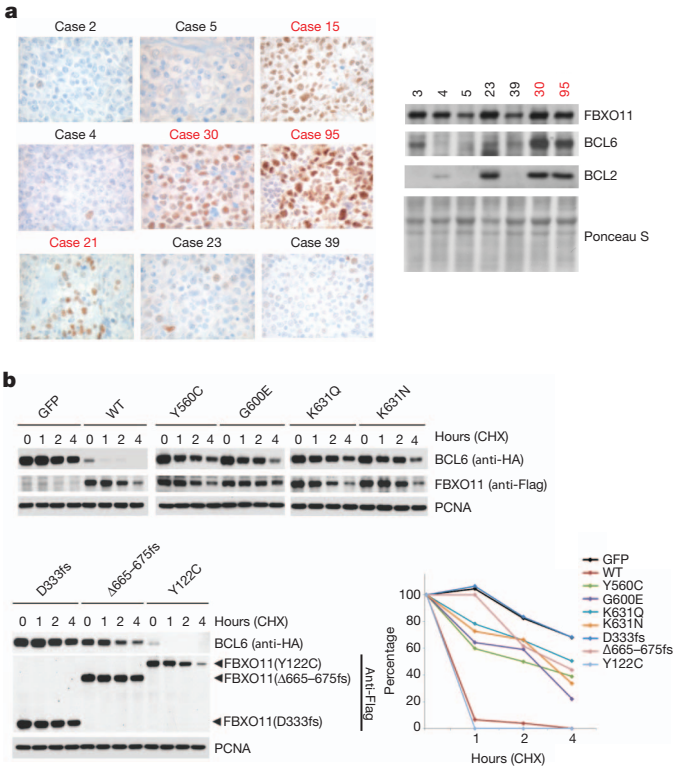


Figure 3 | Human DLBCLs with FBXO11 mutations display increased levels of BCL6, and FBXO11 tumour-derived mutants have impaired abilities to induce BCL6 degradation. **a**, Immunohistochemical stains for BCL6 from eight DLBCLs, including four tumours with mutations in *FBXO11*, are shown (brown) (×100). The right panel shows lysates from human DLBCLs that were analysed by immunoblotting for the indicated proteins. The red colour denotes DLBCL cases with *FBXO11* mutations. **b**, HEK-293T cells were transfected with BCL6 in combination with GFP, Flag-tagged wild type (WT) FBXO11, or the indicated Flag-tagged FBXO11 tumour-derived mutants. Twenty-four hours after transfection, cells were treated with cycloheximide (CHX) and samples were collected at the indicated time points for immunoblotting. The graph shows the quantification of BCL6 over the time course. The intensity of the bands from the experiment shown on the left was measured, and the ratio between the relative levels of BCL6 and PCNA in each time 0 was set as 100%.

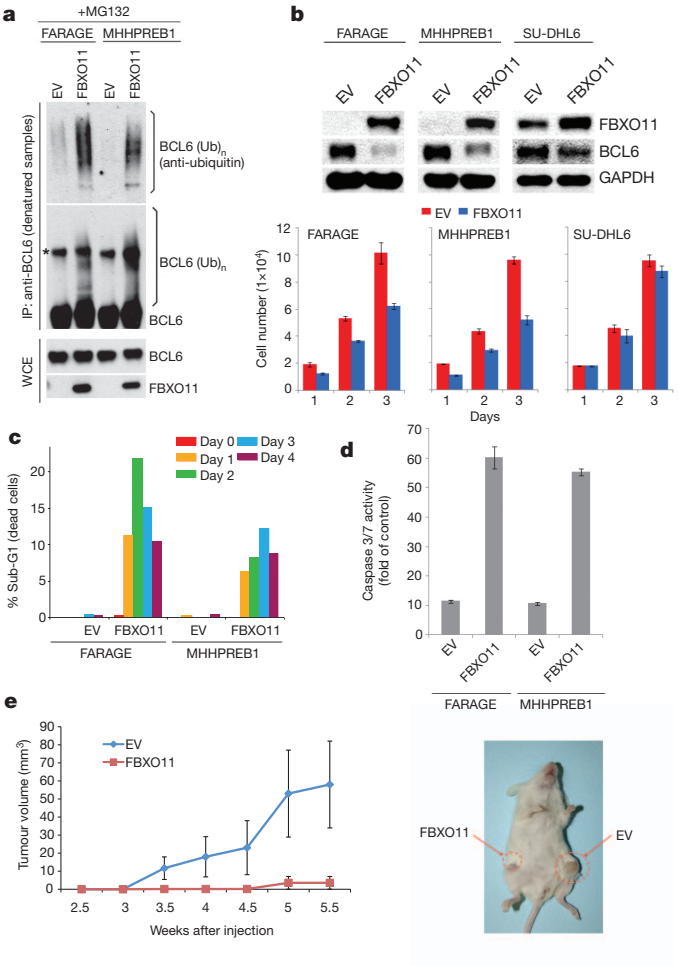


Figure 4 | Expression of FBXO11 in *FBXO11*-null cells promotes BCL6 ubiquitylation and degradation, inhibits cell proliferation and induces apoptosis. **a**, Denatured extracts from doxycycline- and MG132-treated FARAGE and MHHPREB1 cells stably transduced with a doxycycline-inducible FBXO11 construct or an empty virus (EV) were immunoprecipitated with an anti-BCL6 antibody and immunoblotted as indicated. The asterisk denotes a nonspecific band. Immunoblots of whole-cell extracts are also shown at the bottom. **b**, Extracts from doxycycline-treated FARAGE, MHHPREB1 and SU-DHL6 cells stably transduced with a doxycycline-inducible FBXO11 construct or an empty virus were immunoblotted as indicated. The graphs at the bottom show the number of cells measured in three different experiments (±s.d.) at different days after addition of doxycycline. **c**, Cells treated as in **b** for the indicated days were analysed by FACS. The graph represents the percentage of dying cells in the population (sub-G1 DNA content). **d**, Caspase 3 and Caspase 7 activity as measured in four different experiments (±s.d.) by the cleavage of a luminescent substrate in cells treated as in **b** for 2 days. The *y* axis indicates the caspase 3 and caspase 7 activity over cell number. The value given for the caspase activity in non-infected cells was set as 1. **e**, 1×10^7 FARAGE cells, stably transduced with a doxycycline-inducible FBXO11 construct (left flank) or an empty virus (right flank), were re-suspended in matrigel and injected subcutaneously into NOD/SCID mice. Two days after injection, doxycycline was administered in drinking water. Tumour growth was measured using a caliper at the indicated times after injection. $n = 4$ for each group; $P < 0.05$. Error bars indicate standard error of the mean (s.e.m.). The image on the right shows a representative mouse injected with the indicated cells.

mutated in 14.3% of DLBCL cell lines ($n = 7$) and 4% of primary DLBCLs ($n = 100$). Moreover, data available in public databases report *FBXO11* deletions in 8.7% of human DLBCLs ($n = 69$).

Notably, the DLBCL tumours with mutations in *FBXO11* expressed much higher BCL6 levels than control samples, as detected by immunohistochemistry (Fig. 3a). High BCL6 protein levels were also confirmed using immunoblotting when sufficient frozen tissue was available (Fig. 3a).

To investigate whether the *FBXO11* mutations in DLBCLs and OCI-LY1 cells interfere with FBXO11 activity, we generated complementary DNA constructs containing these mutations. Wild-type FBXO11 or FBXO11 mutants were expressed in HEK-293T cells to evaluate their ability to decrease BCL6 stability. Compared to wild-type FBXO11, tumour-derived FBXO11 mutants displayed an impaired ability to promote BCL6 degradation (Fig. 3b and Supplementary Fig. 13). Notably, FBXO11(Y122C), which has wild-type CASH domains, was the only mutant with activity similar to wild-type FBXO11. However, in the original sample this mutation was coupled with a deletion in FBXO11 that creates a frame shift and a chain termination (Δ666–675fs), which impaired the ability of FBXO11 to promote BCL6 degradation (Fig. 3b and Supplementary Fig. 13). We also found that mutations in FBXO11 modify its subcellular localization, preventing co-localization with BCL6 (Supplementary Fig. 14a, b). Moreover, FBXO11 mutants bind BCL6 less efficiently than wild-type FBXO11 (Supplementary Fig. 14c). These results indicate that the high levels of BCL6 observed in DLBCLs with FBXO11 mutations are due to a decreased ability of FBXO11 mutants to induce the proteolysis of BCL6.

To study the role of FBXO11 in tumorigenesis further, we used doxycycline-inducible retroviral expression vectors to express FBXO11 in two cell lines (FARAGE and MHHPREB1) with biallelic *FBXO11* deletion. BCL6 was poorly ubiquitylated in FARAGE and MHHPREB1 cells before FBXO11 reconstitution (Fig. 4a). Moreover, in parallel with a reduction in BCL6 levels, FBXO11 reconstitution inhibited cell proliferation in the *FBXO11*-null cell lines, but not a control line with an intact *FBXO11* locus (Fig. 4b). FBXO11 expression also induced apoptosis, as determined by a significant increase in the percentage of the sub-G1 population and activation of caspases 3 and 7 (Fig. 4c, d). Finally, FARAGE cells transduced with FBXO11 or an empty virus were injected subcutaneously into NOD/SCID mice. Induction of FBXO11 expression with doxycycline inhibited tumour growth, as revealed by measuring tumour mass (Fig. 4e).

Our results show that SCF^{FBXO11} is the ubiquitin ligase for BCL6 and that deregulation of this interaction may represent an oncogenic event. FBXO11 loss or mutation may contribute to DLBCL pathogenesis by allowing the accumulation of BCL6, an oncoprotein that has a critical role in lymphomagenesis. Because the deletions and mutations observed in our DLBCL samples were typically mono-allelic, we propose that FBXO11 is the product of a haplo-insufficient tumour suppressor gene.

METHODS SUMMARY

Biochemical methods. Extract preparation, immunoprecipitation and immunoblotting have previously been described[16]. Whole-cell lysates were generated by lysing cells in SDS lysis buffer (50 mM Tris pH 6.8, 2% SDS, 10% glycerol, 1 mM Na_3VO_4, 1 mM NaF and protease inhibitor mix (Roche)). *In vitro* ubiquitylation assays were preformed as described[16,17]. Briefly, haemagglutinin (HA)-tagged BCL6 was transfected into HEK-293T cells together with either Flag-tagged FBXO11 or an empty vector. Where indicated, MYC-tagged CUL1(1–242) was also transfected. Twenty-four hours after transfection, cells were incubated with MG132 for three hours before harvesting. Anti-Flag M2 agarose beads were used to immunoprecipitate the SCF^{FBXO11} complex. The beads were washed four times in lysis buffer and twice in ubiquitylation reaction buffer (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, and 0.5 mM dithiothreitol). Beads were then used for *in vitro* ubiquitylation assays in a volume of 30 μl, containing 0.1 μM E1 (Boston Biochem), 0.25 μM Ubch3, 0.25 μM Ubch5c, 1 μM ubiquitin aldehyde, 2.5 μg μl^{-1} ubiquitin, and 1× magnesium/ATP cocktail. Samples were incubated for 2 h at 30 °C and analysed by immunoblotting. *In vivo* ubiquitylation assays were preformed as described[18].

Mutation analysis. The complete coding sequences and exon/intron junctions of *FBXO11* were analysed by PCR amplification and sequencing. The reference sequences for all annotated exons and flanking introns of *FBXO11* were obtained from the UCSC Human Genome database using the mRNA accession NM_025133.4. PCR primers (see Supplementary Table 3), located ≥50 bp upstream or downstream of the target exon boundaries, were designed in the Primer 3 program (http://frodo.wi.mit.edu/primer3/). Sequences were compared to the corresponding germline sequences using the Mutation Surveyor Version 2.41 software package (Softgenetics; http://www.softgenetics.com). Synonymous mutations, previously reported polymorphisms (Human dbSNP Database at NCBI, build 130, and Ensembl Database) and changes present in the matched normal DNA were excluded.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 21 March; accepted 28 October 2011.
Published online 23 November 2011; corrected 4 January 2012 (see full-text HTML version for details).

1. Ci, W., Polo, J. M. & Melnick, A. B-cell lymphoma 6 and the molecular pathogenesis of diffuse large B-cell lymphoma. *Curr. Opin. Hematol.* **15,** 381–390 (2008).
2. Staudt, L. M. & Dave, S. The biology of human lymphoid malignancies revealed by gene expression profiling. *Adv. Immunol.* **87,** 163–208 (2005).
3. Baron, B. W. *et al.* The human BCL6 transgene promotes the development of lymphomas in the mouse. *Proc. Natl Acad. Sci. USA* **101,** 14198–14203 (2004).
4. Cattoretti, G. *et al.* Deregulated BCL6 expression recapitulates the pathogenesis of human diffuse large B cell lymphomas in mice. *Cancer Cell* **7,** 445–455 (2005).
5. Cardozo, T. & Pagano, M. The SCF ubiquitin ligase: insights into a molecular machine. *Nature Rev. Mol. Cell Biol.* **5,** 739–751 (2004).
6. Skaar, J. R., D'Angiolella, V., Pagan, J. K. & Pagano, M. SnapShot: F Box Proteins II. *Cell* **137,** 1358 (2009).
7. Niu, H., Ye, B. H. & Dalla-Favera, R. Antigen receptor signaling induces MAP kinase-mediated phosphorylation and degradation of the BCL-6 transcription factor. *Genes Dev.* **12,** 1953–1961 (1998).
8. Phan, R. T., Saito, M., Kitagawa, Y., Means, A. R. & Dalla-Favera, R. Genotoxic stress regulates expression of the proto-oncogene Bcl6 in germinal center B cells. *Nature Immunol.* **8,** 1132–1139 (2007).
9. Benmaamar, R. & Pagano, M. Involvement of the SCF complex in the control of Cdh1 degradation in S-phase. *Cell Cycle* **4,** 1230–1232 (2005).
10. Piva, R. *et al. In vivo* interference with Skp1 function leads to genetic instability and neoplastic transformation. *Mol. Cell. Biol.* **22,** 8375–8387 (2002).
11. Yen, H. C. & Elledge, S. J. Identification of SCF ubiquitin ligase substrates by global protein stability profiling. *Science* **322,** 923–929 (2008).
12. Skaar, J. R. & Pagano, M. Control of cell growth by the SCF and APC/C ubiquitin ligases. *Curr. Opin. Cell Biol.* **21,** 816–824 (2009).
13. Jin, J. *et al.* Systematic analysis and nomenclature of mammalian F-box proteins. *Genes Dev.* **18,** 2573–2580 (2004).
14. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30,** 207–210 (2002).
15. Kato, M. *et al.* Frequent inactivation of A20 in B-cell lymphomas. *Nature* **459,** 712–716 (2009).
16. D'Angiolella, V. *et al.* SCF(Cyclin F) controls centrosome homeostasis and mitotic fidelity through CP110 degradation. *Nature* **466,** 138–142 (2010).
17. Duan, S. *et al.* mTOR generates an auto-amplification loop by triggering the βTrCP- and CK1α-dependent degradation of DEPTOR. *Mol. Cell* **44,** 317–324 (2011).
18. Bloom, J., Amador, V., Bartolini, F., DeMartino, G. & Pagano, M. Proteasome-mediated degradation of p21 via N-terminal ubiquitinylation. *Cell* **115,** 71–82 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank W. Carroll, S. Shaham and L. Staudt for sharing unpublished results; A. Melnick, Y. Sun and Y. Xiong for reagents; L. Cerchietti, Y. Cheng, E. Dehan, V. Donato, N. V. Dorello, S. N. Yang and Z. Yao for advice and/or contributions to this work. M.P. is grateful to T. M. Thor and K. E. Davidson, and L.C. to Zuzana S. for continuous support. This work was supported by grants from the National Institutes of Health to M.P. (R01-GM57587, R37-CA76584 and R21-CA161108) and M.S. (PO1-CA092625), a grant from Susan G. Komen for the Cure to S.D., a Lymphoma Research Foundation Fellowship to J.K.P. and grants from AIRC and ERC (ERC-2009-StG-Proposal No242965-LUNELY) to R.C. M.P. is an Investigator with the Howard Hughes Medical Institute.

Author Contributions S.D., L.C. and J.K.P. planned and performed most experiments and helped to write the manuscript. M.P. coordinated the study, oversaw the results and wrote the manuscript. C.M., P.F.d.C. and R.C. provided the DLBCL tumour samples and some DLBCL cell lines, and performed some experiments. M.R. generated several constructs. B.C. and M.S. provided the HD-SNP data in Supplementary Fig. 10. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to M.P. (michele.pagano@nyumc.org) or R.C. (roberto.chiarle@unito.it).

METHODS

Cell culture and drug treatments. With the exception of OCI-LY7 and OCI-LY10, which were maintained in Iscove's Modified Dulbecco's Medium, all other B-cell lines were grown in RPMI-1640 medium. SK-MEL-28, HeLa, U-2OS and HEK-293T cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM). All media were supplemented with 10% fetal calf serum (FCS), 100 U ml^{-1} penicillin, 100 U ml^{-1} streptomycin and 2 mM L-glutamine. Where indicated, the following drugs were used: MG132 (10 μM), etoposide (20 μM) and cycloheximide (100 μg ml^{-1}).

Transient transfections. HEK-293T cells were transfected using polyethylenimine (PEI, Polysciences). HeLa, SK-MEL-28 and U-2OS cells were transfected using Lipofectamine 2000 (Invitrogen).

Gene silencing by siRNAs and shRNAs. All sequences of siRNAs and shRNAs are listed in Supplementary Table 3. shRNA lentiviruses were produced as described[8]. B cells were infected by re-suspension in the virus-containing supernatants and centrifugation. Twenty-four hours after infection, cells were re-suspended in fresh medium and selected with puromycin.

Antibodies. The following rabbit polyclonal antibodies were used: BCL6 (N-3; Santa Cruz Biotechnology), FBXO11 (Novus Bio), FBXO1 (C-20; Santa Cruz Biotechnology), FBXO5 (Invitrogen), FBXO31 (Bethyl), CUL1 (Invitrogen), FBXL1 (Invitrogen), BCL2 (C21; Santa Cruz Biotechnology), CDH1 (Invitrogen), CDC20 (Invitrogen), α-tubulin (Sigma), HA (Covance), Myc (Bethyl) and Flag (Sigma). Antibodies against FBXO18, FBXW9 and SKP1 were generated in the Pagano laboratory, and those to RBX1 were from Y. Xiong. Rabbit monoclonal antibodies were from Cell Signaling (FBXW1, GAPDH and phospho-ERK1/2 (Thr 202/Tyr 204)). The following mouse monoclonal antibodies were used: BCL6 (5G11; Santa Cruz Biotechnology), BCL6 (PG-B6p, Dako), BCL2 (C2; Santa Cruz Biotechnology), CD10 (56C6, Novocastra), pATM (S1981) (Cell Signaling), PCNA (PC10, Invitrogen), MUM1/IRF4 (MUM1p, Dako), p53 (DO-1, Santa Cruz Biotechnology), ubiquitin (Millipore) and β-actin (Sigma).

***In vitro* binding assay.** HA-tagged BCL6 was *in vitro* transcribed/translated using the TnT T3 Coupled Wheat Germ Extract System (Promega) and purified with anti-HA agarose (Roche). Bead-bound BCL6 was incubated with soluble Flag-tagged FBXO11 (purified from HEK293T cells), re-purified with HA agarose beads and subjected to immunoblotting. In some cases, bead-bound BCL6 was either treated with λ-phosphatase (NEB) or phosphorylated with ERK2 before extensive washing and incubation with Flag-tagged FBXO11. Kinase assays were performed by incubating bead-bound BCL6 at 30 °C for 1 h with 2.5 μCi [γ-^{32}P]-ATP, 2 mM ATP and 200 ng active, recombinant, purified ERK2 (Millipore) in 30 μl kinase buffer (10 mM MOPS, pH 7.6, 1× magnesium/ATP cocktail (Upstate), 10 mM $MnCl_2$, 1 mM dithiothreitol).

Apoptosis and *in vivo* assays. 1×10^5 FARAGE, MHHPREB1, or SU-DHL6 cells stably transduced with a tetracycline-inducible FBXO11 construct or an empty virus were seeded in the presence of doxycycline. Cells were collected at different days, washed in PBS, and fixed in 70% cold ethanol. Cell death was measured by propidium iodide staining and FACS. The percentage of sub-G1 phase cells was calculated using FlowJo software (FlowJo). Apoptosis was determined by measuring the activity of the caspases 3 and 7 using a luminescent substrate (Caspase-Glo 3/7; Promega). *In vivo* tumorigenicity assay was performed as described[15]. All animal procedures followed NIH protocols and were approved by the University Animal Institute Committee.

Immunofluorescence and immunohistochemistry. For indirect immunofluorescence staining, cells were grown on chamber slides, pre-extracted with CSK buffer (0.5% Triton X-100, 1 mM EDTA, 100 mM NaCl, 300 mM sucrose, 3 mM $MgCl_2$, 1 mM Na_3VO_4, 1 mM NaF, 2 mM $CaCl_2$, 10 mM HEPES pH 7.4) and fixed with 4% paraformaldehyde. Cells were permeabilized with PBS/0.2% Triton X-100 and blocked in PBS/0.1% Tween containing 3% BSA before incubation with primary antibodies. Alexa Fluor-conjugated 555 donkey anti-rabbit and Alexa Fluor 488-conjugated goat anti-mouse IgG were used as secondary antibodies. DAPI was used to counterstain DNA. Slides were mounted with Prolong-Gold (Invitrogen). Image acquisition was performed using a Zeiss Axiovert 200 M microscope, equipped with a cooled Retiga 2000R CCD (QImaging) and Metamorph Software (Molecular Devices). Immunohistochemistry was performed as described[19].

Intraperitoneal immunization and preparation of activated B cells. Sheep blood Alsevers (BD) were washed with PBS and re-suspended in PBS at a concentration of 1×10^9 sheep red blood cells (SRBCs) per ml. Mice were then immunized intraperitoneally with 2×10^8 SRBCs in a 200 μl volume. After 5 days, mice were immunized again using 5–10-fold more SRBCs. Spleens were collected 7–10 days after immunization, placed on ice, washed in PBS to remove residual blood, cut into small pieces, crushed, physically dissociated using a Falcon cell strainer, and subjected to hypotonic lysis of erythrocytes.

Preparation of genomic DNA. DNA from DLBCL samples was generated by proteinase K digestion and phenol-chloroform extraction. Genomic DNA from cell lines was prepared using QIAamp DNA kits (Qiagen).

Copy number determination by quantitative RT–PCR of genomic DNA. Genomic DNA was prepared using QIAamp DNA Kits (Qiagen). The DNA was amplified with Absolute Blue QPCR SYBR-Green Mix (Thermo Scientific) using a Light Cycler 480 System (Roche). Primers are listed in Supplementary Table 3. Normalization of gene copy number was performed with additional loci on chromosomes 5, 6 and 12. Genomic DNA from primary human diploid fibroblasts was used as a 2N DNA standard.

Quantitative RT–PCR of cDNA. Total RNA was prepared using RNeasy Kits (Qiagen) and treated with DNase I, and cDNA was generated by reverse transcription. The cDNA was amplified with Absolute Blue QPCR SYBR-Green Mix (Thermo Scientific) using a Light Cycler 480 System (Roche). Primers are listed in Supplementary Table 3. Normalization across samples was performed using the average gene expression of *SDHA*, *GAPDH* and *RPS13*.

Copy number determination by GEO database data analysis. The CEL files used for this study were obtained from the National Center for Biotechnology Information, Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/projects/geo). Raw data were processed, analysed and visualized using Genespring GX 11.5 software (Agilent).

19. Latres, E. *et al.* Role of the F-box protein Skp2 in lymphomagenesis. *Proc. Natl Acad. Sci. USA* **98,** 2515–2520 (2001).

doi:10.1038/nature10670

Open structure of the Ca^{2+} gating ring in the high-conductance Ca^{2+}-activated K^+ channel

Peng Yuan[1], Manuel D. Leonetti[1], Yichun Hsiung[1] & Roderick MacKinnon[1]

High-conductance voltage- and Ca^{2+}-activated K^+ channels function in many physiological processes that link cell membrane voltage and intracellular Ca^{2+} concentration, including neuronal electrical activity, skeletal and smooth muscle contraction, and hair cell tuning[1–8]. Like other voltage-dependent K^+ channels, Ca^{2+}-activated K^+ channels open when the cell membrane depolarizes, but in contrast to other voltage-dependent K^+ channels, they also open when intracellular Ca^{2+} concentrations rise. Channel opening by Ca^{2+} is made possible by a structure called the gating ring, which is located in the cytoplasm. Recent structural studies have defined the Ca^{2+}-free, closed, conformation of the gating ring, but the Ca^{2+}-bound, open, conformation is not yet known[9]. Here we present the Ca^{2+}-bound conformation of the gating ring. This structure shows how one layer of the gating ring, in response to the binding of Ca^{2+}, opens like the petals of a flower. The degree to which it opens explains how Ca^{2+} binding can open the transmembrane pore. These findings present a molecular basis for Ca^{2+} activation of K^+ channels and suggest new possibilities for targeting the gating ring to treat conditions such as asthma and hypertension.

Regulators of K^+ conductance (RCK) domains are ubiquitous among ion channels and transporters in prokaryotic cells[10–14]. Eight RCK domains assemble in the cytoplasm to form a closed-ring structure that changes its diameter on ligand binding, thus enabling the gating ring to regulate allosterically the transmembrane component of the transport protein. In prokaryotic cells, the gating ring most often consists of eight identical RCK domains, arranged as a tetrad of pairs, giving rise to a four-fold symmetric ring with identical top and bottom layers of RCK domains[12,15]. Intracellular ligands such as Ca^{2+} or small organic molecules bind in a cleft between RCK pairs forming the top and bottom layers to affect the shape of the ring[12,15–17].

RCK domains are also found in multicellular eukaryotes in the Slo family of K^+ channels[9,10,18,19]. There, two non-identical RCK domains are encoded in the carboxy terminus of the K^+ channel subunit. A tetrameric K^+ channel thus provides eight RCK domains to make a gating ring, but in contrast to most prokaryotic gating rings with identical RCK domains, the eukaryotic gating ring has one kind of RCK domain (RCK1) forming its top layer and another kind (RCK2) forming its bottom layer (Fig. 1a). The structure of a Ca^{2+}-activated (Slo1 or BK) K^+-channel gating ring in its Ca^{2+}-free, closed, conformation was recently determined (Protein Data Bank accession 3NAF), as was the structure of an RCK1–RCK2 pair in the presence of Ca^{2+} (PDB accession 3MT5), detailing the chemistry of the Ca^{2+}-binding site[9,18]. These studies showed that the BK gating ring indeed has distinct top and bottom layers, and also that a Ca^{2+}-binding site known as the Ca^{2+} bowl is in a different location than the ligand-binding site in prokaryotic gating rings. These unique structural properties suggest that to regulate conduction, the eukaryotic gating ring may undergo conformational changes in a different manner than its prokaryotic counterpart.

We determined the crystal structure of the Ca^{2+}-bound gating ring from the zebrafish BK channel at a resolution of 3.6 Å using an expression construct in which a loop consisting of residues 839–872 was deleted (Fig. 1b, Supplementary Table 1 and Supplementary Figs 1 and 2). This loop was disordered in both the Ca^{2+}-free human BK gating ring and the Ca^{2+}-bound RCK1–RCK2 pair (PDB accessions as above), and its deletion in both the human and zebrafish BK channels has no detectable affect on function[9,18] (Supplementary Fig. 3). Therefore, the relative stability of conformational states of the BK gating ring seem to be unaffected by this loop. Initial phases were determined by molecular replacement using the RCK1–RCK2 Ca^{2+}-bound structure as a search model[18]. The solution showed the presence of eight RCK1–RCK2 pairs (two gating rings) per asymmetric unit of the crystal, which allowed eight-fold non-crystallographic symmetry averaging of electron density maps and eight-fold restraints throughout model refinement. The first step of refinement required rigid-body adjustment of the amino-terminal lobe of RCK1. The final model was refined to working and free residuals (R_{work} and R_{free}) of 26.0% and 28.9%, respectively.

The zebrafish channel is 93% identical in amino-acid sequence to the human BK channel, for which the Ca^{2+}-free, presumably closed, conformation is known[9] (Fig. 1c and sequence alignment in Supplementary Information). The high degree of identity between zebrafish and human BK channels justifies comparison of the two structures to assess potential conformational changes underlying Ca^{2+}-mediated gating. The Ca^{2+}-bound and Ca^{2+}-free gating rings are shown in Fig. 1b, c. In these representations, RCK1, which forms the layer of the gating ring closest to the membrane, is coloured blue and RCK2 is coloured red. In both gating rings, a green asterisk marks the N-terminal residue of RCK1, Lys 343. In a full-length BK channel, the N-terminal residue of RCK1 is connected by a 16-amino-acid linker to the C terminus of the transmembrane inner helix, which forms the pore's gate[20].

The two gating ring structures, Ca^{2+}-bound and Ca^{2+}-free, differ in two main respects. First, in the Ca^{2+}-bound structure electron density attributable to Ca^{2+} is present in the Ca^{2+} bowl at the 'assembly interface', and, second, the RCK1 layer in the Ca^{2+}-bound gating ring is expanded from a diameter of 81 to 93 Å, measured at the position of Lys 343 (Fig. 1d, e and Supplementary Fig. 2c, d). An objective comparison of conformational differences between the Ca^{2+}-bound and Ca^{2+}-free gating rings is achieved by superimposing all Cα atoms of the tetrameric gating rings. Such a superposition shows that the RCK2 components of the ring undergo very little change, and that only modest conformational differences occur at the assembly interface, except locally near the Ca^{2+} bowl (Fig. 2c and Supplementary Fig. 4). The Ca^{2+} bowl structure is nearly identical to that of the Ca^{2+}-bound RCK1–RCK2 pair[18] (Supplementary Fig. 2e). The major conformational change involves the N-terminal lobe of RCK1, which changes its angle with respect to RCK2 (Fig. 2a, b and Supplementary Fig. 5). Viewing an interpolative movie in which the Ca^{2+}-free gating ring is morphed to the Ca^{2+}-bound gating ring, it seems that Ca^{2+} binding causes the N-terminal lobes of the RCK1 units to 'open up' on the membrane-facing surface of the gating ring in a way akin to petals opening on a flower (Supplementary Movie 1, part 1).

We next analysed whether these conformational changes are compatible with prior structural data on the opening and closing of K^+

[1]Laboratory of Molecular Neurobiology and Biophysics, Rockefeller University, Howard Hughes Medical Institute, 1230 York Avenue, New York, New York 10065, USA.

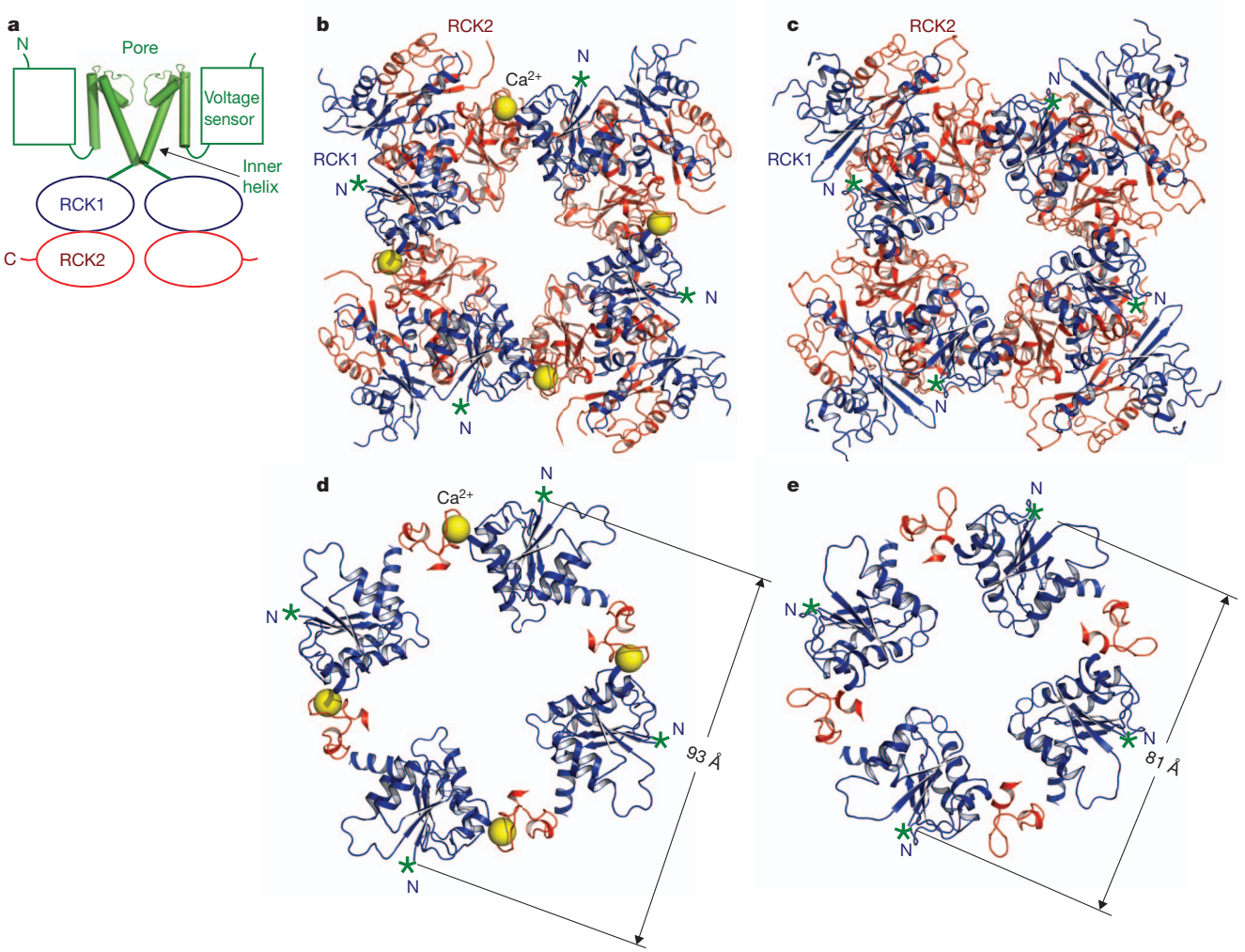


Figure 1 | The BK channel and the gating ring. **a**, Domain topology of the BK channel. For clarity, only two opposing subunits are shown. **b**, Crystal structure of the Ca^{2+}-bound, open, gating ring, with RCK1 in blue and RCK2 in red. The Ca^{2+} ions are shown as yellow spheres. The N termini of RCK1 that connect to the C termini of the inner helices are indicated as green asterisks. **c**, Structure of the Ca^{2+}-free, closed, gating ring, with RCK1 in blue and RCK2 in red (PDB accession 3NAF). **d**, The RCK1 N-terminal lobes (blue) and the Ca^{2+} bowls (red) from the Ca^{2+}-bound, open, gating ring. The diagonal distance between the Cα atoms of the N-terminal residues (Lys 343) is indicated. **e**, The corresponding region to that shown in **d** from the closed gating ring (PDB accession 3NAF).

channels. We started with the open prokaryotic K^+ channel MthK and replaced its gating ring with the Ca^{2+}-bound BK-channel gating ring[15]. This step involved alignment of the four-fold axis followed by minimization of the distance between the N terminus of the BK gating ring and the corresponding position within the MthK gating ring structure, through least-squares minimization of the distances between residues 343 of all four BK subunits and the corresponding residues, 116, of all four MthK subunits. The resulting model of an open BK-channel gating ring with an open K^+ pore is shown in Fig. 3a. The distance between the C-terminal residue of the inner helix of the pore and the N-terminal residue of RCK1, which was not constrained in making the model, is 33 Å. In a second step, we replaced the open MthK pore with the closed KcsA pore by aligning the invariant selectivity filter, and replaced the Ca^{2+}-bound BK-channel gating ring with the Ca^{2+}-free BK-channel gating ring by aligning the nearly invariant RCK2 layer[9,21]. The resulting model of a Ca^{2+}-free BK gating ring with a closed K^+ pore is shown in Fig. 3b. In this closed model, the distance between the C-terminal residue of the inner helix and the N-terminal residue of RCK1, which again was not constrained in making the model, is 32 Å. In this analysis, the similar pore-to-gating-ring distance (~32 Å) in the open model and the closed model suggests that the magnitude of conformational change observed within the gating ring is compatible with the magnitude of change known to occur in the pores of other kinds of K^+ channel (Supplementary Movie 1, part 2). Whether or not the pore of the BK channel undergoes conformational changes similar to or different than those of other K^+ channels remains to be determined[22–26].

Another feature of the model in Fig. 3 is noteworthy. In the published structure of the Ca^{2+}-free gating ring, part of the linker extending centrally from the gating ring towards the pore was present in the crystal structure and visible on the surface of the gating ring[9]. The linker was held by a four-helix bundle fused to the N terminus to facilitate crystallization. Although the linker position in the structure might be affected by this attachment, we note that it is almost coincident with the dashed lines in our closed model (Fig. 3b). The agreement between the observed linker position in the crystal structure of the closed gating ring and its inferred position in the closed conformation of our model is an interesting correlation because the radial direction of the linkers in the model (dashed lines) would naturally allow the expansion of the gating ring to exert an opening force on the pore's inner helices. The idea that the gating ring exerts a direct force on the pore has been proposed on the basis of mutagenesis studies showing that the linker length is critical to efficient Ca^{2+}-mediated gating[19,20].

The structural differences between the BK gating ring and the MthK gating ring underlie distinct ligand-induced conformational changes. A higher degree of molecular symmetry in the MthK gating ring stems from its containing eight identical RCK domains[12,15]. Consequently, the MthK gating ring changes equally on both layers such that Ca^{2+}

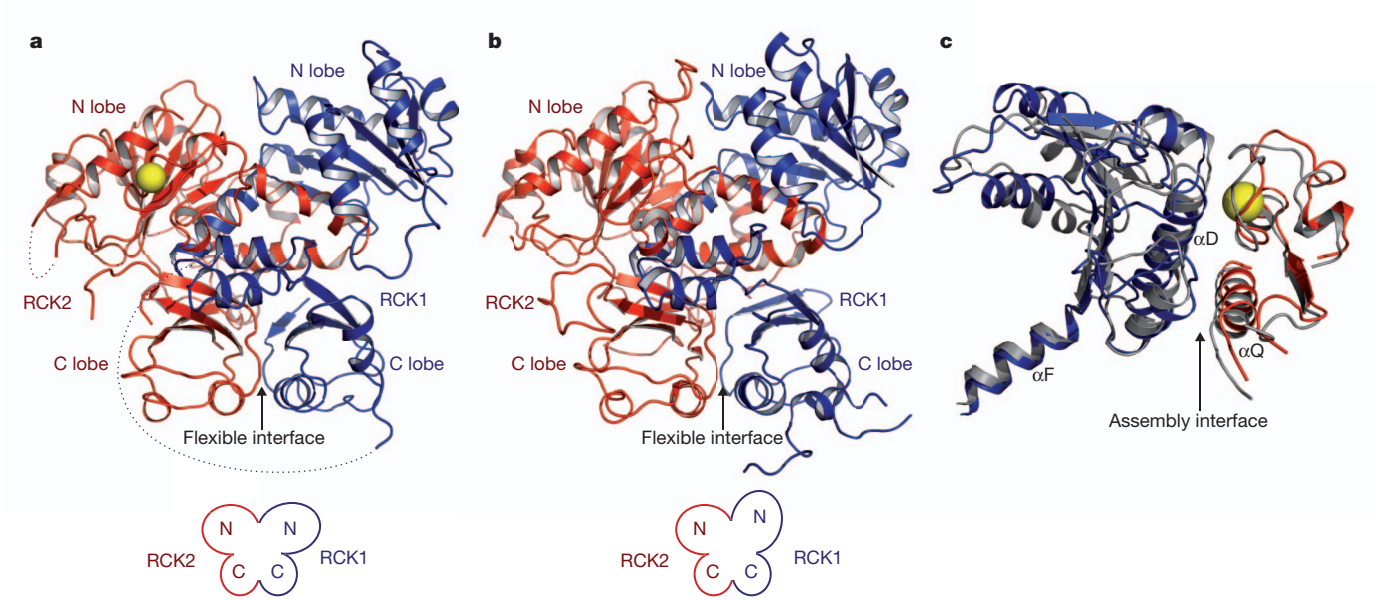


Figure 2 | The flexible interface and the assembly interface. **a**, The flexible interface from the open gating ring, with RCK1 in blue and RCK2 in red. The Ca^{2+} ions are shown as yellow spheres. The N- and C-terminal lobes of RCK1 and RCK2 are labelled. Large disordered regions are indicated by dotted lines. **b**, The flexible interface from the closed gating ring, with RCK1 in blue and RCK2 in red (PDB accession 3NAF). **c**, Superposition of the assembly interfaces from the open and closed gating ring structures. Secondary structural elements αF, αD and αQ are labelled.

binding causes a shrinking of the ring's height and an expansion of its diameter[12] (Supplementary Movie 2). By contrast, in the BK gating ring Ca^{2+} binding causes only the layer facing the membrane to undergo a substantial conformational change (Supplementary Movie 1). This comparison presents a fascinating example in which the evolution of molecular structure has given rise to new or modified mechanical properties within a class of molecules. Precisely how the free energy of Ca^{2+} binding in either the BK or MthK gating rings is transduced into

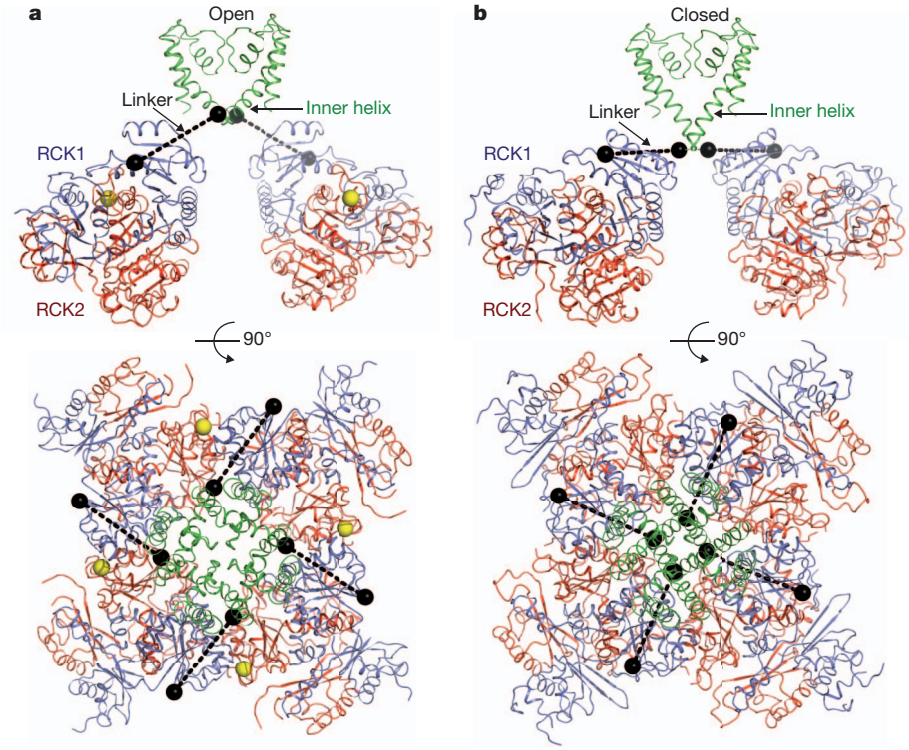


Figure 3 | A Ca^{2+}-gating model for the BK channel. **a**, Views of the BK-channel model with an open gating ring and an open pore. The structure of the Ca^{2+}-bound BK gating ring is docked onto the open MthK channel (PDB accession 1LNQ) by aligning the gating rings. Yellow spheres represent Ca^{2+} ions. The N-terminal residues of RCK1 and the C-terminal residues of the inner helices are shown as black spheres. **b**, Views of the BK channel model with a Ca^{2+}-free gating ring and a closed pore. Here, the Ca^{2+}-bound gating ring in **a** is replaced by the Ca^{2+}-free gating ring (PDB accession 3NAF) and the open MthK pore in **a** is replaced by the closed KcsA pore (PDB accession 1K4C). The inner helix from KcsA is truncated to match the length in MthK. See Supplementary Movie 1.

mechanical work remains an important outstanding question that will require further experimental and theoretical work to understand at a deep level.

The data and analysis presented here define end points in the Ca^{2+}-induced gating conformational transition of the BK gating ring. Given the large conformational change that occurs on Ca^{2+} binding, it should in principle now be possible to identify small molecules that stabilize one conformation or the other. Such small molecules promise to have important and possibly beneficial physiological effects because BK channels regulate smooth muscle tone in pulmonary airways and vascular beds[27–30].

METHODS SUMMARY

The cytoplasmic domain construct of the zebrafish BK channel including residues 341–1,060 with a loop deletion (residues 839–872) was expressed and purified as described previously[18]. However, the size exclusion chromatography was carried out in buffer containing 20 mM HEPES (pH 8.0), 150 mM NaCl and 20 mM dithiothreitol. The protein was concentrated to about 6 mg ml^{-1} and 10 mM $CaCl_2$ was added before crystallization experiments were performed. Crystals were grown at 20 °C using hanging-drop vapour diffusion by mixing equal volumes of protein and a reservoir solution containing 50 mM 2-(*N*-morpholino)ethanesulfonic acid, 4% (w/v) PEG 4000 and 100 mM potassium sodium tartrate (pH 6.3). The crystals belong to space group $P2_12_12_1$ and the unit cell has dimensions of $a = 137.65$ Å, $b = 210.82$ Å and $c = 238.76$ Å with $\alpha = \beta = \gamma = 90°$. Each asymmetric unit contains eight protein subunits forming two gating rings. The structure was determined by molecular replacement using the monomeric Ca^{2+}-bound human BK cytoplasmic domain (PDB accession 3MT5) as the search model. The final model was refined to a resolution of 3.6 Å with $R_{work} = 0.260$ and $R_{free} = 0.289$. Electrophysiology recordings were conducted using patch–clamp on *Xenopus oocytes* expressing the wild-type and mutant channels.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 29 July; accepted 24 October 2011.
Published online 4 December 2011.

1. Robitaille, R., Garcia, M. L., Kaczorowski, G. J. & Charlton, M. P. Functional colocalization of calcium and calcium-gated potassium channels in control of transmitter release. *Neuron* **11,** 645–655 (1993).
2. Fettiplace, R. & Fuchs, P. A. Mechanisms of hair cell tuning. *Annu. Rev. Physiol.* **61,** 809–834 (1999).
3. Nelson, M. T. *et al.* Relaxation of arterial smooth muscle by calcium sparks. *Science* **270,** 633–637 (1995).
4. Brenner, R. *et al.* Vasoregulation by the β1 subunit of the calcium-activated potassium channel. *Nature* **407,** 870–876 (2000).
5. Petkov, G. V. *et al.* β1-subunit of the Ca^{2+}-activated K^+ channel regulates contractile activity of mouse urinary bladder smooth muscle. *J. Physiol. (Lond.)* **537,** 443–452 (2001).
6. Lee, U. S. & Cui, J. BK channel activation: structural and functional insights. *Trends Neurosci.* **33,** 415–423 (2010).
7. Kaczorowski, G. J., Knaus, H. G., Leonard, R. J., McManus, O. B. & Garcia, M. L. High-conductance calcium-activated potassium channels; structure, pharmacology, and function. *J. Bioenerg. Biomembr.* **28,** 255–267 (1996).
8. Salkoff, L., Butler, A., Ferreira, G., Santi, C. & Wei, A. High-conductance potassium channels of the SLO family. *Nature Rev. Neurosci.* **7,** 921–931 (2006).
9. Wu, Y., Yang, Y., Ye, S. & Jiang, Y. Structure of the gating ring from the human large-conductance Ca^{2+}-gated K^+ channel. *Nature* **466,** 393–397 (2010).
10. Jiang, Y., Pico, A., Cadene, M., Chait, B. T. & MacKinnon, R. Structure of the RCK domain from the *E. coli* K^+ channel and demonstration of its presence in the human BK channel. *Neuron* **29,** 593–601 (2001).
11. Albright, R. A., Ibar, J. L., Kim, C. U., Gruner, S. M. & Morais-Cabral, J. H. The RCK domain of the KtrAB K^+ transporter: multiple conformations of an octameric ring. *Cell* **126,** 1147–1159 (2006).
12. Ye, S., Li, Y., Chen, L. & Jiang, Y. Crystal structures of a ligand-free MthK gating ring: insights into the ligand gating mechanism of K^+ channels. *Cell* **126,** 1161–1173 (2006).
13. Bakker, E. P., Booth, I. R., Dinnbier, U., Epstein, W. & Gajewska, A. Evidence for multiple K^+ export systems in *Escherichia coli. J. Bacteriol.* **169,** 3743–3749 (1987).
14. Roosild, T. P., Miller, S., Booth, I. R. & Choe, S. A mechanism of regulating transmembrane potassium flux through a ligand-mediated conformational switch. *Cell* **109,** 781–791 (2002).
15. Jiang, Y. *et al.* Crystal structure and mechanism of a calcium-gated potassium channel. *Nature* **417,** 515–522 (2002).
16. Bellamacina, C. R. The nicotinamide dinucleotide binding motif: a comparison of nucleotide binding proteins. *FASEB J.* **10,** 1257–1269 (1996).
17. Dong, J., Shi, N., Berke, I., Chen, L. & Jiang, Y. Structures of the MthK RCK domain and the effect of Ca^{2+} on gating ring stability. *J. Biol. Chem.* **280,** 41716–41724 (2005).
18. Yuan, P., Leonetti, M. D., Pico, A. R., Hsiung, Y. & MacKinnon, R. Structure of the human BK channel Ca^{2+}-activation apparatus at 3.0 Å resolution. *Science* **329,** 182–186 (2010).
19. Pico, A. R. *RCK Domain Model of Calcium Activation in BK Channels* 43–46. PhD thesis, Rockefeller Univ. (2003).
20. Niu, X., Qian, X. & Magleby, K. L. Linker-gating ring complex as passive spring and Ca^{2+}-dependent machine for a voltage- and Ca^{2+}-activated potassium channel. *Neuron* **42,** 745–756 (2004).
21. 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).
22. Wilkens, C. M. & Aldrich, R. W. State-independent block of BK channels by an intracellular quaternary ammonium. *J. Gen. Physiol.* **128,** 347–364 (2006).
23. Li, W. & Aldrich, R. W. State-dependent block of BK channels by synthesized shaker ball peptides. *J. Gen. Physiol.* **128,** 423–441 (2006).
24. Zhou, Y., Xia, X. M. & Lingle, C. J. Cysteine scanning and modification reveal major differences between BK channels and Kv channels in the inner pore region. *Proc. Natl Acad. Sci. USA* **108,** 12161–12166 (2011).
25. Tang, Q. Y., Zeng, X. H. & Lingle, C. J. Closed-channel block of BK potassium channels by bbTBA requires partial activation. *J. Gen. Physiol.* **134,** 409–436 (2009).
26. Li, W. & Aldrich, R. W. Unique inner pore properties of BK channels revealed by quaternary ammonium block. *J. Gen. Physiol.* **124,** 43–57 (2004).
27. Wu, R. S. & Marx, S. O. The BK potassium channel in the vascular smooth muscle and kidney: α- and β-subunits. *Kidney Int.* **78,** 963–974 (2010).
28. Ghatta, S., Nimmagadda, D., Xu, X. & O'Rourke, S. T. Large-conductance, calcium-activated potassium channels: structural and functional implications. *Pharmacol. Ther.* **110,** 103–116 (2006).
29. Miura, M., Belvisi, M. G., Stretton, C. D., Yacoub, M. H. & Barnes, P. J. Role of potassium channels in bronchodilator responses in human airways. *Am. Rev. Respir. Dis.* **146,** 132–136 (1992).
30. Jones, T. R., Charette, L., Garcia, M. L. & Kaczorowski, G. J. Interaction of iberiotoxin with β-adrenoceptor agonists and sodium nitroprusside on guinea pig trachea. *J. Appl. Physiol.* **74,** 1879–1884 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank staff members at NSLS X29, Brookhaven National Laboratory, for beamline assistance, and members of the MacKinnon laboratory for discussion. We thank P. Hoff and members of the Gadsby laboratory for help with oocyte preparation. R.M. is an investigator in the Howard Hughes Medical Institute. The research is supported by the American Asthma Foundation grant 07-0127.

Author Contributions P.Y. purified and crystallized the protein, collected the X-ray diffraction data, determined the structure and conducted electrophysiology recordings. M.D.L. aided in initial crystallization and electrophysiology experiments. Y.H. provided assistance with protein expression. P.Y. and R.M. designed the research and analysed data. P.Y., M.D.L. and R.M. prepared the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structure have been deposited into the Protein Data Bank under accession code 3U6N. 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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to R.M. (mackinn@rockefeller.edu).

METHODS

Cloning, expression and purification. A synthetic gene encoding the zebrafish BK channel (GI: 189526846) was purchased from Bio Basic Inc. and served as the template for subcloning. The amino-acid sequence is 93% identical to that of the human BK channel. On the basis of previous structural and functional studies of the human BK channel[18], a cytoplasmic domain (CTD) construct including residues 341–1,060 with a loop deletion (residues 839–872) was expressed and purified following a method previously described[18]. However, the size exclusion chromatography was carried out in 20 mM HEPES (pH 8.0), 150 mM NaCl and 20 mM dithiothreitol. The protein was concentrated to about 6 mg ml^{-1} and 10 mM $CaCl_2$ was added to achieve the ligand-bound, open, conformation before crystallization experiments.

Crystallization and structure determination. Crystals of the zebrafish BK CTD were grown at 20 °C using hanging-drop vapour diffusion by mixing equal volumes of protein and a reservoir solution containing 50 mM 2-(*N*-morpholino)ethanesulfonic acid, 4% (w/v) PEG 4000 and 100 mM potassium sodium tartrate (pH 6.3). The crystals grew to a maximum size of about 0.2 mm × 0.2 mm × 0.3 mm within three days. Before being flash frozen in liquid nitrogen, the crystals were briefly transferred to a cryoprotectant solution containing 50 mM 2-(*N*-morpholino)ethanesulfonic acid, 6% (w/v) PEG 4000, 100 mM potassium sodium tartrate, 150 mM NaCl, 10 mM $CaCl_2$ and 30% ethylene glycol (pH 6.3). The crystals belong to space group $P2_12_12_1$ and the unit cell has dimensions of $a = 137.65$ Å, $b = 210.82$ Å and $c = 238.76$ Å with $\alpha = \beta = \gamma = 90°$. Each asymmetric unit contains eight protein subunits.

Diffraction data were measured at beamline X29 of the National Synchrotron Light Source and were processed with the HKL2000 program suite[31]. The scaled data set was anisotropically corrected to resolution limits of 4.0, 3.6 and 3.6 Å along the reciprocal cell directions a^*, b^*, and c^*, respectively, using the diffraction anisotropy server at the University of California, Los Angeles[32]. An isotropic B factor of -29.98 Å^2 was applied to restore the magnitude of the high-resolution reflections diminished by anisotropic scaling. The structure was determined by molecular replacement using the monomeric Ca^{2+}-bound human BK CTD (PDB accession 3MT5) as the search model[18]. Initial phases were improved by eight-fold non-crystallographic symmetry averaging carried out with DM from the CCP4 program suite[33]. Iterative model building was carried out in COOT[34]. Rounds of refinement were performed with CNS[35] and REFMAC[36], and strong eight-fold non-crystallographic symmetry restraints were maintained throughout refinement. The final model was refined to a resolution of 3.6 Å with $R_{work} = 0.260$ and $R_{free} = 0.289$. A few disordered regions were not modelled owing to weak electron density, and the final refined model includes residues 343–571, 581–614, 688–810, 817–837, 875–949, 953–1,024 and 1,029–1,059. Residues for which side-chain density was poorly defined were modelled as alanines. The majority (95.9%) of the residues lie in the most favoured region in a Ramachandran plot, with the remaining 4.1% in the additionally allowed regions. Eight Ca^{2+} ions were modelled in the Ca^{2+} bowls from the two gating rings in the crystal. Data collection and structure refinement statistics are shown in Supplementary Table 1. All structural illustrations were prepared with PYMOL (http://www.pymol.org). The multiple-chain morphing script for CNS[37] was used to generate coordinates for the movies in Supplementary Information.

Electrophysiology. Channel expression in *Xenopus laevis* oocytes and electrophysiological recordings in inside-out patches were carried out as previously described[18]. Wild-type and deletion zebrafish BK constructs were cloned into the pGEM-HE vector[38]. Plasmids were linearized with PciI and capped cRNA was produced by *in vitro* transcription from the T7 promoter using the AmpliCap-Max T7 kit (Epicentre Biotechnologies). *Xenopus laevis* oocytes were injected with ~10 ng of cRNA and recordings were made 48–60 h later. For recording, inside-out patches were excised from freshly devitellinized oocytes using fire-polished glass pipettes with a typical resistance of 1–2 MΩ. In Ca^{2+}-free conditions, the bath solution contained 140 mM K-$MeSO_3$, 20 mM HEPES, 2 mM KCl, 5 mM EGTA (final pH, 7.0). For recordings in 10 μM free Ca^{2+}, the same solution was supplemented with 4.83 mM $CaCl_2$, adjusting the pH to 7.0 (we assumed a K_d of 376 nM for the EGTA–Ca^{2+} complex at pH 7.0). The pipette solution was the same as the Ca^{2+}-free bath solution, but supplemented with 2 mM $MgCl_2$. During recording, the membrane voltage was held at −60 mV and currents were elicited by pulsing the voltage from −120 to +140 mV in 10-mV increments (length of pulse, 20 ms). For plots of conductance versus voltage, measurements were taken from tail currents at –60 mV, 700 μs after the voltage step.

31. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
32. Strong, M. *et al.* Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103,** 8060–8065 (2006).
33. Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallogr. D* **59,** 1131–1137 (2003).
34. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
35. Brunger, A. T. Version 1.2 of the crystallography and NMR system. *Nature Protocols* **2,** 2728–2733 (2007).
36. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
37. Echols, N., Milburn, D. & Gerstein, M. MolMovDB: analysis and visualization of conformational change and structural flexibility. *Nucleic Acids Res.* **31,** 478–482 (2003).
38. Liman, E. R., Tytgat, J. & Hess, P. Subunit stoichiometry of a mammalian K^+ channel determined by construction of multimeric cDNAs. *Neuron* **9,** 861–871 (1992).

doi:10.1038/nature10663

Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically

Seigo Shima[1,2], Martin Krueger[3], Tobias Weinert[4], Ulrike Demmer[4], Jörg Kahnt[1], Rudolf K. Thauer[1] & Ulrich Ermler[4]

The anaerobic oxidation of methane (AOM) with sulphate, an area currently generating great interest in microbiology, is accomplished by consortia of methanotrophic archaea (ANME) and sulphate-reducing bacteria[1,2]. The enzyme activating methane in methanotrophic archaea has tentatively been identified as a homologue of methyl-coenzyme M reductase (MCR) that catalyses the methane-forming step in methanogenic archaea[3,4]. Here we report an X-ray structure of the 280 kDa heterohexameric ANME-1 MCR complex. It was crystallized uniquely from a protein ensemble purified from consortia of microorganisms collected with a submersible from a Black Sea mat catalysing AOM with sulphate[4]. Crystals grown from the heterogeneous sample diffract to 2.1 Å resolution and consist of a single ANME-1 MCR population, demonstrating the strong selective power of crystallization. The structure revealed ANME-1 MCR in complex with coenzyme M and coenzyme B, indicating the same substrates for MCR from methanotrophic and methanogenic archaea. Differences between the highly similar structures of ANME-1 MCR and methanogenic MCR include a F_{430} modification, a cysteine-rich patch and an altered post-translational amino acid modification pattern, which may tune the enzymes for their functions in different biological contexts.

Large amounts (up to 300 Tg per year) of the potent greenhouse gas, methane, are oxidized to CO_2 (reaction 1) in marine sediments by consortia of methanotrophic archaea (ANME-1, ANME-2 or ANME-3)[5–7] and sulphate-reducing bacteria, thus preventing methane release into the atmosphere[1].

$$CH_4 + SO_4^{2-} \rightarrow HCO_3^- + HS^- + H_2O \quad \Delta G^{\circ\prime} = -17\,\mathrm{kJ\,mol^{-1}} \qquad (1)$$

Indirect evidence suggests that AOM with sulphate functions at least in part as the reverse of methanogenesis from CO_2 (refs 2, 8). Accordingly, the enzyme activating methane should be a methyl-coenzyme M reductase (MCR)[9], which contains the Ni porphinoid F_{430} as prosthetic group and reversibly catalyses the conversion of methane (CH_4) and the heterodisulphide of coenzyme M (HS-CoM, 2-mercaptoethane sulphonate) and coenzyme B (HS-CoB, 7-mercaptoheptanoylthreonine phosphate) to methyl-coenzyme M (CH_3-S-CoM) and HS-CoB (reaction 2)[3].

$$CH_4 + \text{CoM-S-S-CoB} \rightleftharpoons CH_3\text{-S-CoM} + \text{HS-CoB} \quad \Delta G^{o} = +30 \pm 10\,\mathrm{kJ\,mol^{-1}} \qquad (2)$$

The involvement of the reverse MCR reaction in AOM with sulphate is based on the following findings: (1) AOM with sulphate is inhibited by bromoethane sulphonate, a known inhibitor of MCR[10]; (2) the methanotrophic archaea that catalyse AOM contain genes homologous to those of methanogenic MCR (abbreviated as ANME MCR) with a sequence identity of around 50%[2,5,6] (for sequences see Supplementary Fig. 1); (3) the methanotrophic archaea contain high concentrations of ANME MCR as revealed by direct purification studies from the consortia[4], immunological localization analysis[11] and identification of F_{430} (refs 4, 12); (4) MCR from methanogens catalyses the oxidation of methane to CH_3-S-CoM at specific rates sufficient to account for those observed *in vivo* for AOM[3].

Despite these well-founded arguments for MCR as the methane-oxidizing enzyme, its catalytic activity and its coenzyme specificity are not measurable as the purified enzyme is isolated in an inactive form. All attempts to reactivate the enzyme, as measured by CH_3-S-CoM reduction to methane with HS-CoB, have failed so far. Coenzyme specificity of ANME-1 MCR might therefore differ from the methanogenic MCR, which is specific for HS-CoM and HS-CoB, and indeed, recent data suggest that methanethiol rather than CH_3-S-CoM might serve as an intermediate in AOM with sulphate[13]. Moreover, ANME-1 MCR shows significant differences compared to methanogenic MCR regarding the prosthetic group F_{430} (ref. 12), the accumulation of cysteines near F_{430} (ref. 12) and post-translational modifications[14].

A structural characterization of ANME-1 MCR by X-ray analysis would provide important contributions to the molecular understanding of AOM; however, its determination has to involve novel methods because methanotrophic archaea cannot be grown in pure cultures and a recombinant production of MCR has not yet been possible because of its unusual F_{430} cofactor and post-translational modifications[14]. Methanotrophic MCR is only available in larger amounts from microbial mats harvested from the Black Sea (see Methods). These mats consist of microbial communities with many closely related ANME-1 organisms (pink mats) and ANME-2 organisms (black mats) and thus of heterogeneous ANME-1 and ANME-2 MCR populations as revealed by 16S RNA analyses[15] and *mcrA* gene sequence analyses, respectively. Fortunately, the major quantity of the ANME-1 and ANME-2 MCR populations in the preparations was separable into two peaks via anion-exchange chromatography. The first peak to elute is composed of at least six different ANME-1 MCR populations (two of them correlate well with the *mcrBGA* gene sequences D1JBK2-4 and Q6MZD1-3) as identified by SDS–polyacrylamide gel electrophoresis, amino-terminal Edman sequence analyses and mass-spectrometric analyses of tryptic digests of the α-subunit (Supplementary Figs 1, 2 and Supplementary Table 1). ANME-1 MCR only contains methylthio-F_{430} as prosthetic group, in contrast to ANME-2 MCR eluted as second peak that only hosts F_{430} (ref. 4). According to the metagenome sequence, four of the six ANME-1 MCR sequences are identical with respect to those tryptic peptides that invariantly contain a modified glycine (thioglycine) and a modified cysteine (S-methyl cysteine) in methanogenic MCR. Their mass-spectrometric sequence analysis revealed the presence of thioglycine but not of S-methyl cysteine in about 30% of the ANME-1 MCR ensemble, and of S-methyl cysteine but not of thioglycine in about 70% (Supplementary Fig. 3). Despite the highly heterogeneous sample solution crystals grew and the resulting crystal structure (Fig. 1a) only revealed ANME-1 MCR D1JBK2-4 with thioglycine substantiated by a mass-spectrometric analysis of the dissolved crystal used for structure determination (Supplementary Fig. 4).

For initial model building we first used the metagenome *mcr* gene sequence ANME-1 Q6MZD1-3 (ref. 4; Supplementary Fig. 1) because the ANME-1 MCR D1JBK2-4 sequences were not available at that

[1]Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch-Strasse 10, D-35043 Marburg, Germany. [2]PRESTO, Japan Science and Technology Agency (JST), Honcho, Kawaguchi, Saitama 332-0012 Japan. [3]Federal Institute for Geosciences and Resources, Stilleweg 2, 30655 Hannover, Germany. [4]Max Planck Institute for Biophysics, Max-von-Laue-Strasse 3, D-60438 Frankfurt, Germany.

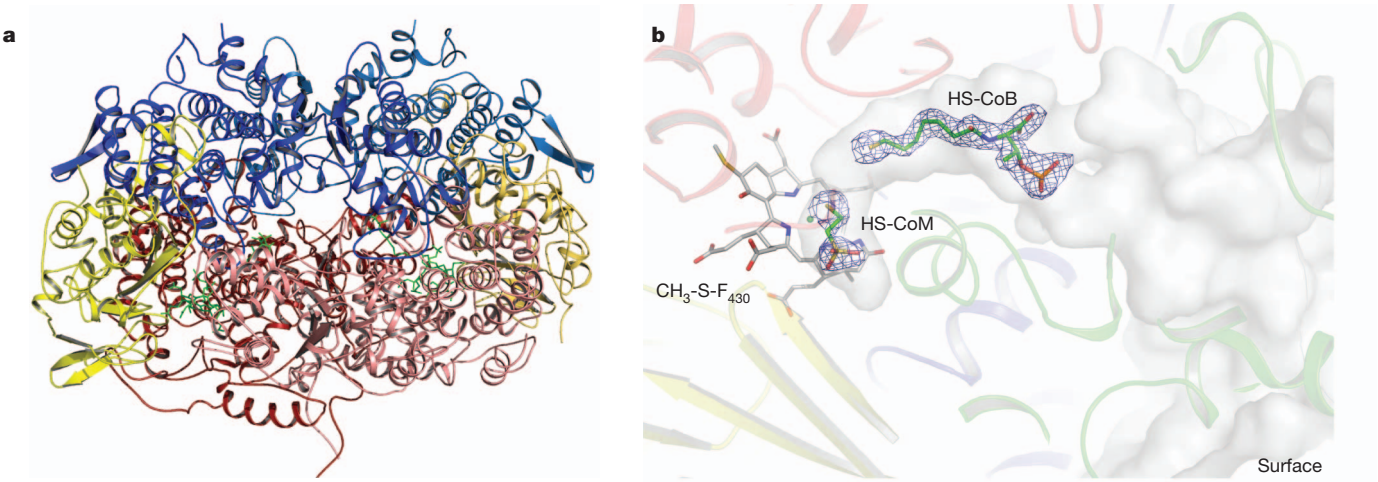


Figure 1 | Global architecture and active site of ANME-1 MCR. **a**, The structure of ANME-1 MCR represents a 280 kDa heterohexameric $(\alpha\beta\gamma)_2$ protein complex (α in red, β in blue, γ in yellow). The structure is virtually identical to that of methanogenic MCR as indicated by an r.m.s.d. value of about 1 Å. Structurally known methanogenic MCRs (*M. marburgensis*, *Methanocaldococcus jannaschii* and *Methanosarcina barkeri*) have a root mean squared deviation of about 0.8 Å (ref. 29). **b**, ANME-1 MCR harbours two coenzymes and the prosthetic group F_{430}; each is located at the end of a 50-Å funnel-shaped channel. HS-CoB binds into the narrowest segment of the channel and HS-CoM between HS-CoB and F_{430} (ref. 16). The $2F_{obs} - F_{calc}$ electron density (in blue) was drawn at a contour level of 1σ.

time[5]. Strong and ambiguous $2F_{obs} - F_{calc}$ and $F_{obs} - F_{calc}$ electron density was observed at side-chain positions which varied between the six MCR sequences and was therefore initially interpreted as a superposition of different ANME-1 MCR populations in the crystal lattice. However, after applying the ANME-1 MCR sequence D1JBK2-4 (Supplementary Fig. 1) the electron density for all side chains could be explained except for Ala 105 and Ala 400. (If not explicitly mentioned the residues are from the α-subunit.) Due to the presence of three αβγ-subunits in the asymmetric unit (allowing three independent analyses after independent refinement), the high quality of the electron density at 2.1-Å resolution and the high number of residues (1,288), the assignment of the sequence was very reliable. A homogeneous MCR population in the crystal is further corroborated by the single well-defined conformation found along the whole polypeptide backbone. The subsequent refinement on the basis of the ANME-1 D1JBK sequence converged to final R/R_{free} factors of 16.4/21.0% at 2.1-Å resolution (Supplementary Table 2). This result once again demonstrates that crystallization is a powerful selection process which can distinguish between nearly identical protein populations. Crystals of high quality composed of several protein populations grow very rarely as already minor sequence differences create slightly different protein surface profiles that prevent the formation of an exactly periodic crystal lattice. If well-diffracting, heterogeneous protein crystals could be obtained, the information from the multiple sequences would be difficult to exploit as only an average electron density is experimentally accessible (no meta-crystallography).

ANME-1 MCR possesses the same overall structure as the methanogenic MCR (Fig. 1a) and contains HS-CoM and HS-CoB inside the active site channel. Their occupancy is approximately 60% and 100%, respectively (Fig. 1b). Their positions and conformations are virtually identical to those found in the methanogenic $MCR_{ox1\text{-}silent}$ structure[16]. HS-CoM is covalently bound to F_{430} with a nickel-sulphur distance of 2.5 Å. The distances between the HS-CoB and HS-CoM sulphurs are 6.2 Å and between the HS-CoB sulphur and the F_{430} nickel 8.6 Å (Fig. 1b). The conformations of the contacting residues and the type of interacting atoms are strictly conserved in structurally characterized methanogenic MCR family members indicating an identical channel profile. These findings basically exclude that ANME-1 MCR uses different substrates than methanogenic MCR. The almost identical active site structures of MCR reflect strong structural restraints imposed by the intricate chemical reaction (2), which is not conclusively elucidated yet due to its exceptional Ni and S chemistry[17,18].

The methanogenic $MCR_{ox1\text{-}silent}$ (Ni^{2+}) form in the crystal is generated via spontaneous reduction from the MCR_{ox1} (Ni^{3+}) form in which the enzyme is present after enrichment. This oxidized form is induced in methanogens growing on H_2 and CO_2 when the concentration of H_2 in the culture is set so low that formation of methane from H_2 and CO_2 is no longer exergonic[8]. Such conditions prevail in microbial sediments catalysing AOM with sulphate and, indeed, MCR_{ox1} (Ni^{3+}) electron paramagnetic resonance (EPR) signals were detected in the Black Sea mat extracts (Supplementary Fig. 5). Consequently, the $MCR_{ox1\text{-}silent}$ form found in the X-ray structure is consistent with ANME-1 MCR being involved in methane oxidation rather than in methane formation in the mat system.

Structural comparison between ANME-1 MCR and other MCRs confirmed that ANME-1 MCR contains a F_{430} variant[4,12] in which a methylthio group is covalently bound to the $C17^2$ atom of the six-membered carbocyclic ring. The profile of the extra electron density accurately fits to the methylthio substituent and confirms the 17^2S isomeric configuration determined by NMR spectroscopy (Fig. 2a). The 17^2S methylthio group of the modified F_{430} in ANME-1 MCR would collide with the bulky side chain of the 2-C methylated glutamine in methanogenic MCR and this residue is therefore replaced by a small valine in all known ANME-1 MCR sequences (Val 419 in ANME-1 MCR D1JBK4, see Fig. 2a and Supplementary Fig. 1). According to metagenome sequences, ANME-2/-3 MCRs possess a glutamine at this position. Therefore F_{430} is unlikely to be modified and this is consistent with experimental observations for methanogenic MCR and ANME-2 MCR[4].

ANME-1 MCR contains, in contrast to the other MCRs, five pronounced cysteines between F_{430} and the protein surface (Cys 422, 423, 434, 361 and 365) (Fig. 2b). Cys 422 and Cys 423 of helix 418–438 are 5–6 Å away from Val 419. Cys 434 forms a disulphide bond to Cys 365 which protrudes from helix 356–371. Cys 361 is exactly positioned on the connecting line between the two cysteine pairs (Fig. 2b). This striking cysteine-rich patch proposed as a redox-relay system[12] might be used for the reduction of ANME-1 MCR from the inactive Ni^{2+} to the active Ni^{1+} oxidation state[19].

In all methanogenic MCRs the peptide oxygen of Gly 464 (ANME-1 MCR sequence) is exchanged by sulphur. The thioglycine is also found in the ANME-1 MCR D1JBK2-4 crystal structure (Fig. 2c) but not in 70% of the ANME-1 MCRs in the sample used for crystallization. Its absence has been observed earlier and was interpreted as a spontaneous loss of the sulphur in the thioxo peptide bond by hydrolysis[14].

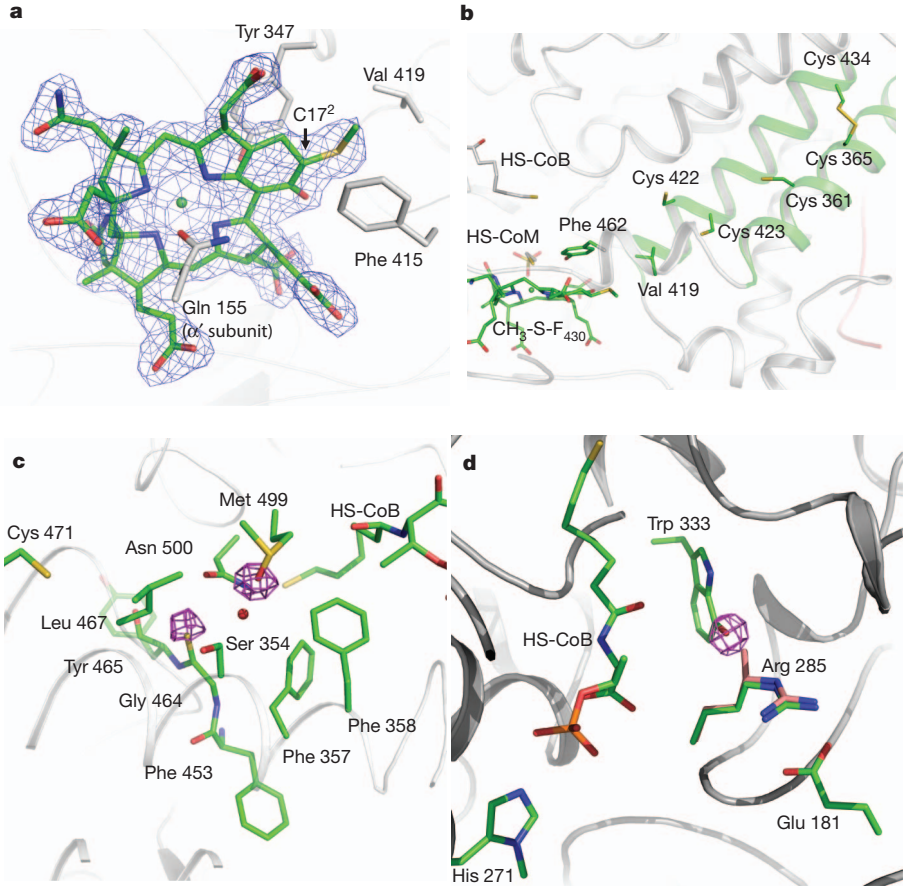


Figure 2 | Specific differences between ANME-1 and methanogenic MCRs. **a**, Methylthio-F_{430}. The Ni-porphinoid F_{430} tetrapyrrole macrocycle is more reduced and extended by a six-membered carbocyclic and a lactam ring compared to other tetrapyrrole containing prosthetic groups. The thiomethyl group of F_{430} is bound to the $C17^2$ atom of the carbocyclic ring in an S configuration consistent with NMR spectroscopic data[12]. The $2F_{obs} - F_{calc}$ electron density (in blue) is drawn at a contour level of 1.5σ. Except for the methylthio group, the binding site and the conformation of F_{430} is strictly conserved within the MCR family. **b**, The cysteine-rich patch. In contrast to the methanogenic and ANME-2/3 MCRs, helices 418–438 and 356–371 of ANME-1 MCR contain five cysteines that are distributed between the protein surface and F_{430}. Cys 434 forms a disulphide bond to Cys 365 that is 15 Å away from the Cys 422-Cys 423 pair. The distance between the latter thiol sulphurs is 4.9 Å, both being approximately 10 Å apart from F_{430}. **c**, Thio-Gly 464 and its microenvironment. The strictly conserved thio-Gly 464 is surrounded by Met 499 that seems to be post-translationally modified in ANME-1 MCR. The $F_{obs} - F_{calc}$ electron density (in magenta) contoured at 5σ is based on a refinement that omits the post-translational modifications. Cys 471 is S-methylated in methanogenic MCR but not in ANME-1 MCR D1JBK2-4. These changes induce a rearrangement of the Leu 467 side chain which is in contact to the sulphur of thioglycine. A methyl group modelled to the Cys 471 sulphur would slightly interfere with Leu 467 in the conformation of ANME-1 MCR D1JBK2-4. **d**, Post-translationally modified Trp 333. Arg 285 is C-methylated in methanogenic MCR (carbons highlighted in salmon) but not in ANME-1 MCR. Perhaps for compensation the adjacent Trp 333 is modified in ANME-1 MCR. We modelled a hydroxyl group into the extra electron density (magenta, contour level 5σ) as C_7-OH-tryptophan was recently found in a methylamine dehydrogenase[30]. Adjacent to Arg 285 is N-methyl His 271 that is conserved in methanogenic and ANME-1 MCR as identified by mass-spectrometric[14] and structural data.

Thio-Gly 464 might be crucial for the enzymatic mechanism by formation of a thioketyl anion radical[14] and/or by cis/trans isomerization of the thioxo bond[20]. The microenvironment of thio-Gly 464 slightly differs between ANME-1 and methanogenic MCR mainly because of the presence of a serine instead of a glycine at residue 354 of the β-subunit (ANME-1 MCR sequence), the absent post-translational S-methyl group at Cys 471 and a new post-translationally modified Met 499 in ANME-1 MCR D1JBK2-4. The sulphoxide form of Met 499, detected by the structural and mass-spectrometric data, has to be interpreted with care because methionine can be spontaneously oxidized to sulphoxide during the aerobic analysis procedure. The additional oxygen of Met 499 points towards a solvent molecule that interacts with the thio-Gly 464 sulphur (Fig. 2c).

Arg 285 is invariantly C-methylated in methanogenic MCRs but not in ANME-1 MCR D1JBK2-4, as indicated by the presented structural data and by mass-spectrometric analysis of the respective tryptic peptide. Unexpectedly, the lacking methyl group in ANME-1 MCR seems to be compensated by a modification at the adjacent Trp 333 side chain. The extra group covalently bound to C_7 of the indole ring points exactly to the position of the methyl group of the equivalent C-methyl-arginine in methanogenic MCR (Fig. 2d). It might be speculated that the absent C-methyl groups in ANME-1 MCR compared to 2-methyl-Gln 400 and 5-methyl-Arg 271 (*Methanothermobacter marburgensis* sequence) in methanogenic MCR are due to the lack of a C-methylation machinery[21].

ANME archaea are phylogenetically related to methanogenic archaea but ANME-1 less than ANME-2/-3 (refs 5, 7, 15) and this is reflected in the structural features of their MCR. Because ANME-2/-3 organisms also catalyse the oxidation of methane, the structural differences in MCR do not primarily affect the basic enzymatic process, but rather express a divergent evolutionary process and/or the fine-tuning of the enzyme for distinct functional requirements. The found compensatory character of some of the some post-translational modifications might be the result of a complex phylogenetic development. The location of the ANME-1 and ANME-2 organisms in the core and exterior part of the Black Sea mats, respectively[22], might be an indicator

for different living conditions (such as different methane concentrations and redox potentials) that require an adaptation on the structural level of MCR.

METHODS SUMMARY

Black Sea microbial mat material was collected during a RV *Poseidon* cruise POS 317-2 with submersible *JAGO* in August 2004 (Project BEBOP, University of Hamburg) at 44° 46′ N, 31° 60′ E (ref. 23) and maintained active in artificial sea water medium (28 mM SO_4^{2-}) under CH_4 and CO_2 (5:1) at 8 °C. The exterior part of the mat of the chimney-like structures is black and harbours mainly consortia of the ANME-2 with sulphate-reducing bacteria. Its interior part is pink and dominated by ANME-1 cells[22]. Thirty grams (wet mass) of the pink interior part of the mats were suspended in 60 ml of 50 mM MOPS/KOH pH 7.0, decanted from settled rock particles, ultrasonicated and purified as described[4]. ANME-1 MCR (4 mg from 10 ml cell extract) was crystallized by the sitting drop method using 25% PEG 400, 0.1 M Tris pH 8.5 and 0.2 M Li_2SO_4 as reservoir solution. After about three years a few crystals with dimensions of 50 × 100 × 200 μm were discovered in two drops. Data of the ANME-1 MCR crystals were collected at the Swiss-Light-Source PXII beamline and processed with XDS[24]. The statistics are listed in Supplementary Table 2. The structure was solved by the molecular replacement method using the structure of MCR from *M. marburgensis* (PDB code 1HBN) as search model[25]. The side chains were manually exchanged with COOT[26] by using the metagenome sequence D1JBK2-4 (Supplementary Fig. 1). Refinement was initially carried out using REFMAC5[27] and then continued with PHENIX[28]. All side chain densities matched the D1JBK2-4 sequence except Ala 105 and Ala 400 where longer side chains are required. All other density ambiguities could be explained by the normal conformational flexibilities of a certain number of side chains. The post-translational modification profile seen in the X-ray structure could in part be validated by mass-spectrometric analysis of tryptic peptides from the ANME-1 MCR sample used for crystallization[14]. Figures 1 and 2 were generated with PYMOL (Schrodinger).

Received 14 April; accepted 19 October 2011.
Published online 27 November 2011.

1. Reeburgh, W. S. Oceanic methane biogeochemistry. *Chem. Rev.* **107,** 486–513 (2007).
2. Hallam, S. J. *et al.* Reverse methanogenesis: testing the hypothesis with environmental genomics. *Science* **305,** 1457–1462 (2004).
3. Scheller, S., Goenrich, M., Boecher, R., Thauer, R. K. & Jaun, B. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465,** 606–608 (2010).
4. Krüger, M. *et al.* A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* **426,** 878–881 (2003).
5. Meyerdierks, A. *et al.* Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. *Environ. Microbiol.* **12,** 422–439 (2010).
6. Knittel, K. & Boetius, A. Anaerobic oxidation of methane: progress with an unknown process. *Annu. Rev. Microbiol.* **63,** 311–334 (2009).
7. Losekann, T. *et al.* Diversity and abundance of aerobic and anaerobic methane oxidizers at the Haakon Mosby Mud Volcano, Barents Sea. *Appl. Environ. Microbiol.* **73,** 3348–3362 (2007).
8. Thauer, R. K. Anaerobic oxidation of methane with sulfate: on the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO_2. *Curr. Opin. Microbiol.* **14,** 292–299 (2011).
9. Thauer, R. K. & Shima, S. Methane as fuel for anaerobic microorganisms. *Ann. NY Acad. Sci.* **1125,** 158–170 (2008).
10. Nauhaus, K., Treude, T., Boetius, A. & Krüger, M. Environmental regulation of the anaerobic oxidation of methane: a comparison of ANME-I and ANME-II communities. *Environ. Microbiol.* **7,** 98–106 (2005).
11. Heller, C., Hoppert, M. & Reitner, J. Immunological localization of coenzyme M reductase in anaerobic methane-oxidizing archaea of ANME 1 and ANME 2 type. *Geomicrobiol. J.* **25,** 149–156 (2008).
12. Mayr, S. *et al.* Structure of an F430 variant from archaea associated with anaerobic oxidation of methane. *J. Am. Chem. Soc.* **130,** 10758–10767 (2008).
13. Moran, J. J. *et al.* Methyl sulfides as intermediates in the anaerobic oxidation of methane. *Environ. Microbiol.* **10,** 162–173 (2008).
14. Kahnt, J. *et al.* Post-translational modifications in the active site region of methyl-coenzyme M reductase from methanogenic and methanotrophic archaea. *FEBS J.* **274,** 4913–4921 (2007).
15. Knittel, K., Losekann, T., Boetius, A., Kort, R. & Amann, R. Diversity and distribution of methanotrophic archaea at cold seeps. *Appl. Environ. Microbiol.* **71,** 467–479 (2005).
16. Grabarse, W. *et al.* On the mechanism of biological methane formation: structural evidence for conformational changes in methyl-coenzyme M reductase upon substrate binding. *J. Mol. Biol.* **309,** 315–330 (2001).
17. Scheller, S., Goenrich, M., Mayr, S., Thauer, R. K. & Jaun, B. Intermediates in the catalytic cycle of methyl coenzyme M reductase: isotope exchange is consistent with formation of a σ-alkane–nickel complex. *Angew. Chem. Int. Edn Engl.* **49,** 8112–8115 (2010).
18. Dey, M., Li, X., Kunz, R. C. & Ragsdale, S. W. Detection of organometallic and radical intermediates in the catalytic mechanism of methyl-coenzyme M reductase using the natural substrate methyl-coenzyme M and a coenzyme B substrate analogue. *Biochemistry* **49,** 10902–10911 (2010).
19. Thauer, R. K. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. *Microbiology* **144,** 2377–2406 (1998).
20. Satzger, H. *et al.* Photoswitchable elements within a peptide backbone-ultrafast spectroscopy of thioxylated amides. *J. Phys. Chem. B* **109,** 4770–4775 (2005).
21. Grove, T. L. *et al.* A radically different mechanism for S-adenosylmethionine-dependent methyltransferases. *Science* **332,** 604–607 (2011).
22. Krüger, M. *et al.* A novel, multi-layered methanotrophic microbial mat system growing on the sediment of the Black Sea. *Environ. Microbiol.* **10,** 1934–1947 (2008).
23. Michaelis, W. *et al.* Microbial reefs in the Black Sea fueled by anaerobic oxidation of methane. *Science* **297,** 1013–1015 (2002).
24. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26,** 795–800 (1993).
25. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Cryst.* **40,** 658–674 (2007).
26. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
27. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53,** 240–255 (1997).
28. Afonine, P. V. *et al. phenix.model_vs_data*: a high-level tool for the calculation of crystallographic model and data statistics. *J. Appl. Cryst.* **43,** 669–676 (2010).
29. Grabarse, W., Mahlert, F., Shima, S., Thauer, R. K. & Ermler, U. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation. *J. Mol. Biol.* **303,** 329–344 (2000).
30. Jensen, L. M., Sanishvili, R., Davidson, V. L. & Wilmot, C. M. In crystallo posttranslational modification within a MauG/pre-methylamine dehydrogenase complex. *Science* **327,** 1392–1394 (2010).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the Max Planck Society, the Fonds der Chemischen Industrie, the Deutsche Forschungsgemeinschaft (SPP 1319, ER 222/5-1, KR 3311/6-2) and the projects MUMM, and BEBOP (Biogeochemistry and Microbiology of Bioherms Prospering in the Black Sea funded by the University of Hamburg). S.S. was financed by the PRESTO program, Japan Science and Technology Agency (JST). We are indebted to W. Michaelis and R. Seifert for access to samples. We thank the crews of RV *Poseidon* with *JAGO* for help during field work, M. Brefort and L. Känel for the electron spin resonance measurement and Edman sequencing, respectively. We are grateful to H. Michel for continuous support, the staff of the PXII beamline at the Swiss-Light-Source, Villigen for assistance during data collection, and K. Davies for critical reading of the manuscript.

Author Contributions S.S., R.K.T. and U.E. designed the experiments, interpreted the data and wrote the paper. The mat sample from the Black Sea was collected and maintained by M.K. The enzyme was purified and characterized by M.K. and S.S. Crystallization was performed by S.S. and U.D.; U.E. collected the X-ray data. T.W. refined the structure and found that the primary structure conformed with that of the D1JBK2-4 sequences in the metagenome. J.K. performed the mass-spectrometric analyses.

Author Information Atomic coordinates and structure factors of ANME-1 MCR have been deposited in the Protein Data Bank with the code 3SQG. 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 this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to S.S. (shima@mpi-marburg.mpg.de) or U.E. (ulrich.ermler@biophys.mpg.de).