nature

www.nature.com/nature

Vol 456 | Issue no. 7219 | 13 November 2008

Danger and opportunity

The response to the financial crisis needs to go beyond the immediate pressures. Policy-makers must seize this moment to solidify the science and innovation required for sustained economic growth.

"While innovation is commonly associated with growth, it is now more correctly pinned to survival." That was one conclusion from a meeting convened in Dubai last weekend by the World Economic Forum, a body best known for its annual summit in Davos, Switzerland. It is correct: with an economic crisis of unknown proportions looming, more emphasis on science and innovation — not less — will be crucial to achieving a sustained recovery.

Over the past few decades the world has built up a formidable science and innovation complex. Among countries in the Organisation for Economic Co-operation and Development, funding for research and development (R&D) alone almost doubled from US$468 billion in 1996 to $818 billion in 2006 — although growth has been slowing since 2001, particularly in the United States. Globalization has given rise to major new players in research, such as China and India, as well as to transnational webs of R&D centres and company subsidiaries. And, perhaps most important, many countries have developed increasingly sophisticated innovation policies. These include support for education at every level; for robust capital markets; for reliable energy, transportation and communications infrastructures; and for all the other tightly interlinked factors that contribute to innovation, and which in turn affect what happens in the economy.

The present financial crisis is placing multiple strains on this web, as a Special Report on page 155 makes clear. National funding for research is under pressure as governments face falling tax revenues and the need for massive financial-stimulus measures. Universities' funding and endowments are shrinking. The crisis in confidence among banks could spread, creating a risk-averse investment culture that would starve high-tech start-up firms and other businesses of capital.

Taken together, such threats could prove a setback to the progress made in building global innovation systems. Policy-makers should instead take every step possible to reinforce those systems.

In coming months they will have to balance many legitimate demands on the public purse, including the need to relieve their citizen's immediate economic pain, and resuscitate economic demand with tax cuts and public spending. Such needs will become acute if the global recession is deep and prolonged, especially in countries with already weak economies, and will have to take priority.

Future investments

Nonetheless, policy-makers also have a responsibility to take the long view. High-tech companies have long known that protecting R&D is key to coming out of a recession strong; they tend to cut everything else first. It's encouraging that many governments — large and small — are now taking a similar view: education and research are not merely costs to be cut, but investments in long-term economic recovery. Whether governments live up to this rhetoric is another matter, but they would be irresponsible to undertake large stimulus efforts without also providing for long-term growth to reduce the resulting deficits.

Also encouraging is the emerging appreciation that short- and long-term measures to address the economic crisis need not be mutually exclusive. Often they can and should be considered simultaneously. After trillion-dollar rescues of the banking system, for example, many countries are indeed contemplating massive economic stimulus packages. But instead of crafting conventional, spending-oriented plans focused solely on tax cuts for individuals and spending increases, policy-makers should look for ways to give the economy a quick shot in the arm while also boosting investment that spurs productivity, growth and innovation.

"Education and research are not merely costs to be cut, but investments in long-term economic recovery."

China seems to be thinking in those terms. Last weekend, Beijing announced a 4-trillion-yuan (US$586-billion) stimulus package, at least some of which — details remain vague — will go on research tax credits and investments in infrastructure. The United States, which is debating a stimulus package of similar magnitude, may do likewise: Barack Obama, the US president-elect, campaigned on the need for major investments in clean energy, education, health and infrastructure, most of which will have a strong research component.

Innovative thinking

Good specific ideas are emerging. On 29 October, for example, the Information Technology & Innovation Foundation, a think tank based in Washington DC, outlined eight possibilities in the IT realm. Among them were increased grants for universities to buy research equipment in 2009, a tax credit for investments in energy-efficient equipment in 2009 and billions of dollars to buy computers and broadband for low-income families with children at home. In the area of clean energy, meanwhile, others have suggested the use of stimulus money to spur demand for new technologies, much as consumer subsidies to use solar energy in Germany have helped drive the growth of solar start-ups — not only in that country, but also in the United States and China.

Indeed, the crisis is providing multiple opportunities for governments to take decisive action on innovation, and policy-makers should champion those possibilities as they hammer out stimulus packages. Given the deeply intertwined nature of the modern global innovation system, they should act collaboratively — as they have already begun to do in addressing the problems of the global financial system. International cooperation at this level has been all too rare in recent years. But there is every reason to hope that the experience provided by this crisis will be a model for future efforts to combat challenges such as climate change and water shortage by a world aptly described by Obama in his victory speech as "connected by our own science and imagination". ■

Situations vacant

Italy's universities should be free to hire who they want — and should be accountable for the result.

It took violent street demonstrations to force the Italian government to backtrack on its proposal to enact — hot on the heels of a hefty budget cut — a major reform of the nation's universities through decree. Last week, education and research minister Mariastella Gelmini agreed instead to put her planned reform through normal legislative procedures, which, unlike a decree, will involve parliamentary debates and, hopefully, consultation with the universities.

But on 6 November, Gelmini rushed through part of the reform in a decree anyway. With a round of *concorsi* — the national competitions to select academic staff — due to start within days, Gelmini introduced a relatively minor change in the procedures of the committees that select the staff. The not-so-minor result is that those *concorsi*, for 1,800 professorships, will be delayed by at least three months. And if the change is challenged in court, as it may well be, the delay could stretch beyond a year — at a time when Italian universities have already been unable to recruit new professors for more than four years.

The cumbersome *concorso* system does not need such tinkering, it needs to be abandoned. Imagine if the Massachusetts Institute of Technology (MIT) in Cambridge had to tell Washington whenever it had an vacancy; then wait for the administration to collate enough empty posts nationwide to warrant opening a competition; and then allow academics from all universities to elect a national, discipline-related committee to choose the candidate — a committee on which only one MIT representative could sit.

Such centralized recruitment has been a feature of Italian universities for the best part of a century. It wasn't until the 1990s that universities gained sufficient control of their budgets to decide how many professors they wanted to recruit, even if they couldn't choose the successful candidates. The government refuses to grant them that last authority partly because politicians fear that, left to themselves, some universities would appoint professors on the basis of their local political and personal connections rather than their scientific merit. There is good reason for this worry: it happens even within the *concorso* system.

Nonetheless, Italy's universities should be allowed to recruit whomever and however they want — with the all-important proviso that they also be evaluated on their academic performance. If the best-performing universities received more state support, and the underperformers received less, the incentive to play politics when hiring would plummet.

Italy's previous, centre-left government paved the way for such a system just before it fell in April, when it passed a law to create ANVUR, an agency to evaluate university and academic performance. Gelmini simply needs to complete the establishment of ANVUR, get it working and put an end to the *concorso* system. Her predecessor had recognized that it would take a while to establish the new institution and so had set in motion the current, now frozen, round of *concorsi* to allow university life to go on. Gelmini was wrong to interfere with it.

A level of reform is clearly needed. Because Italian universities don't have to take responsibility for any recruitment decisions, some have become lax, bloated and lazy. But reforms need to be done with a strong, knowledgeable and clever hand — something that Gelmini has so far failed to provide. ■

Science by litigation

A company's lawsuit against researchers should not be allowed to intimidate others.

The curious case of Biopure versus Natanson pits a struggling biotechnology company against a biomedical researcher in a libel case (see *Nature News* doi:10.1038/news.2008.1219; 2008). Historically, such attacks on the scientific literature have been given short shrift by the courts. The case is dangerous nevertheless.

Biopure of Cambridge, Massachusetts, is attacking an analysis performed by Charles Natanson, a senior investigator at the US National Institutes of Health, and his colleagues, and published in *The Journal of the American Medical Association* (*JAMA* — C. Natanson *et al. J. Am. Med. Assoc.* **299,** 2304–2312; 2008). Natanson's article was a meta-analysis: a statistical lumping together of several small clinical trials. Such analyses can be complex and problematic. A bad meta-analysis has the power to needlessly frighten or groundlessly assure consumers, but a good meta-analysis can protect the public from a previously unrecognized — or unpublicized — safety risk. Regardless of quality, a well publicized meta-analysis can cost a firm millions in lost sales.

Biopure alleges that this analysis was not sufficiently sound to warrant one of the paper's conclusions: that a blood substitute produced by the company is not safe.

Natanson and his colleagues combined data for different blood substitutes that share a mechanism of action; Biopure says this means that the results could not be applied specifically to its product. The researchers also lacked access to critical data — a common lament of the meta-analyser, and an inevitable result of companies refusing to disclose the results of their clinical trials. *JAMA*'s editors — who have declined to comment — and their reviewers presumably felt that the evidence as described justified the conclusions.

There is a traditional forum, in science, to air such grievances: the journal itself. And indeed, after the article was published, *JAMA* published numerous critiques from readers and the authors' response. Biopure elected to put its questions about scientific quality into the hands of the law. It claims that it has suffered financial loss and that its corporate mission, financed to the tune of $600 million, has been put at risk.

The US legal system has historically treated scientific publications with respect under the US constitution's first amendment, which protects freedom of the press. That does not mean that scientists are exempt from laws on libel and slander simply because they are scientists. But those involved should be aware that this action potentially threatens public trust in the very industry of which Biopure is a part, which is founded on open scientific analysis and debate. Researchers (and their editors) must never forget what is at stake in a meta-analysis, but above all must not be intimidated by this action. ■

RESEARCH HIGHLIGHTS

Clones of the dead

Proc. Natl Acad. Sci. USA doi:10.1073/pnas.0806166105 (2008)

Mice stored at –20 °C for 16 years have been cloned, purportedly raising the possibility that extinct animals could be 'resurrected' from frozen tissue samples.

Researchers had previously cloned mice from frozen cells stored in chemicals that protect against the damaging effects of freezing. In the new study, Teruhiko Wakayama at RIKEN in Kobe, Japan, and his colleagues generated embryonic stem cells using nuclei harvested from mice that were frozen whole and without protective chemicals. They then transferred nuclei from the stem cells into unfertilized eggs that had had their nuclei removed. Surrogate mothers implanted with these eggs went on to produce several healthy offspring.

The finding paves the way for cloning extinct species such as the woolly mammoth, the authors claim.

NATL ACAD. SCI.

ASTRONOMY

Star birth

Astrophys. J. **687,** 1004–1018 (2008)

Star formation in galaxies is a tug-of-war between opposing forces. Clouds of gas collapse, feeding fledgling stars. These can, in turn, prevent new stars from forming by blowing away the nurturing gas with their stellar winds. Star clusters are thus expected to lie at the centre of gas voids.

Yet in a nearby dwarf galaxy called IC 2574 there is little correlation between the locations of stars and observed gas voids, according to observations made with the Large Binocular Telescope on Mount Graham, Arizona, by Anna Pasquali of the Max Planck Institute for Astronomy in Heidelberg, Germany, and her colleagues. The energy balances between the expanding gas voids are consistent with recent bouts of star formation, suggesting that this tug-of-war is not well understood.

METEOROLOGY

Prediction with plants

Water Resour. Res. doi:10.1029/2007WR006514 (2008)

Monsoons are driven in large part by contrasts between land and sea temperatures, which are key to their prediction. However, other factors are also at play, such as soil moisture; higher moisture makes the transfer of heat from the land to the air through evaporation easier.

Eungul Lee, at the University of Wisconsin, Madison, and his colleagues show through analysis of satellite and other records that including this effect can improve monsoon forecasts. Taking into account the vegetation growth — which has a strong influence on soil moisture — in the months preceding the northern and southern East Asian summer monsoons allowed them to improve the reliability of forecasts by a factor of two for the northern and three for the southern monsoon.

ANIMAL BEHAVIOUR

Got the scent

Naturwissenschaften doi:10.1007/s00114-008-0465-x (2008)

Honeybees that find nectar tell the rest of the hive about it by dancing. But bumblebees instruct nestmates through smell, find Mathieu Molet and his colleagues at Queen Mary, University of London.

They exposed bumblebees (*Bombus terrestris*; pictured below) to anise flower scent, in some cases combined with a pheromone that has been linked to bee foraging. Bees leaving home to find food followed the anise odour whether or not the pheromone was present in the nest.

Foraging bumblebees, the authors argue, bring pollen and nectar into the nest and their peers thus learn what to search for by the target's scent. The pheromone the finders release has no specific teaching role; instead it increases the foraging behaviour of nestmates. Co-author Nigel Raine suggests dosing bumblebee nests with it and specific fragrances to improve commercial crop pollination.

N. RAINE

MOLECULAR BIOLOGY

Micro-conspirator

Cell **135,** 437–448 (2008)

An existing cancer drug may be effective against a specific type of nerve tumour because of its ability to block a particular metabolic pathway.

Wade Clapp at the Indiana University School of Medicine in Indianapolis, Luis Parada at the University of Texas Southwestern Medical Center in Dallas and their colleagues studied mice with tumours that derived their cancerous nature from a mutation in a gene called *Nf1*. For the tumours to grow, a biochemical pathway, c-kit, must be activated in bone-marrow cells with only one healthy copy of the gene, and immune cells called 'mast' cells from the bone marrow must have

migrated close to the tumour cells.

The authors suppressed the tumours' growth with the c-kit-inhibiting drug imatinib mesylate (Gleevec). Imatinib was then given to a child with a life-threatening *NF1*-dependent tumour, which subsequently shrank by about 70%.

CHEMICAL BIOLOGY

Clotting by quorum

Nature Chem. Biol. doi:10.1038/nchembio.124 (2008)

In some invertebrates, such as horseshoe crabs, the presence of bacteria can directly trigger blood clotting, which stops infection from spreading. But the initiation of vertebrate blood clotting was thought to require a more complex system of biochemical signalling.

Rustem Ismagilov of the University of Chicago in Illinois and his co-workers suggest otherwise. The team found that clusters of *Bacillus anthracis*, which causes anthrax, and *Bacillus cereus* initiated coagulation in mouse and human blood within minutes, yet dispersed bacteria of the same species did not. The bacteria directly reacted with clotting enzymes called coagulation factors and the blood clotted only when these factors reached a critical density — which happened when the bacteria formed clusters.

QUANTUM PHYSICS

Entangled accuracy

Nature Phys. doi:10.1038/nphys1112 (2008)

Quantum information is usually sent by light that is 'entangled'. This means that properties of photons sent between the two parties are quantum mechanically linked, and that the information changes when it is 'read', thus providing theoretically perfect data security. But entanglement is delicate and can easily be disrupted by factors such as atmospheric turbulence — so the receiver cannot tell whether a message was intercepted, or if the weather got in the way.

Ulrik Andersen of the Technical University of Denmark and his co-workers think that they can. They sent entangled light pulses through a simulated atmosphere, finding that when the pulse's amplitude fell within a certain range, it remained entangled.

The authors hope that the technology will help to improve the long-distance transfer of quantum information.

MOLECULAR BIOLOGY

Chewing the fat

Science **322,** 957–960 (2008)

Certain stem cells may link ageing and obesity, according to Gary Ruvkun and his colleagues at the Massachusetts General Hospital in Boston. When they halted the development of germline stem cells — those that make reproductive cells — from the worm *Caenorhabditis elegans*, the animals suddenly began storing less fat.

The authors pinpointed a gene, *K04A8.5*, that encodes a fat-burning enzyme, and showed that this gene's expression is boosted in fat storage tissues in worms that lack germline stem cells. Elevating the expression of *K04A8.5* decreased fat storage in worms carrying mutations that conferred increased longevity, whereas reducing it using a technique called RNA interference shortened their lifespans.

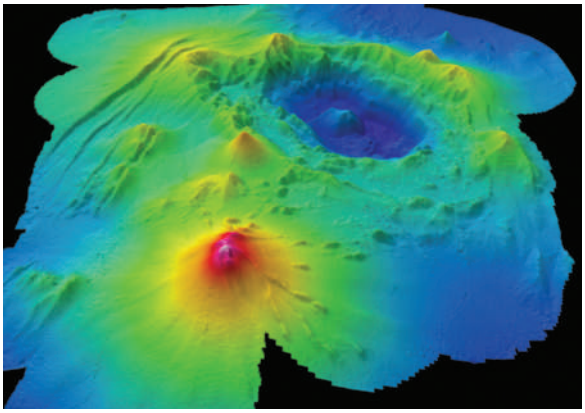
GEOSCIENCES

Submarine slippage

Geochem. Geophys. Geosys. doi:10.1029/2008GC002113 (2008)

Monowai Cone (pictured below) is an undersea volcano north of New Zealand. It is probably the world's most active, making it a good model system for studying how submarine eruptions cause landslides, and how both of these can trigger tsunamis. Bill Chadwick at Oregon State University in Newport and his colleagues have measured Monowai Cone's topology in 1998, 2004 and 2007 and have compared the timing of its build-up and collapse with data from the Polynesian Seismic Network.

Between 1998 and 2004 there were nine swarms of '*T* waves' — indicators of explosive eruptive activity — of which the biggest was linked to a collapse. But another large collapse occurred between 2004 and 2007 when no anomalous *T*-waves were recorded — possibly owing to limitations of the monitoring network. The authors say that this is the first study of its type and that more effort should be spent monitoring undersea volcanoes.



JOURNAL CLUB

Douglas Natelson
Rice University, Houston, Texas

A physicist foresees a new era in electronics.

A material's electronic properties depend largely on its density of mobile charge carriers (electrons and holes). The most common way of tuning that density is 'doping'. This involves carefully adding atoms or molecules that donate or take up electrons from the surrounding material. But doping comes with a downside: these added impurities themselves become charged, so they scatter mobile charge carriers and muddy the predictability of the material's electronic properties.

How to avoid doping? Look to Julius Edgar Lilienfield. In 1925, he proposed what is now called the 'field effect', in which the material of interest functions as one electrode of a capacitor. When a voltage is applied to the other electrode, equal and opposite charge densities accumulate on the sample material. The density of charge carriers can be varied as it is in doping, but not to the same extent. Nonetheless, the field effect has an everyday role in transistors — which are the fundamental parts of consumer electronics.

Another of Lilienfield's inventions, the electrolytic capacitor, holds the key to much higher field-effect charge densities, which could have dramatic consequences. Researchers at Tohoku University in Sendai, Japan, recently used a polymer electrolyte to achieve sufficiently large charge densities at a strontium titanate surface to generate superconductivity (K. Ueno *et al. Nature Mater.* **7,** 856–858; 2008). This has been seen before in doped strontium titanate, but the electrolytic capacitor approach avoids the disorder inherent in doping.

By using mobile ions in an electrolyte to attract charges in the sample, this quirky capacitor can build up charge densities approaching those of chemically doped electronic materials such as high-temperature superconductors. This opens up the possibility of transistor-like devices that can work with very low voltages.

Discuss this paper at **http://blogs.nature.com/nature/journalclub**

NEWS

Climate first for Obama transition team

Appointments to key energy positions should reveal the new president's priorities.

After eight years out in the cold, Democrats in Washington DC are rushing to warm themselves by the White House fireplace. And although research advocates may have to wait to see major science-policy positions in the new administration filled, a sense of the next president's priorities and style of governance is starting to emerge.

Barack Obama, who defeated John McCain on 4 November for the presidency, swiftly set up a transition team to advise on the hundreds of jobs at senior levels in government that the administration will need filled, many of them, ideally, before the presidential inauguration on 20 January 2009. Running the team are John Podesta, the last of President Bill Clinton's four White House chiefs of staff; Valerie Jarrett, a long-time Obama confidante who is chief executive of The Habitat Company, a Chicago real-estate business; and Peter Rouse, who serves as Obama's chief of staff in the Senate. On 6 November Obama appointed congressman Rahm Emanuel (Democrat, Illinois) to be his White House chief of staff. Announcements of Cabinet positions are expected over the next couple of weeks. The secretaries of the treasury and defence are considered the highest priority, and former Harvard University president Lawrence Summers may reassume the treasury position he held under Clinton. Robert Gates, the incumbent defence secretary, may be asked to stay on, a move that would confirm the team's stated willingness to look at Republicans and independents as well as Democrats.

Podesta, who trained in psychology as an undergraduate before going on to study law, founded the liberal think tank Center for American Progress (CAP) in 2003. The Washington-based group includes a number of leading scholars in science, technology and the environment, including Thomas Kalil — another Clinton-White-House veteran, now special assistant to the chancellor for science and technology at the University of California, Berkeley — who served as a science adviser to the Obama campaign.

The transition team declines interviews, but the CAP may provide some clues to how certain issues may fare under an Obama presidency — chief among them energy and climate. In a book published this week, *Change for America: A Progressive Blueprint for the 44th President*, the think tank outlines its case for a 'national energy council' that would coordinate energy policy among the various agencies and departments with a stake in such issues.

The creation of such a council, under an 'energy czar', would not have the same radical restructuring effect on government as that intended in the United Kingdom's recent creation of a cabinet-level Department of Energy and Climate Change, but it would provide a coordinating role not seen before. "I think it's a great idea," says Bill Becker, a climate policy expert at the University of Colorado in Denver. "A council like this can ensure that comprehensive climate action is delegated to the agencies and that they are coordinated in carrying out a climate action plan."

"The top question on many researchers' minds will be the next presidential science adviser."

In the Clinton White House, Vice-President Al Gore took on a de facto role as climate czar. Incoming Vice-President Joe Biden seems unlikely to follow him in this, although he has worked extensively on climate issues as chair of the Senate foreign-relations committee, and it would not be surprising to see him playing a role in the climate negotiations slated for Copenhagen in December 2009, as Gore did in Kyoto. Tim Wirth, a former senator who led climate efforts in the US State Department during the first Clinton administration, says that Podesta, as chief architect of the plan, would automatically leap to the top of any list of people who might assume the suggested czarship — although whether that is what Podesta wants remains unclear.

Other advisers who might be in the frame for the job, or for a seat on the council, would be Dan Kammen, of the University of California, Berkeley, and possibly John Holdren, of Harvard University. For energy- and environment-related cabinet appointments, such as secretary of energy and head of the Environmental Protection Agency (EPA), insiders poo-poo some of the high-profile names that have been bandied about for these jobs, including lawyer and environmental activist Robert F. Kennedy Jr, and Arnold Schwarzenegger, the Republican governor of California. More likely contenders

The congressional outcome

In last week's elections, 35 Senate seats and all 435 seats in the House of Representatives were contested. Democrats increased their majorities in both houses, although they did not attain the 60 seats in the Senate that would allow them to circumvent some opposition stalling tactics.

Bart Gordon (Democrat, Tennessee) and Ralph Hall (Republican, Texas) will continue to serve as ranking majority and minority members of the House Committee on Science and Technology. The three scientists who serve in the House of Representatives — Democrat Rush Holt of New Jersey, Republican Vern Ehlers of Michigan, and Democrat Bill Foster of Illinois — were all re-elected. But Democrat Nick Lampson of Texas, a big supporter of NASA, lost his seat after the boundaries of his district were redrawn.

Rearrangements, rather than comings and goings, may have the most effect on the Hill. In the House, liberal Representative Henry Waxman of California is trying to wrest control of the energy and commerce committee from John Dingell of Michigan, the longest-serving member in the House, whose district includes a lot of America's troubled, subsidy-seeking auto industry. Waxman might push climate policy more fervently; Dingell, though, would be better placed to get industry on board. In the Senate, 90-year-old Robert Byrd of West Virginia has relinquished control of the appropriations committee, which allocates funding to science and other agencies. He will be replaced by Daniel Inouye of Hawaii, 84, also a Democrat, and Inouye's position as chair of the Commerce, Science and Transportation Committee goes to Jay Rockefeller of West Virginia. Inouye, who wanted to be a surgeon before losing an arm in military service, has a demonstrated and geographically unsuprising interest in ocean research. **A.W.**

John Dingell.

S. J. FERRELL/CONGRESSIONAL QUARTERLY/GETTY IMAGES

Barack Obama could make key decisions sooner rather than later.

D. GUTTENFELDER/AP

Executive orders

P. MCCARTEN/UPI PHOTO/NEWSCOM

John Podesta (right), co-chair of the transition team, is already talking about how President-elect Obama might overturn executive orders previously signed by President George W. Bush. Talking to Fox News on 9 November, Podesta called Bush's restrictions on stem-cell research and his opening of public lands to oil and gas drilling "not in the best interest of the country". Bush's stem-cell order, signed on 9 August 2001, prohibits federal funding from being used for research on human embryonic stem-cell lines derived after that date. Obama aides have reportedly already drafted wording to lift the order.

In other good news for stem-cell research, Michigan's Proposal 2 ballot initiative (see *Nature* **455,** 1154–1155; 2008) was approved by voters; Michigan researchers will now be able to legally derive human embryonic stem-cell lines. **A.W.**

for EPA chief include New Jersey environmental chief Lisa Jackson and Massachusetts energy secretary Ian Bowles. Pennsylvania governor Ed Rendell (Democrat) has been mentioned as a potential secretary of energy.

Whoever gets the EPA job could have a strong influence on how climate policy is implemented; last year, the Supreme Court ordered the agency to regulate greenhouse gases under the Clean Air Act. Obama campaigned on a promise to reduce greenhouse-gas emissions by 80% from 1990 levels by 2050 using a cap-and-trade system that would generate US$150 billion for clean-energy technologies; Robert Sussman — a former EPA deputy administrator who is now serving on the transition team — told *Nature* earlier this year that this would be a priority if Obama were elected. (Sussman declined interview requests last week.) In a recent interview with *New York* magazine, Emanuel pointed out that of the four areas an Obama administration would seek to reform — financial regulation, tax, health care and energy — energy was one where "you can do some things immediately".

Beyond the potential energy appointees, the top question on many researchers' minds will be the next presidential science adviser and head of the Office of Science and Technology Policy. Obama has said his presidential science adviser will have the official status of an assistant to the president — something the current science adviser, John Marburger, has not been granted. In the past, presidents have made this appointment anywhere from six weeks before the inauguration — Richard Nixon appointing Lee DuBridge — to five months after it — Bush appointing Marburger. Many researchers will want Obama to emulate Nixon in this (although not to go on to abolish the post completely, as Nixon did in a fit of pique in 1973), and on the campaign trail he said he would fill the position quickly.

There is no shortage of people offering advice on the subject. Last week, the Union of Concerned Scientists (UCS), an advocacy group based in Cambridge, Massachusetts, released a five-point checklist for science appointments. Francesca Grifo, director of the scientific integrity programme at the UCS, says that incoming appointees can do a few simple things to assure their staff members that science will have a place at the policy table; as an example she cites William Ruckelshaus, who on taking charge of the EPA in 1983 wrote a 'fishbowl' memo stating that all work would be done in complete transparency. Whether that is sufficient remains to be seen. "The big event is when the first scientist at an agency has research results that are controversial and tries to speak out," says Grifo. "What's going to happen to that first scientist?"

Other open positions in the Obama administration include directorship of the National Institutes of Health (NIH), from which Elias Zerhouni stepped down on 31 October, leaving former deputy Raynard Kington in charge as acting director. The NIH is under the umbrella of the Department of Health and Human Services. Various names have been floated for the Cabinet-level position of chief of this department, including former presidential candidate Howard Dean, a physician and former Democratic governor of Vermont, where he greatly improved access to health care, and former South Dakota senator Tom Daschle.

One major position is not just open but new. Obama has promised to appoint the first-ever government chief technology officer to look after digital infrastructure, electronic delivery of government services and other high-tech issues. Names rumoured to be under discussion include more or less everyone in the Internet business that people outside it have ever heard of, including Eric Schmidt, the chief executive of Google; Jeff Bezos, the founder of Amazon.com; Bill Joy, formerly of Sun Microsystems; Vint Cerf, one of the Internet's original architects; and Steve Ballmer of Microsoft. Technology advisers have played a prominent role in the early days after Obama's election — Schmidt was standing by his side during his first press conference as president-elect, and is also part of the economics transition team. ■

Alexandra Witze, with additional reporting by Jeff Tollefson

Have your say on who should assume the top science positions at http://tinyurl.com/5sjd4e.

K. HYNDMAN

FISH GET FRESH
Saltwater sculpin evolve for estuary life.
www.nature.com/news

Modified genes spread to local maize

Transgenes from genetically modified (GM) maize (corn) crops have been found in traditional 'landrace' maize in the Mexican heartland, a study says. The work largely confirms a similar, controversial result published in *Nature* in 2001 (ref. 1) and may re-ignite the debate in Mexico over GM crops.

The paper reports finding transgenes in three of the 23 locations that were sampled in 2001, and again in two of those locations using samples taken in 2004. Written by a team led by Elena Álvarez-Buylla of the National Autonomous University of Mexico (UNAM) in Mexico City, the study will be published in the journal *Molecular Ecology*.

In 1998, the Mexican government outlawed the planting of GM maize to protect its approximately 60 domesticated landraces and their wild relatives. But newspaper reports suggest that farmers have planted at least 70 hectares of GM maize crops in the northern state of Chihuahua, and it is unclear what repercussions this may have.

Only about 25% of the maize planted in Mexico comes from commercially sold seed; the majority is saved from harvest to harvest. That's why, says Álvarez-Buylla, researchers need to pin down whether transgenes really have made it into local crops. "It is urgent to establish rigorous molecular and sampling criteria for biomonitoring at centres of crop origination and diversification," the team writes.

Allison Snow, a plant ecologist from Ohio State University in Columbus, led a team that reported[2] in 2005 it could not detect transgenes in maize from regions sampled by the original *Nature* paper. She calls the new work "a very good study, with positive signs of transgenes".

"It is good to see this," adds Ignacio Chapela, the ecologist from the University of California, Berkeley, who was senior author on the *Nature* publication. "But it took seven years."

E. ÁLVAREZ-BUYLLA

Mexico's ban on genetically modified maize has not stopped transgenes getting into traditional crops.

Testing times

The original paper caused a storm of controversy[3–5]. Critics pointed out some technical errors, including problems with the type of PCR used to amplify the genetic sequences, although Chapela and his co-author David Quist stood by their conclusions[6]. Others questioned whether the critics were influenced by their association with the biotechnology industry, which they denied. In the end, *Nature* published an editor's note saying there was insufficient evidence to justify the original publication. Advocates of GM crops widely, and erroneously, called this a retraction.

A second round of criticism was sparked in 2005, after the Snow paper reported no evidence for transgenes in Mexican maize. Some criticized this article as being statistically inconclusive and lacking representative samples[7], which the authors disputed[8].

Álvarez-Buylla's team set out to resolve the issue by conducting genetic tests on thousands of maize seed and leaf samples for evidence of two transgenes: a gene promoter from the 35S cauliflower mosaic virus, and the nopaline synthase terminator, NOSt. The team found transgenes in about 1% of more than 100 fields it sampled, including some sampled by Quist and Chapela in 2001.

"The importance of the study is not the impact of the transgenes themselves, but the fact that their spread has occurred so easily."

Jose Sarukhán, a biologist at the UNAM and a member of the US National Academy of Sciences, recommended the Álvarez-Buylla article for publication in *Proceedings of the National Academy of Sciences*. It was rejected; in a letter to the authors on 14 March this year, the journal's editor-in-chief Randy Schekman, a professor at the University of California, Berkeley, wrote that the biology and genetics didn't warrant publication, and that a reviewer had pointed out the report could "gain undue exposure in the press due to a political or other environmental agenda". Sarukhán responds: "I saw no reason why it should not be published."

Norman Ellstrand, a plant geneticist at the University of California at Riverside, called the study intriguing. "The importance of the study is not in the impact of the transgenes themselves," he says, "but in the fact that their spread has occurred so easily in a country where the planting of transgenic maize has not occurred for several years."

However, the new paper doesn't confirm an important conclusion from the original *Nature* paper — whether the transgenes had been integrated into landrace genomes and passed along to progeny plants. Álvarez-Buylla suspects this may be the case, but she's not interested in pursuing another round of politically charged battles — and will leave that work to others. ■

Rex Dalton

1. Quist, D. & Chapela, I. *Nature* **414,** 541–543 (2001).
2. Ortiz-Garcia, S. *et al. Proc. Natl Acad. Sci. USA* **102,** 12338–12343 (2005).
3. Suarez, A. V. *et al. Nature* **417,** 897 (2002).
4. Metz, M. & Futterer, J. *Nature* **416,** 600–601 (2002).
5. Kaplinsky, N. *et al. Nature* **416,** 601–602 (2002).
6. Quist, D. & Chapela, I. *Nature* **416,** 602 (2002).
7. Cleveland, D. A. *et al. Environ. Biosafety Res.* **4,** 197–208 (2005).
8. Ortiz-Garcia, S. *et al. Environ. Biosafety Res.* **4,** 209–215 (2005).

PUNCHSTOCK

SECURITY SLOW-DOWN
Scientist-vetting scheme faces delays in Britain.
www.nature.com/news

China asks world to step up on climate

BEIJING

The United Nations Framework Convention on Climate Change (UNFCCC) should set up an inter-government agency to help developing countries tackle climate change, international representatives agreed last week at a meeting in Beijing.

The meeting, convened by the UN and China's National Development and Reform Commission (NDRC), was meant to prepare the ground for next month's summit in Poznań, Poland, where negotiations on a treaty to replace the Kyoto Protocol will kick off. Many believe the very existence of the Beijing conference is a sign that China is seeking a more active role in climate-change talks, in contrast to its previously low profile.

In his opening address, Chinese premier Wen Jiabao said that the developed world needs to move more aggressively to transfer environmentally sound technologies and know-how to other countries, as specified by the UNFCCC. "Many developing countries are undergoing rapid economic growth and urgently need access to clean technologies to make their development environmentally friendly," he said. "There has been no substantial progress so far."

At a UN conference in Bali last December, developing countries agreed to work harder to reduce greenhouse-gas emissions provided they receive "steady, predictable and reliable" financial and technological support. "It's critical that the details of how this could function are worked out," says Yvo de Boer, executive secretary of the UNFCCC.

In 2001, the convention created an Expert Group on Technology Transfer, but this is largely an advisory committee, with no power to implement its recommendations. "It's a lion without teeth," says William Kojo Agyemang-Bonsu of Ghana's Environmental Protection Agency.

Under the framework proposed in brainstorming sessions at the Beijing conference, the new inter-government agency would be an independent body able to make and implement decisions and monitor compliance. It would oversee and verify mitigation targets of developing countries, identify barriers to technology transfer, and propose countermeasures. Developed countries would commit to providing it with a steady stream of income for its primary operating budget, possibly supplemented with money from the private sector and other sources.

R. JONES/REX FEATURES

China's greenhouse-gas emissions have caught up with those of the United States.

Last month, Gao Guangsheng, a senior official for climate-change policy at the NDRC, said that developed countries should commit 1% of their gross domestic product (GDP) to help developing countries tackle climate change. As measured for all the countries in the Organisation for Economic Co-operation and Development, that would clock in at $284 billion. Ana Maria Kleymeyer of Argentina's secretariat of environment and sustainable development calls that commitment reasonable. "Developed nations have the obligation of clearing up the mess they have made over the past centuries," she says.

> "Developed nations have the obligation of clearing up the mess they have made over the past centuries."

Still, many representatives from developed countries say they first want to see developing countries make a firm commitment to emissions-reduction targets. "We are talking about a lot of money here," says Jukka Uosukainen of Finland's environment ministry and chair of the Expert Group on Technology Transfer. "There is no point in transferring low-carbon technologies to a country that is actively promoting a high-carbon economy."

For instance, many would like to see China cap its own greenhouse-gas emissions. On 29 October, China acknowledged for the first time that its greenhouse emissions have caught up with those of the United States and are unlikely to fall any time soon. It declined to give specific figures or say whether these have exceeded US levels, which last year reached 1.6 billion tonnes of carbon emitted from burning fossil fuels.

Su Wei, director of the climate-change department at the NDRC, says China has invested heavily in mitigation strategies and remains committed to reducing energy consumption and major pollutant discharge per unit of GDP by 20% and 10%, respectively, from 2005 levels by 2010. "However, it's not time to set an absolute emission cap yet," he says, noting that the emissions level per person in China is only a fifth of that in the United States. And even with such reductions, overall energy consumption would still be higher in 2010 than in 2005 because of the sheer growth of China's economy.

Many questions remain. At the Beijing conference, some discussions covered ways of generating market incentives for technology transfer and reforming the patent system so that poor countries can benefit from technological innovations. Also raised was the fact that market approaches won't solve other neglected areas of development, such as ensuring clean fuel supplies and better agricultural practices.

Conference organizers are putting together a detailed report on overcoming barriers to technology transfer that will be discussed at the Poznań conference. ■

Jane Qiu

WHO congress backs traditional medicine

The World Health Organization (WHO) has endorsed an international agreement that promises to give traditional medicines a foothold in health systems around the world.

At its first Congress on Traditional Medicine in Beijing on 8 November, the WHO unveiled the 'Beijing Declaration'. The declaration calls on member states of the United Nations to formulate policies for ensuring the safe and effective use of traditional medicines, to create systems for licensing practitioners, and to encourage communication between Western and traditional practitioners.

L. DERIMAIS/EPA/CORBIS

Chinese medicine is coming in from the cold.

China in particular has long been keen on the international acceptance of traditional medicines. But researchers have had difficulty pinning down and standardizing the active ingredients, as these are often found in plant extracts that have regional or seasonal variations. Because of this, drug regulatory agencies around the world have hesitated to approve their use.

Proposed reform to animal testing rules draws fire

The European Commission last week unveiled a proposal to reform the regulation of scientific experiments that use animals.

The commission's suggested update to the 20-year-old 86/609/EEC directive will now be considered by the European Parliament, probably in 2009. However, business and academic groups are pushing for a number of changes, voicing concern over the draft's proposals to limit the reuse of the same animal in a series of experiments, and to curb the use of specimens captured from the wild.

Primate researchers are particularly worried by elements of the proposed directive that could prevent any basic research on non-human primates that does not relate to either the survival of the species or to serious clinical conditions in humans. All experiments on great apes would be banned, except during extreme circumstances such as pandemics.

Contaminants identified in plastic lab equipment

Thousands of scientists could be unwittingly ruining their experiments merely by using standard plastic lab equipment, according to a study published in *Science*. Andrew Holt of the University of Alberta in Edmonton, Canada, found that a disinfectant and a lubricant used in the manufacture of plastic pipette tips, tubes and micro-plates could leach into common solvents, including water (G. R. McDonald *et al. Science* 322, 917; 2008).

Holt and his team determined that the contaminants are potent enzyme inhibitors, and skewed their drug-activity assays. "People in any big science department will say that they don't trust plastic in some experiments," says Holt. But the compounds responsible had never been identified before, he adds. Eppendorf, a manufacturer of some of the plastic products, says that "so far, we have not experienced any product problems with our customers due to these substances".

For a longer version of this story, see http://tinyurl.com/5z3hau.

Marine census nets wealth of unknown ocean life

M. RAUSCHERT

Giant sea stars and bacteria that can grow to several centimetres long are among more than 5,600 marine species that scientists have discovered during work on the first comprehensive global marine-life census. The census, which began in 2000 and should be completed by 2010, has also provided evidence that many of the world's deep-sea octopuses evolved from a common ancestor, of which a close relative, *Megaleledone setebos* (pictured), still lives in the Southern Ocean.

The most recent findings from the 2,000-strong international team of marine scientists behind the effort were presented at the World Conference on Marine Biodiversity in Valencia, Spain, which began on 11 November.

Research agency focuses on first-time applicants

Elias Zerhouni, former director of the US National Institutes of Health (NIH), spent his last day in the job helping young scientists. On 31 October, he created a formal regulation out of what has been a de facto policy for the past two years aimed at reversing a steady decline in the number of young investigators securing the agency's bedrock R01 research grants.

The regulation means that first-time grant applicants won't be competing directly against seasoned investigators. Instead, the agency will "wherever possible" conduct initial peer review on applications from new investigators separately.

The agency expects that the policy will mean that at least 1,650 new investigators will receive awards in 2009 — up from a decade-long nadir of 1,354 grants in 2006. Since 1980, the average age at which researchers receive their first NIH grant has climbed from 37 to 42.

Billion-dollar bid for stem-cell treatments

Genzyme, a biotechnology company based in Cambridge, Massachusetts, announced last week that it will invest up to US$1.4 billion in two adult stem-cell therapies being developed at Osiris Therapeutics.

Osiris, which is headquartered in Columbia, Maryland, is conducting late-stage clinical trials using mesenchymal stem cells harvested from bone marrow and grown in large-scale cultures. Genzyme will pay $130 million up front for the rights to commercialize the products and will fund 40% of the cost of future phase III and IV clinical trials, with the remainder of the sum as milestone and royalty payments.

The first therapy, Chondrogen, is intended to treat osteoarthritis in the knee. The second, Prochymal, has already been approved by the US Food and Drug Administration as a last-ditch treatment for children with a potentially deadly immune response to tissue transplantation called graft-versus-host disease.

SPECIAL REPORT



M. LIMA/AFP/GETTY IMAGES; CHENZHUO/REX FEATURES; R. DREW/AP

Science in the meltdown

The research enterprise faces many uncertainties in the looming global recession — but it also has many strengths that may help it weather the storm. *Nature* investigates.

The crisis that swept across the world's financial markets this autumn is widely regarded as the worst since the 1930s. The global economic downturn that helped precipitate the crash, and will be duly amplified by it, is widely expected to be the worst in a generation, at least. The effect this will have on the research enterprise will depend crucially on how the world's governments respond to the crisis — on what stimulus they think is necessary, and on what long-term commitments they may be willing to cut to deal with present pain (see page 141).

Whatever those responses, there are likely to be adverse consequences for many research undertakings, from university departments to corporate laboratories. The cost of capital, and the timescales of return expected on investment, will change the outlook for industry. Changed economic expectations may have an impact on the way investments in education are seen. In this special report, *Nature* looks at the current impacts and future trends in academia, industry and government — and at the particular problems and opportunities in energy research.

Industry

If history is any guide, the worldwide research enterprise could survive this downturn comparatively well. Total research and development (R&D) spending by both government and industry rose at a fairly steady rate through major recessions in the early 1980s and 1990s, as well as through the 'dot-com' bust in the early 2000s (see graph, page 156). So did the total number of researchers (see graph, top of page 158). Certainly these trends are far less volatile than those seen in the stock markets (see graph, page 157).

However, it remains to be seen whether that pattern will persist. And in any case, such aggregate figures can hide a great deal of turmoil and uncertainty felt by those living through the crisis. Nowhere is this more obvious than in the pharmaceutical and biotechnology industries. Large companies with well-established product lines seem to be in relatively good shape for the moment; sales may well drop during the downturn, but will not go to zero. Most major pharmaceutical companies have fairly copious cash — enough to be looking for acquisitions among the less fortunate. Meanwhile, the few larger biotech companies that already have products in the market — such as Amgen in Thousand Oaks, California, which reported strong earnings last month — are weathering the storm well.

But, the financial crisis has been brutal for the biotechnology industry. Falling stock prices — the Amex Biotechnology Index has lost about 25% of its value since mid August — and the global credit crunch have hit companies around the world. For example, Bavarian Nordic of Kvistgaard, Denmark, which supplies smallpox vaccines to the US biodefence stockpile, said last month that it is shoring up its defences against a hostile takeover based on its low share price. The gloom is not universal. At least some of the firms involved with personal genomics are optimistic, as investors see it as an area poised for quick market growth. And the sequencing business is not in crisis either.

"We're in a pretty fortunate position" and relatively immune from the economic crisis, says Jay Flatley, chief executive of Illumina in San Diego, California, which makes microarrays and high-speed sequencing machines.

But, says Cliff Reid, chief executive of Complete Genomics, a sequencing company in Mountain View, California, "we're a really nice home in a really bad neighbourhood".

Moreover, venture capital is becoming scarce, as venture-capital firms shift more towards 'feeding their older children': reinvesting in established companies that promise near-term payoffs, while turning away from riskier prospects that promise only long-term results. "It's going to be tough sledding for companies needing to raise capital in the next 6 to 12 months," says Glen Giovannetti, global biotech leader at Ernst & Young in Boston, Massachusetts. So the smaller biotech companies are tightening their belts and turning to the more established firms as likely saviours. "Clearly there will be a lot of companies knocking on the doors of Pfizer, Lilly and Merck to see if they want to collaborate," says Alan Lewis, chief executive of Novocell, a stem-cell engineering company in San Diego.

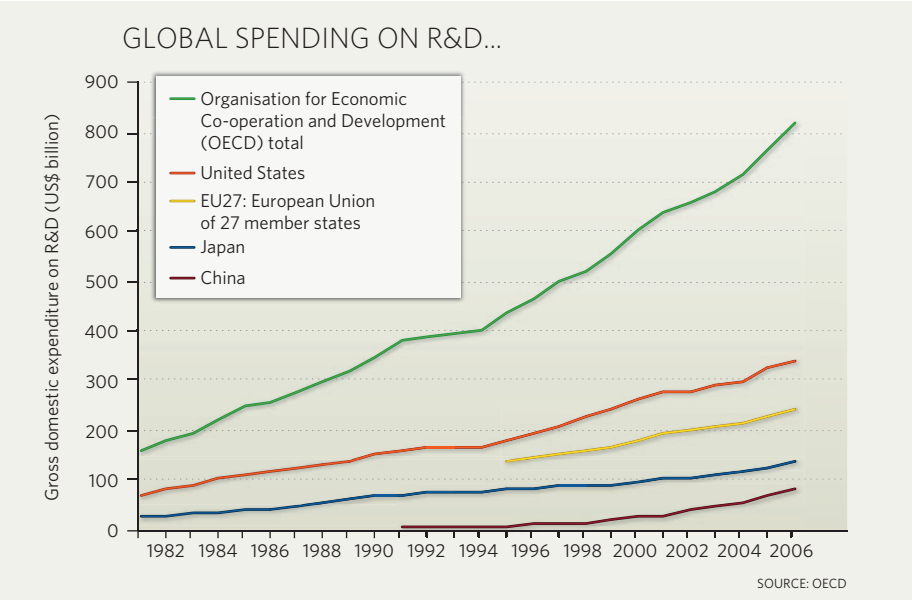
It's not all one-way traffic. Drug companies are looking more and more to biotechnology companies to help replenish their dwindling pipelines. So one likely outcome of the downturn is a flurry of alliances and acquisitions, as companies combine forces.

A similar story is seen in the emerging clean-energy industries. Companies could flourish in some potentially high-growth areas. For example, Michael Holman, an analyst for Lux Research, a market-research firm in New York, foresees continued growth in the

development of batteries for hybrid vehicles. And recent years have been highly profitable for solar-energy firms — a fact that has led Wacker, an international chemical giant based in Munich, Germany, to announce a plan to build a €760-million (US$977-million) polysilicon plant at its Nünchritz site. The company remains bullish despite the downturn, pointing out that the growth of photovoltaics has led to a shortage of polysilicon in the market.

Overall, however, the falling price of oil since last summer has taken some of the fizz out of green energy, as has the recognition that some earlier expectations were unrealistic. The return on investment for most non-solar-energy companies — especially biofuels — has ranged from poor to abysmal. According to New Energy Finance, a consultancy based in London, the prices of green-energy stocks worldwide have fallen 61% since the beginning of 2008. Just as in the biotech industry, there is less capital available for start-ups, says Jonathan Kestenbaum, chief executive of the UK National Endowment for Science, Technology and the Arts. And small start-ups are more likely to get a buy-out than to raise money in an initial public offering.

This story is repeated in industry after industry. The online business and technology news provider Xconomy in Cambridge, Massachusetts, reports that there is less venture capital everywhere — and energy and biotech are actually two of the bright spots. Xconomy predicts that there could be trouble ahead for Internet-start-up deals and capital-intensive technology operations such as telecommunications and semiconductor companies.



Universities and government

In the United States, unlike in most other countries, the great research universities are often private, not public — and those that are public are funded by individual US states. In either case, they face a unique set of challenges from the downturn.

On the one hand, they are somewhat buffered by an almost guaranteed set of customers: unless the economy melts down catastrophically, a new crop of undergraduates will show up every year. And the science- and technology-related graduate schools might even flourish. US National Science Foundation data show that enrolments have typically gone up (see graph, bottom of page 158) during previous economic downturns, when the forbiddingly tight job market led students to continue their education, or caused laid-off older workers to go back to school.

On the other hand, private universities are in much the same position as research foundations such as the Wellcome Trust. They depend heavily on investment returns from their endowments, which can drop precipitously during downturns (see *Nature* **455,** 712–703; 2008). All universities also face a reduction in charitable donations and falling income for corporate-sponsored research. At the same time, students' need for financial aid is increasing.

"All of these factors contribute to perhaps the tightest financial outlook we have seen in decades," wrote John Hennessy, the president, and John Etchemendy, the provost, at Stanford University in Palo Alto, California, in an e-mail circulated to the faculty members and staff on 30 October. They warned faculty members and staff to plan for roughly a 5% reduction in the $800 million general-funds budget for each of the next two years, and a "quite modest" salary programme.

Long-term planning becomes very difficult in a volatile environment, and institutions become reluctant to take on large new programmes — or new faculty members. Instead, many universities besides Stanford have been taking pre-emptive action to conserve resources in the near term. On 30 September, Robert Brown, president of Boston University, called for an immediate hiring freeze and a moratorium on new construction projects. In early November, Brown University in Providence, Rhode Island, and Cornell University in Ithaca, New York, imposed similar measures. Drew Faust, president of Harvard University, has said that the planning and development of the university's Allston campus will be reassessed (see *Nature* **454,** 686–689; 2008).

"Less investment in science would only turn this crisis into a long-term problem."
— Janez Potočnik

US states have also been cutting back on university support. On 21 October, the Regents of the University of California approved a 2008–09 operating budget that incorporated an effective 5% drop in the state's $3-billion contribution to the university — and the university's president, Mark Yudof, warned that more cuts were possible.

Elsewhere, the effect on universities — and research in general — will depend on the attitude of each national government: will they see money spent on research and education as an expense to be cut, or as a strategic investment for long-term economic growth?

So far, at least, most governments around the world have embraced the latter view, and pledged to keep their investment steady. For example, on 17 October, at the opening of an ultra-modern technology hub named Fusionopolis, Singapore Prime Minister Lee Hsien Loong noted concerns about his country's already well-advanced recession — and then declared: "The government remains fully committed to investing in R&D, in order to develop a key capability that will keep our economy competitive in the long term."

"What we need is more research and more collaboration," says Janez Potočnik, the European Union research commissioner. "Less investment in science and innovation would only turn this hopefully manageable crisis into a long-term structural problem for Europe."

Depending on the severity of the coming recession, however, those ambitions could soon be diminished by reality. Many countries are already deferring new initiatives, and proposing flat or decreasing budgets. "If there is less money available, there will only be losers, no winners," says Dieter

Imboden, president of the Research Council of the Swiss National Science Foundation.

In the United Kingdom, for example, research funding was generally flat last year, even before the downturn, so scientists are expecting hard cuts in the 2009 budget. "The only unknown is how deep and for how long," says Marjorie Wilson, a member of the Natural Environment Research Council, who is also pro-dean for research in the faculty of environment at the University of Leeds, UK. However, Paul Drayson, the newly appointed minister for science and innovation, has declared that he will "vigorously" defend the science budget against cuts; it remains to be seen how effectively he can do that.

"The only unknown is how deep and for how long."
— Marjorie Wilson

Elsewhere in Europe, experts think it is unlikely that the downturn will lead to any immediate major cuts in public science budgets. Most notably, says Potočnik, the European Union's €50-billion Seventh Framework Programme, the largest single research funding programme in the world, is expected to continue unscathed to its planned completion in 2013.

And in some European countries, including Germany, Switzerland, Sweden and the Netherlands, research budgets are buffered by multi-year science plans that call for annual rises in the budgets of national funding agencies, as well as for research organizations such as the German Max Planck Society. Indeed, Germany and Sweden recently announced that they would invest more in science and education despite the cloudy economic outlook.

Nonetheless, a deep and long-lasting economic recession could leave its mark on science throughout Europe, most likely through the effects of cuts or budget freezes on individual grants for basic research. Countries such as Iceland and Hungary, which have been hit hard, are thought to be particularly vulnerable in that respect. Russia, for its part, is taking the official view that the financial crisis is largely a problem for others. Expecting large economic growth to continue despite falling oil and gas prices, government officials said last month that all public budgets, including science, are safe.

In Asia, most of the R&D powerhouses are following Singapore's path by sticking to mid or long-term strategies, often laid out in five-year plans. Indeed, if continued support for science in Asia is coupled with cutbacks elsewhere, it might further accelerate a trend in which accomplished Asian scientists find better opportunities back home than in the West.

Japan, for example, has dipped in and out of recession since an asset bubble burst in the late 1980s. But Hayashi Towatari, the head of the science ministry's science and technology policy bureau, points out that Japan has maintained one of the highest rates of R&D investment as a proportion of gross domestic product (GDP) in the world — 3.62% in 2007. The science budget is currently flat, he says, and seems unlikely to grow anytime soon given the economic turmoil. But even so, he says, Japan is committed to the idea that science will build the economy (see *Nature* **455**, 1030–1031; 2008).

China has little exposure to the securities that have poisoned global financial institutions, although it has begun to see declining growth rates as economic turmoil around the world spoils its export markets. On 9 November, the government unveiled an economic stimulus programme totalling four trillion yuan (US$586 billion). Among other things, it will feature a tax deduction for corporate spending. But the Chinese government is still intent on bolstering its R&D spending as laid out in the current five-year plan for science and technology, which covers the period from 2006 to 2010.

...RISES FAR MORE STEADILY THAN THE STOCK MARKET

SOURCE: NASDAQ

feared additional burdens for their domestic industries (see *Nature* **455,** 1008; 2008). And the crisis may have had a role in the Canadian elections as voters backed the incumbent Conservative Steven Harper over Liberal opposition leader Stephane Dion, who had been pushing for carbon taxes to put Canada's economy on a greener path.

But the news isn't all bad. The United Kingdom recently adopted the world's first law mandating substantial long-term emissions cuts. The bill, which would cut emissions by 80% from 1990 levels by 2050, passed only after being strengthened to include a provision on emissions from aviation.

Moreover, the United States — the largest sticking point in the global-climate negotiations — seems to have resoundingly reversed course with the election win of Obama and the strengthening of the Democratic Party. Few expect a congressional passage of a carbon-trading scheme to be easy, even with larger Democratic majorities; there are too many interests to be reconciled. But Capitol Hill is very aware of the potential pay-off. Greenhouse-gas regulations could create a vast new revenue stream that could be tapped to pay for clean technology or raided for other priorities.

And, Democrats (and many Republicans) have long been sympathetic to the argument that clean energy could generate millions of new jobs, serving as a boost to the ailing economy, and have not backed away from funding such measures. Last month, Congress tacked $17 billion in tax incentives for clean energy onto the $700-billion Wall Street bailout bill; they also enacted a $25-billion loan-guarantee programme to help car makers to retool their facilities for the production of fuel-efficient vehicles.

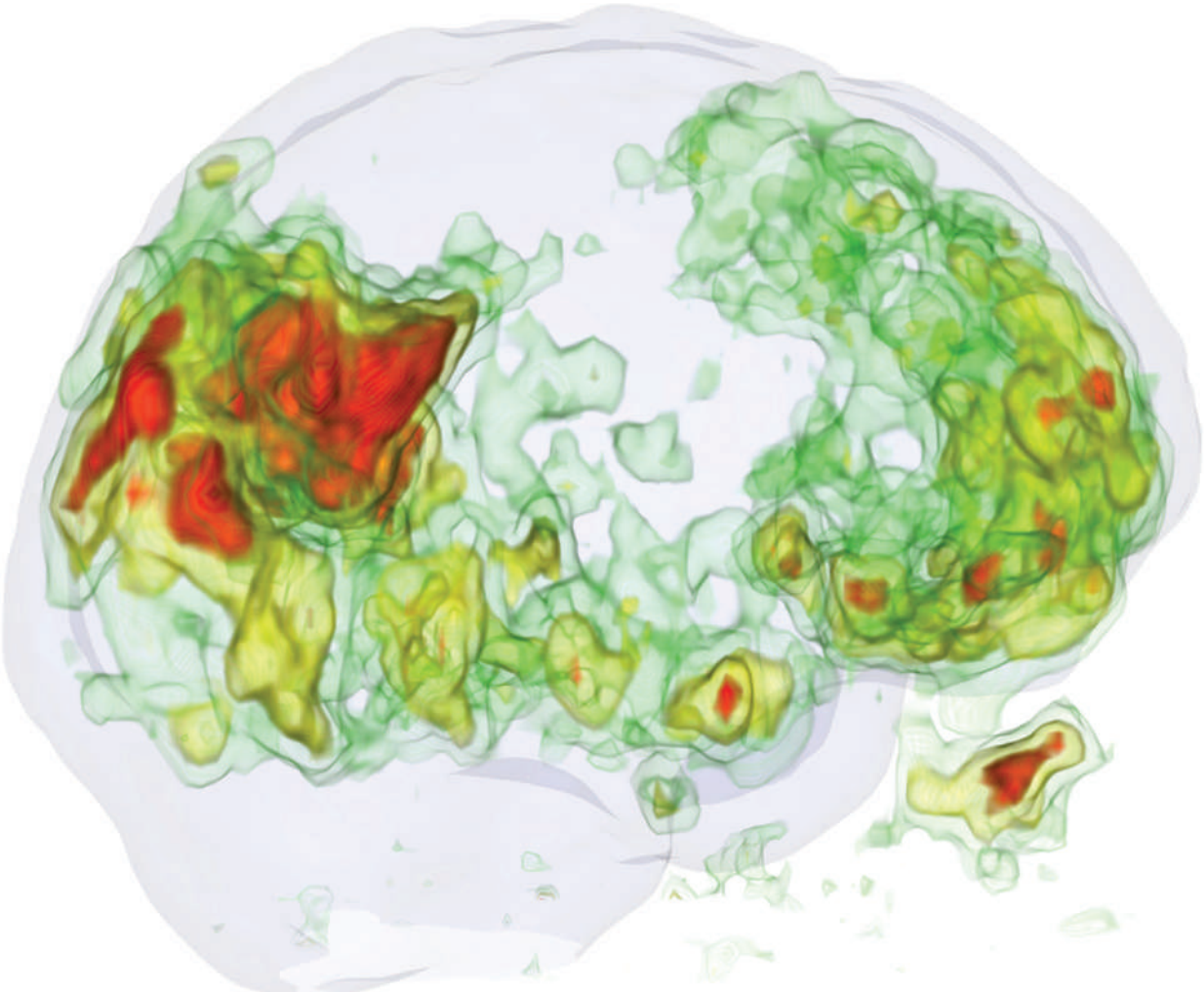
Of course, declining tax revenues and a growing national deficit could make such investments increasingly difficult. The financial sector's problems could also make it tougher for wind, solar and other green-energy companies to find financing for major projects. In addition, the price of carbon credits on the European Union's carbon trading scheme has fallen by roughly 40% since July — which translates into less investment in reducing emissions. And the plunge in oil prices — from $147 per barrel in July to just over $60 per barrel this week — is adding a whole new set of uncertainties.

Cheaper energy will put more money in the hands of consumers, and might well soften the looming recession. But it will also reduce demand for alternatives and could threaten investments in advanced-energy technology — particularly if wary investors want a faster pay-off. In some respects, this is exactly what happened after the price of oil collapsed in the 1980s.

Still, the difference this time is the concern over global warming, which isn't going away. Alden Meyer, director of strategy and policy for the Union of Concerned Scientists in Washington DC, points out that policy-makers will eventually have their hands forced by problems associated with global warming if they don't make the difficult choices now. "The reality is that there is no cost-free future." ■

M. Mitchell Waldrop, with reporting by Declan Butler, Rex Dalton, Erika Check, David Cyranoski, Natasha Gilbert, Eric Hand, Katharine Sanderson, Quirin Schiermeier, Jeff Tollefson and Meredith Wadman.

See Editorial, page 141.



Z. CSELÉNYI/ASTRAZENECA/KAROLINSKA INST. PET

THE PLAQUE PLAN

Neuroscientists are pretty sure they know what causes Alzheimer's disease, but their theory has not yet given rise to effective drugs. **Alison Abbott** asks what's wrong.

It is sad and bewildering to watch a parent slowly fade away, the years of authority dissolving into childlike moodiness. But it is an experience that more and more people will face as the number of elderly people with Alzheimer's disease climbs to the 80 million or so expected by 2040 (ref. 1).

If a strong hypothesis were enough to help, then treatments should be on the shelf. For years researchers have been confident that the disease is caused by protein 'plaques' and 'tangles' that eat away at the brain — and pharmaceutical companies have spent untold millions identifying and testing therapies that interfere with these structures.

But medical need and mechanism are not, it seems, enough. It has come as an acute disappointment that the first clinical trials of therapies designed to deal with the plaques have failed to help patients. In July this year, participants at the annual Alzheimer's Association meeting in Chicago, Illinois, heard that two major trials, one a phase III including 1,600 people, the other a phase II involving 234 people, failed to improve the mental state of patients even though they had looked promising in earlier studies. Another phase III trial was abandoned in 2007 when it was similarly unsuccessful. Something seems in need of a rethink, but what? The presumed biological mechanism — or the trials themselves?

Researchers say that they remain confident about their hypothesis. Their concerns lie in the way that the trials have been done. Because diagnosis is relatively rudimentary, some of the recruited patients may have had a different sort of dementia from Alzheimer's, and all of them may have had such advanced disease that their brains were irreversibly damaged. "We diagnose people when they already show dementia, not at a more helpful earlier stage," says Bruno Dubois, a neurologist at the Hôpital de la Salpêtrière in Paris. "It is like diagnosing someone with Parkinson's disease when they are bedridden, not when they develop a tremor." This means that the trial results are not so much negative as uninterpretable.

> "It is like diagnosing someone with Parkinson's when they are bedridden, not when they develop a tremor."
> — Bruno Dubois

Researchers and pharmaceutical companies had anticipated these problems, but hoped that the trials would show some therapeutic effect anyway. They pushed ahead because of the pressing medical need and the potential therapeutic market. At the same time, scientists have started work to improve the next round of trials by launching major international research programmes to identify biological markers of disease progression, such as those that expose plaques in living brains (pictured above). Such biomarkers might eventually reveal invisible equivalents to the Parkinson's tremor so that the disease can be recognized before dementia sets in and so that researchers can observe directly whether a

drug is doing the biological work expected of it. Researchers are already incorporating the most promising biomarkers into many of the clinical trials now under way.

Only with biomarkers in hand will scientists be able to seriously test the theory of Alzheimer's on which they have staked so much. "The theory is not proven, and we have to be open to the fact that it could be wrong," says Paul Aisen, a neurologist at the University of California, San Diego, who runs the US Alzheimer's Disease Co-operative Study for the National Institute of Aging in La Jolla, California. "But the trials so far have simply not tested the hypothesis."

Deep roots

The hypothesis in question has roots that can be traced back more than a century to Alois Alzheimer, a physician and neuropathologist in Munich, Germany, who followed the five-year mental decline of one of his patients, Auguste Deter. When Deter died in 1906, Alzheimer scrutinized her brain using a staining technique developed by his colleague Franz Nissl. He saw a mess of clumped proteins in the form of plaques, and tangles of fibrils like snarled-up thread. These characteristic features, which until recently could be seen only at autopsy, still provide the most secure diagnosis for the disease, together with neuropsychiatric assessments of memory, language and other cognitive functions.

When molecular biologists started to dissect the structures that Alzheimer observed, they found that plaques are made up of small, sticky protein fragments, or peptides. These peptides, the most common of which is called β-amyloid, are cleaved from a larger amyloid precursor protein (APP) by the enzymes β-secretase and γ-secretase. Neurofibrillary tangles, inside neurons, are now known to be made of an insoluble protein called tau. The prevailing theory holds that Alzheimer's disease is initiated when APP — the normal function of which is unknown — is converted to toxic amyloid peptides that are taken up and released by plaques. This stresses cells in a way that promotes so much phosphorylation of tau that it forms tangles. The process is known as the amyloid cascade. As the plaques and tangles spread from the hippocampus into the cerebral cortex, the people affected gradually lose their memory, their logic and eventually the ability to speak and move. In this weakened state, patients typically die of pneumonia.

The only drugs available for Alzheimer's patients aim to treat symptoms, trying to chemically stimulate areas of the brain that are already damaged by the time dementia sets in. They are marginally effective at best. So in the past five years, academic groups and pharmaceutical companies have focused on preventing the plaques and tangles from forming, or dismantling those already there.

So far, these strategies have generated a string of failures. Aisen was principal investigator on the North American trial of a drug called tramiprosate (Alzhemed) being developed by Neurochem, a biopharmaceutical company based in Montreal, Canada. It was the first trial based on the amyloid-cascade hypothesis to reach this advanced stage of testing and as such was greatly anticipated by the community. Tramiprosate is a small chemical that inhibits the formation of plaques in animal tests by binding to β-amyloid. Aisen's team followed up more than 1,000 patients for longer than 18 months but did not detect any change to their mental states and Neurochem has since pulled out of further clinical testing.

"The amyloid theory is not proven, and we have to be open to the fact that it could be wrong." — Paul Aisen

The two trials that failed this year used different approaches to attack amyloid, but with no greater success. Myriad Genetics' compound tarenflurbil (Flurizan) was designed to prevent the build-up of amyloid plaques by inhibiting γ-secretase. Myriad says that the drug might not have reached appropriate concentrations in the brain and the company abandoned it when it failed to show benefit. Elan Pharmaceuticals' bapineuzumab, a humanized monoclonal antibody that binds directly to β-amyloid and mimics an immune system attack on plaques, fared little better.

Perhaps the cruellest blow came when Clive Holmes of the Moorgreen Hospital in Southampton, UK, reported this July[2] on patients six years after they received an experimental amyloid peptide vaccine developed by Elan. The vaccine was designed to teach the immune system to recognize amyloid as being foreign so that it would launch an attack against the protein. Post-mortem analyses showed that almost all the patients had stripped-down amyloid plaques, despite most of them having progressed to severe dementia before they died.

Paul Aisen suspects that some patients in Alzheimer's trials have other types of dementia.

UC SAN DIEGO SCHOOL OF MEDICINE

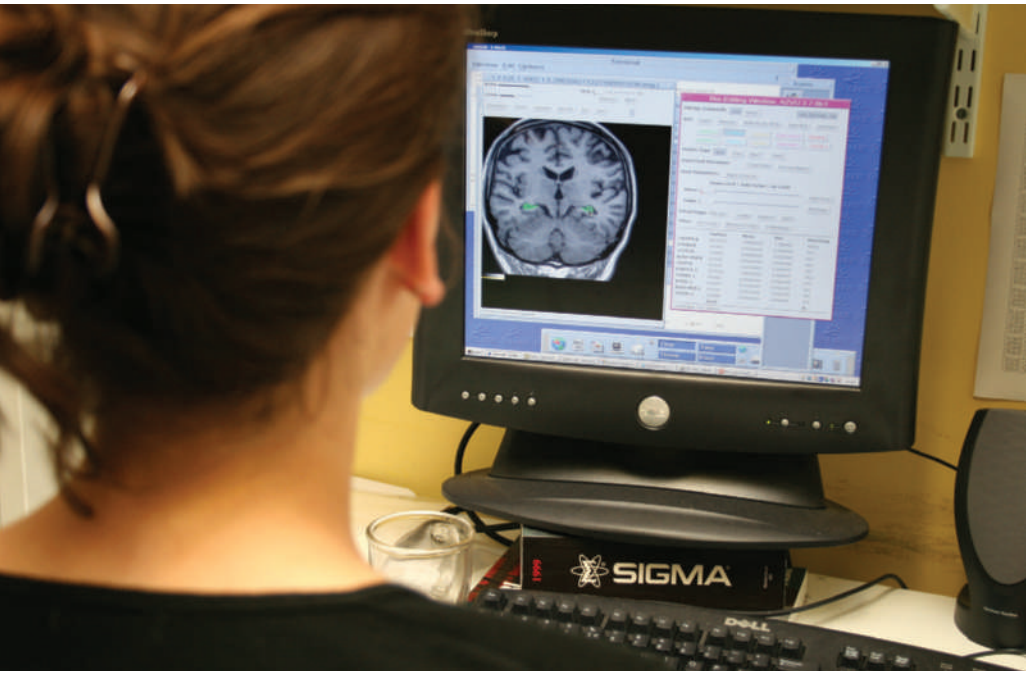
Some researchers were left wondering whether too much emphasis had been placed on removing the plaques. "It was only a small number of patients," says Karen Duff, an Alzheimer's researcher at Columbia University in New York, "but one has to question the amyloid-cascade theory, which dictates that clearance of the plaques should be correlated with disease improvement." She notes that tau tangles were still present in the autopsied brains. "This, for me, showed that far too much attention has been given to amyloid proteins, whereas other targets such as tau itself and other pathological features have been neglected."

Protein principles

Researchers have very strong reasons for believing that plaques and tangles are the cause of Alzheimer's symptoms, rather than just markers of it. Mutated versions of APP can cause an early-onset, hereditary form of the disease in which amyloid peptide has a greater propensity to accumulate in the brain years before symptoms develop[3]. Mice engineered to produce mutant APP that increases production of β-amyloid develop Alzheimer's-like symptoms that can be reversed by reducing tau expression[4].

And even if plaques and tangles do cause the symptoms, it does not mean that removing plaques and tangles will treat them. These structures may damage the brain in ways that cannot be repaired by simply dissolving the offending protein accumulations. "It could well be that there has been too much focus on early events of the amyloid cascade," admits Simon Lovestone, a psychiatrist and neuroscientist at London's Institute of Psychiatry. "But let's not throw the baby out with the bath water — there is just too much convincing preclinical research that supports the amyloid-cascade hypothesis."

Rather than rethinking the hypothesis, researchers are rethinking the trials. Aisen says that a major problem for the tramiprosate trial "was the huge variability in rates of decline among patients". In retrospect, he estimates that one-fifth of the patients recruited might not have had Alzheimer's at all, and so could not have been helped by therapy geared to disrupt the amyloid build-up typical of the disease. Only 50–70% of dementia cases are



The volume of the brain's hippocampus may serve as a marker of Alzheimer's progression.

E. PENDERS-GULJÉ

the Alzheimer's type; the rest have dementias that have other pathological indicators and neuropsychiatric profiles.

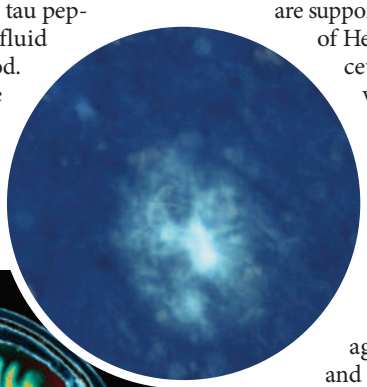
Aisen also says that 18 months may not have been long enough to pick up any clinical effect on the memory or behaviour of patients. "One of the huge difficulties of disease-modifying treatments in general is that they aim to slow the rate of progression of the disease and this is what trials have to measure," he says. With a disease that advances slowly, as Alzheimer's does, there may be only a small difference in how fast it progresses between patients in the treatment group and those in the control group. Furthermore, any difference may take a long time to become apparent. That means that trials need to recruit large numbers of patients and follow them for a long time — and that costs a lot of money.

Early start

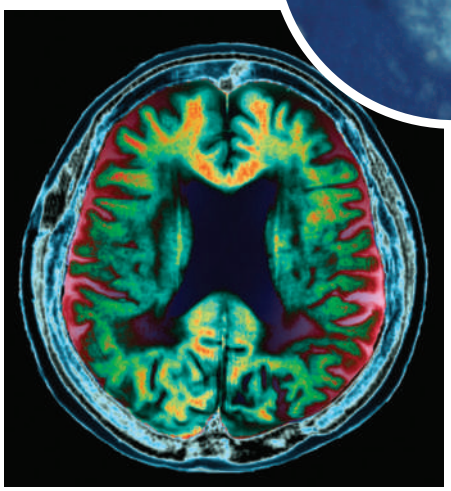
Yet another problem is that disease-modifying drugs are best deployed in the early stages of disease, and it is not currently possible to identify Alzheimer's before dementia sets in. Some of the promising preclinical results are based on mouse models of Alzheimer's that are treated before symptoms develop. By the time patients are recruited into trials, they are already taking such an array of therapies, including memory-enhancing drugs and antidepressants, that variability in both subject and control groups confounds analysis further.

The only way round these problems, everyone agrees, is to develop reliable biomarkers to help follow the progress of disease and treatment. In fact, the search has been on for several years.

There are two main biomarker approaches. One is to image the brain to determine its structure, its activity and how much amyloid it contains. The other involves measuring concentrations of amyloid and tau peptides in the cerebrospinal fluid and perhaps even in the blood. These techniques could be used to select those with true Alzheimer's for clinical trials. They might also help neurologists to follow the disease as it advances and monitor whether a test drug is having a biological effect (such as removing amyloid) even if the patient has no measurable change in mental state. The major challenge for researchers now is to validate the proposed biomarkers by showing that the detection methods are consistent and reliable in thousands of patients from different clinical centres and that the measurements correlate with biological or clinical endpoints.



J.-C. REVY/PHOTOLIBRARY.COM



ZEPHYR/SPL

Protein plaques (above) contribute to the destruction of brain tissue in Alzheimer's disease.

The community's perceived need for Alzheimer's biomarkers is reflected in the amount being spent to pursue them. Lovestone is principal investigator of one of the largest consortia of academics and pharmaceutical companies. Known as AddNeuroMed, this €8.6-million (US$11-million) five-year scheme is one of only two pilot projects supported by the European Union's €2-billion Innovative Medicines Initiative, planned to remove research bottlenecks in drug-development processes.

In the United States, Alzheimer's biomarkers are being sought as part of the Biomarkers Consortium launched in October 2006, and by the $60-million Alzheimer's Disease Neuroimaging Initiative. The neuroimaging initiative began in 2004 and is following the neuropsychiatric progression of 200 healthy elderly people, 400 patients with mild cognitive impairment and 200 with Alzheimer's disease using various test biomarkers. Both consortia are supported by the National Institutes of Health and by several pharmaceutical companies. These and various other initiatives under way in Japan, Australia and China, coordinate with each other carefully, sharing protocols and data and having regular conference calls, not least because those doing clinical trials are keen for drug regulatory agencies such as the US Food and Drug Administration and the European Medicines Agency to accept their validity.

Magnetic resonance imaging is already widely used to monitor key affected parts of the brain, such as the medial temporal lobe, as they shrink in nearly all patients with Alzheimer's disease, but in fewer than 30% of the normal ageing population. Changes in brain volume are believed to correlate well with memory performance. Positron-emission tomography (PET) imaging, used to measure the rate of metabolism in different brain areas, can also be used to detect Alzheimer's, although there is large variability between centres. A new PET technique developed by Steven DeKosky from

INST. PSYCHIATRY, KING'S COLLEGE LONDON

Simon Lovestone wonders whether too much focus has been placed on early parts of the amyloid cascade.

the University of Pittsburgh in Pennsylvania and his colleagues uses special radioactive ligands that bind to amyloid peptides[5] to visualize plaques directly, allowing them to be tracked as the disease progresses, and as treatment interferes with them.

Concentrations of amyloid peptides, particularly one called amyloid-61-42, are low in the cerebrospinal fluid of patients with Alzheimer's disease compared with healthy controls because the plaques are thought to suck them out of circulation, and concentrations of tau protein and phosphorylated tau are high. The concentrations and ratios of these three proteins can be useful indicators of whether a test drug is having a biological effect in the brain, and are already being used in some of the 30 or so trials that are now under way for treatments based on the amyloid-cascade hypothesis. These include a phase III study by Eli Lilly in Indianapolis, Indiana, of a β-secretase inhibitor and further studies of bapineuzumab.

None of these methods is as yet sensitive or specific enough to act as a biomarker for Alzheimer's on their own. In a paper published last year[6], 19 authors representing centres in Europe, North America and Japan, proposed that diagnosis of Alzheimer's should include at least one of the biomarkers under development in addition to new memory assessment techniques that can differentiate more precisely between different dementias. The authors hope that this approach might be applied routinely to diagnose a condition known as 'mild cognitive impairment', which often precedes full-blown Alzheimer's. If it were possible to identify patients with mild cognitive impairment who were highly likely to progress to Alzheimer's, they could form the basis of much more useful cohorts in dinical trials, says Dubois, who was lead author on the paper.

Eventually, and this is the trickiest part, a set of proven biomarkers may be able to act as surrogates for the disease, just as CD4-positive cells are used as a surrogate for HIV progression. When antivirals were first tested against HIV, their efficacy was measured based on the number of days a patient spent in hospital, or on time until death, says Michael Weiner, director of the Center for Imaging of Neurodegenerative Diseases at the Veteran Administration Medical Center campus of the University of California, San Francisco. Once CD4 was shown to correlate with clinical outcome, "they didn't have to wait for people to die any more — and this is what we want in Alzheimer's". Surrogate endpoints might allow clinical trials to be shorter and cheaper.

"Biomarkers are really changing our approach to studying and, hopefully, treating the disease."
— Dale Schenk

Surrogate scepticism

But regulatory agencies need a lot of persuasion to accept a new surrogate. In the past year, the validity of even very established biomarkers — such as blood cholesterol for cardiovascular health and blood glucose for diabetes — have been called into question when trials showed they were poor indicators of morbidity or mortality[7]. Marisa Papaluca-Amiti, a biomarkers expert at the European Medicines Agency, says that different centres still show a high level of variability for Alzheimer's biomarkers, especially those performing imaging. "We'll be persuaded when the scientific community gives systematic validation data that are reproducible everywhere," she says. "We are not going to see a surrogate marker for Alzheimer's anytime soon."

Most pharmaceutical companies seem to be unwavering in their pursuit of the amyloid target, with several beginning new trials and few giving up on those that are running. Even though Elan's early-phase bapineuzumab trials did not perform well across the board, post-hoc analysis indicated some improvement in a subgroup of patients who, counterintuitively, did not carry a risk gene for the disease called *APOE4*. That was more than enough to give the company confidence to carry on, says Dale Schenk, Elan's chief scientific officer. He says that the company has set up four phase III trials, involving some 4,000 patients, of which two trials will evaluate those who carry an *APOE4* mutation, and two will evaluate patients without it.

Schenk believes that biomarkers will also make all the difference. Some patients in the bapineuzumab studies are being evaluated with the new PET technique for imaging plaques in living brains. "Biomarkers are really changing our approach to studying and, hopefully, treating the disease," Schenk says. "These technologies are promising in their ability to show us whether amyloid is being cleared from the patient's brain while the treatment is going on."

Some serious attention is now being given to tau as a dinical target. One dinical trial under way by the Singapore-based company TauRx Therapeutics is testing whether methylthioninium, commonly known as methylene blue, can clear tangles. "I believe that eventually we'll find we need different therapeutic strategies for different stages of the disease, including symptomatic treatments as well as disease-targeted treatments," says Bob Holland, vice-president for neuroscience at AstraZeneca in Wilmington, Delaware. "But amyloid-targeted therapies are highly like to be a big part of it."

Those with Alzheimer's, and the people caring for them, face difficult tests every day. But the biggest tests of the amyloid hypothesis are yet to come. ■

Alison Abbott is *Nature*'s senior European correspondent.

1. Ferri, C. P. *et al. Lancet* **366,** 2112–2117 (2005).
2. Holmes, C. *et al. Lancet* **372,** 216–233 (2008).
3. Lippa, C. F., Nee, L. E., Mori, H. & St George-Hyslop, P. *Lancet* **352,** 1117–1118 (1998).
4. Roberson, E. D. *et al. Science* **316,** 750–754 (2007).
5. Ikonomovic, M. D. *et al. Brain* **131,** 1630–1645 (2008).
6. Dubois, B. *et al. Lancet Neurol.* **6,** 734–746 (2007).
7. Ledford, H. *Nature* **452,** 510–511 (2008).

CORRESPONDENCE

Italy: 'draconian' new law galvanizes demonstrations

SIR — Your News story 'New law threatens Italian research jobs' (*Nature* **455,** 840–841; 2008) and Editorial 'Cut-throat savings' (*Nature* **455,** 835–836; 2008) both describe the dire situation for research scientists in Italy. But that's only half the story.

The 4,000 researchers you cite belong only to research centres and include only researchers eligible to apply for permanent positions. Most temporary researchers in Italy are employed at universities under one of 20 or more different types of contract applicable to postgraduate and postdoctoral scholars. Neither the university administrators nor the ministry of education keeps track of this population, which the Italian network of temporary researchers, Rete Nazionale Ricercatori Precari, believes could consist of more than 60,000 people.

The present government's funding cuts to research and higher education are so drastic that Italy will fall short of the 2000 Lisbon Treaty requirements for government support of research and development across the European Union.

Although funds are available to hire 4,000 researchers, a law approved in August this year will withdraw this funding. Researchers already slated for jobs may lose the chance of a permanent position.

This law has galvanized demonstrations across Italy against the government's draconian measures. Students, technicians, permanent staff and temporary researchers have met to discuss the law and its consequences, with several faculties officially opposing the law. A petition has been set up at http://tinyurl.com/6otega.

Italian universities must change to become more responsive to international standards of research. But this governmental reform is not the change the universities need, assassinating as it will research and higher education in Italy.

Rete Nazionale Ricercatori Precari
www.ricercatoriprecari.org
e-mail: bologna@ricercatoriprecari.org
Author's name and address supplied

Italy: restoring axed funds is not enough to cure all ills

SIR — The Italian government's proposed diversion of funds from scientific research to help boost the economy — discussed in your Editorial 'Cut-throat savings' (*Nature* **455,** 835–836; 2008) — is lamentable. According to various independent surveys based on ISI Web of Science data, Italy's research productivity is already low and excellence is the exception rather than the rule.

One problem is stagnation. All Italy's universities are treated as equals and all academic staff are comparably paid. They jointly decide on all recruitment and promotion: if a position is available in an Italian zoology department, all the zoology professors will elect a commission that decides who shall take the post.

Another problem is that research resources are taken up by academics who only teach, rather than doing internationally recognized research. All have to teach 250 hours a year — fine if they are also doing research, but ridiculous if that is all they do. This may explain why it is hard to ascertain who just teaches and who (a numerous minority, but still a minority) does both teaching and research. It also explains why there is a marked resistance to the evaluation of scientific output, particularly from the unproductive areas. In the rare cases where evaluation is carried out, this is done entirely on impact factors, driving disciplines such as taxonomy towards extinction. (Paradoxically, Italian taxonomy is of international standard.)

The government may hope to create a new system in which teaching-only universities or departments, with no doctorate curriculum or research labs, are distinct from universities or departments that do both teaching and research. This idea has been discussed, but it would be risky and, knowing Italy, unlikely to be carried out effectively.

Restored funding alone would not be enough: we need to specify what we want for our university system. If the majority only teach, and their interests prevail in the allocation of resources, those who do both teaching and research will be the losers — along with Italy's research productivity.

Our brightest young postdocs leave for other countries because Italy holds no future for them. Money would be better spent on educating them if we then gave them the chance to work productively here.

Ferdinando Boero Dipartimento di Scienze e Tecnologie Biologiche e Ambientali, Universitá del Salento, 73100 Lecce, Italy
e-mail: boero@unile.it

Peer-reviewed parrot studies speak for themselves, as he did

SIR — I take exception to Clive Wynne's Review of my book *Alex & Me: How a Scientist and a Parrot Discovered a Hidden World of Animal Intelligence — and Formed a Deep Bond in the Process* (*Nature* **455,** 864–865; 2008). As a memoir for a general audience, it does not contain in-depth technical detail. But Wynne's questioning of the underlying science is answered by my publications in the peer-reviewed literature — one of which won the Frank A. Beach comparative psychology award (I. M. Pepperberg and J. D. Gordon *J. Comp. Psychol.* **119,** 197–209; 2005).

Detailed controls were included against inadvertent cueing during 30 years' study of the cognitive and communicative abilities of Alex, the African grey parrot; during testing, I never knew what Alex would be asked to identify in any trial. Because tests covered several different topics, neither Alex nor evaluators could target a particular subset of labels. Individuals never tested Alex on the labels they taught.

Wynne writes that the human ear is easily tricked, suggesting that Alex's responses were often not actual human labels. But incoherent responses were discarded as erroneous, and inter-observer reliability trials (with individuals from a different university listening to a video sound track) gave agreement above 90%. Sonagraphic analysis showed that Alex's vocal patterns closely matched his human trainers' and were not "jumbled speech sounds".

Alex's idiosyncratic labels are overlooked in Wynne's review. The bird would use these irrespective of how a colleague or stranger might cue him otherwise. For example, toy cars and squares were always classed respectively as "truck" and "four-corner".

Replication failure was not a problem. Published data on my younger subjects detail a level of replication. Birds were presented with a barrage of training techniques, mostly different from those used with Alex, to determine which aspects of Alex's training engendered his success. These birds, with a quarter of Alex's exposure to the single, successful method, could not, by definition, achieve as much as Alex during the same time frame.

Wynne has no basis for implying that my methods might be flawed — other than a possible inherent scepticism, which, ironically, is a common bias discussed at length in the book.

Irene M. Pepperberg Vision Lab, Department of Psychology, Harvard University, Cambridge, Massachusetts 02138, USA
e-mail: impepper@wjh.harvard.edu

Contributions may be sent to correspondence@nature.com, and will be edited.

COMMENTARY

The innovative brain

'Hot' decision-making, involving the evaluation of reward and punishment, is essential to the entrepreneurial process and may be possible to teach, argue **Barbara Sahakian** and her coauthors.

The creativity and innovation shown by entrepreneurs is a crucial source of employment and economic growth. Although there is potential for considerable profit in making the decision to 'go it alone', rather than working for an existing company, these individuals accept considerable risk. Many new ventures fail, with possible consequences for finances, reputation, self-esteem and even family stability.

Psychological and biomedical research has traditionally considered risk-taking an abnormal expression of behaviour, as exemplified by its association with substance abuse and bipolar mania[1]. We propose that entrepreneurs represent an example of highly adaptive risk-taking behaviour, with positive functional outcomes in the context of stressful economic decision-making. This 'functional impulsivity' may have evolutionary value as a means of seizing opportunities in a rapidly changing environment.

Here, we present preliminary neurocognitive data from matched groups of entrepreneurs and managerial controls, which highlight the potential of this approach to examine advantageous risk-taking. Our 16-member entrepreneurial group from the 'Silicon Fen'[2] cluster of high-tech companies around Cambridge, UK, consists of individuals who have founded at least two companies. Seventeen non-founding managers lacking the experience of venture creation were recruited as a comparison group who were matched to the entrepreneurs for demographic characteristics including age and intelligence, estimated with the National Adult Reading Test.

All the participants completed a computerized neurocognitive assessment that measured various aspects of their decision-making abilities. Classical models of decision-making from psychology and economics focus on the computation of expected values for each decision option, based on the relative gains and losses of the associated outcomes. But these models fail to accommodate emotional states when computing decisions, as well as emotional reactions to gains and losses, which can greatly alter our reasoning processes[3]. A growing body of evidence from functional neuroimaging and studies of patients with brain injuries points to the involvement of distinct processes in risky and risk-free decision-making. Referred to as 'hot' and 'cold' processes, these appear to be localized to distinct regions of the brain's frontal lobes[4].

All in the mind

We measured cold processes using the one-touch Tower of London task (Fig. 1), which requires the participant to mentally rearrange an array of coloured balls. They are required to select the minimum number of moves to complete the problem from a display across the foot of the screen. The problem difficulty varies between one and six moves, with harder problems requiring greater planning, working memory and decision-making skills.

Performance on this task reflects cold processes that are dependent upon the dorsolateral prefrontal cortex[5]. These are emotionally neutral decisions with no weighing of rewards and punishments. Cold processes govern real-life decisions such as planning the details involved in opening of a new company branch.

Risky or hot decisions, on the other hand, involve an evaluation of reward and punishment, with emotional significance attached to both outcomes. For example, the decision between financing one of several potentially excellent but risky business opportunities is a hot decision; it is difficult to ignore emotions. Hot processes are dependent on the medial and orbital sectors of the prefrontal cortex[6], which lie above the orbit of the eyes and connect with limbic emotion-related areas such as the amygdala.

Hot processes were measured using the Cambridge Gamble Task[7] (Fig. 2). A participant is presented with an array of ten boxes coloured either red or blue. A token is hidden randomly in one of the boxes, and the participant must decide whether the token is in a red box, or a blue box. The ratio of red-to-blue boxes is manipulated to present different levels of uncertainty of winning.

Participants then choose a bet to stake on the colour decision. A displayed points total reminds participants how well they are doing. Thus, on each trial, subjects must balance the chance of gaining more points with a high bet against the chance of being incorrect and losing their wager.

Our groups of entrepreneurs and managers showed comparable performance on the Tower of London test of cold processes, with no differences in the number of solutions solved at the first attempt. On the Cambridge Gamble Task, both groups were able to make high-quality decisions, selecting the majority colour at least 95% of the time. The remaining 5% is likely to be accounted for by 'gambler's fallacy' in which subjects try to second guess the computer by choosing the less likely colour. However, when subjects were introduced to the hot components of the task, differences were observed. We found

B. MELLOR

that entrepreneurs behaved in a significantly riskier way, betting a greater percentage of their accrued points (63%) than their managerial counterparts (51%). Both groups adjusted their wagering according to the likelihood of success (dependent on the ratio of red-to-blue boxes). The only performance difference was the amount that was bet.

Interestingly, this risk-taking performance in the entrepreneurs was accompanied by elevated scores on personality impulsiveness measures and superior cognitive-flexibility performance. We conclude that entrepreneurs and managers do equally well when asked to perform cold decision-making tasks, but differences emerge in the context of risky or emotional decisions.

The pattern of performance seen on the gambling task in entrepreneurs reflects a behavioural index of risk-seeking or risk tolerance. Greater rewards (as well as greater losses) are available for those who bet more. A large normative study of the Cambridge Gamble Task found that the betting behaviour decreases with age across the lifespan[8]. Comparing our recent data against this normative sample shows that the entrepreneurs (mean age 51) are comparable to the young adults aged 17–27, whereas the managers (mean age 50.5) resemble their age-appropriate group (Fig. 3). These cognitive processes are intimately linked to brain neurochemistry, particularly to the neurotransmitter dopamine. Using single-dose psychostimulants to manipulate dopamine levels, we have seen modulation of risky decision-making on this task[9]. Therefore, it might be possible to enhance entrepreneurship pharmacologically.

The dark side?

Entrepreneurs demonstrate creativity through the development of positive financial opportunities for society as well as themselves. Although we know little about the psychopathological 'dark side' of entrepreneurship[10], we would argue that the effects seen with our gambling task reflect functional impulsivity[11], a subtype of impulsiveness that may enable impulsive individuals to capitalize on environmental niches. This functional impulsiveness of entrepreneurs combined with enhanced cognitive flexibility is a winning combination.

Our results raise important questions: if these impulsive risk-taking traits can be beneficial, can they be taught or otherwise imparted to the potential entrepreneur? What does it take to make an entrepreneur — is it an inherited, inbuilt characteristic, or is it acquired, and if so, can it be acquired by anyone? What specific mechanisms lead to the risk tolerance that we have seen in our study? Are entrepreneurs conscious risk seekers, or do they place less weight on risk in their decision-making?

There are courses that teach the 'know-how' of entrepreneurship and, within this, teach how to mitigate risk through extensive business planning and market research. Yet given this new evidence, courses teaching the opposite — risk-tolerance in both behaviour and personality — may also be desirable. Indeed, this may be easier to teach in young adults, in whom higher risk-taking is likely and age-appropriate. We know that perception of risk is particularly sensitive to framing effects; therefore, training potential entrepreneurs to reframe their decisions may encourage greater entrepreneurial activity.

Similarly, negative perception of risk, such as that involved in starting a new venture, could be mitigated by social/cultural norms. One of the beneficial effects of entrepreneurial clusters in regions such as Silicon Fen may be that the increased networking and contact amongst the entrepreneurs works to create a culture that normalizes a more risk-tolerant type of decision-making. One might argue that business skills can be acquired relatively easily from multiple sources. But if one is to boost entrepreneurial activity, and if risk tolerance and the ability to cope with ambiguity are barriers to achieving a cultural shift in society, then more focus is needed on understanding how risk and emotion are integrated into decision-making. ■

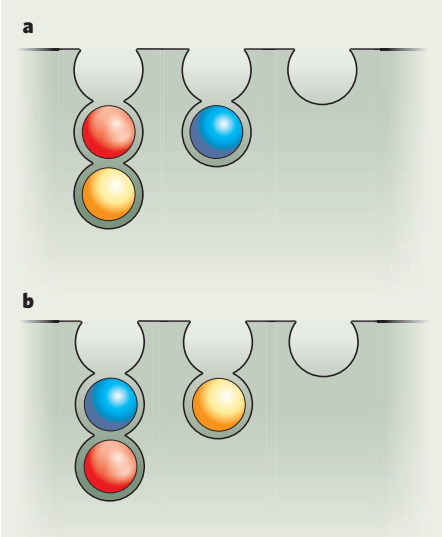


Figure 1 | How many moves? The Tower of London task measures 'cold' decision making, asking subjects to calculate the minimum number of moves required to get the array in **b** to match that in **a**. This is a five-move problem.

Figure 2 | Place your bets. To evaluate risk-taking, 'hot' processes in the Cambridge Gamble Task (part of the Cambridge Neuropsychological Test Automated Battery), a randomly hidden token is in a blue or a red box. Subjects choose a colour and a stake to bet.

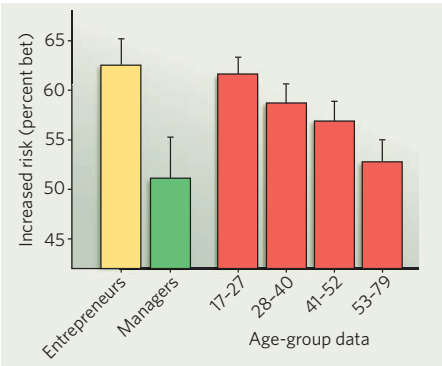


Figure 3 | Risk taking. Entrepreneurs bet a higher percentage of their points than managers. This measure decreases with age in a large normative sample. Managers resemble their age-appropriate counterparts, whereas entrepreneurs show a risk-tolerance similar to the youngest group.

Andrew Lawrence, **Luke Clark**, **Jamie Nicole Labuzetta** and **Barbara Sahakian** are at the MRC-Wellcome Trust Behavioural and Clinical Neuroscience Institute, in the Departments of Experimental Psychology and Psychiatry, University of Cambridge, Cambridge CB2 3EB, UK.
e-mail: lc260@cam.ac.uk.
Shai Vyakarnum is at the Centre for Entrepreneurial Learning, Judge Business School, University of Cambridge, Cambridge CB2 1QA, UK.

1. Paulus, M. P. *Science* **318,** 602–606 (2007).
2. Myint, Y. M., Vyakarnum, S. & New, M. J. *Strategic Change* **14,** 165–177 (2005).
3. Bechara, A., Damasio, H., Tranel, D. & Damasio, A. R. *Science* **275,** 1293–1295 (1997).
4. Roiser, J. P., Rubinsztein, J. S. & Sahakian, B. J. *Psychiatry* **2,** 43–47 (2003).
5. Baker, S. C. *et al. Neuropsychologia* **34,** 515–526 (1996).
6. Rogers, R. D. *et al. J. Neurosci.* **19,** 9029–9038 (1999).
7. Clark, L. *et al. Brain* **131,** 1311–1322 (2008).
8. Deakin, J., Aitken, M., Robbins, T. & Sahakian, B. J. *J. Int. Neuropsychol. Soc.* **10,** 590–598 (2004).
9. DeVito, E. E. *et al. Biol. Psychiatry* **64,** 636–639 (2008).
10. Hisrich, R., Langan-Fox, J. & Grant, S. *Am. Psychol.* **62,** 575–589 (2007).
11. Dickman, S. J. *J. Pers. Soc. Psychol.* **58,** 95–102 (1990).

Competing financial interests: declared (see online article for details).

To comment on this article and others in our innovation series, visit http://tinyurl.com/5uolx2.

BOOKS & ARTS

Meditating on consciousness

Two books exploring the relationship between Buddhism and science reveal surprising synergies — and hint that insights into the brain may come from studying the religion's practices, finds **Michael Bond**.

Buddhism and Science: A Guide for the Perplexed
by Donald S. Lopez
University of Chicago Press: 2008, 280 pp. $25, £13

Mind and Life: Discussions with the Dalai Lama on the Nature of Reality
by Pier Luigi Luisi, with Zara Houshmand
Columbia University Press: 2008. 232 pp. $24.95, £14.95

In the troubled relationship between science and religion, Buddhism represents something of a singularity, in which the usual rules do not apply. Sharing quests for the big truths about the Universe and the human condition, science and Buddhism seem strangely compatible. At a fundamental level they are not quite aligned, as both these books make clear. But they can talk to each other without the whiff of intellectual or spiritual insult that haunts scientific engagement with other faiths.

The disciplines are compatible for two reasons. First, to a large degree, Buddhism is a study in human development. Unencumbered by a creator deity, it embraces empirical investigation rather than blind faith. Thus it sings from the same hymn-sheet as science. Second, it has in one of its figureheads an energetic champion of science. The current Dalai Lama, spiritual leader of Tibetans, has met regularly with many prominent researchers during the past three decades. He has even written his own book on the interaction between science and Buddhism (*The Universe in a Single Atom*; Little, Brown; 2006). His motivation is clear from the prologue of that book, which Donald Lopez cites in his latest work *Buddhism and Science*: for the alleviation of human suffering, we need both science and spirituality.

Not all scientists are convinced by the need for this dialogue, and some are profoundly suspicious. When the Society for Neuroscience invited the Dalai Lama to give the inaugural lecture at its 2005 annual meeting in Washington DC, more than 500 researchers signed a petition objecting. They claimed it was inappropriate for a religious leader to address a scientific meeting, and that the study of empathy and compassion and how meditation affects brain activity, on which the Dalai Lama had been invited to speak, was too flaky to be taken seriously (see *Nature* **436,** 1071; 2005).

J. BOURG/REUTERS/CORBIS

The Dalai Lama is keen for Buddhists and scientists to interact.

It is unclear if the petitioners' motives were also political — many of the signatories were of Chinese origin — but their concerns over scientific integrity are hard to justify. Science has nothing to fear from a religious leader who has declared that should science prove some Buddhist concept wrong, "then Buddhism will have to change". Lopez, whose book is more a history of the discourse between Buddhism and science than an examination of how the two inform each other, makes much of the Dalai Lama's doctrinal flexibility. He suggests that this stems partly from the Tibetan leader's desire to show that his religion is not the primitive superstition that many nineteenth-century European writers — and modern Chinese communists — have described. Perhaps so, but it must also derive from the Buddhist desire to know reality and not hide behind false assumptions about the world or our own nature.

During the past two decades, the Dalai Lama has directed his enthusiasm for modern scientific knowledge largely through the Mind and Life Institute, a non-profit organization based in Boulder, Colorado, which promotes dialogue and research partnerships between science and Buddhism. The institute holds regular private conferences at which scientists and Buddhists explore their respective views of the world. Pier Luigi Luisi's book *Mind and Life* is a first-hand account of one of these meetings, held over a week in 2002 in Dharamsala, India, where the Dalai Lama lives. The discussion, entitled 'the nature of matter, the nature of life', covered everything from particle physics to the evolution and nature of consciousness. One glance at the guest list, which included the Nobel prize-winning physicist Steven Chu and the biologist Eric Lander, should dispel any doubts about how seriously the scientific community takes such collaborations.

Luisi, a biologist himself, does a fine job of capturing the ebb and flow of debate and the delicate dynamics of cross-cultural interaction. Some of the dialogue is riveting, in particular when the participants come up against each other's characterization of reality. When Chu describes how physicists measure the properties of simple particles as if the particles were independent entities, there is a buzz of scepticism from the monks in the Tibetan benches; his depiction conflicts with the Buddhist idea that it makes sense to consider something only in terms of the parts or properties that make it up. What the monks want to know is whether, say, the electron "is really something out there that has those properties?" Chu responds: "we don't actually ask that question!"

What does science get from such an exchange? At the least, it encourages alternative ways of thinking about reality. Yet there is one area in which dialogue between Buddhists and scientists could lead to genuine advances in understanding: the study of consciousness. Here, Buddhism offers something

that science lacks — a tried and tested way of observing and altering, through careful attention to meditation, the subjective workings of the mind. Neuroscientists can show how the practices used by meditators result in physiological changes in the brain, but as several of the Dharamsala conference participants attest in the book, neuroscience does not yet have the tools to explore the various states of consciousness they experience. Buddhism seems to offer a kind of science of introspection.

As a research exercise, the East–West discourse on consciousness sounds harmonious, but at a deep level, it is anything but. Both Luisi and Lopez identify this as an area of great conceptual divergence. Whereas cognitive science's best guess is that consciousness is an emergent property of neuronal organization, Buddhists see it at some pure subtle level as not contingent on matter at all, but deriving instead from "a previous continuum of consciousness" — the Dalai Lama's words — that transcends death and has neither beginning nor end. That is hard to test. Furthermore, it seems impossible for anyone to grasp such Buddhist notions of consciousness without experiencing them, because there is no way yet of quantifying them — and that means years of meditation. As Chu says in *Mind and Life*: "It's like a physicist explaining electromagnetic waves to someone who doesn't know mathematics."

"Buddhism seems to offer a kind of science of introspection."

Despite this, Luisi's depiction leaves you with the impression that if cognitive scientists and Buddhists can learn a little more of each other's language, they might be on to something. Consciousness aside, the book is stimulating whatever your field or expertise, because it is likely to offer a way of looking at the world that you had not tried. Readers will also get a short, sharp primer into the nature of fundamental particles and the origins of life, and the philosophy behind Buddhist ethics. Lopez's book, by contrast, is more likely to excite those seeking an in-depth analysis of Buddhism's historical relationship with science. He purports to offer a 'guide for the perplexed' — presumably those who are perplexed that the two disciplines should be compatible at all. His scholarly treatment should provide succour, yet he gets off to a sticky start by pondering what it means to group the words 'Buddhism' and 'science' in the same phrase, concluding that it depends on "what one means by Buddhism, what one means by Science, and, not insignificantly, what one means by and". It would take more than a week in Dharamsala to unpack that one. ■

Michael Bond is a writer based in London.
e-mail: michael.bond@btinternet.com

Seeing is behaving

Living Control Systems III: The Fact of Control
by William T. Powers
Benchmark Publications: 2008. 250 pp.

The field of behavioural science, combining psychology, sociology and neuroscience, has diversified over the past century such that there is a desperate need for an integrative theory. William T. Powers, medical physicist and engineer, proposes that 'control' is the unifying process. *Living Control Systems III* is the latest in an influential but contentious series of works in which Powers presents his theory of perceptual control and illustrates its explanatory power.

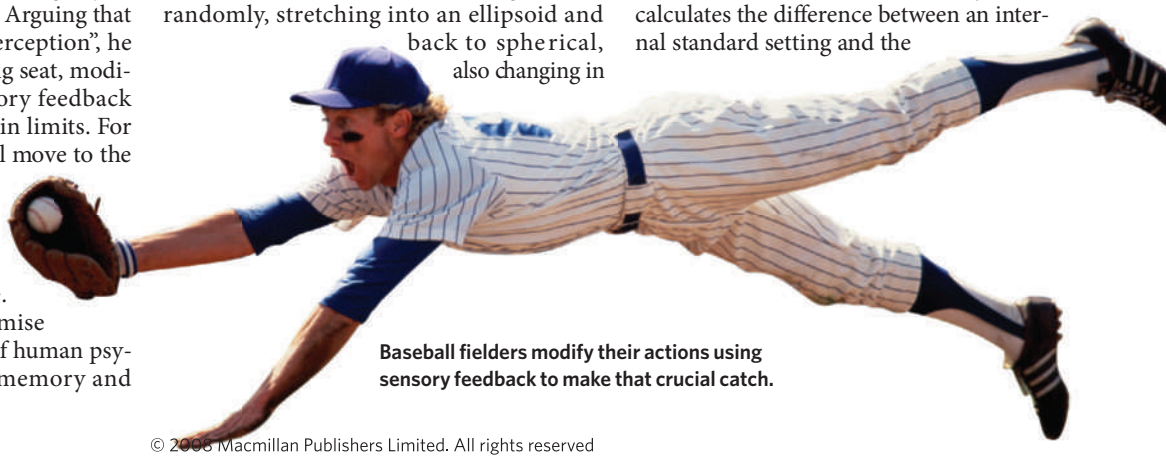
Powers innovatively applies the principles of 'control engineering' — as used in devices such as amplifiers and cruise-control systems — to the management of perceptual variables, such as our ability to track a moving object or maintain a sense of comfort. Arguing that "behaviour is the control of perception", he puts the organism in the driving seat, modifying its action through sensory feedback to control its experiences within limits. For example, a baseball fielder will move to the optimum position to catch a ball by maintaining an image of the ball moving at a constant velocity, relative to the playing field, on the retina of his eye. Powers builds on his basic premise to account for the complexity of human psychology, including learning, memory and skill acquisition, through the operation of multilevel hierarchies of control systems.

Recent developments within computing allow Powers to bring his theory to life. The book is organized around an accompanying compact disc containing 13 computer simulations of perceptual control. They span from the tracking of moving targets and the simulation of balance, to three-dimensional models of arm coordination and the emergence of crowd behaviour. Powers uses a combination of common sense reasoning, philosophical argument and mathematical models to make his case. Throughout, the style is engaging yet authoritative.

"Behaviour is the control of perception and not a learned response."

After setting the historical context, Powers turns to the interactive computer demonstrations. A sphere on the screen changes shape randomly, stretching into an ellipsoid and back to spherical, also changing in placement and projected angle. The viewer is asked to control the shape, location or the orientation of the figure by moving the computer mouse, the movement of which is recorded and plotted. The graph makes it clear that the user's behaviour exactly opposes the random changes in the sphere's perceived properties. The more that the computer tries to squash its shape, the more the user moves the mouse to return it to a spherical form. According to Powers, this shows that behaviour is controlling perception and is not a learned response to the environment. By contrast, most traditional theories explain behaviour as the result of a learned association between a stimulus and a response; for example, as a rat might be conditioned to press a lever for a food reward. Such theories, Powers proposes, do not provide an adequate explanation of behaviour.

The basic unit of control is the negative feedback loop. This is a mathematical system that calculates the difference between an internal standard setting and the



D. MADISON/GETTY IMAGES

Baseball fielders modify their actions using sensory feedback to make that crucial catch.

perception of that same quantity, and uses this difference to drive the behaviour that reduces the discrepancy. A newly hatched duckling, for instance, would have a setting for how near it needs to be to its parent. When the parent is further away, the difference drives the duckling's behaviour, perhaps making it run, to restore the error to zero. According to Powers, this continuous process of perception, comparison and action goes on in many systems simultaneously within any living animal.

Because these adjustments involve many systems, conflict between them is common. An animal might need to regulate its proximity to safety and an optimum level of novelty or stimulation in its environment, for instance. But going closer to novel environments can take the animal further from the safety of its lair. Powers proposes that two systems in conflict have their set points determined by a higher level system. In order for the conflict to cease, these settings need to change. He suggests that a random, trial-and-error change in these settings takes place until they no longer cause conflict — a process known as reorganization. An example is the mechanism by which the bacterium *Escherichia coli* heads towards increasing levels of certain chemicals in a surrounding fluid, known as chemotaxis.

Living Control Systems III is an interactive tutorial rather than a coffee-table introduction, and occasionally, but not crucially, demands mathematical knowledge. Increasingly, academic textbooks are becoming more electronic and interactive. But this book's dual format is inconvenient: it would be easier to absorb the material if it was fully electronic. The computer demonstrations are powerful, and the novice reader may get a good grasp of the theory simply by skipping to the penultimate chapter to experiment with a range of novel displays of perceptual control.

Perceptual control theory has spawned a diverse range of applications within psychotherapy, education and artificial intelligence over the past 50 years. It has been lauded by Carl Rogers, the creator of person-centred counselling; Thomas Kuhn, the philosopher of science; and self-regulation theorists such as Charles Carver. Yet it has still to be accepted within psychology, perhaps because it requires psychologists to accept the same level of precision as in physics, biology and engineering. Whether this animated third volume will prompt that leap is an open question, but it succeeds in being a sophisticated yet colourful demonstration of a contentious theory. ■

Warren Mansell is a senior lecturer in the School of Psychological Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, UK.
e-mail: warren.mansell@manchester.ac.uk

Change begins at home

Climate Change: The Threat to Life and a New Energy Future
American Museum of Natural History, New York City
Until 16 August 2009

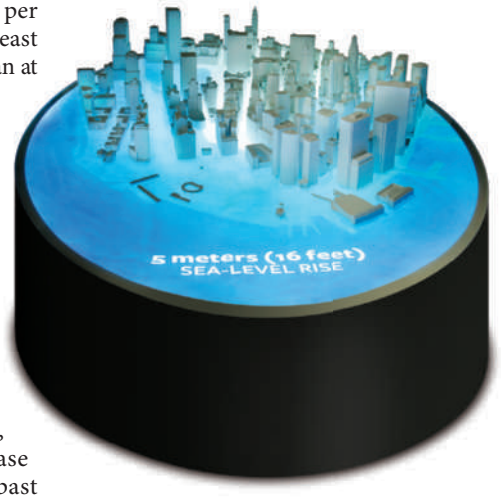
A thin red line snakes along a wall, rising from below knee level to a height of three metres. The long glowing tube plots global carbon dioxide concentrations from the year 1600 — when the world's population was less than 600 million, goods were transported by wheelbarrow and horse-drawn vehicles, and the estimated atmospheric CO_2 concentration was 274 parts per million — until the year 2000. By that time, the world's population had reached more than 6 billion, aeroplanes were plying the skies, and atmospheric CO_2 had topped 369 parts per million. Today, that concentration is at more than 385 parts per million, a level not seen on Earth for at least 800,000 years. Earth is also hotter now than at any time in the past four centuries.

With this grand flourish, the American Museum of Natural History in New York opens its new climate-change exhibition, which runs until August 2009. With displays, interactive videos and dioramas, *Climate Change* clearly and comprehensively provides the facts on global warming. It dispels doubts about the reality of the crisis — of which there are still plenty among the American public.

The exhibition stresses that humans are responsible for the recent rise in CO_2, and that this is linked to a 0.8 °C increase in average global temperatures in the past 100 years. A video installation explains how greenhouse gases such as CO_2 trap heat, and shows how global temperatures and atmospheric CO_2 are rising in tandem. A small model of a coal-powered steam engine, invented by Thomas Newcomen in 1712, ties the rise in CO_2 to the start of the Industrial Revolution. A giant, jagged, one-tonne slab of coal reminds us that this dirty fuel, the burning of which generates nearly twice as much CO_2 as natural gas, still provides 40% of the world's electrical energy.

The effects of climate change are already obvious: longer, drier droughts in the southwestern United States and the North African Sahel, spring arriving earlier, warmer seas that kill coral — a bleached and lifeless model reef sits in one corner — and shrinking Arctic sea ice. This last consequence has forced polar bears to spend more time on land, and a diorama depicts a stuffed polar bear foraging for food on a pile of snow-sprinkled rubbish — plastic bottles, a discarded television set, a doorless refrigerator. If that sad scene spurs its viewers to change their lifestyles, then the curators suggest plenty of ways to do so. Touch-screen videos invite visitors to click on their current car and see how much carbon it emits. Drive 99 kilometres in a sports utility vehicle every day, and in a year it will cough out more than 14,000 kilograms of CO_2 — enough to fill up two-and-a-half hot-air balloons. The same daily distance in a hybrid-fuel economy car will pump out one-third of that amount.

Many other options are offered for combating climate change: ditch the car and take public transport; avoid bottled water; switch to energy-saving light bulbs; line-dry clothes; fly less; install low-flow showers; pay for wind power; plant trees and shrubs. The exhibition also considers broader solutions such as carbon capture and storage; wind power — already supplying Denmark with more than 20% of its electricity; hydropower; geothermal and solar energy. More than 25,000 solar panels are now sold every year in Kenya, where they are used to power television sets and recharge mobile-phone batteries.



D. FINNIN/AMNH

Rising seas may threaten cities such as New York.

A final wall display invites visitors to write or draw their own solutions on postcards. "We need nuclear energy" is followed by "Windmills and solar power, not nuclear — until we have a plan for nuclear waste". Another exhorts us to "Promote science education — Study Nature!". But my favourite card offers one simple word: "Vote". ■

Josie Glausiusz is a journalist based in New York.
e-mail: josiegz@earthlink.net

S. BAKER/GETTY IMAGES

The two cultures of Wall Street

Physicists on Wall Street and Other Essays on Science and Society
by Jeremy Bernstein
Springer: 2008. 195 pp. £19.99

Released in the middle of the greatest financial crisis in the United States for 70 years, the timing of this book could hardly be better. *Physicists on Wall Street and Other Essays on Science and Society* presents a diverse collection of writings by essayist Jeremy Bernstein, starting with the world of finance.

Physicists are famous for building black boxes that can trade on financial markets with little or no human intervention, and for developing the arcane formulas that set the prices of strange financial instruments such as mortgage-backed securities derivatives. The title essay follows the story of Emanuel Derman, a particle physicist who moved from academia to become a quantitative analyst, or 'quant', on New York's Wall Street, and who wrote an excellent memoir called *My Life as a Quant* (Wiley, 2004). Derman is a good subject because he is articulate and his experience is typical of that of many physicists who moved to Wall Street during the 1980s and 1990s. This curious migration came about as a result of the coincidence of a bad job market in physics, the proliferation of quantitative mathematical methods in finance, and the contrast between salaries on Wall Street and those in academia. This is interesting from an anthropological viewpoint because the two cultures are completely different. Moving from a physics department to an investment bank is like moving from a suburb in California to a tribe in the Amazon. Except that this particular tribe is enormously rich and controls the financial well-being of the world.

Although he is a regular contributor to *The New Yorker*, Bernstein's writing can be hard to follow. He strives for a conversational style, but often comes across like someone speaking quickly into a tape recorder. Abrupt segues and free associations are the norm. For example, part way through the title essay he suddenly switches from Derman's story to a discussion of Long Term Capital Management (LTCM), the infamous US hedge fund that nearly brought down the global financial system in 1998.

Although LTCM also employed quantitative analysts, the hedge fund was run by economists rather than physicists, including two winners of the Nobel Prize in Economics. Bernstein skims over this point, perhaps because he does not want to draw attention to it. This confusion affected the publisher too, with the book's dust jacket describing "how some physicists who developed interesting economic models based on derivatives and hedge funds almost led the country into bankruptcy".

As a physicist who spent eight years running a quantitative hedge fund, I must protest. Economists, not physicists, were responsible for the financial near-meltdown that was triggered by the irresponsible practices of LTCM. Quantitative hedge funds tend to divide into those run by economists and those run by scientists from other disciplines, such as physics, maths or computer science. For example, perhaps the most successful hedge fund, the New York-based firm Renaissance, has some 70 researchers, none of whom is an economist. By contrast, LTCM's only connection to physics is that Derman once applied for a job there and did not get it.

This distinction is not just a matter of professional pride and disciplinary boundaries. Bernstein misses an important point: economists and physicists traditionally approach the problem of risk control in different ways. Risk control is the art of determining the likelihood of large and unexpected price changes happening in the future. It is well known that extremely large changes, and financial crashes in particular, are more frequent than would be expected from a 'normal' statistical distribution. Physicists tend to favour a 'power law' mathematical description to model the heavy tails of these distributions, giving a pessimistic view of the likelihood of large price movements. By contrast, the economists who led LTCM spoke about price movements in terms of standard deviations, a terminology that is only relevant for normal distributions. This demonstrates that they were not thinking about the problem in the right way.

As a result, LTCM dramatically underestimated the likelihood of large adverse price moves. They took on enormous levels of borrowed capital, known as leverage — for every $1 they had under management, they reputedly owned $40 of bonds. When things were good this resulted in enormous profits, but when things went bad they lost all of the money under management. LTCM controlled a lot of assets, and their mistake might have brought down the global bond markets if Alan Greenspan, then the Chairman of the US Federal Reserve, had not forced a consortium of large banks who were their clients to bail them out.

Unfortunately, dramatic underestimates of risk are still prevalent on Wall Street. The recent subprime mortgage crisis is a good example. Without a proper understanding of extreme risks, people follow cycles in which leverage slowly creeps up during good times, until a big loss happens and leverage requirements drop to levels that are too low, exacerbating or even causing a financial crash. Excessively low levels of leverage, such as those we are experiencing now, make recovery even harder. Far-thinking economists, such as John Geanakoplos of Yale University in New Haven, Connecticut, are beginning to understand this process and suggest possible solutions. An answer in the meantime would be to simply impose sensible risk requirements. Wall Street should follow the conservative approach to risk control that arises from properly modelling risks as power laws.

'Physicists on Wall Street' is one of many essays in this volume, which spans a wide range of topics. These include the German atomic-bomb project, attempts to use nuclear weapons to power space craft, the Warsaw ghetto, deciphering early Etruscan writing, interpretations of quantum physics in the novels of Michel Houellebecq, a proof of an important conjecture in algebraic topology, and scientists gambling in Las Vegas. Each essay focuses on one individual or on a few people, often involving Bernstein's own ample and varied personal experience — in these cases, the essays work well. Unfortunately, in many cases, he is not involved and has not done enough background research. His definition of leverage in the first essay is wrong and his diatribe about the dangers of hedge funds is naive. Many of the essays are essentially reports about books that the reader could digest on his or her own.

Bernstein's book contains a lot of interesting science, and many entertaining stories about the people responsible. It is a pity that, at times, the barrier to understanding the science within the essays is larger than it should be. ■

J. Doyne Farmer is a professor at the Santa Fe Institute, Santa Fe, New Mexico 87501, USA.
e-mail: jdf@santafe.edu

Mapping the cerebral globe

A detailed sketch by architect Christopher Wren reveals his surprising contribution to neuroscience, explain **Martin Kemp** and **Nathan Flis**.

It would have been better for Christopher Wren's reputation as "that rare and early prodigy of universal science", in the words of his friend the diarist John Evelyn, if he had not subsequently proved to be an architect of genius. The towering presence of his St Paul's cathedral and magnificently varied London churches, erected in the wake of the 1666 fire, overshadow his achievements in a wide range of sciences that include mathematics, astronomy, microscopy, instrument design, natural philosophy, natural history, anatomy and physiology.

During his years at the University of Oxford, UK, the precocious Wren was part of the circle of thinkers, investigators and experimentalists who were to lay the foundations of the Royal Society. They intended that no part of the natural world or the cosmos should escape systematic analysis, with a precision that can broadly be called mathematical. It is obvious how astronomical studies should fit within this description, but less clear how the results of anatomical dissection could match that ideal. Wren's illustrations for Thomas Willis's 1664 book *Cerebri Anatome* (*The Anatomy of the Brain*) demonstrate how this could be accomplished (see example, pictured right).

A good sense of how the seventeenth-century concept of 'curiosity' fitted with the precise sciences is provided by Bishop John Wilkins, who chaired the founding meeting of the Royal Society. Evelyn, that most assiduous chronicler of his contemporaries, described how Wilkins had constructed "transparent apiaries, which he had built like castles and palaces, and so order'd them one upon another so as to take the honey without destroying the bees".

Wilkins also exhibited a "wide variety of shadows, dyals, perspectives, and many other artificial, mathematical, and magical curiosities". Most of the devices were contrived by Wilkins himself, but others were credited to "that prodigious young scholar Mr Christopher Wren, who presented me with a piece of white marble, which he had stain'd with a lively red very deepe, as beautiful as if it had been natural".

WELLCOME LIBRARY, LONDON

One of Wren's mathematical inventions was a perspective machine, or perspectograph, an ingenious device that allows a draftsman, squinting through a sight, to trace the lineaments of any object or landscape. One result was his drawing entitled 'Prospect of Windsor Castle from the North', engraved by Wenceslaus Hollar for Elias Ashmole's 1672 book *Institution, Laws and Ceremonies of the Most Noble Order of the Garter*.

Wren's perspectograph was probably used to trace the topography of the brain, its nerves and blood vessels, just as he used a telescope coupled to a micrometer to map the heavens and create the first lunar globe. However, it would be wrong to think that Wren used his instruments and great skill as a draftsman to create the equivalent of a photograph of a dissection — his representations are a succinct synthesis of what Willis and his collaborators had deduced about the structure of the brain over of period of time. To arrive at this point, they exploited the latest techniques of injection to preserve and dye the brain's tissues and vessels. As Willis explained, "let a dyed liquor, and contained in a large squirt or pipe, be injected upwards in the trunk [of an artery] of one side: after once or twice injecting, you shall see the tincture or dyed liquor to descend from the other side by the trunk of the opposite artery".

The resulting illustration (pictured) is, in effect, a complex act of visual pointing, in which key features are elucidated in their plastic and spatial disposition. Of particular concern was the vascular system at the base of the brain, most notably the looped arterial anastomosis, or network of arteries. This had been noticed by earlier writers, but was here given its clearest visual description.

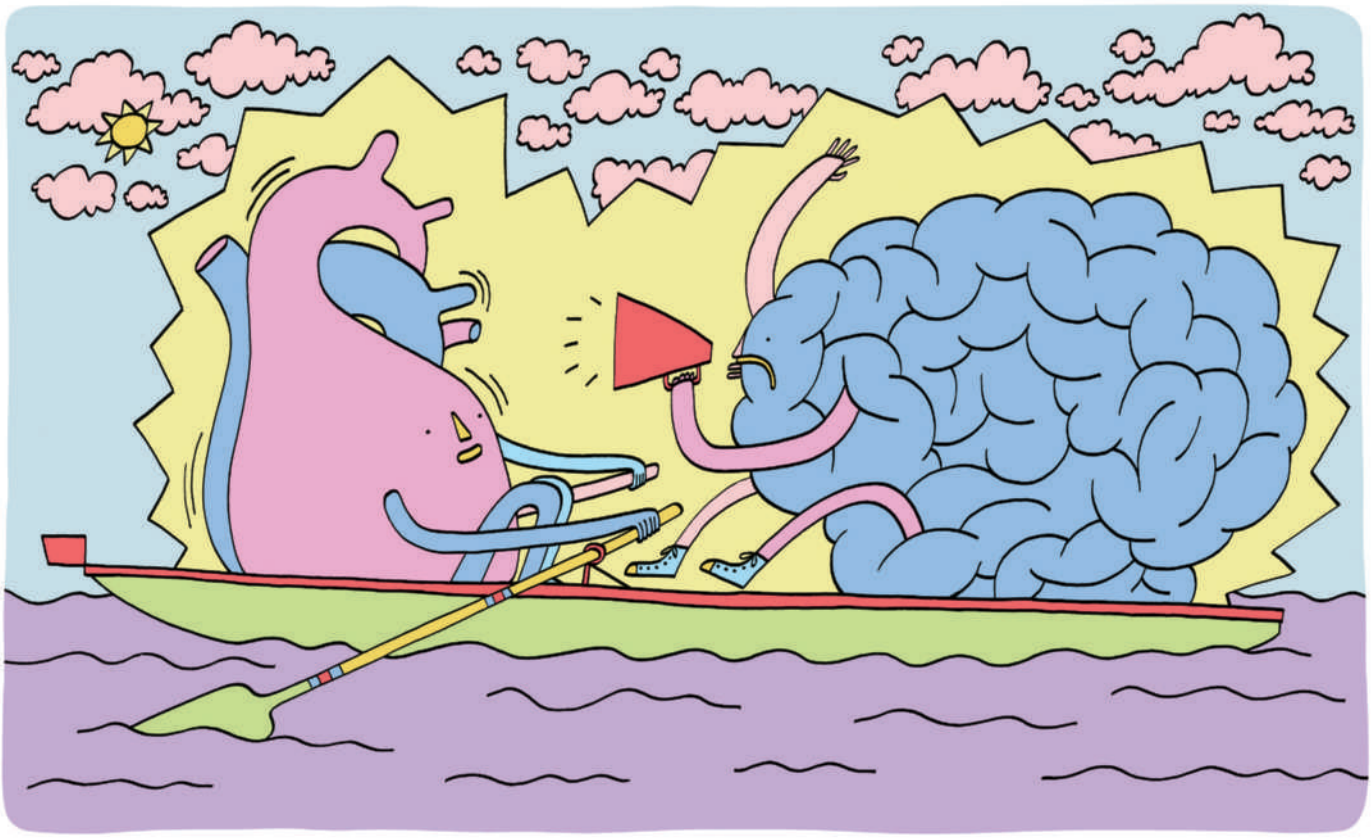
It was the potency of Wren's illustrations, with their precise 'mathematical' character, that led to the anastomosis being designated the Circle of Willis. Willis's renumbering of the cranial nerves still survives and, for good measure, he also coined the term 'neurologie' in his *De Anima Brutorum* of 1672.

Willis's *Cerebri Anatome* was dedicated to Archbishop Gilbert Sheldon, for whom Wren was constructing the eponymous Sheldonian theatre in Oxford. The wide interior is spanned by a ceiling devoid of visible beams, using a system of trusses that Wren devised, perhaps already in collaboration with Robert Hooke, with whom he worked on St Paul's. Hooke, also a distinguished architect, was a notable scientist and creator of *Micrographia* in 1665, the greatest of all books of microscopic illustration.

Given such conjunctions between architecture and science, it is appropriate that Willis should express unbounded admiration for the way that the "brain, like a castle, divided into many towers or places of defence, is thereby made the stronger and harder to be taken". The cranium and its precious brain were, after all, the greatest designs of nature, the supreme and most prudent architect of all. ■

Martin Kemp is emeritus professor in history of art at the University of Oxford, Oxford, UK, and author of *The Human Animal*. **Nathan Flis** is a doctoral student at the University of Oxford.

ESSAY



M. HODSON

Why do intelligent people live longer?

We must discover why cognitive differences are related to morbidity and mortality, argues **Ian Deary**, in order to help tackle health inequalities.

Ten years ago, on 16 October 1998, I presented findings that people from Aberdeen with higher childhood IQs — measured at age 11 in the Scottish Mental Survey of 1932 — were significantly more likely to survive to age 76. It was at a psychology seminar at Glasgow Caledonian University, UK. For one audience member, the finding did not go down well. "So, you're saying that the thick die quick?" It was not a point of clarification; it was an accusation. The temperature in the room rose as the questioner railed against a result he found insulting and wanted to invalidate. Hadn't intelligence tests been discredited?

Actually, no. Scores from cognitive-ability tests (also known as intelligence tests or IQ tests) have validity that is almost unequalled in psychology[1]. A general cognitive-ability factor emerges from measures of diverse mental tasks, something that hundreds of data sets since 1904 have replicated. People's rankings on intelligence tests show high stability across almost the whole lifespan, are substantially heritable and are associated with important life outcomes — including educational achievements, occupational success and morbidity and mortality. More thumping confirmatory studies of the link between intelligence and mortality have appeared since our first work. One of these contains nearly a million Swedish men tested at around age 19 during military induction and followed for almost 20 years[2]. It shows a dear association: as intelligence test scores go up the scale, so too does the likelihood of survival over those two decades.

When we attempted to publish our original study, we came across a different complaint: the journals to which we submitted our initial findings said they found the link obvious. It was already well known that health inequalities are associated with different social backgrounds. That was deemed the likely explanation for the finding. But it has since been shown that childhood social class does not account for the association between childhood intelligence and later mortality[3].

Intelligence can predict mortality more strongly than body mass index, total cholesterol, blood pressure or blood glucose, and at a similar level to smoking[4]. But the reasons for this are still mysterious. That needs to change. Reducing health inequalities is a priority, and to do that we need to determine their causes.

Catchy headlines

We weren't the first to find a link between individually assessed intelligence and mortality. In 1992 a report in the journal *Personality and Individual Differences* showed that mental test scores of Australian soldiers at the time of induction for the Vietnam War were associated with deaths up to early middle age[5]. The finding should have been headline grabbing, but it has been slow to gain a popular and academic reaction. It wasn't until our 2001 paper in the *British Medical Journal*[6] that the media began to pay attention — 'Brainy Kids Live Longer' proclaimed Scotland's *Daily Record*; 'The Higher Your IQ, the Longer You're Likely to Live' said Britain's *Daily Mail*. And the papers became more frequently cited.

Cognitive epidemiology has since become

established as a distinct area of study, and a systematic review has established the association between early-life intelligence and mortality across different populations, in different countries, and in different epochs[3].

Four explanations

The field has focused on four non-exclusive possibilities for the link between intelligence and death. First, what occurs to many people as an obvious pathway of explanation, is that intelligence is associated with more education, and thereafter with more professional occupations that might place the person in healthier environments. Statistical adjustment for education and adult social class can make the association between early-life intelligence and mortality lessen or disappear[2,7]. But not always.

Moreover, cause and effect among intelligence, education and social class is not settled. It may be that a person with more and better education achieves a higher IQ score; but a child with a high IQ score is more likely to undergo more years of education, attain higher qualifications and go on to a better job. Thus, adjusting for education and social class in the intelligence–mortality association could be an over-adjustment — it might weed out some of the very influence of intelligence that we are trying to detect. Linda Gottfredson, in the School of Education at the University of Delaware in Newark, has proposed that intelligence is a fundamental cause' of the association between education, social class and health[8]. Her gauntlet-throwing-down paper — which has just won the prestigious George A. Miller award from the American Psychological Association — is provocative to epidemiologists, who tend to look outside rather than within the individual for causes of health inequalities. Those with contrasting ideas in this area need a chance to debate their theories and work towards a solution.

> "Why do we die when we do, and to what extent is this question tractable?"

Second, people with higher intelligence might engage in more healthy behaviours. Evidence is accruing that people with higher intelligence in early life are more likely to have better diets, take more exercise, avoid accidents, give up smoking, engage in less binge drinking and put on less weight in adulthood.

But this too doesn't seem to be the whole story. In an assessment known as the US Vietnam Experience Study[9], higher intelligence test scores at induction were associated with a lower likelihood of developing metabolic syndrome in middle age — a condition combining factors such as obesity, high blood pressure and impaired glucose metabolism. In the same study, higher intelligence was also associated with lower likelihood of dying in 15 years of further follow-up. Adjusting for metabolic syndrome reduced the association of intelligence with total mortality by up to 10%, and with cardiovascular disease mortality by up to 32%.

Third, mental test scores from early life might act as a record of insults to the brain that have occurred before that date. These insults — perinatal events, or the result of illnesses, accidents or deprivations before the mental testing — might be the fundamental cause behind both intelligence test scores and mortality risk. So far, little evidence supports this. Both birth weight (commonly used as a marker of fetal development) and parental social class (used as a marker of early-life circumstances) are correlated with intelligence test scores. But, when the associations between intelligence and mortality are adjusted for these factors, the association remains almost unaltered. Perhaps subsequent work may find better indicators of early-life tribulations that have more explanatory power.

Fourth, mental test scores obtained in youth might be an indicator of a well-put-together system. It is hypothesized that a well-wired body is more able to respond effectively to environmental insults. This 'system integrity' idea has a parallel in the field of ageing, where some data suggest that bodily and cognitive functions age in concert.

Some supporting evidence comes from the finding that simple reaction speed — the time taken to press a button when a stimulus appears — can displace intelligence test scores as an even better predictor of mortality risk[10]. Reaction-time tasks do not require complex reasoning, and so are unlikely to be improved by education. A major job for the field is to discover better markers for system integrity and to test their explanatory power for mortality.

Unclear chain

Although intelligence plays a part in health behaviours and health outcomes that contribute to specific causes of death, a clear chain of causation from intelligence to health outcomes and then to death has not emerged. Different types of mortality, including cardiovascular disease, homicide and suicide, seem to demand their own explanations for being associated with early-life intelligence. Those who found the intelligence–death association 'obvious' must think again.

The new University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, of which I am director, opened on 1 September 2008. High among its aims is to provide a forum and infrastructure to unpick the extent to which cognitive and other effects underlie different causes of mortality. Psychologists, epidemiologists, neuroscientists, behavioural geneticists, statisticians and others will interact in our programme, providing complementary skills to understand the causal pathways.

The field has benefited from many varied data sets, none of which was initiated with the aim of asking why intelligence foretells death. Considerable credit for identifying such data goes to David Batty, an epidemiologist interested in the causes of chronic disease and a colleague of mine since 2002, who was responsible for establishing links with the Swedish conscripts study and the Vietnam Experience Study, amongst others, to test our theories.

High among our priorities is to find more such relevant data sets, especially some involving twins. Members of our centre are now analysing twin studies from Minnesota and Denmark that seem, in the early stages, likely to be informative about the intelligence–mortality link.

There is also a search for other, non-cognitive psychological characteristics that are associated with living longer. For example, it seems that, independently of any association with intelligence, being more dependable or conscientious in childhood is also significantly protective to health. Children who scored in the top 50% of the population for intelligence and dependability were in one study[11] more than twice as likely to survive to their late sixties as children scoring in the bottom half for both.

Why do we die when we do, and to what extent is this question tractable? This fundamental curiosity drives our team and others on to better scientific studies. More than satisfying our curiosity, the search has practical implications. The influence of intelligence on mortality isn't fated; intelligence does not unalterably spin, measure and cut the thread of life. The things that people with higher intelligence have and do that make them live longer may be found and, we hope, shared, towards the goal of better and more equal health. ■

Ian Deary is director of the University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, Edinburgh EH8 9JZ, UK.
e-mail: i.deary@ed.ac.uk

1. Deary, I. J. *Intelligence: A Very Short Introduction* (Oxford Univ. Press, 2001).
2. Batty, G. D. *et al. Epidemiology* (in the press).
3. Batty, G. D., Deary, I. J. & Gottfredson, L. S. *Ann. Epidemiol.* **17,** 278–288 (2007).
4. Batty, G. D., Shipley, M. J., Gale, C. R., Mortensen, L. & Deary, I. *Heart* doi:10.1136/hrt.2008.149567 (2008).
5. O'Toole, B. I. & Stankov, L. *Pers. Individ. Dif.* **13,** 699–716 (1992).
6. Whalley, L. J. & Deary, I. J. *Br. Med. J.* **322,** 819–822 (2001).
7. Hart, C. *et al. Psychosom. Med.* **65,** 877–883 (2003).
8. Gottfredson, L. S. J. *Pers. Soc. Psychol.* **86,** 174–199 (2004).
9. Batty, G. D. *et al. Diabetologia* **51,** 436–443 (2008).
10. Deary, I. J. & Der, G. *Psychol. Sci.* **16,** 64–69 (2005).
11. Deary, I. J., Batty, G. D., Pattie, A. & Gale, C. R. *Psychol. Sci.* **19,** 874–880 (2008).

NEWS & VIEWS

NEUROSCIENCE

Light moulds plastic brains

Stefan Thor

In tadpoles, the number of neurons expressing the neurotransmitter dopamine increases on exposure to light. Such plasticity might allow animals to match their brains' response to environmental stimuli.

Nervous systems are known to adapt to environmental inputs. But such plasticity has been thought to involve modifications of neural circuits and communication between neurons via synaptic junctions — as in learning and memory — rather than alterations in the numbers of distinct classes of neuron. On page 195 of this issue, Dulcis and Spitzer[1] challenge this view. They show that, when frog tadpoles (*Xenopus laevis*) are exposed to bright light, the number of dopamine-secreting (dopaminergic) neurons in the animals' brains increases, allowing them to adapt more rapidly to subsequent exposure to light.

Anyone who has caught wild frog tadpoles from a pond has probably been surprised to see the captive animals turn pale after a couple of hours. This rapid change in pigmentation allows tadpoles to better blend in with their surroundings, reducing their risk of becoming prey. A distinct neural circuit controls this process. Specifically, light-induced signals from the eye are relayed to a brain region called the suprachiasmatic nucleus, which contains dopaminergic neurons. From there, signals pass on to another region containing neurons that secrete melanocyte-stimulating hormone to trigger pigment cells in the skin (Fig. 1a). This circuit works in an alternating manner such that, in response to light, positive inputs from the eye to the suprachiasmatic nucleus trigger increased dopamine release. High dopamine levels then provide negative inputs to the hormone cells, resulting in reduced hormone secretion and so decreased pigmentation of the peripheral skin.

The pigmentation response is modulated by previous experience, because prolonged or repeated exposure to bright light results in tadpoles adapting more rapidly to subsequent exposures[2]. Changes in this response and its underlying circuitry have been studied extensively[2–5], and were believed to primarily involve plasticity at the level of synaptic connections and signals passing through the circuit. Dulcis and Spitzer[1] reveal that this adaptation also involves a rapid increase in the number of dopaminergic neurons within the circuitry.

The speed of the response that the authors observe is remarkable — when exposed to only two hours of light, tadpoles that had been raised in the dark exhibited a doubling of dopaminergic neurons within the suprachiasmatic nucleus. What's more, the newly emerging dopaminergic neurons seem to affect the pigmentation process on subsequent exposures to light by reducing pigmentation more rapidly (Fig. 1b). The authors traced these neurons' axonal processes and found that they project onto the hormone-releasing neurons. They next ablated the baseline dopaminergic neurons using specific drugs to show that this treatment completely abolishes light adaptation. But when animals with ablated dopaminergic neurons were exposed to light on subsequent occasions, dopaminergic neurons that had appeared after drug treatment could restore light adaptation.

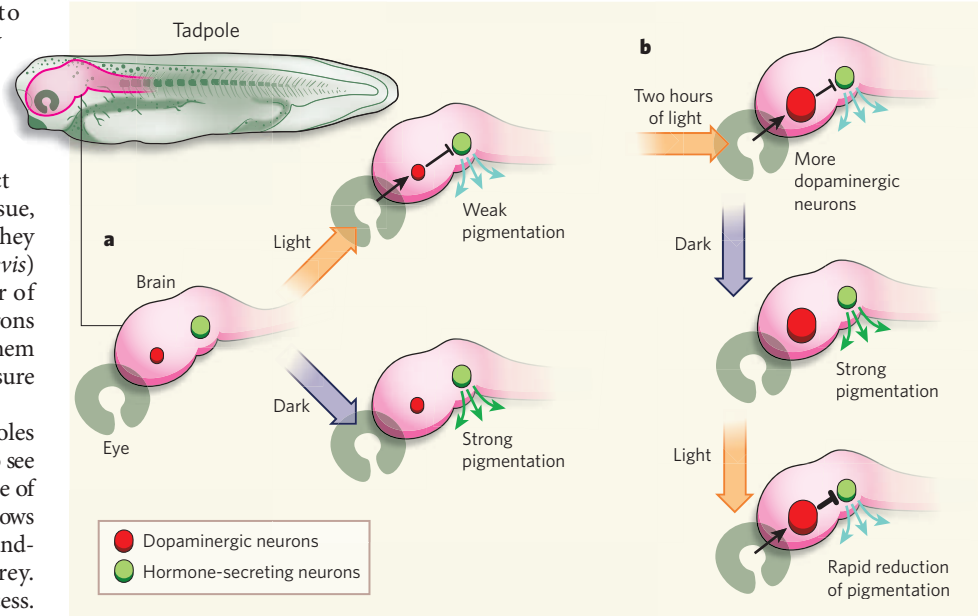


Figure 1 | Plasticity in the tadpole pigmentation circuit. a, Tadpoles adjust their surface pigmentation in response to the surrounding light conditions using a neuronal circuit involving signals from the eye to the dopaminergic neurons in the brain's suprachiasmatic nucleus (red). These neurons inhibit pigment-hormone release from melanocyte-stimulating cells (green). **b**, Dulcis and Spitzer[1] find that a longer exposure to light — about two hours — induces the generation of extra dopaminergic neurons in the suprachiasmatic nucleus of tadpoles. Return to darkness still allows for the darker pigmentation. Yet, on subsequent exposure to light, the presence of the new and pre-existing dopaminergic cells triggers a more rapid reduction in pigmentation. Thus, in the developing tadpole, the extra dopaminergic neurons provide an adaptive advantage.

Where do these 'new' dopaminergic neurons come from? Do they result from a change in the type of neurotransmitter secreted by pre-existing neurons, or are they generated *de novo*? Earlier work revealed that the mammalian brain (even that of adult mammals) constantly generates new neurons[6,7], and that this process can be enhanced in response to environmental cues[8–10]. For instance, adult laboratory mice living in an enriched environment — large cages containing running wheels, nesting material and toys — have increased numbers of neurons in specific brain areas, particularly those crucial for spatial orientation[10]. Likewise, songbirds add and remove neurons to certain brain regions on a seasonal basis, a mechanism that acts to match brain anatomy to appropriate seasonal behaviour[11]. In the case of light

NATURE|Vol 456|13 November 2008

adaptation in tadpoles, however, Dulcis and Spitzer find no evidence for new cells being generated within the suprachiasmatic nucleus. Given the rapid appearance of the extra dopaminergic neurons, this observation was perhaps expected: it is unlikely that additional neurons could be generated *de novo* within the relatively short time frame of only two hours. Instead, it seems that pre-existing neurons expressing a different neurotransmitter now co-produce dopamine.

Dulcis and Spitzer's findings advance the idea that external sensory inputs modulate a specific response by regulating the population size of specific neuronal subtypes — those that are involved in controlling the physiological response to the input — in the brain. From a broader perspective, their observation that pre-existing neurons can switch on the expression of an additional type of neurotransmitter adds to the growing list of ways in which brain plasticity can arise: weakening or strengthening of communication between neurons, formation of new connections, and the recent findings[6–10] that additional neurons of certain types can be added to the system *de novo*. An issue that Dulcis and Spitzer do not address, however, is whether the particular type of brain plasticity they observe is limited to developing tadpoles, or whether it also applies to adult frogs — and mammals, for that matter.

In humans, dysfunction of signalling cascades mediated by dopamine may be an essential element of seasonal affective disorder, also known as winter depression[12]. Here, a way forward might be to analyse patients using positron emission tomography (PET), a technique that is routinely used to visualize dopaminergic cells[13]. Although PET images are of limited resolution, a dramatic increase in the number of dopamine neurons — possibly in response to seasonal changes in day length or diurnal changes in light intensity — may be detectable. If plasticity mediated by changes in the pattern of neurotransmitter production applies to humans, it is likely to open fresh avenues towards combating neurological diseases. ■

Stefan Thor is in the Department of Clinical and Experimental Medicine, Linkoping University, Linkoping SE-58183, Sweden.
e-mail: stefan.thor@liu.se

1. Dulcis, D. & Spitzer, N. C. *Nature* **456,** 195–201 (2008).
2. Roubos, E. W. *Comp. Biochem. Physiol.* **118,** 533–550 (1997).
3. Kramer, B. M. R. *et al. Microsc. Res. Tech.* **54,** 188–199 (2001).
4. Tuinhof, R. *et al. Neuroscience* **61,** 411–420 (1994).
5. Ubink, R., Tuinhof, R. & Roubos, E. W. *J. Comp. Neurol.* **397,** 60–68 (1998).
6. Chen, J., Magavi, S. S. P. & Macklis, J. D. *Proc. Natl Acad. Sci. USA* **101,** 16357–16362 (2004).
7. Magavi, S. S., Leavitt, B. R. & Macklis, J. D. *Nature* **405,** 951–955 (2000).
8. van Praag, H., Christie, B. R., Sejnowski, T. J. & Gage, F. H. *Proc. Natl Acad. Sci. USA* **96,** 13427–13431 (1999).
9. van Praag, H. *et al. Nature* **415,** 1030–1034 (2002).
10. Kempermann, G., Gast, D. & Gage, F. H. *Ann. Neurol.* **52,** 135–143 (2002).
11. Nottebohm, F. *Brain Res. Bull.* **57,** 737–749 (2002).
12. Lam, R. W., Tam, E. M., Grewal, A. & Yatham, L. N. *Neuropsychopharmacology* **25** (Suppl.), S97–S101 (2001).
13. Perlmutter, J. S. & Moerlein, S. M. *Q. J. Nucl. Med.* **43,** 140–154 (1999).

ANALYTICAL CHEMISTRY

Home diagnostics to music

Jeffrey S. Erickson and Frances S. Ligler

Fed up with sitting in the doctor's surgery among all those sneezy patients, waiting for the results of a health check? With the latest technology, you could one day perform bioassays on your home compact-disc player.

Biosensors are under commercial development for all sorts of applications — detecting human and animal pathogens, measuring clinical markers for heart attack and cancer, and monitoring environmental pollutants. The most effective way to put biosensors into the hands of potential customers, especially those with limited budgets, might be to modify the technology so that it can run on everyday devices, rather than on specialized apparatus. Popular electronic gadgets such as palm-sized computers, flatbed scanners and DVD players could potentially be converted into low-cost, battery-operated, user-friendly biosensors. Reporting in *Analytical Chemistry*, Li *et al.*[1] describe a notable advance in this direction: a system for running assays on standard compact discs that uses free CD-quality-analysis software to detect and quantify biological binding events.

With data-reading optical elements and electronics already incorporated into CD players, it is easy to imagine that sample delivery systems could be integrated into these devices to provide a system for running assays on discs. In particular, researchers developing microfluidics technology have shown that centrifugal force can be used to move fluids through compartments in a spinning disc[2,3]. The first such systems required external devices to read out results from disc-based assays, but the best readers would be standard-issue machines such as those found in homes or cars.

Although Li *et al.*[1] are not the first to run bioassays on a disc, or to read results using a standard CD player, they are the first to have integrated most of the elements required to do both. Building on previous findings[4–7], they report a gentle and widely applicable method for attaching biorecognition molecules — which can bind to a wide variety of medically important target molecules — to the surface of a CD. This method involves oxidizing the polymeric surface of the disc to form chemical groups to which biotin molecules can be tethered. The surface-bound biotin molecules bind specifically to streptavidin proteins, which can subsequently be used to immobilize various biorecognition molecules. The authors designed their system so that, in the presence of a target molecule, a surface-bound complex forms that includes gold nanoparticles. The nanoparticles reflect light and thus can generate an optical signal. But in practice, the nanoparticles are too small to be detected by CD players. Li *et al.* therefore added an extra step to their protocol, in which larger silver particles are grown around the gold nanoparticles, providing a detectable proxy for the binding of target molecules.

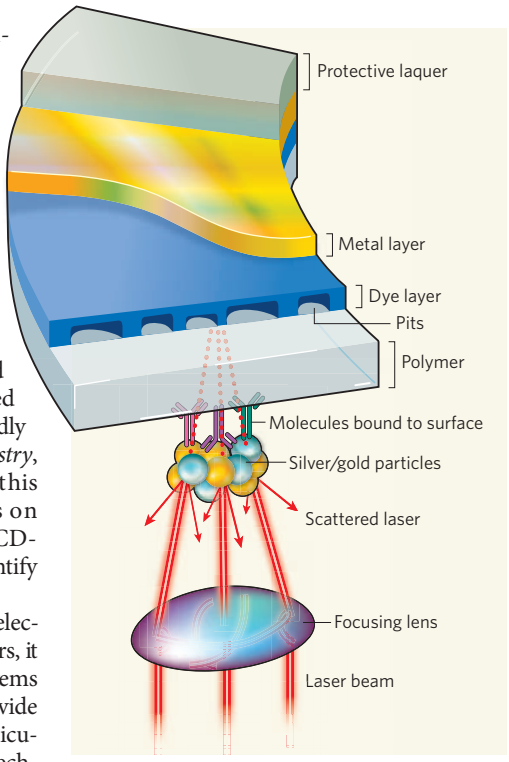


Figure 1 | Musical bioassays. Each track on a CD stores data as a series of pits in a light-sensitive dye layer that is backed by a reflective metal layer. CD players read the data by scanning the dye's surface through a polymer layer. Li *et al.*[1] have developed a system for running binding assays on standard CD players. 'Capture molecules' attached to the polymer surface of a CD bind to specific molecular targets. Gold nanoparticles attached to the targets nucleate the formation of larger silver particles, which prevent the laser beam of a CD player from reading the data on the disc. This creates an error in the retrieved data, which can be detected by the CD player. The authors use free, CD-quality-analysis software to detect and quantify binding events across the surface of the CD. (Figure modified from ref. 1.)

Perhaps the real breakthrough in this work[1] is the digital readout protocol. The data on every pre-recorded audio CD are encoded in such a way that errors can be detected and corrected by CD players, using a standard algorithm. For example, scratches up to about 8.5 millimetres long can be compensated for during playback. The silver particles that are associated with

the formation of target complexes in Li and colleagues' system obscure data pre-recorded on the CD, and so can be detected as errors (Fig. 1). Furthermore, because each block of data on a CD is associated with a physical position on the disc, these errors can be tracked to their locations. Li *et al.* therefore used free diagnostic software to analyse the error information measured by the CD player reading the disc. By plotting the rate of error detection against the distance from the centre of the CD, the formation of target complexes at all points on the CD's surface can be quantified.

Methods for acquiring data from a disc-based assay using a standard CD reader[8] or specially written error-determination software[9] have been reported previously, but Li and colleagues' use of freely available software opens up such assays for widespread use. And by demonstrating that three different software packages can be used for detecting spots of bound targets, the authors show that their technique need not be restricted by the availability of any single error-detection program.

Using their biochemical and analytical strategy[1], the authors performed assays to detect the binding of single-stranded DNA to complementary strands and to detect the binding of antibodies to antigens. The detection rate increased with the concentration of the targets, suggesting that the assays can indeed produce quantitative results. After carrying out comparison experiments, Li *et al.* claim that the sensitivity of their technique is tenfold higher than that of fluorescence-based methods. This assessment may not be valid, as they placed ten times more target molecule in the microfluidic channels for the CD reader than in the channels for their fluorescence assays — but this is a minor criticism.

Only one piece of the puzzle now remains to be slotted into place: microfluidic structures must be integrated into CDs for sample processing and reagent storage in a format that can be run in a standard CD player. Figuring out how to do this without compromising the detector's function is a tough problem. An additional transparent layer would need to be attached to the disc to either contain or constrain microfluidic channels, buffers and freeze-dried reagents. Potential fabrication methods include injection moulding, embossing or casting[10], but any such modifications to the original disc must be done in a way that does not damage the underlying metal or light-sensitive layer.

If fluidics were integrated into the disc by removing material from the thick polymer layer, the sensitivity of the technique might actually be increased: the binding events would occur closer to the reflecting layer of the disc, on which the detecting laser is focused. But local changes in the thickness and refractive index of the polymer layer might also alter the focal distance of the laser, reducing sensitivity. Furthermore, light scattering off the walls of the microfluidic structures might produce signals that are difficult to distinguish from binding signals.

One solution would be to make all the surfaces forming the microfluidic channels from materials that have the same refractive index, and then to fill the channels with a fluid of similar refractive index. This is much easier said than done. Alternatively, the fluidic layer could be made to be removable. This would necessitate a two-spin assay — one to run the assay, and another, without the fluidic layer, to detect the results. The removal of the fluidic layer between spins would have to be performed in a way that doesn't damage the sensing surface of the disc, but such handling by the user would best be avoided.

Clearly, there is much work to be done, yet the integration of bioassays into consumer electronic devices has a bright outlook. Perhaps we really will eventually be able to recreate the experience of a doctor's surgery at home: not only will we be able to run a quick and inexpensive biological test in our own living room, but the analytical device will also provide the waiting-room background music. ■

Jeffrey S. Erickson and Frances S. Ligler are at the Center for Bio/Molecular Science & Engineering, Naval Research Laboratory, 4555 Overlook Avenue SW, Washington DC 20375-5348, USA.
e-mail: frances.ligler@nrl.navy.mil

1. Li, Y., Ou, L. M. L. & Yu, H.-Z. *Anal. Chem.* **80,** 8216–8223 (2008).
2. Madou, M. J., Lee, L. J., Daunert, S., Lai, S. & Shih, C.-H. *Biomed. Microdevices* **3,** 245–254 (2001).
3. Duffy, D. C., Gillis, H. L., Lin, J., Sheppard, N. F. Jr & Kellogg, G. J. *Anal. Chem.* **71,** 4669–4678 (1999).
4. Kido, H., Maquieira, A. & Hammock, B. D. *Anal. Chim. Acta* **411,** 1–11 (2000).
5. Morais, S., Carrascosa, J., Mira, D., Puchades, R. & Maquieira, A. *Anal. Chem.* **79,** 7628–7635 (2007).
6. Bañuls, M.-J., González-Pedro, V., Puchades, R. & Maquieira, A. *Bioconjugate Chem.* **18,** 1408–1414 (2007).
7. Lange, S. A. *et al. Angew. Chem. Int. Edn* **45,** 270–273 (2005).
8. Potyrailo, R. A. *et al. Anal. Chem.* **78,** 5893–5899 (2006).
9. La Clair, J. J. & Burkart, M. D. *Org. Biomol. Chem.* **4,** 3052–3055 (2006).
10. Lee, L. J. *et al. Biomed. Microdevices* **3,** 339–351 (2001).

MARINE BIOLOGY

Genes in the glass house

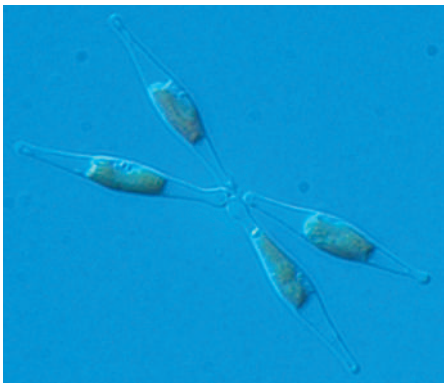
Ronald P. Kiene

Sequence data on a second species of diatom provide abundant insights into the evolution and metabolic capabilities of this group, as well as into mechanisms of gene acquisition and diversification.

Diatoms are one of life's big success stories. These silica-depositing microalgae have been abundant for at least the past 180 million years. Today they contribute an astonishing 40% of total ocean primary production, and about 20% of global primary production[1,2], making them central players in the global carbon and silica cycles.

The main lineages of diatoms are the centrics and the pennates, the two having begun to diverge from their common ancestor about 90 million years ago[3]. The full genome sequence of a member of the centric lineage, *Thalassiosira pseudonana*, was reported[4] in 2004, but since then we have been waiting for an equivalent treatment of one of the highly diverse pennates. That conspicuous gap has been filled by Bowler *et al.* (page 239 of this issue)[5] with a study of the genome sequence of *Phaeodactylum tricornutum* (Fig. 1).

It is no easy task to find the genes in the seemingly endless string of base pairs that constitutes a whole genome. To aid in mapping the *P. tricornutum* genome, Bowler *et al.* used more than 130,000 available expressed sequence tags; these are short DNA sequences that are derived from messenger RNA and so represent parts of genes that have been expressed. This and other information provided support for the existence of 10,402 genes in *P. tricornutum*, which is slightly fewer than the 11,776 genes identified in its centric cousin, *T. pseudonana*. Despite the similarity in overall numbers, the two species have only 57% of their genes in common, indicating an exceptional level of divergence since the split of the lineages.



A. DE MARTINO & C. BOWLER

Figure 1 | *Phaeodactylum tricornutum* — a quartet of the model pennate diatom.

Diatoms belong to a larger group called the chromalveolates (Fig. 2, overleaf). Along with their relatives, including the haptophytes, dinoflagellates and oomycetes, they are thought to have evolved from a secondary symbiosis between a photosynthetic red alga and a heterotrophic host[6]. The red alga involved is believed to have arisen from a primary symbiosis between a heterotroph and

a cyanobacterium. In both instances, then, partnerships were forged between an organism that required organic substrates for energy generation (the heterotroph) and an organism that could produce such substrates from inorganic materials and light.

To date, there has been debate as to whether the secondary symbiosis occurred just once or at several points within the different lineages[6,7]. Genes associated with two of the main cell constituents of *P. tricornutum*, the chloroplast and mitochondrion, now provide some helpful clues.

For example, the genomic analyses strongly support the idea that the diatom chloroplast came from red algae: 171 genes of *P. tricornutum* are of red algal origin, and a total of 108 of the red algal genes occur in both centric and pennate diatoms. Of those 108 genes, 11 are also found in the oomycete *Phytophthora sojae* (a water mould)[8], supporting the idea that diatoms and oomycetes shared a common ancestor possessing a red algal chloroplast that was later lost in the development of oomycetes. Furthermore, the suite of mitochondrial genes in *P. tricornutum* is nearly the same as that in the mitochondria of the earlier-diverging haptophytes and cryptophytes (Fig. 2). This provides some support for the chromalveolate hypothesis — the idea that all these groups arose from the same ancient secondary symbiosis event[6]. But there are still enough questions about the details of these phylogenetic associations to ensure that vigorous debate about the various evolutionary hypotheses will continue[9].

The evolution of *P. tricornutum*, a eukaryote with the characteristic membrane-bound nucleus, seems to have been advanced by its acquisition of at least 587 genes (more than 5% of the total complement) from prokaryotic organisms — mostly from bacteria but with some from archaea too. Although this level of horizontal gene transfer is common among prokaryotes[10], the number of prokaryote genes in *P. tricornutum* is unusually high for a eukaryote, and is suggestive of long-term intimate associations between bacteria and diatoms that led to transfer of useful genetic capabilities. The prokaryote genes in *P. tricornutum* support pathways for the use of organic carbon and nitrogen, and they include genes that encode the machinery for the diatom urea cycle, which are also present in *T. pseudonana*.

Some of the prokaryote gene products seem to be involved in the perception of light and may be part of cell-signalling systems. More than 300 of these prokaryote genes are found in both centric and pennate diatoms, indicating that many of them were acquired before these groups split. But only 14 prokaryote sequences are shared between *P. tricornutum* and the oomycete *P. sojae*, indicating that most of the bacterial sequences were acquired after the divergence of oomycetes and the diatom lineage about 700 million years ago. It will be interesting to learn whether many bacterial genes occur in other stramenopile algae

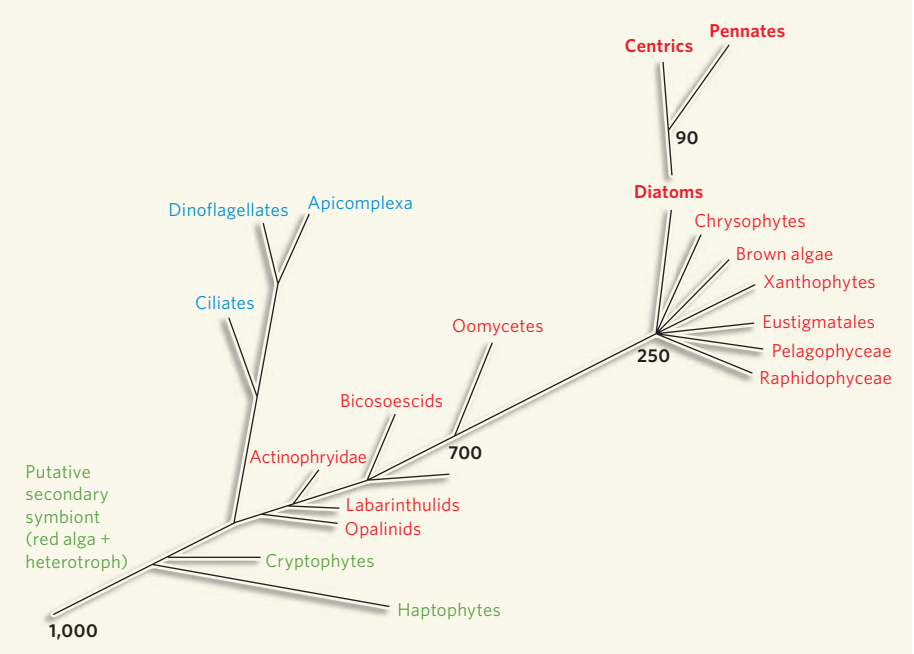


Figure 2 | Relationships among the chromalveolates, to which diatoms belong. Two subgroups are the alveolates (blue) and the stramenopiles, including the diatoms (red). The diatoms themselves consist of two main lineages; examples of the genomes of both of them, the centric[4] and the pennate[5], have now been sequenced. The numbers are indications of approximate divergence times in millions of years ago. (Graphic based on Fig. 1 of ref. 13.)

(Fig. 2), and whether the same bacterial genes have been retained or different ones have been acquired by the different lineages.

Of the more than 10,000 genes in each of the two diatom genomes, 1,338 are unique to the diatoms and will attract interest from researchers wanting to find out what sets this group apart from others. A defining characteristic of diatoms is their intricate silica frustules, or tests, which serve as multifunctional cell walls[11]. These exquisite glass houses are the result of nanofabrication of polymerized silicic acid laid on an organic matrix of long-chain polyamines such as spermine and proteins called silaffins. The resulting silica frustule is coated with a glycoprotein called frustulin. Not surprisingly, the *P. tricornutum* genome contains many genes involved in frustulin synthesis and polyamine metabolism. But there is evidence of just a single silaffin-like protein, indicating either that only one is needed or that other proteins (and their genes) in this family await discovery. There are many genes associated with cell-cycle regulation, which may be related to the need to coordinate the division of cells with rigid and unequal silica valves.

Further analysis of the genome should reveal many more diatom-specific functions, and as other relatives along this phylogenetic tree are sequenced we should learn more about when and how these genetic traits were acquired. Comparative analysis of diatom genomes may also reveal the basis for certain differences, such as the production of the biogeochemically important metabolite dimethylsulphoniopropionate (DMSP) by *T. pseudonana* but not by *P. tricornutum*[12].

Finally, Bowler *et al.*[5] provide evidence that diatom-specific genes in *P. tricornutum* and *T. pseudonana* seem to be evolving faster than other genes in these diatoms or in eukaryotes in general. Mechanisms for this rapid gene diversification vary. They include expansion of certain gene families; gains of non-protein-coding — intron — sequences (mainly in *T. pseudonana*); and, in *P. tricornutum*, an exceptionally high frequency of retrotransposons, which are mobile genetic elements that probably accelerated gene fragmentation. The rapid gene evolution in diatoms may help to explain their extensive diversification, and probably contributed to their niche specialization and ultimate success in the global ecosystem[2]. ■

Ronald P. Kiene is in the Department of Marine Sciences, University of South Alabama, and the Dauphin Island Sea Lab, Mobile, Alabama 36688, USA.
e-mail: rkiene@disl.org

1. Falkowski, P. G., Barber, R. T. & Smetacek, V. *Science* **281,** 200–206 (1998).
2. Falkowski, P. G. *et al. Science* **305,** 354–360 (2004).
3. Kooistra, W. H. C. F., Gersonde, R., Medlin, L. K. & Mann, D. G. in *Evolution of Primary Producers in the Sea* (eds Falkowski, P. G. & Knoll, A. H.) 207–249 (Academic, 2007).
4. Armbrust, E. V. *et al. Science* **306,** 79–86 (2004).
5. Bowler, C. *et al. Nature* **456,** 239–244 (2008).
6. Cavalier-Smith, T. *J. Eukaryotic Microbiol.* **46,** 347–366 (1999).
7. Bodyl, A. *J. Phycol.* **41,** 712–719 (2005).
8. Tyler, B. M. *et al. Science* **313,** 1261–1266 (2006).
9. Keeling, P. J. *et al. Science* **306,** 2191b (2004).
10. Woese, C. R. *Proc. Natl Acad. Sci. USA* **99,** 8742–8747 (2002).
11. Milligan, A. J. & Morel, F. M. M. *Science* **297,** 1848–1850 (2002).
12. Keller, M. D., Bellows, W. K. & Guillard, R. R. L. in *Biogenic Sulfur in the Environment* (eds Saltzman, E. & Cooper, W. J.) 167–182 (Am. Chem. Soc., 1989).
13. Fehling, J. *et al.* in *Evolution of Primary Producers in the Sea* (eds Falkowski, P. G. & Knoll, A. H.) 75–107 (Academic, 2007).

NATURE|Vol 456|13 November 2008

QUANTUM PHYSICS

Swift control of a single spin

Keiichi Edamatsu

For now, quantum information processing systems remain a dream. Step by step, however, progress towards that goal is being made, with one promising route involving a novel means of manipulating electron spin.

The basic quantity of magnetic recording — the working principle of a computer's hard disk — is an electron's spin. Although the technology for magnetic recording is reaching recording densities as high as 1 terabit per square inch (ref. 1), storing a single bit of information still involves around 10^5 electron spins. Future quantum information processing systems that use the electron's spin as a unit of quantum information[2,3] — a quantum bit or qubit — will require the qubit to be stored in a single spin and manipulated on a timescale in which the coherence of the spin is preserved. Press *et al.*[4] (page 218 of this issue) report that, using ultrafast laser pulses, they have controlled and observed the spin of a single electron in a semiconductor during the spin's coherence time.

The quantum state of a qubit that is based on an electron's spin can be described by a vector, known as a Bloch vector, in a sphere (a Bloch sphere), as shown in Figure 1. An arbitrary single-qubit gate operation is expressed in terms of the rotation of the Bloch vector. In general, the process is split into three rotation steps, called Euler rotations — for example, two rotations about the *x* axis and one about the *z* axis. But how can spin-state rotations with arbitrary angles about the two axes be achieved? The rotation about the *z* axis is implemented by applying a static magnetic field along this axis. This field induces the energy separation between the spin-up and spin-down states, known as Zeeman splitting, and the spin state precesses about the *z* axis with an angular frequency that is proportional to the amplitude of the field — a phenomenon known as Larmor precession. The rotation about the *x* axis is achieved by applying an oscillating (microwave) driving field that is resonant with the energy separation between the spin-up and spin-down states. This technique is called electron spin resonance (ESR). The coherent interaction between the spin state and the oscillating field results in the periodic rotation of the spin state about the *x* axis, and is called Rabi oscillation.

Although the rotation of a single electron spin has been successfully demonstrated using ESR[5,6], the time required to achieve rotation with this technique is rather long — typically longer than a few nanoseconds. For quantum information processing to remain effective, rotation must be performed within a timescale that is much shorter than the spin's decoherence time.

One way of increasing the spin-state rotation is to use ultrafast laser pulses. In place of the microwave field used in ESR, a circularly polarized laser pulse is applied along the *x* axis (Fig. 1). The laser induces transitions between the two spin states through intermediate excited states in a process called stimulated Raman adiabatic passage (STIRAP)[7,8]. The effect of STIRAP is to rotate the spin about the *x* axis, as in ESR.

Another challenge for spin-based quantum information processing is to establish a technique for observing the spin state of a single electron. Until now, optical[5,9,10] and electrical[6,11] methods have done the job. Optical pumping[5,10] in particular, which is a standard technique in atomic systems[12], can be a sensitive tool for single-spin detection. In this method, one observes the emission (or absorption) of a photon following spin-selective optical excitation by a pumping laser. A further advantage of this technique is that the same laser can be used to initialize the spin state.

Press *et al.*[4] have exploited ultrafast laser control and optical pumping to manipulate a single electron spin in a semiconductor using charged quantum dots. Quantum dots are nanometre-sized, artificially fabricated semiconductor structures in which electrons are confined in all three dimensions. The ground state of a neutral quantum dot has no net spin because it forms a singlet state with equal numbers of spin-up and spin-down electrons. When an extra electron, or an electron 'hole', is added to a neutral quantum dot, it will acquire a net charge and a spin. This can be achieved, for instance, by introducing impurities in the semiconductor, a method known as doping.

The above techniques can be used in combination with an optical microscope to optically control and observe the spin state in a single quantum dot. In Press and colleagues' experiment[4], the spin state was initialized to the spin-up state by optical pumping. The spin state was then rotated about the *x* axis by a laser pulse through STIRAP. By changing the intensity of the rotating laser pulse, the authors obtained a remarkable result: they observed up to six and a half rotations (periodic Rabi oscillations) of a single spin state.

In a second experiment, the authors observed an effect known as Ramsey interference. The initialized spin-up state was first rotated by 90° about the *x* axis, and then rotated by an arbitrary angle about the *z* axis using Larmor precession. After a certain period, it was rotated back by −90° about the *x* axis. In terms of Euler rotations, the total process corresponds to performing a rotation about the *y* axis. The spin state projected onto the *z* axis was then observed using optical pumping and photon

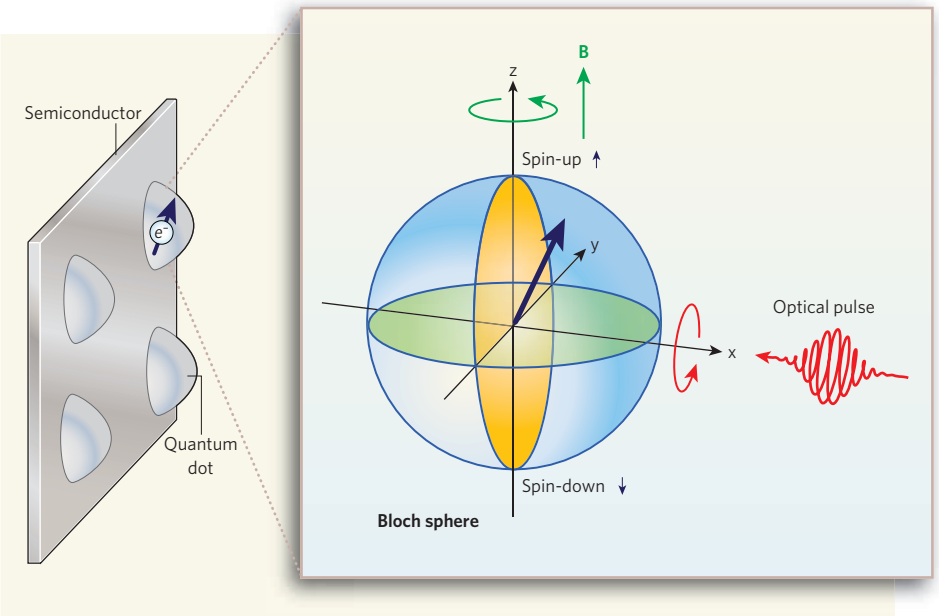


Figure 1 | Bloch sphere of an electron's spin. A semiconductor quantum dot containing an extra electron acquires a net spin. The quantum state of the electron's spin is represented by a vector (bold arrow) from the origin of the Bloch sphere to a point on its surface: the spin-up and spin-down states are at the north and south poles, respectively; and the spins that correspond to equal superpositions of the spin-up and spin-down states are in the equatorial plane. The spin rotation about the *z* axis is achieved by applying a static magnetic field (B) along the *z* axis; the spin rotation about the *x* axis is produced by a circularly polarized optical pulse injected along the *x* axis. Press *et al.*[4] demonstrate that arbitrary Euler rotations of the spin state can be accomplished by combining these two processes.

detection. The result was the observation of Ramsey fringes of interference with amplitudes that decay within about 200 picoseconds. This experiment is the first clear proof-of-principle demonstration of complete control of the single-spin state using an ultrafast laser.

There is no experiment that doesn't have a few 'buts'. First, the measurement of the spin state was obtained from a large, time-averaged ensemble of events, not from a single-shot measurement, a feat that has already been achieved using an electrical method[11]. Second, the amplitude of the Rabi oscillation fell with increasing number of rotations because of incoherent processes induced by the laser rotating the spin. Decoherence in the Ramsey fringes also seems to occur quite rapidly; the short coherence time is attributed mainly to the continuous optical pumping, and could be made longer in future experiments. If the spin coherence time in quantum dots is extended to a few microseconds[13], 10^5 single-qubit gate operations could occur within this time[4].

We have reached a stage at which we can manipulate and observe a single electron spin: albeit not perfectly, we have obtained arbitrary single-qubit gates of spins. The next step will be to realize scalable two-qubit gates, which, together with the single-qubit gates, can form a universal set for quantum computing[2,14]. Another challenge is to interface electron-spin-based qubits with other qubits, such as photons or nuclear spins, so that we can use appropriate qubits for different tasks, such as processing, communicating and storing quantum information. ■

Keiichi Edamatsu is at the Research Institute of Electrical Communication, Tohoku University, Sendai 980-8577, Japan.
e-mail: eda@riec.tohoku.ac.jp

1. http://techon.nikkeibp.co.jp/english/NEWS_EN/20080930/158806/
2. Loss, D. & DiVincenzo, D. P. *Phys. Rev. A* **57,** 120–126 (1998).
3. Hanson, R. & Awschalom, D. D. *Nature* **453,** 1043–1049 (2008).
4. Press, D., Ladd, T. D., Zhang, B. & Yamamoto, Y. *Nature* **456,** 218–221 (2008).
5. Jelezko, F. *et al. Phys. Rev. Lett.* **92,** 076401 (2004).
6. Koppens, F. H. L. *et al. Nature* **442,** 766–771 (2006).
7. Bergmann, K., Theuer, H. & Shore, B. W. *Rev. Mod. Phys.* **70,** 1003–1025 (1998).
8. Chen, P. *et al. Phys. Rev. B* **69,** 075320 (2004).
9. Berezovsky, J. *et al. Science* **320,** 349–352 (2008).
10. Atatüre, M. *et al. Science* **312,** 551–553 (2006).
11. Elzerman, J. M. *et al. Nature* **430,** 431–435 (2004).
12. Blatt, R. & Zoller, P. *Eur. J. Phys.* **9,** 250–256 (1988).
13. Greilich, A. *et al. Science* **313,** 341–345 (2006).
14. Petta, J. R. *et al. Science* **309,** 2180–2184 (2005).

ION CHANNELS

The voltage-sensor quartet

J. R. Bankston and R. S. Kass

Decoding the workings of voltage-gated sodium channels is crucial because their mutation leads to severe disease and their activity is modulated by toxins and drugs. An innovative approach now allows such investigations.

Voltage-gated ion channels are pore-forming transmembrane proteins that open and close (gate) in response to changes in transmembrane voltage, enabling carefully controlled movement of ions across cell membranes. The precise gating of these channels underlies various biological phenomena, including the generation and propagation of electrical impulses in nerve and muscle cells, the secretion of hormones and neurotransmitters, and regulation of heart muscle contraction as well as of skeletal and vascular smooth muscles. On page 202 of this issue, Bosmans *et al.*[1] deconstruct sodium channels to decipher the function of their four voltage sensors.

Sodium channels, like calcium, potassium and cyclic-nucleotide-gated channels, are composed of a tetramer of similar domains (I–IV), with each domain having six transmembrane segments. In both sodium and potassium channels, the segments within each domain can be divided into two basic units. The S5 and S6 segments of all four domains collectively constitute the pore, which allows rapid permeation of ions — more than 100 per second — across the cell membrane. By contrast, transmembrane segments S1–S4 of each domain act as voltage sensors. Within the S4 segments of each domain, positively charged amino-acid residues are interspersed with hydrophobic residues, and these positive charges play an essential part in voltage sensing[2]. Moreover, the structure of a potassium channel shows that S4 and the most extracellular part of S3 (S3b) form a voltage-sensing 'paddle', which moves across the membrane as a unit to open the channel and allow ions to flow[3,4]. Analysis of the functional roles of the voltage-sensing structures in sodium channels, however, has lagged behind, in part owing to the difficulty in studying these domains individually.

Bosmans and colleagues[1] take an original approach to investigate the gating machinery of sodium channels. Inserting individual voltage-sensor paddle domains of two different sodium channels into a potassium channel, they found that, despite significant differences in sequence, the sodium-channel paddle motifs can function as part of a potassium channel (Fig. 1, overleaf). Thus, the authors could use potassium channels as functional templates for studying the behaviour of each of the four sodium-channel voltage sensors. They investigated two aspects of S4 structure/function: how these motifs respond to changes in membrane potential, and their unique binding sites for various toxins.

Each of the four voltage sensors of a sodium channel has specific kinetics and roles in channel gating. In response to depolarization of the cell membrane, voltage sensors of domains I–III move outwards to open the channel gate and so activate the channel. Inactivation follows when the domain-IV voltage sensor, whose movement lags behind the fast motion of the other three sensors, moves outwards and puts the channel in a non-conducting state that will not reopen during sustained membrane depolarization. For reasons that are not entirely understood, this paddle gets trapped in the channel-inactivating position. The slow recovery of this sensor after membrane repolarization potentially reflects the detachment of inactivation machinery, but could also simply reflect the intrinsic nature of this paddle and the environment through which it moves. Whether this voltage paddle also somehow contributes to activation is debated[5].

Examining the kinetics of voltage-sensor movement in a chimaera of a potassium channel and the voltage paddle from domain IV of a sodium channel, Bosmans *et al.* show that slow movement of the sensor is simply an intrinsic property of this paddle, rather than a slowing of the paddle's motion by its association with the inactivation machinery. This intriguing observation is a first step towards a closer examination of the relationship between motion and inactivation of the domain-IV voltage sensor in sodium channels.

Toxins from various organisms, such as the tarantula and scorpion, target the voltage sensors of sodium channels, affecting the regulation of channel gating. Many of these toxins are also used as tools for examining sodium-channel gating, and understanding their precise site of action is important. So Bosmans *et al.* used their paddle-swapping approach to identify the binding sites of such toxins. They isolate the specific voltage sensors targeted by many of these toxins and draw two main conclusions.

First, interactions between toxins and the voltage sensors differ between sodium-channel subtypes. For instance, a toxin that interacts with domains I and II in the skeletal-muscle sodium channel may interact only with domain I in a neuronal channel. Furthermore, the specific channel residues that are essential for toxin binding are unique to each paddle–toxin pair, although the effect of a toxin on the channel mostly depends on which paddle it affects and not on the specific residues in

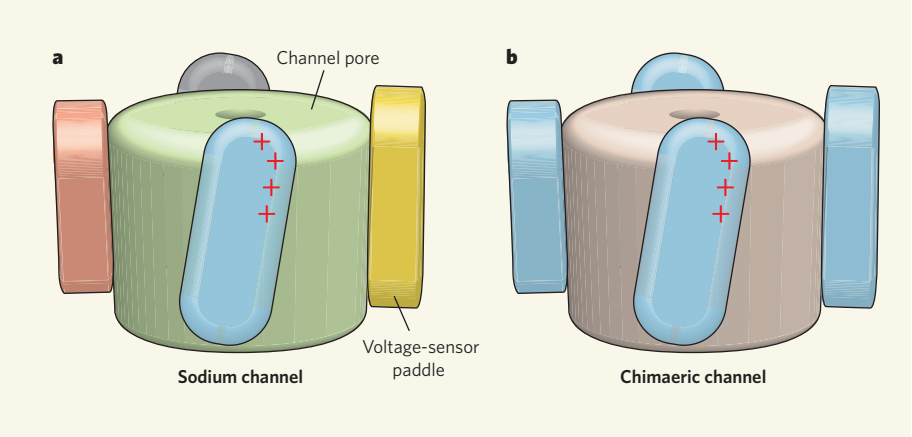


Figure 1 | Chimaera channels. a, Studying voltage sensors of sodium channels has been difficult, partly because these channels are made from four domains, each with a unique voltage paddle. **b**, Bosman *et al.*[1] describe a method for looking at voltage sensors of sodium channels in isolation by inserting each paddle in all four positions in a potassium channel. (Adapted from ref. 6.)

the paddle where the interaction occurs.

Second, the way the specific paddle targeted affects the channel is consistent with current understanding of sodium-channel gating. A toxin that binds to the paddles from domains I–III shifts the voltage range over which the channel opens, making the channel less likely to open over their typical voltage range, whereas a toxin that interacts with domain IV interferes with the channel's ability to become inactive. This observation could be relevant for designing drugs that selectively target voltage-sensor paddles to treat conditions such as epilepsy, long QT syndrome type 3 and many forms of periodic paralysis. These disorders are associated with a multitude of biophysical alterations in sodium-channel gating — including shifts in the voltage range for channel activation or defects in its inactivation — and therefore require specific and targeted treatment strategies. So Bosmans and colleagues' model system[1], in which targeted drugs and toxins can be tested for how they interact with the unique quartet of voltage sensors in sodium channels, not only contributes to our understanding of channel gating, but will also help design new treatments.

More generally, the authors' method can be extended to paddles from other channels — including the many other sodium channels, some of which gate quite differently from one another — or to make different sets of chimaeras using other transmembrane segments of sodium-channel domains. What's more, distinct channel variants with subtle differences in gating properties contribute to such diverse physiological functions as sensing pain and controlling muscle contraction. This approach should therefore greatly advance dissection of the molecular components underlying these subtle but physiologically essential functions. ■

J. R. Bankston and R. S. Kass are in the Department of Pharmacology, College of Physicians and Surgeons, Columbia University, New York, New York 10032, USA.
e-mail: rsk20@columbia.edu

1. Bosmans, F., Martin-Eauclaire, M.-F. & Swartz, K. *Nature* **456,** 202–208 (2008).
2. Yellen, G. *Q. Rev. Biophys.* **31,** 239–295 (1998).
3. Jiang, Y. *et al. Nature* **423,** 33–41 (2003).
4. Jiang, Y. *et al. Nature* **423,** 42–48 (2003).
5. Chanda, B., Asamoah, O. K. & Bezanilla, F. *J. Gen. Physiol.* **123,** 217–230 (2004).
6. Jiang, Y. *et al. Nature* **432,** 33–41 (2003).

MICROSCOPY

Watching catalysts at work

Alexis T. Bell

Solid catalysts speed up many industrial chemical reactions and steer them towards making desired products. A microscopy technique could reveal the changes in composition that catalysts undergo as they perform.

Solid catalysts are used in the production of almost every chemical compound — from plastics and fuels to pharmaceuticals — and for removing environmental pollutants. They typically consist of a porous support onto which are dispersed nanoscale particles of the active catalyst (usually a metal, metal oxide or metal sulphide). Other compounds might also be added to the mix, such as promoters to modify the reactivity of the catalysts, or structural stabilizers[1]. The size, shape and connectivity of these components affect catalytic activity, stability and selectivity (the ability of the catalyst to make a specific product). Nanoscale imaging of catalysts is therefore required to understand how each of these factors changes during a chemical reaction. Reporting on page 222 of this issue, de Smit *et al.*[2] describe how a highly focused beam of X-rays can be used to acquire chemical maps and images of working heterogeneous catalysts at nanometre spatial resolution.

Many methods have been used to try to achieve high-resolution imaging of catalysts, which is a long-standing goal. Transmission electron microscopy (TEM) enables the structure of catalyst components to be observed with a spatial resolution of 0.1 nanometres, and can yield elemental maps of a catalyst with a resolution of 1 nm (ref. 3). Such information is usually acquired under vacuum conditions, although reactors have been developed that can analyse reactions *in situ* at pressures of up to 10^{-2} bar (about one-hundredth of atmospheric pressure) and at temperatures of up to 600 kelvin. But the acquisition of TEM images at pressures of 1 bar and above — conditions more typical of those used in industry for catalysed reactions — remains extremely challenging[4]. Other techniques for imaging the structure and chemical composition of catalysts, such as a combination of scanning transmission electron microscopy with electron energy-loss spectroscopy, also offer a spatial resolution of about 1 nm, although here too, images are usually taken in a vacuum[5].

Another method has recently emerged as a contender in this field. Scanning transmission X-ray microscopy (STXM) offers exciting possibilities for imaging catalysts at nanometre resolution under a broad range of reaction conditions[6]. In this technique, a beam of monochromatic, low-energy X-rays (soft X-rays) is focused to a spot-size of 10–20 nm, using a zone plate — a device that focuses light using diffraction. The sample under investigation is then scanned by the X-ray beam (in fact, the beam is held stationary and the sample is moved relative to it), and the absorption of the beam by the sample is measured. This is repeated for X-rays of different energies, so that absorption can be plotted as a function of X-ray energy.

Because the whole sample is scanned, the relationship between absorption and beam energy can be mapped out across the sample. When such maps are produced using X-rays at the absorption edge of a particular element — that is, at the particular wavelength at which the element absorbs maximum energy from the X-rays — an image of the spatial distribution of that element is obtained. Alternatively, a plot of absorption against X-ray energy at a given spot provides an X-ray absorption

near-edge spectrum; this can be compared with the absorption spectra of standard compounds to deduce the local composition of the material under study. The observed absorption spectrum can thus be thought of as a composite of the individual spectra for components of the catalyst, with the contribution of each component weighted according to its concentration. This means that the spectral map can be used to determine the distribution of pure components throughout the sample.

The STXM technique is ideal for studying heterogeneous catalysts, but has some practical problems associated with it. Soft X-rays are readily absorbed by matter. This means that, for STXM to work, the catalyst particles must be very small, and the distance travelled by the X-rays in the reactor cell — the X-ray gas path length — must be short (less than 100 micrometres). The first STXM study[7] of a catalyst was restricted to experiments using diluted gases at temperatures up to only 533 K. Nevertheless, the technique proved its worth by characterizing the reduction and oxidation of tiny particles of copper oxide (CuO) dispersed on silicon dioxide (silica, a commonly used support for catalysts). The report of de Smit *et al.*[2] describes the design, construction and operation of a much improved STXM cell, which has a gas path length of only 50 micrometres, and which can operate at 1 bar and at temperatures of up to 773 K.

The authors used their reactor cell to examine changes in a catalyst for the Fischer–Tropsch synthesis — a reaction in which 'synthesis gas' (a mixture of carbon monoxide and hydrogen) is converted into hydrocarbons. The Fischer–Tropsch synthesis has been known since the 1920s, but is currently attracting renewed interest as a means of producing transportation fuels from coal, natural gas or biomass. The particular catalyst studied by de Smit *et al.* consists of small iron oxide particles dispersed on silica. However, these are expected to be converted into other compounds after the catalyst has been reduced with hydrogen, and after exposure to synthesis gas.

The authors simulated the reaction conditions of the Fischer–Tropsch synthesis in their cell, and studied the catalyst using X-rays at the absorption edges of carbon, oxygen and iron. In this way, they showed that the iron oxide in the freshly prepared catalyst exists in a single form (known as α-Fe_2O_3), but that, after exposure to hydrogen under the reaction conditions, the particles are reduced to yield metallic iron and a different iron oxide (Fe_3O_4); the authors also observe some formation of an iron silicate (Fe_2SiO_4). After reaction in synthesis gas, the Fe_3O_4 is further converted to iron and Fe_2SiO_4, and the iron in turn reacts to form a compound with carbon (iron carbide). The authors also see evidence for the build-up of hydrocarbon compounds. Observing how the composition of a catalyst alters with changes in reaction conditions and reaction time can reveal insight into the factors controlling catalyst activity and stability.

The current findings[2] demonstrate the potential of STXM for *in situ* chemical imaging of catalysts at the nanometre scale. The spatial resolution of the technique (15 nm) is impressive, but is still not high enough to give a truly atomic-scale view of a catalyst's structure. Improvements in X-ray optics and imaging methods, however, should allow higher spatial resolution, opening the way to a deeper understanding of the structure and composition of multi-component catalysts, and the changes they undergo during a reaction. ■

Alexis T. Bell is in the Department of Chemical Engineering, University of California, Berkeley, Berkeley, California 94720-1462, USA.
e-mail: bell@cchem.berkeley.edu

1. Bell, A. T. *Science* **299,** 1688–1691 (2003).
2. de Smit, E. *et al. Nature* **456,** 222–225 (2008).
3. Gai, P. L. *Topics Catal.* **8**(1–2), 97–113 (1999).
4. Creemer, J. F. *et al. Ultramicroscopy* **108,** 993–998 (2008).
5. Hitchcock, A. P. *et al. Micron* **39,** 311–319 (2008).
6. Kilcoyne, A. L. D. *et al. J. Synchrotron Radiat.* **10,** 125–136 (2003).
7. Drake, I. J. *et al. Rev. Sci. Instrum.* **75,** 3242–4247 (2004).

BIOCHEMISTRY

Flexible peptide assembly

Jan C. M. van Hest

A jack of all trades is a master of none, as the saying goes. But a protein has been discovered that shuns specialism, and that multitasks to give flexibility to its biosynthetic repertoire.

The biological machinery used by microorganisms to synthesize complex, antibiotic peptides is often compared to Henry Ford's car assembly line: different protein modules line up, each with their own specific synthetic job[1–3], and the molecule under construction is passed from one module to the next, accruing new parts until the desired compound is finally made. This is a sophisticated approach for constructing tailor-made peptides, but building large, multi-protein factories requires quite an investment of resources from the microorganisms involved. This raises the question of whether alternative, more efficient approaches have been adopted by nature — for example, the use of compact protein factories in which one module can execute several operations.

Hamano and colleagues[4] report in *Nature Chemical Biology* the first discovery of a protein that does exactly that. They find that the antibiotic polypeptide poly-ε-lysine (ε-PL) is constructed by a single protein structure that selects the correct building blocks, activates them and then couples them together.

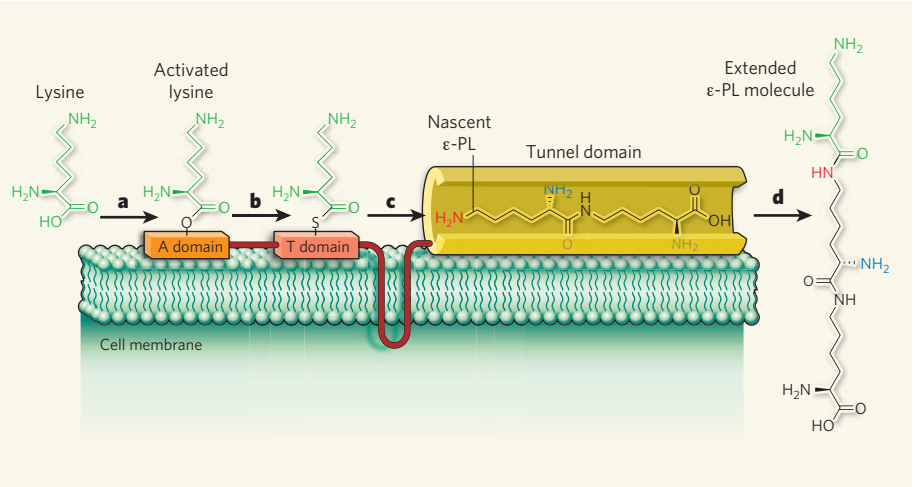


Figure 1 | Antibiotic assembly. Hamano and colleagues[4] have identified the enzyme that synthesizes the polypeptide antibiotic ε-PL, which consists of many lysine amino acids linked together through their side chains. The authors propose the following mechanism for its synthesis. **a**, The first part of the enzyme (the A domain) specifically selects lysine as a building block for ε-PL, and activates the amino acid so that it is ready to react. **b**, The activated lysine is transferred to the next part of the enzyme (the T domain), which holds it in the correct orientation for reaction with the nascent ε-PL. **c**, ε-PL fragments are held in an adjacent tunnel-shaped cavity, in such a way that the side-chain amino group (red), rather than the main-chain amino group (blue), reacts with the activated lysine. **d**, The extended ε-PL molecule may be released from the cavity, or can remain in place, ready to couple with another lysine molecule. For simplicity, the product shown contains three lysine units, but ε-PL actually contains molecules that have 25–35 lysines.

As a potent antibiotic, ε-PL has garnered much attention, and is used in some countries as a food preservative[5]. Constructed from lysine amino acids, its distinctive characteristic is that the lysines are coupled together through their side chains, rather than through groups in their main chains (as occurs between amino acids during normal ribosomal protein synthesis). This mode of coupling suggests the involvement of an enzyme assembly line — a non-ribosomal peptide synthetase (NRPS), to use the technical jargon.

Another quirk of ε-PL is that it is actually a cocktail of peptides, the components of which differ in the number of lysines that make up each molecule. The normal range of chain lengths is 25–35 amino acids. The origin of this chain-length diversity is a topic of debate: it could be the result of degradation of longer ε-PL chains, or the effect of an unusually flexible NRPS synthesis. To solve the mystery, Hamano and colleagues[4] analysed different cell extracts from an ε-PL-producing strain of the bacterium *Streptomyces albulus*. They found one insoluble fraction in which ε-PL was formed, and from this they purified a single active enzyme, which they named ε-PL-synthetase (Pls).

Using an extensive biochemical-characterization procedure, Hamano and colleagues identified within Pls several different domains that are involved in peptide synthesis (Fig. 1). One of these domains is responsible for selecting lysine amino acids (rather than any of the many other naturally occurring amino acids), and for 'activating' the carboxylic acid group in the lysine so that it can form a peptide bond. The authors dubbed this the A domain, after the analogous region in traditional NRPS enzymes. The activated lysine molecule is then connected to another domain, which holds it in the right position for coupling to the nascent ε-PL molecule. Again, by analogy with NRPSs, the authors designated this the T domain.

The recognizable A and T domains comprise the first half of Pls, but the remaining part of the protein acts differently from normal peptide assembly lines. Once activated, lysines cannot be passed on to a neighbouring protein module (as they would be in an NRPS), because there is only one protein involved in the synthesis. Instead, they react with an ε-PL molecule, which is positioned in such a way that the activated lysine can couple only with the side-chain amine of ε-PL. This part of the process is still not completely understood. The most plausible mechanism is that the second part of the Pls protein forms a long, slender tunnel or cavity, which is occupied by the growing ε-PL molecule. Unlike in NRPSs, the nascent peptide is not covalently connected to the protein. After the coupling step with the activated lysine, the extended ε-PL peptide can either be released or remain bound in the cavity. This means that the number of lysines coupled together is not strictly controlled, and would explain the diversity of chain lengths found in ε-PL.

When the working mechanisms of traditional NRPS complexes and the Pls protein are compared, some important differences can be observed. First, each NRPS assembly line allows the synthesis of a complex and perfectly defined peptide, composed of many different amino acids that can form bonds using groups in their main chains or in their side chains. The biological machinery required to create such complex structures is finely tuned, and isn't flexible enough to incorporate amino acids into the product other than those it has evolved to accept. This fine tuning also makes it difficult to modify NRPSs to make analogues of naturally occurring antibiotic peptides (which is desirable for drug discovery).

But the peptide synthesized by Pls lacks the structural complexity of traditional NRPS peptides, because it is composed solely of lysines connected through their side chains. Pls therefore needs only one amino-acid selection domain and one amino-acid activation domain (the A and T domains, respectively). Hamano and colleagues[4] also show that there is some limited flexibility in substrate selection by the enzyme: it can incorporate a few amino acids other than lysine. Furthermore, because the peptide sequence and general domain architecture of Pls are known, it should be possible to change its selectivity even further by protein engineering. Thus, we might one day be able to exploit the biosynthetic flexibility of Pls to make polypeptides with interesting biological activity — not only antibiotics, but potentially also molecules for drug delivery and gene therapy.

So there is more to Pls than meets the eye. At first glance, it might seem less impressive than NRPSs — after all, to stretch the analogy with assembly lines further, if NRPSs produce the peptide equivalent of sophisticated cars, then Pls manufactures relatively simple bicycles. But NRPSs can make only black Model T Fords; Pls meanwhile can produce everything from unicycles to tandems, and, as Hamano and colleagues have shown, might also offer a greater choice of colours. ■

Jan C. M. van Hest is in the Department of Organic Chemistry, Institute for Molecules and Materials, Radboud University Nijmegen, Heyendaalseweg 135, Nijmegen 6525 AJ, the Netherlands.
e-mail: j.vanhest@science.ru.nl

1. Llewellyn, N. M. & Spencer, J. B. *Nature* **448,** 755–756 (2007).
2. Sieber, S. A. & Marahiel, M. A. *Chem. Rev.* **105,** 715–738 (2005).
3. Staunton, J. & Weissman, K. J. *Nat. Prod. Rep.* **18,** 380–416 (2001).
4. Yamanaka, K., Maruyama, C., Takagi, H. & Hamano, Y. *Nature Chem. Biol.* doi:10.1038/nchembio.125 (2008).
5. Oppermann-Sanio, F. & Steinbuchel, A. *Naturwissenschaften* **89,** 11–22 (2002).

NEUROSCIENCE

Cool songs

Chris M. Glaze and Todd Troyer

Cooling a specific cluster of neurons in songbirds' brains slows song tempo without changing other acoustic features. This clever technique could be used for understanding neural control of other complex behaviours.

Complex behaviours, ranging from speech and typing to dancing and swimming, all require careful coordination of dozens and often hundreds of different muscles. The resulting behaviour frequently shows hierarchical structure, with sequences composed of primitive movements, each of which is composed of even more basic movements. How is all this coordinated? Researchers have long postulated that the timing of behavioural components is encoded in specialized brain circuits. Without such circuits, movement would quickly become disorganized, like an orchestra performing without a conductor. On page 189 of this issue, Long and Fee[1] provide a remarkable example of such specialization in the brain of the zebra finch by using an equally remarkable experimental approach.

The zebra finch's song consists of a rich set of acoustic features produced with very precise timing. Yet localizing the brain regions that coordinate these features has been tricky. A traditional approach is to measure how neural lesions or electrical stimulation alter song output. But with complex behaviours such as birdsong, such perturbations will often initiate a cascade of disturbance that causes the action to fall apart, making it impossible to isolate the function of the brain area that was initially perturbed.

To circumvent this problem, Long and Fee put a new twist on an old technique. Localized cooling has long been used as a method to selectively and reversibly inactivate particular brain areas[2]. Long and Fee, however, applied only slight cooling, focusing on nucleus HVC — an area of the bird's brain crucial to song production. They find that cooling HVC in this way stretches the temporal fabric of the song, slowing tempo but leaving acoustic features such as pitch largely unchanged. Moreover, the degree of cooling determined the amount of slowing, allowing the authors to effectively 'dial in' a desired tempo.

NATURE|Vol 456|13 November 2008

The results challenge previous notions about how HVC contributes to the hierarchical arrangement of song acoustics. Songs are organized into motifs, consisting of stereotyped sequences of syllables; syllables in turn are composed of one or more vocalizations with distinct acoustic features called notes. Early studies had suggested[3,4] that this organization mapped directly onto different brain areas, with nucleus HVC driving syllable sequence and timing, and RA (the robust nucleus of the arcopallium), a region directly downstream of HVC, determining how each syllable is produced. Long and Fee reasoned that if this were the case, cooling HVC would only stretch the times between syllable onsets, lengthening silent gaps, and would leave syllable lengths unchanged (Fig. 1a,b). However, they find that cooling HVC stretches syllables and gaps alike (Fig. 1c), whereas cooling RA does not affect the song tempo. It therefore seems that nucleus HVC plays a central part in controlling song timing at all levels of the song hierarchy.

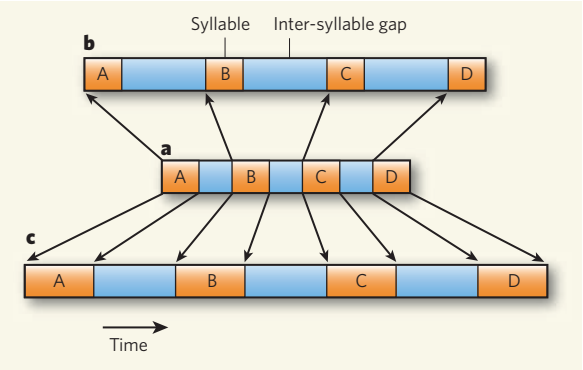


Figure 1 | Effects of cooling on song. **a**, A normal song can be divided into intermittent syllables separated by inter-syllable gaps. **b**, It was thought that the song nucleus HVC controls only the onset times of syllables, and so its cooling would lengthen the inter-syllable gaps without affecting syllable lengths. **c**, Long and Fee[1] show that, in fact, HVC influences timing throughout the song as its cooling stretches both syllables and the gaps between them.

The authors also find that the lengthening of any given song segment in response to cooling is proportional to its original length. Studies of motor control in humans have shown a similar scaling of change in timing with the duration of movement, and several researchers have proposed that this scaling property is a signature of motor-planning circuits[5,6]. Previous recordings in HVC suggested[7] that song is driven by the serial activation of distinct sets of neurons every 5–10 milliseconds. The most likely explanation is that these neuronal populations are chained together in series, and together constitute a 'motor tape' that is read by RA to produce specific combinations of song features. Slow down the tape and you slacken the song tempo. The longer the song segment, the more motor tape it uses and so the longer it will become on cooling.

But there is a major complication with the hypothesis that timing is localized in HVC: there are actually two HVCs, one in each hemisphere of the bird's forebrain. So are there really two motor tapes producing song? What if each has a slightly different tempo during a song rendition, just by chance? Long and Fee examined this by cooling each HVC separately, and found that, indeed, both contribute to tempo, with the left HVC being dominant during some parts of the song, whereas the right HVC prevails during others.

What keeps the two HVCs from drifting completely out of alignment? There are no direct connections between the two halves of the brain at the level of HVC and RA, so synchronizing signals must travel through bilateral circuitry in the brainstem and thalamus that is downstream of RA (Fig. 2). This information can then be fed back up to both HVCs. Intriguingly, similar circuits that loop through the brainstem are known[8] to be instrumental in coordinating other complex behaviours, such as learned eye-movement sequences in primates. But figuring out how such circuits work can be mind-bending: during the song, the input of signals to a given nucleus — whether HVC, RA or those within the brainstem — can be influenced by activities anywhere else within the circuit, including an earlier output of the same nucleus itself. How can one particular area initiate timing?

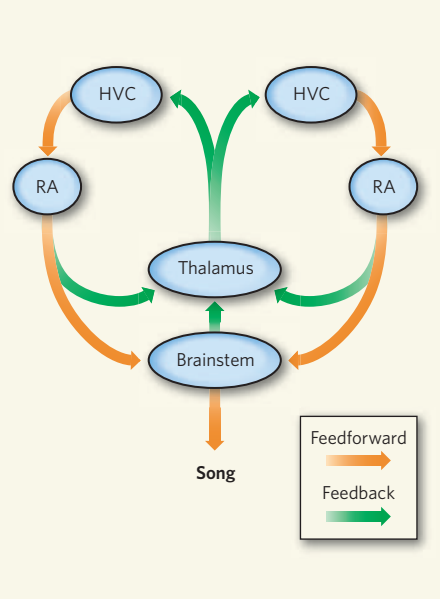


Figure 2 | Song circuitry. As shown in this simplified diagram, the feedforward pathway mediating singing in birds consists of inputs from nucleus HVC to nucleus RA to motor neurons in the brainstem nuclei. RA also sends feedback signals to HVC by way of the brainstem and thalamus. Long and Fee[1] show that cooling either left or right HVC can slow song, whereas cooling RA alone has no influence. As there are no direct connections between the left and right hemispheres at the level of HVC or RA, feedback signals passing through the brainstem or thalamus have an essential role in synchronizing the two sides of the song motor circuit.

Simultaneous recordings in both HVCs of singing birds have indicated[9] that synchronization may occur at discrete points in a song, such as syllable onsets. Drawing on this finding, Long and Fee propose that synchronizing signals originate in both HVCs and rapidly cycle through the brainstem, crossing over to the other side, to ensure that each HVC continues on the same beat. But it is also possible that the brainstem itself initiates synchronization. If so, cooling particular brainstem areas might also slow tempo, albeit with different patterns of stretching — such as that shown in Figure 1b. Other studies have suggested[10] that the influence of brainstem feedback signals is not restricted to specific times during the song. So another possibility is that brainstem feedback acts throughout the song to help keep the two HVCs in step.

The zebra finch has provided a wonderful model for studying a temporally complex behaviour. By demonstrating that nucleus HVC can influence song tempo at all levels of song organization, Long and Fee have taken a crucial step towards deciphering the motor code for song production. Although the findings leave open many questions about the coordination of neural signals during song production, the cooling method itself points in an exciting direction for future studies: by slowing down the neural processing in one brain area, simultaneous electrophysiological recordings from other areas can be used to track how distortion of timing propagates through the system. Given that mild cooling has been shown to slow down neural processing in various systems[11], 'cool and slow' has tremendous potential for understanding the temporal coordination of brain activity. ■

Chris M. Glaze is in the Department of Biology, University of Pennsylvania, 312 Leidy Laboratories, Philadelphia, Pennsylvania 19104, USA. Todd Troyer is in the Department of Biology, University of Texas San Antonio, One UTSA Circle, San Antonio, Texas 78249, USA.
e-mails: cglaze@sas.upenn.edu; todd.troyer@utsa.edu

1. Long, M. A. & Fee, M. A. *Nature* **456,** 189–194 (2008).
2. Mark, V. H., Chato, J. C., Eastman, F. G., Aronow, S. & Ervin, F. R. *Science* **134,** 1520–1521 (1961).
3. Vu, E. T., Mazurek, M. E. & Kuo, Y. *J. Neurosci.* **14,** 6924–6934 (1994).
4. Yu, A. C. & Margoliash, D. *Science* **273,** 1871–1875 (1996).
5. Wing, A. M. & Kristofferson, A. B. *Percept. Psychophys.* **14,** 5–12 (1973).
6. Ivry, R. B. & Richardson, T. C. *Brain Cogn.* **48,** 117–132 (2002).
7. Hahnloser, R. H. R., Kozhevnikov, A. A. & Fee, M. S. *Nature* **419,** 65–70 (2002).
8. Sommer, M. A. & Wurtz, R. H. *Science* **296,** 1480–1482 (2002).
9. Schmidt, M. F. *J. Neurophysiol.* **90,** 3931–3949 (2003).
10. Ashmore, R. C., Wild, J. M. & Schmidt, M. F. *J. Neurosci.* **25,** 8543–8554 (2005).
11. Janssen, R. *Neurosci. Biobehav. Rev.* **16,** 399–413 (1992).

The origin of hydrogen around HD 209458b

Arising from: M. Holmström *et al.* *Nature* 451, 970–972 (2008)

Using numerical simulation, Holmström *et al.*[1] proposed a plausible alternative explanation of the observed Lyman-α absorption that was seen during the transit of HD 209458b (ref. 2). They conclude that radiation pressure alone cannot explain the observations and that a peculiar stellar wind is needed. Here we show that radiation pressure alone can in fact produce the observed high-velocity hydrogen atoms. We also emphasize that even if the stellar wind is responsible for the observed hydrogen, to have a sufficient number of atoms for charge exchange with stellar wind, the energetic neutral atom (ENA) model also needs a significant escape from the planet atmosphere of similar amplitude as quoted in ref. 2.

The simulation of ref. 1 is aimed at reproducing the observed absorption spectrum in Lyman-α with 15 ± 4% absorption between −130 and 100 km s^{-1} (refs 2, 3). A mechanism is needed to produce hydrogen atoms at these high velocities exceeding the planet escape velocity. We previously proposed that hydrogen atoms in the exosphere are naturally accelerated by the stellar radiation pressure[2,4]; however, Holmström *et al.*[1] concluded that radiation pressure alone cannot explain the observation. Nonetheless, in their work, the strength of the radiation pressure has been artificially reduced to a value 2 to 5 times lower than the solar value, whereas the observed Lyman-α line strength and profile shows that it is significantly larger than the solar value. The low radiation pressure assumed by Holmström *et al.*[1] is valid only at high radial velocity. However, if low radiation pressure is assumed, high velocities are not reached, which therefore explains the different conclusion reached by Holmström *et al.*[1]. We believe that the treatment of the link between the radiation pressure and radial velocity needs to be corrected.

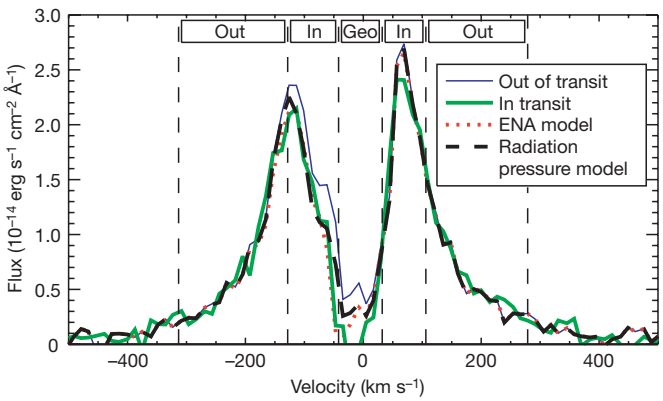


Figure 1 | Comparison of modelled with observed Lyman-α profiles as in Fig. 3 of ref. 1. The thin blue line and the thick green line are for the observed profile before and during the transit, respectively. The red dotted line is for the modelled profile from the ENA model with reduced radiation pressure. The black dashed line is for the modelled profile computed using a model with radiation pressure. This last profile fits the data well with a χ^2 of 45 for 40 degrees of freedom. As the profile for the radiation pressure model is similar to the one for the ENA model, neither possible model can be favoured. Radiation pressure cannot be excluded as an explanation for the observed spectrum. Geo, the wavelength domain contaminated by the geocoronal airglow.

To show that the radiation pressure can explain the observed spectrum, we calculated the modelled Lyman-α profile with radiation pressure alone, in the same way as done in Fig. 3 of ref. 1 for the ENA model. This calculation is done taking into account the strength and profile of the Lyman-α line, and the corresponding variation of radiation pressure as a function of radial velocity. Planetary and stellar gravities are also included. These differences explain the different results obtained with radiation pressure alone in the two models. The result plotted in Fig. 1 shows that the resulting profiles are similar in the two models (radiation pressure alone and ENA with reduced radiation pressure), and neither possible model can be favoured. Radiation pressure cannot be excluded as an explanation of the observed spectrum.

Although we agree that the ENA model is a plausible scenario, we do not believe that ENAs can explain the observations better than a classical scenario with radiation pressure. The ENA model requires extraordinary conditions for the wind parameters (high temperature and low velocity) which are not constrained by any other observations, whereas the radiation pressure as measured in the Lyman-α spectrum can self-consistently explain the observations.

A. Lecavelier des Etangs[1,2], A. Vidal-Madjar[1,2] & J.-M. Désert[1,2]
[1]CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98bis boulevard Arago, F-75014 Paris, France.
[2]UPMC Université Paris 6, UMR 7095, Institut d'Astrophysique de Paris, 98bis boulevard Arago, F-75014 Paris, France.
e-mail: lecaveli@iap.fr

Received 9 May; accepted 27 August 2008.

1. Holmström, M. *et al.* Energetic neutral atoms as the explanation for the high-velocity hydrogen around HD 209458b. *Nature* **451**, 970–972 (2008).
2. Vidal-Madjar, A. *et al.* An extended upper atmosphere around the extrasolar planet HD209458b. *Nature* **422**, 143–146 (2003).
3. Vidal-Madjar, A. *et al.* Exoplanet HD 209458b (Osiris). Evaporation strengthened. *Astrophys. J.* **676**, L57–L60 (2008).
4. Vidal-Madjar, A. & Lecavelier des Etangs, A. "Osiris" (HD209458b), an Evaporating Planet. *ASP Conf. Series* **321**, 152–159 (2004).

doi:10.1038/nature07402

Holmström et al. *reply*

Replying to: A. Lecavelier des Etangs, A. Vidal-Madjar & J.-M. Désert *Nature* 456, doi:10.1038/nature07402 (2008)

Lecavelier des Etangs *et al.*[1] object to the conclusion by Holmström *et al.*[2] that radiation pressure alone cannot explain the Lyman-α absorption observed[3] during transits of HD 209458b. We agree that hydrogen atoms can be accelerated to large velocities by radiation pressure. However, with our model we cannot reproduce the observed spectrum, as shown in the Supplementary Information and Fig. 3 of ref. 2.

To support the hypothesis that radiation pressure alone can explain the observation, Lecavelier des Etangs *et al.* show a modelled spectrum that fits well with the observed spectrum[1]. Thus, there is a

difference between the two models, and it should be possible to resolve this discrepancy upon the publication of a full description of the model used by Lecavelier des Etangs *et al.*[1].

Furthermore, Lecavelier des Etangs *et al.*[1] state that the energetic neutral atom (ENA) model needs a significant escape from the planet's atmosphere. This would seem to be incorrect, as the only requirement for ENA production is that a sufficient number of hydrogen atoms are available for charge exchange with the stellar wind. This does not put any strong constraints on the escape of the planet's atmosphere. ENA production will occur independently of a large or small thermal escape rate, but the focus of Holmström *et al.*[2] was not to study the escape rate in detail.

Lecavelier des Etangs *et al.*[1] also suggest that the stellar wind conditions derived from the model of Holmström *et al.*[2] are extraordinary. We agree that a stellar wind velocity of 50 km s^{-1} is low compared to the solar wind. This might be an effect of the simplified stellar wind flow model that we used (constant stellar wind outside an obstacle). A more detailed model that includes the stellar wind flow around the obstacle might change the estimation of the stellar wind velocity. This is a topic of future investigations.

M. Holmström[1], A. Ekenbäck[1], F. Selsis[2,3], T. Penz[4], H. Lammer[5] & P. Wurz[6]

[1]Swedish Institute of Space Physics, PO Box 812, SE-98128 Kiruna, Sweden.
e-mail: matsh@irf.se
[2]Laboratoire d'Astrophysique de Bordeaux (CNRS, Université Bordeaux 1), BP 89, F-33270 Floirac, France.
[3]Centre de Recherche Astrophysique de Lyon (CNRS, Université de Lyon, Ecole Normale Supérieure de Lyon), 46 Allée d'Italie, F-69007 Lyon, France.
[4]INAF—Osservatorio Astronomico di Palermo, Piazza del Parlamento 1, I-90134 Palermo, Italy.
[5]Space Research Institute, Austrian Academy of Sciences, Schmiedlstrasse 6, A-8042 Graz, Austria.
[6]Physikalisches Institut, University of Bern, Sidlerstrasse 5, CH-3012 Bern, Switzerland.

1. Lecavelier des Etangs, A., Vidal-Madjar, A. & Désert, J.-M. The origin of hydrogen around HD 209458b. *Nature* **456**, doi:10.1038/nature07402 (2008).
2. Holmström, M. *et al.* Energetic neutral atoms as the explanation for the high-velocity hydrogen around HD 209458b. *Nature* **451**, 970–972 (2008).
3. Vidal-Madjar, A. *et al.* An extended upper atmosphere around the extrasolar planet HD209458b. *Nature* **422**, 143–146 (2003).

doi:10.1038/nature07403

ARTICLES

Using temperature to analyse temporal dynamics in the songbird motor pathway

Michael A. Long[1] & Michale S. Fee[1]

Many complex behaviours, like speech or music, have a hierarchical organization with structure on many timescales, but it is not known how the brain controls the timing of behavioural sequences, or whether different circuits control different timescales of the behaviour. Here we address these issues by using temperature to manipulate the biophysical dynamics in different regions of the songbird forebrain involved in song production. We find that cooling the premotor nucleus HVC (formerly known as the high vocal centre) slows song speed across all timescales by up to 45 per cent but only slightly alters the acoustic structure, whereas cooling the downstream motor nucleus RA (robust nucleus of the arcopallium) has no observable effect on song timing. Our observations suggest that dynamics within HVC are involved in the control of song timing, perhaps through a chain-like organization. Local manipulation of brain temperature should be broadly applicable to the identification of neural circuitry that controls the timing of behavioural sequences and, more generally, to the study of the origin and role of oscillatory and other forms of brain dynamics in neural systems.

Motor behaviours are built out of a sequence of movements that evolve through time. From the most basic, such as locomotion, to the most complex, such as playing the piano, the timing of movements is crucial. For some simple oscillatory behaviours, in which the movement evolves on a single timescale, it has been possible to identify the particular neurons and biophysics that control the temporal dynamics of the behaviour—for example pacemaker neurons in the stomatogastric ganglion[1] or the oscillator network that controls swimming in the leech[2]. However, it is not known what mechanisms underlie more complex learned behaviours that have structure on many timescales.

Birdsong has a remarkably precise and hierarchically organized temporal structure[3,4] mediated by a number of distinct, well-studied motor nuclei[5,6] (Fig. 1a), allowing for an unprecedented view into the central control of motor timing. Adult zebra finches generate a 0.5–1.0-s song motif that is repeated a number of times during a bout of singing[7]. The motif itself is made up of song syllables—individual bursts of sound that are approximately 100 ms in length and occur in a precise order. Syllables are highly stereotyped and often contain complex acoustic structure that can evolve rapidly (10-ms timescale). The duration of song elements at all timescales is stereotyped; trial-to-trial fractional variations in song timing are roughly 1% (refs 8–10)z.

It is not known whether different brain regions are responsible for the timing of motifs, syllables, and subsyllabic structure. Two forebrain nuclei in particular have been implicated in the control of the temporal structure of song: HVC and RA. HVC projects to RA, which in turn projects to the vocal motor neurons[11] as well as midbrain vocal control and brainstem respiratory areas[12]. Previous electrophysiological studies have found evidence that these brain regions contribute to song structure in a hierarchical manner[13,14] and have suggested that the dynamics underlying the generation of different song timescales may reside in different brain regions. For instance, syllable-timescale dynamics have been suggested to occur in HVC[15], whereas subsyllable-timescale dynamics may arise in RA[14,16,17]. Additionally, portions of the midbrain and respiratory areas project back to HVC through thalamic nucleus Uvaeformis (Uva)[18,19], raising the further possibility that syllables, which are tightly linked to respiratory patterning[20], may be timed by respiratory oscillator circuits[9,21]. With current techniques, however, it has been difficult to test ideas about the origin of dynamics that underlie the temporal control of song.

Localizing temporal dynamics with temperature

We set out to localize temporal dynamics within the song control system, taking advantage of the fact that the speed of brain processes is strongly temperature dependent[22–24]. The aim was to produce localized mild heating or cooling[25,26], rather than inactivation, for which cooling has also been used[27]. The basic logic of our experiments is as follows. If the circuitry in a particular brain area is involved in controlling song timing, then cooling that area should slow the song. Furthermore, if the neural control of song is organized with a dynamical hierarchy (that is, different song timescales are controlled by biophysical dynamics in different brain areas), it should be possible to differentially alter the behavioural timescales by individually manipulating the temperature in these areas.

Dynamics in HVC

We started by bilaterally manipulating the temperature of nucleus HVC. We designed a device, based on the Peltier effect, that is capable of rapidly heating or cooling HVC in a spatially restricted manner (Fig. 1a–c, Supplementary Fig. 1). Song timing was strongly affected by changes in HVC temperature. At colder temperatures, song motifs were produced more slowly than control songs (Fig. 1d, Supplementary Fig. 2). All birds ($n = 10$) showed a significant increase in motif duration during cooling (ranging from 16.9 to 44.9%; Fig. 1e). Fractional change in motif duration (dilation) was found to vary approximately linearly with temperature in the range from 0 to $-6.5\,^{\circ}\mathrm{C}$ (0 to 1 A in terms of current; Fig. 1f). The slope of this relation was used as a simple metric of temperature-dependent song dilation, which we refer to as stretch (measured in per cent per degree Celsius; see Supplementary Methods). The stretch metric in different birds ranged from -1.89 to $-3.97\%\,^{\circ}\mathrm{C}^{-1}$, with a mean of $-2.83 \pm 0.22\%\,^{\circ}\mathrm{C}^{-1}$. Changes in song speed during cooling were immediate and persisted for an hour or more (Supplementary Fig. 1d). Notably, temperature changes in HVC had only a small effect on the acoustic structure of the song (Supplementary Fig. 4).

[1]McGovern Institute for Brain Research, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

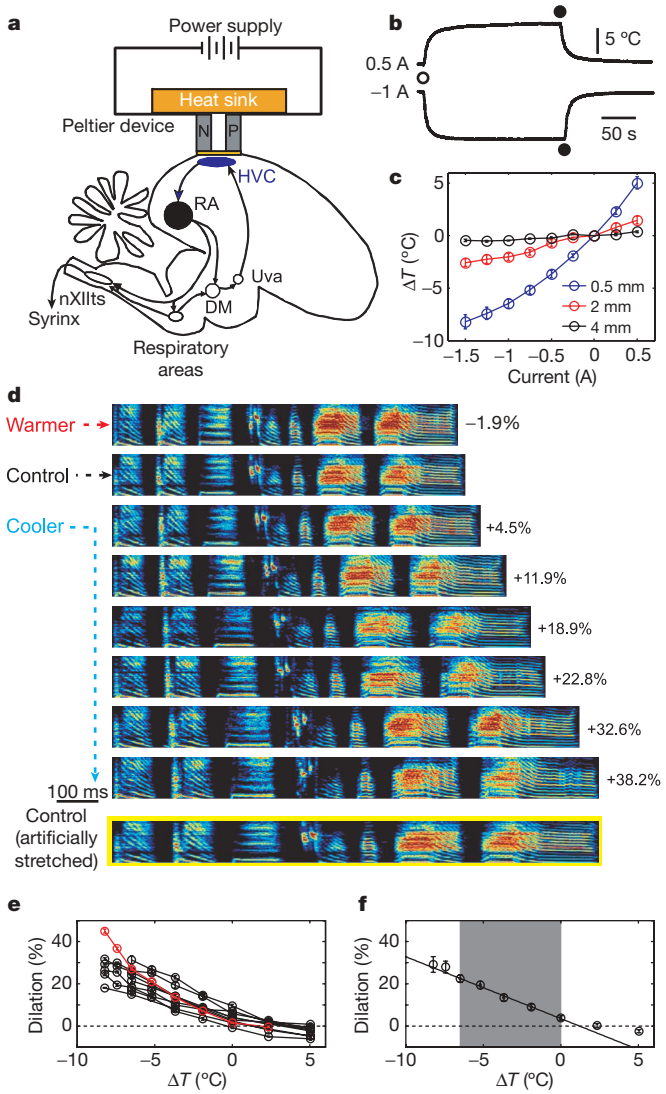


Figure 1 | Changes in HVC temperature affect song duration. **a**, The Peltier device and relevant parts of the song production pathway. N and P, semiconductor elements; DM, dorsomedial nucleus of the intercollicular complex; nXIIts, tracheosyringeal part of the hypoglossal nucleus. **b**, Temperature change in HVC as a function of time after onset (open circle) of the indicated current through the Peltier device (top, heating; bottom, cooling); current switched off at filled circles. **c**, Calibration curves for brain temperature changes (ΔT) at various depths under the Peltier device ($n = 4$). **d**, Representative sonograms (frequency, 1–9 kHz) recorded from bird no. 3 with HVC heated (0.25 A) and cooled (0.25 to 1.5 A in 0.25-A steps), showing percentage song dilation relative to control. Hotter colours represent greater sound intensity. Bottom, spectrogram of the control motif shown artificially stretched. **e**, Percentage change in duration (dilation) of song motif versus change in temperature, relative to the pre-implantation song ($n = 10$). Red, bird no. 3. **f**, Motif dilation averaged over all ten birds. The shaded area represents the range over which the song stretch metric (slope) was calculated. All error bars, s.e.m.

To quantify whether HVC cooling slows the song even at the shortest timescale of subsyllabic structure, we used a standard dynamic time warping algorithm based on the correlation of sound features of the control song with the cooled song (Supplementary Methods, Supplementary Fig. 5), and also directly measured the duration of subsyllabic elements. The average dilation of subsyllabic structure for each song syllable was computed at each temperature condition (0 to −6.5 °C; Fig. 2a), and the slope (the stretch metric) of the dilation as a function of temperature was computed for each syllable (Fig. 2d). The mean stretch for subsyllabic structure was found to be $-2.88 \pm 0.12\%\ ^{\circ}\mathrm{C}^{-1}$, which differs significantly from zero (t-test,

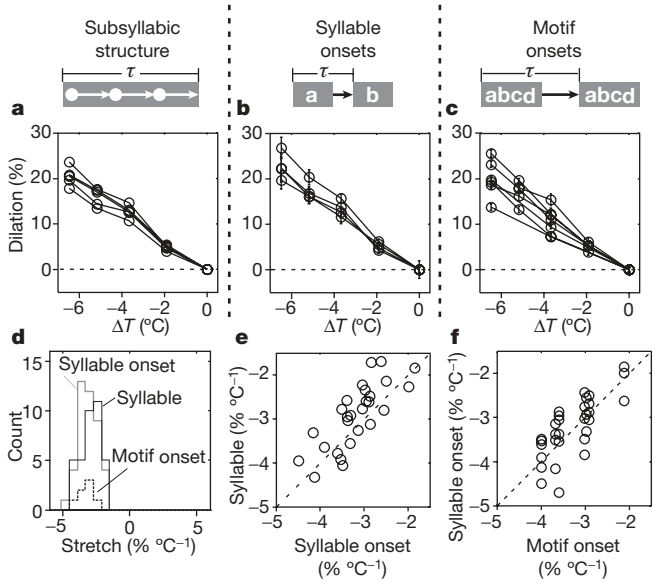


Figure 2 | HVC cooling slows the song at all timescales. **a**, Dilation of subsyllabic structure versus HVC temperature change for all five syllables of bird no. 8. **b**, Dilation of syllable-onset intervals for the bird no. 8. **c**, Dilation of motif-onset intervals for all seven birds that produced concatenated motifs at all temperatures. All error bars, s.e.m. **d**, Distribution of stretch metrics for the entire data set, including syllables (36 syllables, eight birds), syllable onsets (43 syllables, nine birds) and motif onsets (seven birds). **e**, **f**, Stretch of syllable-onset interval was strongly correlated with subsyllabic stretch (**e**) and motif-onset stretch (**f**) (for further details, see Supplementary Information).

$P < 10^{-6}$). This observation suggests that biophysical dynamics in HVC are involved in controlling song timing on a fine timescale.

We then considered the control of syllable onsets. In principle, respiratory circuits projecting to HVC (for example through Uva) could act as a 'clock' that autonomously controls the initiation of syllables, in which case cooling HVC should have little effect on syllable onsets. Alternatively, the onset of a syllable may be linked to the completion of the previous syllable, in which case the interval between onsets should increase during cooling, as the duration of each syllable increases. In fact, the intervals between syllable onsets were significantly dilated by an average of $-3.05 \pm 0.11\%\ ^{\circ}\mathrm{C}^{-1}$ (t-test, $P < 10^{-6}$; Fig. 2b, d). This is not consistent with a model in which syllable timing (or the timing of singing-related respiration) is autonomously controlled by circuit dynamics in respiratory circuits or any other area upstream or downstream of HVC.

The stretch of syllable-onset intervals was significantly correlated with the stretch of the syllables within the intervals ($r^2 = 0.607$, slope of 0.92 ± 0.16; Fig. 2e), and more weakly correlated with the stretch of other syllables ($r^2 = 0.47$; see Supplementary Materials). In other words, for syllables that had a larger stretch than average, the onset interval to the following syllable also had a larger stretch than average. This is consistent with a model in which the onset of each syllable may be causally linked to, or triggered by, the end of the previous syllable.

The silent gaps between syllables were also significantly dilated by HVC cooling ($-3.70 \pm 0.32\%\ ^{\circ}\mathrm{C}^{-1}$; t-test, $P < 10^{-12}$, median of $-3.29\%\ ^{\circ}\mathrm{C}^{-1}$), suggesting that biophysical dynamics in HVC are involved in the timing of gaps. The cooling-related stretch of gaps was slightly larger than that observed for syllables (paired t-test, $P < 0.05$, median gap stretch was 12% larger than median syllable stretch). Similarly, the average stretch of a syllable-onset interval was slightly larger than the stretch of the syllable contained within that interval (mean paired difference, $0.26 \pm 0.088\%\ ^{\circ}\mathrm{C}^{-1}$; paired t-test, $P < 0.001$, median interval stretch was 3.4% larger than median syllable stretch; Supplementary Fig. 6). These observations imply that the circuit mechanisms involved in initiating song syllables may be different from those involved in generating structure within song

syllables, as has been suggested by measurements of the variability in timing of gaps and syllables in natural singing[8,9].

An analogous argument can be made for the timing of motif onsets; if there is a 'motif clock' outside HVC that independently controls the intervals between motif onsets, then cooling HVC should have little effect on motif-onset intervals. In fact, motif-onset intervals were significantly dilated by an average of $-3.19 \pm 0.24\%\ ^{\circ}C^{-1}$ (t-test, $n = 7$, $P < 10^{-5}$; Fig. 2c, d, f), which is not consistent with a model in which motif onsets are timed autonomously by circuit dynamics outside HVC.

Dynamics in RA

Although the HVC cooling experiments strongly suggest the involvement of HVC in generating the fine temporal structure within syllables, they do not rule out some involvement of other brain areas. In particular, circuit dynamics[14,16,17] and connectivity[28,29] within RA, as well as reciprocal connections from RA to HVC[30], have been implicated in the generation of these short timescales. In general, these models would predict that song timing can be affected by manipulating circuit dynamics in RA. Here we directly test this prediction by bilaterally cooling RA during singing. We use a Peltier device similar to that used for HVC, but with attached gold probes (330-μm diameter) that were implanted into RA to facilitate thermal conduction (Fig. 3a, Supplementary Figs 7, 8). At a distance of 200 μm from the probe, the distance estimated to be the farthest extent of RA neurons, we observed a temperature drop of 10 °C at the maximum current used. We also found that the RA cooling device produced a slight temperature change in HVC of roughly 30% of the temperature change in RA at each current level (Fig. 3b).

As expected, because of the residual effect of the RA cooling probe on HVC temperature, we observed a slight increase in motif duration at higher cooling currents ($n = 4$; Fig. 3c, Supplementary Fig. 9a). The effect of the residual HVC cooling on motif duration can be accurately predicted by the results of the HVC cooling experiments (Fig. 1f) and fully accounts for changes in motif duration produced by the RA probe (Fig. 3d). Thus, after incorporating a correction for HVC temperature changes, we find no evidence that changes in RA temperature affect song motif duration ($P > 0.20$; Fig. 3e) or the timing of subsyllabic structure (Fig. 3f), suggesting that dynamics in RA may not contribute significantly to song timing, at least by any mechanism that is sensitive to temperature changes in the range we were able to achieve here.

The fact that RA cooling had so little effect on song structure led us to wonder whether our temperature manipulation had any effect on the neuronal properties in RA. In non-singing birds, RA neurons spontaneously generate tonic, regular spiking[14], possibly associated with an intrinsic subthreshold membrane potential oscillation[31]. We measured the spiking frequency of single units in RA in an anaesthetized preparation while changing the temperature using the RA cooling device. Cooling produced a rapid, roughly linear decrease in RA neuron tonic spiking rate (19 cells from seven birds, slope of $0.85\ Hz\ ^{\circ}C^{-1}$; Fig. 4a–c and Supplementary Fig. 9b) that resulted in a near cessation of spontaneous spiking at the coldest temperatures ($\Delta T = -10\ ^{\circ}C$). Our observation that cooling RA by 10 °C produces a 2.5-fold reduction in the intrinsic oscillation frequency of RA neurons, yet has no detectable effect on song structure or timing, implies that these oscillations are not likely to be a source of dynamics underlying song production.

In contrast, the incidence of high-frequency spontaneous bursts in RA (Fig. 4d, top), known to be driven by synaptic input from HVC under anaesthesia and during sleep[32,33], does not show a significant trend with temperature ($P > 0.6$; Fig. 4d). The bursts exhibited only a slight cooling-related decrease in firing rate ($5.6\ Hz\ ^{\circ}C^{-1}$; Supplementary Fig. 9c). RA is thus capable of a robust response to burst input from HVC, even at temperatures low enough to substantially suppress tonic spiking.

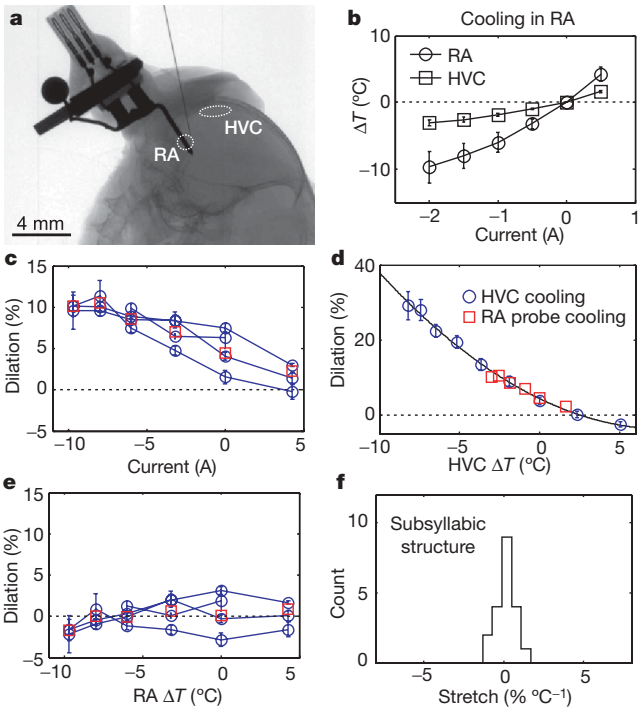


Figure 3 | Effects of RA temperature change on song timing. **a**, X-ray image of the implanted RA cooling device and approximate locations of HVC and RA. **b**, Temperature in RA (200 μm from cooling probe) and HVC as a function of RA probe current. We note that the RA probe produces some cooling in HVC. **c**, Change in motif duration as a function of RA probe current ($n = 4$; red squares, mean). **d**, Average change in motif duration (red squares) during RA probe cooling or heating, plotted as a function of HVC temperature. Also plotted is the average change in motif duration (blue circles) as a function of HVC temperature measured in the HVC cooling experiment (Fig. 1f). **e**, Change in motif duration as a function of RA temperature, corrected for the effect of HVC temperature change. All error bars, s.e.m. **f**, Stretch of subsyllabic elements for the population of RA cooled birds ($n = 4$, 20 syllables), corrected for HVC temperature change.

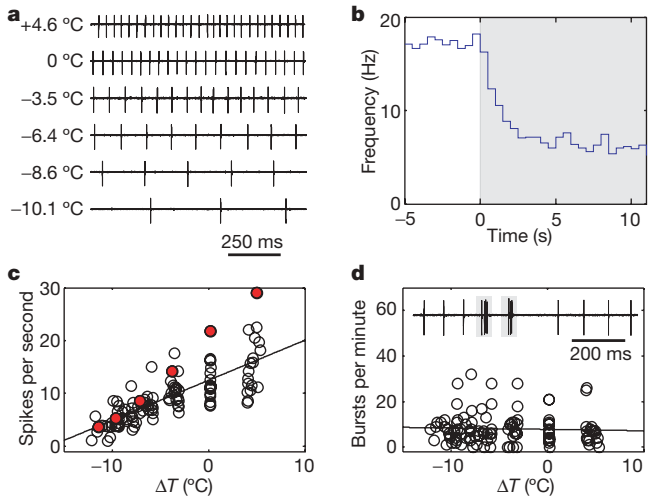


Figure 4 | Effects of RA temperature change on RA spiking activity. **a**, An example of the tonic spiking activity of an RA neuron in an anaesthetized bird for various temperature changes. **b**, Average firing rate response (25 trials) to the application of 1-A cooling current to the RA probe. **c**, Average tonic spiking rate versus temperature for all recorded neurons (19 cells, seven birds). Filled-red circles are from the example shown in **a**. **d**, Spike train showing tonic spiking and spontaneous bursts (top) and incidence of bursts (defined as an instantaneous firing rate greater than 100 Hz) for all neurons (bottom).

The control of song timing by HVC: lateralization

The HVC and RA cooling experiments highlight the centrality of HVC in controlling song timing. HVC is a bilateral structure, and it is natural to wonder how the two HVCs are coordinated during singing. Bilateral multi-unit recordings in HVC have revealed brief episodes of correlated activity across hemispheres that occur before the onset of each syllable (and at some acoustic transitions within long complex syllables)[34,35], probably mediated by feedback pathways from RA to midbrain areas and bilaterally back to HVC[18,19]. Do these episodes reflect actual bilateral synchronization of the HVCs? If the two HVCs were, hypothetically, synchronized only at the beginning of the motif, after which they operated independently, cooling HVC in only one hemisphere (Fig. 5a) should cause the two HVCs to become misaligned in time by more than a whole syllable by the end of the motif (compare control to bilaterally cooled song; Fig. 5c), causing song degradation. However, we found that unilateral cooling of HVC did not produce song degradation, but resulted in slowed songs of normal acoustic structure ($n = 4$; Fig. 5b, c, Supplementary Fig. 10), ruling out any model in which hemispheric synchronization occurs only at motif onsets.

Do both HVCs contribute to song timing? Given previous observations of hemispheric dominance in songbirds[5,36], it is conceivable that one HVC acts as a 'master clock' for song timing and that the other follows as a slave. In fact, in all birds ($n = 4$) we observed that cooling either HVC alone caused song slowing intermediate to that seen for bilateral cooling (Fig. 5b, Supplementary Fig. 10), ruling out the possibility that song timing is controlled by dynamics in a single hemisphere.

One possible explanation for the observation of intermediate slowing is that the left HVC may control the timing of some parts of the song and the right HVC may control the timing of other parts of the song. In fact, we found that cooling the right HVC or left HVC produced less uniform stretching of the song in comparison with bilateral cooling, as predicted by this model (bilateral non-uniformity s.d. of 0.67% °C^{-1}, unilateral non-uniformity s.d. of 1.33% °C^{-1}; Fig. 5d–j). This inhomogeneity in stretch during unilateral cooling is comparable to the mean stretch (~1.2% °C^{-1}), suggesting that a large fraction of song elements were not stretched by unilateral cooling (bilateral cooling, 14% not significantly stretched; left only, 71%; right only, 43%). Furthermore, some elements that were not stretched by right HVC cooling were strongly stretched by left HVC cooling, and vice versa (Fig. 5d–h, Supplementary Fig. 10). Consistent with this, stretch during left cooling was significantly anticorrelated with stretch during right cooling ($P = 0.026$; Fig. 5k; see Supplementary Methods). Thus, it appears that there may be some switching of the control of song timing between the two HVCs on the timescale of song syllables or long subsyllabic elements.

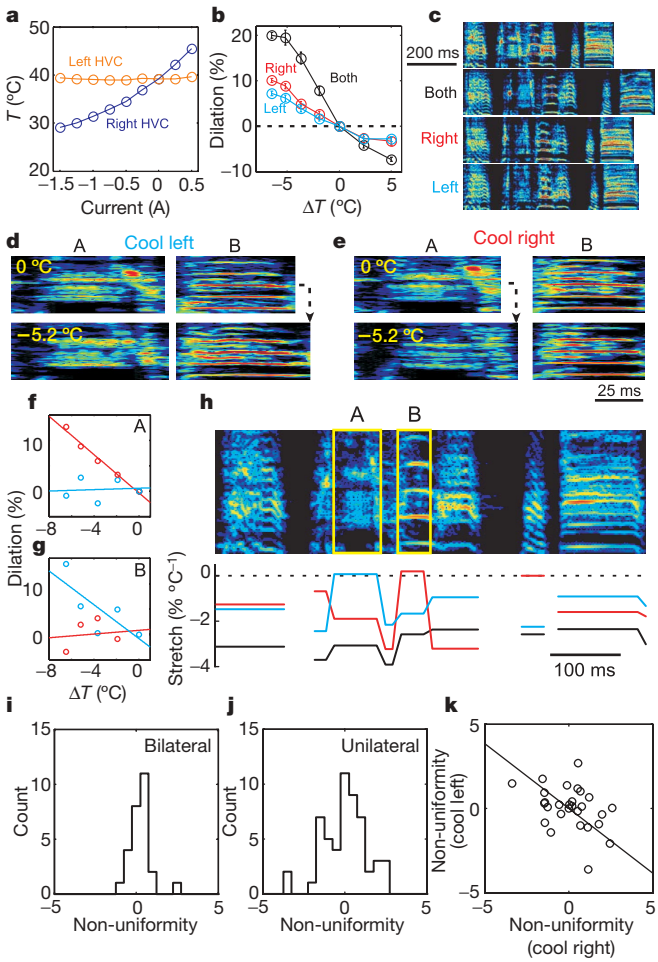


Figure 5 | Effect of unilateral HVC cooling on song timing. **a**, Simultaneous temperature measurements from HVC in both hemispheres when the Peltier device was configured for right HVC cooling. **b**, Change in motif duration as a function of HVC temperature change during unilateral and bilateral cooling in bird no. 11. Error bars, s.e.m. **c**, Spectrograms of song motif during control, bilateral, left and right HVC cooling. **d**, Selective dilation of subsyllabic element B, but not A, during left HVC cooling. **e**, Selective dilation of element A, but not B, during right HVC cooling. **f**, **g**, Dilation of subsyllabic element A (**f**) and B (**g**) during left (blue) and right (red) HVC cooling. **h**, Stretch metric of identified song segments during cooling of left (blue), right (red) or both (black) HVCs. **i**, **j**, Distributions of stretch non-uniformity values during bilateral (**i**) and unilateral (**j**) cooling. **k**, Non-uniformity values during left and right cooling show significant anticorrelation ($P = 0.026$). Solid line shows the first principal component of the distribution.

A chain model of song dynamics

The results of our HVC and RA cooling experiments do not support a view in which the dynamics underlying song timing are divided at different timescales between different brain areas. One possible model of dynamics in the song control system that is consistent with all of our observations is strongly anticipated by the singing-related firing patterns of HVC neurons that project to RA. These neurons burst extremely sparsely during singing, each generating a single brief (~6-ms) burst of spikes at a particular moment in every repetition of the song motif[37]. In addition, different HVC neurons burst at different times throughout the song. We have proposed that, as a population, these HVC neurons form domino-like chains of activity that control the timing of the song[38]. The HVC cooling experiments suggest that the dynamics underlying the sparse sequential activation of HVC neurons reside at least partly within HVC; if the sparse HVC bursts were driven by an independent upstream sequence generating circuit, cooling HVC should not have affected the speed of the song. An interesting possibility is that such sequential chains of activity arise, in part, from a chain-like synaptic organization within HVC[39–43]. In this case, cooling HVC may simply increase the time it takes each neuron to burst following activity in a preceding neuron, thus introducing an accumulating delay that slows down the chain.

Although the cooling results rule out models in which brain areas outside HVC independently or autonomously control song timing on any timescale, they do not exclude the involvement of other brain areas important for song production, in particular feedback projections from RA, through the brainstem/midbrain and Uva back to HVC[19,21]. These projections could form an integral part of the connectivity that generates sequential bursts in HVC. Consider the timescale on which this feedback might operate. In principle, every burst in HVC could be driven by fast, rapidly cycling feedback through this loop, rather than by intrinsic connections within HVC. Cooling anywhere in this feedback loop, including RA, should introduce accumulating delays as well. The fact that we do not observe song slowing from cooling in RA suggests that the feedback circuitry may not

operate at this rapidly cycling timescale, but less frequently during the song.

Thus, one interesting possibility is that HVC may contain multiple independent chains[8] (or modules), which may be associated with syllables or long subsyllabic elements[34]. These modules, each of which could run autonomously by virtue of circuit dynamics within HVC, may then be linked together in time by the feedback connections from RA through the thalamus and back to HVC (Supplementary Fig. 11). The feedback circuitry may act to detect the end of one syllable and rapidly and bilaterally initiate the next gap and syllable, thus simultaneously continuing the song sequence and resynchronizing the two HVCs. In this case, cooling HVC would slow the production of a song syllable, because such production is generated by dynamics within HVC, and would also delay the onset of the next syllable, because this onset is linked to the termination of the previous syllable by feedback circuitry.

A module of chain-like activity in HVC may produce a song syllable as follows. During singing, RA neurons generate a complex but highly stereotyped sequence of spike bursts[14,44]. We have previously suggested that these RA bursts are driven rapidly and on a moment-to-moment basis by bursting inputs from HVC[37,38]. In this view, because RA burst patterns simply follow the timing set by HVC, we would expect RA cooling to have a minimal effect on the timing of RA activity, whereas slowing the chain in HVC would necessarily slow the sequence of bursts in RA. Furthermore, if structures downstream of RA (brainstem motor neurons and syringeal muscles) respond rapidly to descending drive from RA, perhaps on the timescale of the fastest song modulations (10–20 ms)[38],we would expect that slowing the sequence of bursts in RA might slow the song yet have a minimal effect on song acoustic structure. Thus, a simple model of chain-like dynamics in HVC that drives a fast response in RA and downstream structures is consistent with the electrophysiological data and the HVC and RA cooling experiments.

Here we have used local manipulation of brain temperature to identify components within the avian song system that control the timing of a complex behavioural sequence. This approach may be broadly useful in localizing specialized brain circuits that control the timing of other behaviours. We have also used temperature changes to test ideas about the contribution to song production of oscillatory dynamics in the song control pathway, an approach that should be generally applicable to localizing the biophysical origin of oscillatory and other forms of brain dynamics, and for studying their role in brain function.

METHODS SUMMARY

Subjects. Subjects were adult zebra finches (>120 days post-hatch) obtained either from our colony or from an outside distributor (Preferred Birds). All animal procedures were approved by the committee on animal care at the Massachusetts Institute of Technology.

Cooling devices. We used a small (0.7-g) custom-built thermoelectric device based on the Peltier effect to cool HVC and RA. The HVC cooling device was constructed from two 1 mm × 2 mm gold cooling elements that bilaterally contacted the surface of the dura overlying the left and right HVCs. The temperature change in HVC was spatially restricted (Fig. 1c), producing a maximal change of only 0.5 °C in RA, the nearest brain region known to be involved in song production in the zebra finch[5,45]. The current could be switched to flow bilaterally or unilaterally through only the left or only the right cooling element. The RA cooling devices were equipped with a gold spike implanted into RA to facilitate heat transfer.

RA electrophysiology. A craniotomy was made over RA under isoflurane anaesthesia (1.5%). The borders of RA were identified electrophysiologically and the cooling device implanted at the centre. Single neurons were isolated using carbon fibre electrodes (Carbostar-1, Kation Scientific), with a signal-to-noise ratio of greater than 10:1.

Temperature measurements. Temperatures were measured using small thermocouples (5SRTC-TT-K-40-36, Omega). For HVC calibration, three thermocouples were inserted in one hemisphere, under anaesthesia, at respective depths of 0.5, 2.0 and 4.0 mm beneath the gold pad. In some birds, an additional probe was placed in the contralateral HVC (at a depth of 0.5 mm). Once inserted, the thermocouples were secured with dental acrylic and the bird was placed in a cage and allowed to awaken. At each Peltier current level, three minutes were allowed for the brain to reach a steady-state temperature before measurements were taken from all thermocouple locations.

Received 22 May; accepted 23 September 2008.

1. Marder, E. & Bucher, D. Understanding circuit dynamics using the stomatogastric nervous system of lobsters and crabs. *Annu. Rev. Physiol.* **69,** 291–316 (2007).
2. Stent, G. S. *et al.* Neuronal generation of the leech swimming movement. *Science* **200,** 1348–1357 (1978).
3. Williams, H. Birdsong and singing behavior. *Ann. NY Acad. Sci.* **1016,** 1–30 (2004).
4. Konishi, M. Birdsong: from behavior to neuron. *Annu. Rev. Neurosci.* **8,** 125–170 (1985).
5. Nottebohm, F., Stokes, T. M. & Leonard, C. M. Central control of song in the canary, Serinus canarius. *J. Comp. Neurol.* **165,** 457–486 (1976).
6. Nottebohm, F., Kelley, D. B. & Paton, J. A. Connections of vocal control nuclei in the canary telencephalon. *J. Comp. Neurol.* **207,** 344–357 (1982).
7. Zann, R. A. *The Zebra Finch: A Synthesis of Field and Laboratory Studies* 196–246 (Oxford Univ. Press, 1996).
8. Glaze, C. M. & Troyer, T. W. Temporal structure in zebra finch song: implications for motor coding. *J. Neurosci.* **26,** 991–1005 (2006).
9. Cooper, B. G. & Goller, F. Physiological insights into the social-context-dependent changes in the rhythm of the song motor program. *J. Neurophysiol.* **95,** 3798–3809 (2006).
10. Sossinka, R. & Böhner, J. Song types in the zebra finch (Poephila guttata castanotis). *Z. Tierpsychol.* **53,** 123–132 (1980).
11. Vicario, D. S. Organization of the zebra finch song control system: II. Functional organization of outputs from nucleus Robustus archistriatalis. *J. Comp. Neurol.* **309,** 486–494 (1991).
12. Wild, J. M. Descending projections of the songbird nucleus robustus archistriatalis. *J. Comp. Neurol.* **338,** 225–241 (1993).
13. Vu, E. T., Mazurek, M. E. & Kuo, Y. C. Identification of a forebrain motor programming network for the learned song of zebra finches. *J. Neurosci.* **14,** 6924–6934 (1994).
14. Yu, A. C. & Margoliash, D. Temporal hierarchical control of singing in birds. *Science* **273,** 1871–1875 (1996).
15. Solis, M. M. & Perkel, D. J. Rhythmic activity in a forebrain vocal control nucleus *in vitro. J. Neurosci.* **25,** 2811–2822 (2005).
16. Chi, Z. & Margoliash, D. Temporal precision and temporal drift in brain and behavior of zebra finch song. *Neuron* **32,** 899–910 (2001).
17. Laje, R. & Mindlin, G. B. Diversity within a birdsong. *Phys. Rev. Lett.* **89,** 288102 (2002).
18. Striedter, G. F. & Vu, E. T. Bilateral feedback projections to the forebrain in the premotor network for singing in zebra finches. *J. Neurobiol.* **34,** 27–40 (1998).
19. Ashmore, R. C., Renk, J. A. & Schmidt, M. F. Bottom-up activation of the vocal motor forebrain by the respiratory brainstem. *J. Neurosci.* **28,** 2613–2623 (2008).
20. Goller, F. & Cooper, B. G. Peripheral motor dynamics of song production in the zebra finch. *Ann. NY Acad. Sci.* **1016,** 130–152 (2004).
21. Ashmore, R. C., Wild, J. M. & Schmidt, M. F. Brainstem and forebrain contributions to the generation of learned motor behaviors for song. *J. Neurosci.* **25,** 8543–8554 (2005).
22. Thompson, S. M., Masukawa, L. M. & Prince, D. A. Temperature dependence of intrinsic membrane properties and synaptic potentials in hippocampal CA1 neurons *in vitro. J. Neurosci.* **5,** 817–824 (1985).
23. Volgushev, M., Vidyasagar, T. R., Chistiakova, M. & Eysel, U. T. Synaptic transmission in the neocortex during reversible cooling. *Neuroscience* **98,** 9–22 (2000).
24. Volgushev, M., Vidyasagar, T. R., Chistiakova, M., Yousef, T. & Eysel, U. T. Membrane properties and spike generation in rat visual cortical cells during reversible cooling. *J. Physiol. (Lond.)* **522,** 59–76 (2000).
25. Arbas, E. A. & Calabrese, R. L. Rate modification in the heartbeat central pattern generator of the medicinal leech. *J. Comp. Physiol. A* **155,** 783–794 (1984).
26. Katz, P. S., Sakurai, A., Clemens, S. & Davis, D. Cycle period of a network oscillator is independent of membrane potential and spiking activity in individual central pattern generator neurons. *J. Neurophysiol.* **92,** 1904–1917 (2004).
27. Ferster, D., Chung, S. & Wheat, H. Orientation selectivity of thalamic input to simple cells of cat visual cortex. *Nature* **380,** 249–252 (1996).
28. Herrmann, K. & Arnold, A. P. The development of afferent projections to the robust archistriatal nucleus in male zebra finches: a quantitative electron microscopic study. *J. Neurosci.* **11,** 2063–2074 (1991).
29. Spiro, J. E., Dalva, M. B. & Mooney, R. Long-range inhibition within the zebra finch song nucleus RA can coordinate the firing of multiple projection neurons. *J. Neurophysiol.* **81,** 3007–3020 (1999).
30. Roberts, T. F., Klein, M. E., Kubke, M. F., Wild, J. M. & Mooney, R. Telencephalic neurons monosynaptically link brainstem and forebrain premotor networks necessary for song. *J. Neurosci.* **28,** 3479–3489 (2008).
31. Mooney, R. Synaptic basis for developmental plasticity in a birdsong nucleus. *J. Neurosci.* **12,** 2464–2477 (1992).
32. Janata, P. & Margoliash, D. Gradual emergence of song selectivity in sensorimotor structures of the male zebra finch song system. *J. Neurosci.* **19,** 5108–5118 (1999).

33. Hahnloser, R. H., Kozhevnikov, A. A. & Fee, M. S. Sleep-related neural activity in a premotor and a basal-ganglia pathway of the songbird. *J. Neurophysiol.* **96,** 794–812 (2006).
34. Schmidt, M. F. Pattern of interhemispheric synchronization in HVc during singing correlates with key transitions in the song pattern. *J. Neurophysiol.* **90,** 3931–3949 (2003).
35. Schmidt, M. F., Ashmore, R. C. & Vu, E. T. Bilateral control and interhemispheric coordination in the avian song motor system. *Ann. NY Acad. Sci.* **1016,** 171–186 (2004).
36. Williams, H., Crane, L. A., Hale, T. K., Esposito, M. A. & Nottebohm, F. Right-side dominance for song control in the zebra finch. *J. Neurobiol.* **23,** 1006–1020 (1992).
37. Hahnloser, R. H., Kozhevnikov, A. A. & Fee, M. S. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* **419,** 65–70 (2002).
38. Fee, M. S., Kozhevnikov, A. A. & Hahnloser, R. H. Neural mechanisms of vocal sequence generation in the songbird. *Ann. NY Acad. Sci.* **1016,** 153–170 (2004).
39. James, W. *Psychology: The Briefer Course* 1–17 (Dover, 2001).
40. Li, M. & Greenside, H. Stable propagation of a burst through a one-dimensional homogeneous excitatory chain model of songbird nucleus HVC. *Phys. Rev. E* **74,** 011918 (2006).
41. Jin, D. Z., Ramazanoglu, F. M. & Seung, H. S. Intrinsic bursting enhances the robustness of a neural network model of sequence generation by avian brain area HVC. *J. Comput. Neurosci.* **23,** 283–299 (2007).
42. Glaze, C. M. & Troyer, T. W. Behavioral measurements of a temporally precise motor code for birdsong. *J. Neurosci.* **27,** 7631–7639 (2007).
43. Abeles, M. *Coriconics: Neural Circuits of the Cerebral Cortex* 208–226 (Cambridge Univ. Press, 1991).
44. Leonardo, A. & Fee, M. S. Ensemble coding of vocal control in birdsong. *J. Neurosci.* **25,** 652–661 (2005).
45. Cardin, J. A., Raksin, J. N. & Schmidt, M. F. Sensorimotor nucleus NIf is necessary for auditory processing but not vocal motor output in the avian song system. *J. Neurophysiol.* **93,** 2157–2166 (2005).

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

Acknowledgements We thank D. Aronov, T. Gardner, J. Goldberg, L. Las, B. Ölveczky, S. Seung and M. Wilson for their comments on earlier versions of this manuscript. This work is supported by funding from the US National Institutes of Health to M.S.F. (MH067105) and to M.A.L. (DC009280) as well as funding from the Human Frontiers Science Project.

Author Contributions M.A.L. and M.S.F. both contributed to all aspects of this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.S.F. (fee@mit.edu).

ARTICLES

Illumination controls differentiation of dopamine neurons regulating behaviour

Davide Dulcis[1] & Nicholas C. Spitzer[1]

Specification of the appropriate neurotransmitter is a crucial step in neuronal differentiation because it enables signalling among populations of neurons. Experimental manipulations demonstrate that both autonomous and activity-dependent genetic programs contribute to this process during development, but whether natural environmental stimuli specify transmitter expression in a neuronal population is unknown. We investigated neurons of the ventral suprachiasmatic nucleus that regulate neuroendocrine pituitary function in response to light in teleosts, amphibia and primates. Here we show that altering light exposure, which changes the sensory input to the circuit controlling adaptation of skin pigmentation to background, changes the number of neurons expressing dopamine in larvae of the amphibian *Xenopus laevis* in a circuit-specific and activity-dependent manner. Neurons newly expressing dopamine then regulate changes in camouflage colouration in response to illumination. Thus, physiological activity alters the numbers of behaviourally relevant amine-transmitter-expressing neurons in the brain at postembryonic stages of development. The results may be pertinent to changes in cognitive states that are regulated by biogenic amines.

Electrical activity has a key role in neurotransmitter specification[1–4] and postsynaptic receptor expression[5–7] in differentiating neurons. Moreover, postsynaptic receptors can be selected to match the transmitters expressed presynaptically[8,9]. L-type calcium channels have been implicated in the developmental expression of dopamine (DA), because a loss-of-function mutation in tottering (*tg*) mutant mice[10] leads to abnormal expression of tyrosine hydroxylase (TH), the rate-limiting DA-synthetic enzyme[11]. TH is ectopically co-expressed with GABA (γ-aminobutyric acid) in cerebellar Purkinje neurons in these mutants. These studies led us to hypothesize that activity driven by environmental stimuli will cause neurons to express an additional transmitter, without changing the number of neurons or affecting their original transmitter phenotype. The neurons would be able to perform a new function in addition to their default role. Hijacking activity-dependent mechanisms regulating transmitter expression could be a useful step towards the development of clinical treatments that drive transmitter specification in selected classes of neurons.

Here we examine the role of illumination in the specification of DA in neurons of the ventral suprachiasmatic nucleus (VSC) of *Xenopus laevis*. The neuroanatomy of the circuit regulating adaptation of skin pigmentation to background and the camouflage behaviour of adult *Xenopus* have been established[12–15]. We describe background adaptation in larvae and show that suppressing or enhancing electrical activity by misexpressing ion channels during embryonic development leads to widespread changes in the number of TH-expressing neurons in the brain. In contrast, changes in activation of the retinohypothalamic projection at larval stages of development alter the number of TH-inhibitory neurons specifically in the VSC. We demonstrate that enhancing activity by brief light exposure stimulates a subclass of VSC neurons, characterized by neuropeptide Y (NPY) transmitter and LIM1 (also known as LHX1) and LIM2 (also known as LHX5) transcription factor expression, to acquire the additional DA phenotype. We find that these newly TH/NPY neurons retain their physiological target, the melanotrope cells, and drive DA-dependent, behaviourally appropriate changes in skin pigmentation.

Dopaminergic neurons regulate pigmentation

The regulation of pigmentation is a well-conserved behaviour among vertebrate species. Mammals undergo gradual changes in skin colour under hormonal control[16]; many other vertebrates including teleosts[17] and amphibians[18] display rapid physiological colour changes in response to the same hormones. This behaviour is mediated by simple neuronal circuitry in Amphibia: glutamatergic retinal ganglion cells project to dopaminergic suprachiasmatic melanotrope inhibitory neurons (SMINs), which innervate cholinergic melanocyte-stimulating-hormone (MSH)-releasing cells (melanotrope cells, Fig. 1a)[13,18]. We find that epithelial pigmentation of amphibian larvae is regulated by the level of ambient illumination, either by incident light or by background light level, which leads to aggregation and dispersion of melanin granules within melanocytes[19] as in teleosts. Pigmentation appears lighter with bright illumination or a white background and darker with dim illumination or a black background (Fig. 1b and Supplementary Fig. 1) after as little as 10 min exposure. The same circuit controls this behaviour in the adult[13,20].

To identify the contribution of DA neurons to regulation of pigmentation we exposed larvae to sulpiride, a DA D_2 receptor antagonist, because co-expressed GABA and NPY can contribute to regulation of this behaviour[12,14]. Sulpiride treatment blocked changes in skin pigmentation in response to altered illumination (Fig. 1c and Supplementary Fig. 2a, c), demonstrating the predominant role of DA neurons in this pathway in *Xenopus* larvae. The number of TH VSC neurons was not affected by sulpiride (Supplementary Fig. 2b). Immunostaining for pro-opiomelanocortin (POMC) and NPY in combination with the nuclear marker 1,5-bis{[2-(di-methylamino)ethyl]amino}-4,8-dihydroxyanthracene-9,10-dione (DRAQ5) revealed that the number of POMC melanotrope cells was also unchanged (Supplementary Fig. 2d, e).

We determined whether D_2 receptor activation in the eye is involved in the circuit-regulating adaptation to background illumination, because DA has complex effects on the retina[21–23]. The blanching effect of a D_2 receptor agonist, quinpirole hydrochloride, on dark-adapted larvae remaining in the dark persisted to the same extent after bilateral

[1]Neurobiology Section, Division of Biological Sciences and Center for Molecular Genetics, Kavli Institute for Brain and Mind, UCSD La Jolla, California 92093-0357, USA.

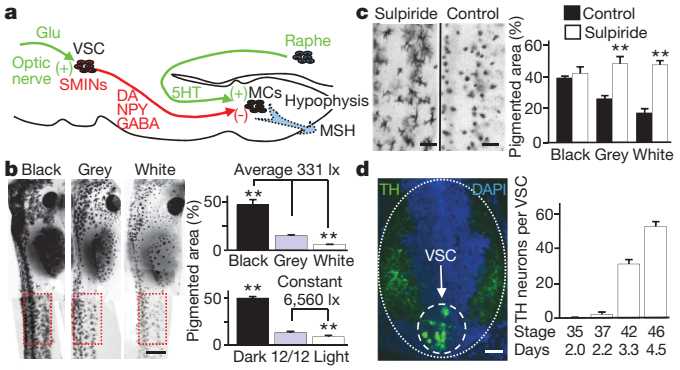


Figure 1 | Dopaminergic VSC neurons regulate skin pigmentation. **a**, Diagram illustrating the neuronal circuit controlling this behaviour. Glu, glutamate; MC, melanotrope cells. **b**, Raising animals in different light levels changes pigmentation. Left, phenotypes on different backgrounds; right, pigmentation adaptation assayed in the boxed regions to the left after different background or incident illumination. **c**, Activity of DA SMINs is necessary for regulation of pigmentation. Left, 2 h white adaptation in a larva treated with 10 nM sulpiride for 30 min compared to control; right, pigmentation at various backgrounds in sulpiride-treated larvae and controls. For **b** and **c**, $N > 6$ stage-42 larvae raised in the dark. **d**, VSC neuron numbers increase during development of larvae raised in 12 h light/12 h dark on a grey background. Left, transverse section through the stage-42 diencephalon (dotted oval) and VSC (dashed circle); right, appearance of TH neurons during development. $N \geq 6$ larvae for each stage. Scale bars: **b**, 1 mm; **c**, 200 μm; **d**, 50 μm. Error bars, s.e.m.; $^{**}P<0.001$.

eye enucleation in the dark (Supplementary Fig. 3a, b), suggesting that retinal D_2 receptor activation plays little part in the regulation of this adaptation. The D_1 receptor agonist and antagonist, SKF-38393 and SCH23390, had the opposite effect on this behaviour, darkening and lightening pigmentation, respectively (Supplementary Fig. 3c, d). Other DA nuclei do not participate in this circuit. The pineal gland, a light sensor controlling circadian oscillation of melatonin that initiates slow changes in pigmentation, is not involved in rapid background adaptation[24]. These results demonstrate participation of DA VSC neurons in the pathway controlling this camouflage behaviour.

We analysed LIM1/2 and PAX6 transcription factor expression during development to identify molecular signatures of TH VSC neurons. These neurons (Fig. 1d) were first detected at 2 days of development, and by 3 days they formed a nucleus with a core that ranged from 10 to 30 neurons depending on the illumination conditions in which the larvae were raised. VSC neurons were recognized by their expression of the LIM1/2 transcription factor and the absence of expression of PAX6 during the first 4 days of development[25,26] (Fig. 2a, b, control).

Activity-dependence of dopaminergic differentiation

To discover whether the temporal and spatial parameters of activity have a role in selective neuronal recruitment for neurotransmitter respecification, we compared the effect of sustained and global changes of activity throughout the central nervous system (CNS) with brief changes in circuit-specific activation. To test the general activity-dependence of DA expression, we investigated the effect of ion channel overexpression on the number of DA neurons. We first confocally imaged intracellular calcium with the Fluo-4AM indicator to determine whether neurons in the brain generate spontaneous calcium spikes. Neurons in the developing hypothalamus showed calcium spikes similar to those regulating transmitter specification in the posterior neural tube[3]; fixing and immunostaining these preparations identified inactive neurons that express TH (Fig. 2c and Supplementary Figs 4 and 5). These results suggested that, as in the spinal cord, calcium spikes precede the appearance of the transmitter, and raised the possibility that here too calcium spike activity regulates transmitter expression.

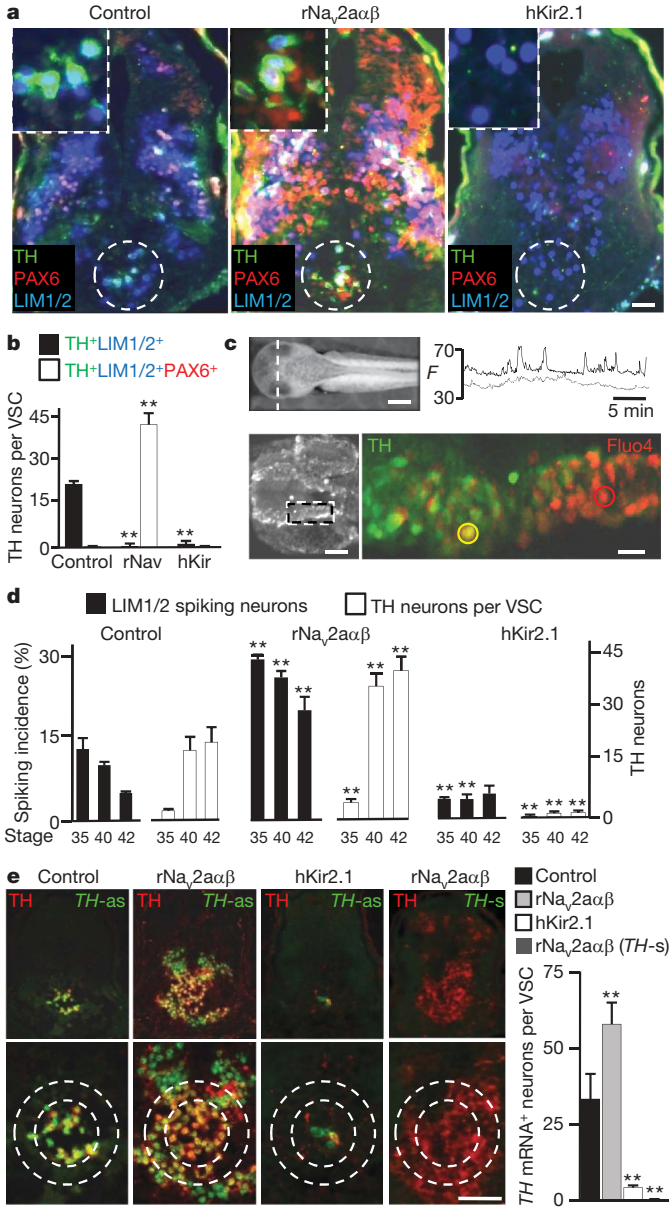


Figure 2 | Dopaminergic differentiation is activity-dependent. **a**, Molecular markers identify VSC neurons (dashed circles) in transverse sections in a control larva and in larvae after sodium ($rNa_V2a\alpha\beta$) or potassium (hKir2.1) channel overexpression (stage 42). The VSC is shown at higher magnification (insets). **b**, Altering spike activity drives proportional changes in the number of TH neurons. Shown is quantification of data in **a**. $N \geq 6$ larvae. **c**, Neurons generate calcium spikes in the developing brain. Left, transverse section (dashed line) through the hypothalamus (stage 35) and Fluo-4AM-loaded tissue (below); confocal imaging and staining are from the inset. Right, digitized fluorescence (F) of representative cells circled in the panel below; traces show calcium spike activity recorded from dorsal (top trace, red circle) TH^- and ventral (bottom trace, yellow circle) TH^+ cells. $N = 3$ larvae. **d**, Across developmental stages the number of TH neurons is inversely proportional to the incidence of spiking in LIM1/2 neurons in controls and after alterations of activity. **e**, *TH* mRNA expression is altered by ion channel overexpression. VSC double labelling with TH antibodies and *TH*-antisense probe (*TH*-as) in transverse brain sections (top row) and enlargements of the VSC (dashed circles, bottom row) in larvae injected with cascade blue (left panel, control), overexpressing sodium channels (second panel) or potassium channels (third panel). Antisense-probe-binding in sodium-channel-overexpressing larvae (*TH*-s, fourth panel). Quantification is shown to the right. a–e, Larvae were raised on a 12 h/12 h day/night cycle on a grey background. $N \geq 6$ larvae. Scale bars: **a**, 50 μm; **c**, 500 μm (left), 30 μm (right); **e**, 100 μm. Error bars, s.e.m.; $^{**}P<0.001$.

To test this hypothesis we overexpressed inward rectifier potassium channels (human Kir2.1) or voltage-gated sodium channels (rat $Na_v2a\alpha\beta$) at early developmental stages to suppress or enhance calcium spike generation[3]. Increasing activity increased the number of TH neurons in and around the VSC, whereas suppressing activity decreased the number of TH neurons (Fig. 2a, b). Similar changes were observed for the dorsolateral suprachiasmatic nucleus (data not shown). Changes in the incidence of calcium spiking (percentage of neurons spiking per h) during development were inversely correlated with changes in TH expression, establishing a quantitative connection between the two (Fig. 2d). Increasing activity also induced expression of PAX6, but not LIM1/2 (data not shown), and newly TH neurons were recruited from the larger PAX6 and stable LIM1/2 pool of cells (Fig. 2a, b). *In situ* hybridization showed that increasing calcium spike generation leads to widespread expression of *X. laevis TH* messenger RNA in cells not normally expressing it; decreasing calcium spike generation leads to a decrease in the number of cells expressing these transcripts (Fig. 2e). Thus, the changes in DA phenotype are regulated transcriptionally. These findings indicate that spontaneous calcium spike activity controls the number of VSC DA neurons to regulate circuit function and skin pigmentation, because these neurons inhibit melanotrope cells[14], whereas global ion channel misexpression has widespread effects that seem less specific.

Light-dependence of dopamine expression

To determine whether newly dopaminergic neurons can be selectively recruited in specific circuits we investigated the effect of activating the retinohypothalamic projection on neurotransmitter expression of VSC neurons. Raising animals in different levels of illumination and background changed the number of TH VSC neurons (Fig. 3a), and light on a white background was most effective in inducing an increase. Exposure of dark-raised animals to this illumination protocol for as little as 2 h led to the appearance of TH neurons in the annular region surrounding the core (Fig. 3b). By immunostaining neurons with antibodies to DA, the DA transporter (DAT) and vesicular monoamine transporter (VMAT2), we found that they are all co-regulated with TH: all four markers were undetected in annular neurons of larvae kept in the dark and expressed in the NPY annular neurons of larvae that were exposed to 2 h light (Supplementary Figs 6–8). During the respecification process these annular DA neurons are immunoreactive for NPY and for LIM1/2 but not for PAX6 (Fig. 3b–e and Supplementary Fig. 9). Thus, the more restricted physiological stimulus, in contrast to ubiquitous ion channel overexpression, recruits newly dopaminergic neurons from a restricted population. The absence of PAX6 induction may be due to the difference in the extent of altered activity. Increasing light exposure led to the appearance of *TH* mRNA in annular NPY neurons. This transmitter respecification was reversible; decreasing light exposure reduced the extent of expression of *TH* mRNA in neurons in both the annulus and the core of the VSC (Fig. 3f), and TH protein was no longer detected in 90% of TH annular neurons when larvae were kept in the dark for >2 h (Supplementary Fig. 10).

Light-dependent acquisition of TH by NPY annular neurons in the VSC of dark-adapted larvae produces a new set of DA neurons that could regulate pigmentation, because NPY neurons are reported to innervate melanotrope cells in the adult[13] and inhibit MSH production in a sustained manner[14]. There was no change in the total number of 4,6-diamidino-2-phenylindole (DAPI)- or LIM1/2-stained nuclei, implying that neither cell proliferation nor migration account for the observed change in the total number of DA neurons (Supplementary Fig. 11a, b). 5-bromodeoxyuridine (BrdU) labelling and TdT-mediated dUTP nick end labelling (TUNEL) assays did not reveal apoptosis or cell birth in these cells (Supplementary Fig. 11c, d). The number of melanotrope cells remained constant, whereas the number of POMC melanotrope cells increased after light exposure and decreased in the dark, associated with changes in intensity of

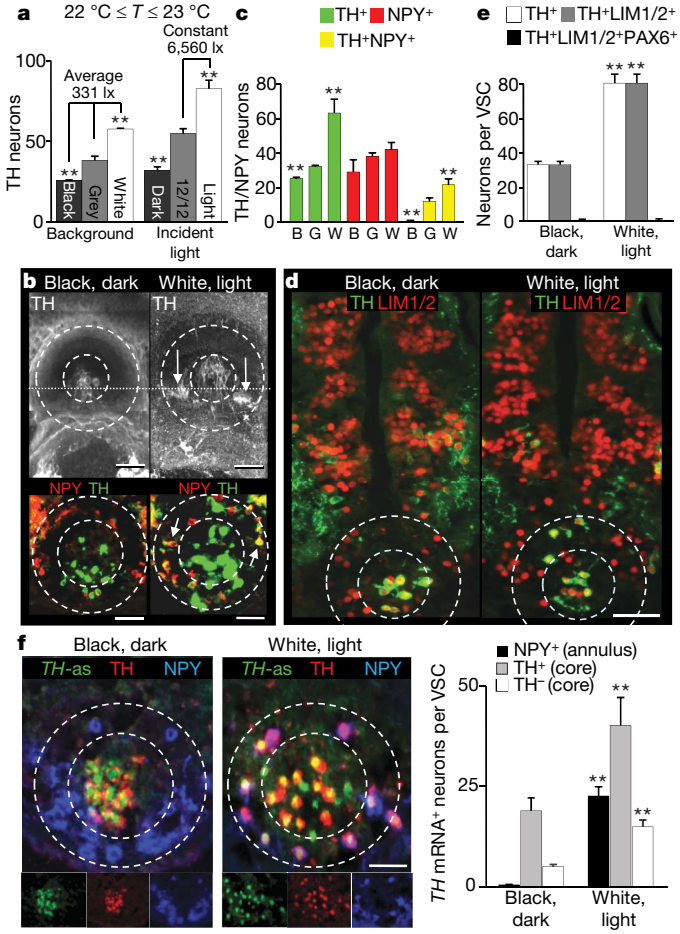


Figure 3 | Illumination changes the number of dopaminergic neurons selectively in the VSC. a, Raising animals in different light levels changes the number of VSC TH neurons. $N \geq 9$ stage-42 larvae. **b**, Whole mounts (top) and transverse sections (bottom) show the TH core (dashed inner circle) and annular neurons (arrows, between dashed circles) in 2-h-black- and white-adapted stage-42 larvae. Dashed line indicates section orientation. **c**, Two-hour exposure to different background illumination (B, black; G, grey; W, white) changes the number of TH/NPY neurons in the VSC (core plus annulus). **d**, Core and newly TH annular neurons from 2-h-white-adapted larvae are from the LIM1/2 pool; sections as in **b**. **e**, Exposure to different backgrounds for 2 h changes the number of LIM1/2 neurons expressing TH. **f**, *TH* mRNA expression is altered by light or dark adaptation. Triple-labelling of transverse VSC sections (dashed circles) with TH, NPY and *TH*-as probe in 2-h-dark/black-adapted and 2-h-light-/white-adapted larvae. Merged images (top row) and separate channels (bottom row). Quantification is presented to the right. **b–f**, Stage-42 larvae raised in the dark. **c**, **e**, **f**, $N \geq 6$ larvae. Scale bars: **b**, 100 μm (top), 60 μm (bottom); **d**, 100 μm; **f**, 60 μm. Error bars, s.e.m.; $^{**}P<0.001$.

POMC staining. Similar light-dependent accumulation of POMC in melanotrope cells occurs in white-background-adapted adult frogs[27].

Suppressing activity eliminated DA expression elicited by light. Binocular eye enucleation that eliminates retinal input abolished both white background adaptation and light-induced changes in the number of TH neurons (Fig. 4a). We then used a pharmacological approach to test the involvement of electrical activity and calcium signalling. We first implanted 80-μm agarose beads loaded with calcein acetoxymethylester (calcein-AM) tracer in the roof of the hypothalamus above the VSC to evaluate the time course of drug delivery. Twenty-four hours after implantation, the spread of tracer from the bead was localized within a radius of ~100 μm that included the VSC (Supplementary Fig. 12). Larvae implanted with beads loaded with the sodium channel blocker tetrodotoxin (TTX) or with the calcium buffer 1,2-bis-(o-aminophenoxy)-ethane-N,N,N′,N′-tetraacetic acid,

tetraacetoxymethyl ester (BAPTA-AM) (similar in molecular weight and diffusion profile to calcein-AM) showed no increase in number of either TH neurons or NPY/TH annular neurons in the VSC after 2 h illumination in white light following 24 h in the dark (Fig. 4b, c).

Function of light-induced dopamine neurons

To assess the behavioural potential of newly dopaminergic neurons we determined whether they project to appropriate targets. After adaptation in the light for a minimum of 2 h, annular neurons, identified by LIM1/2 and NPY expression, co-expressed TH in their somata and axonal projections to the melanotrope cells (Fig. 5a and Supplementary Fig. 13). Localization of TH/NPY immunoreactivity adjacent to melanotrope cells, identified by staining for POMC[28], showed that NPY terminals are present in dark-adapted larvae and become TH/NPY in white-adapted larvae (Fig. 5b). POMC was not detected in a subpopulation of melanotrope cells that received only NPY terminals, but appeared after white adaptation when closely apposed nerve terminals became both NPY- and TH-positive (Fig. 5b, c). Moreover, D_2 receptors were not observed at NPY terminals in dark-adapted larvae but were evident at TH/NPY terminals of white-adapted larvae (Fig. 5d and Supplementary Fig. 14). These results suggest that the inhibitory output from the increased number of TH-expressing neurons in the VSC stimulates POMC storage in a larger number of melanotrope cells (Fig. 5e, f) and in turn reduces MSH release, which may account for accumulation of POMC in the melanotrope cells of white-background-adapted adult frogs[27].

To determine whether the larger population of DA neurons enhances the light sensitivity of changes in pigmentation, we compared white-adaptation of larvae raised in the dark to those raised in the light. These larvae were black-background-adapted for 30 min to

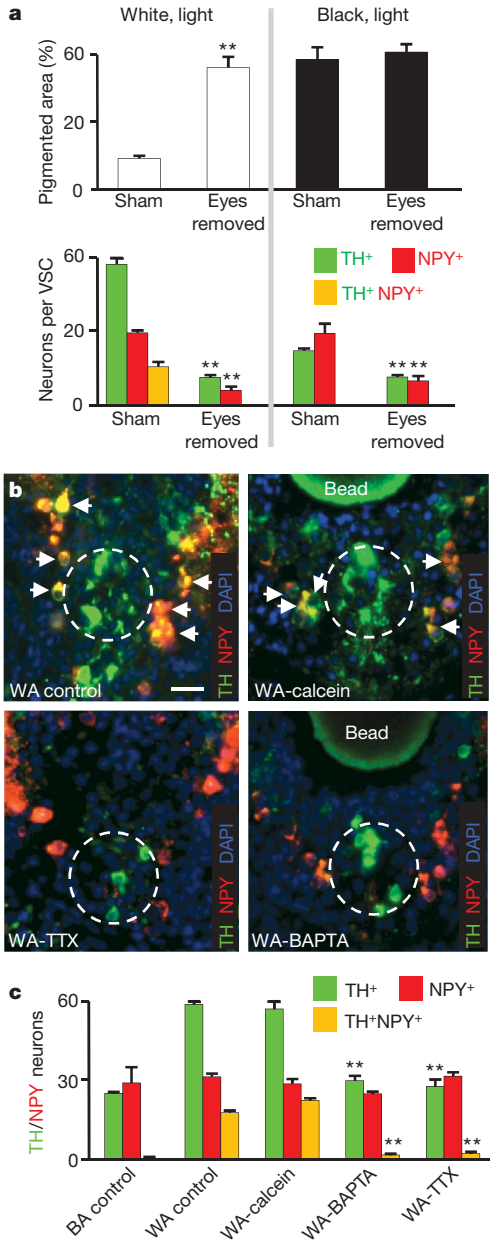


Figure 4 | Blocking physiological activity eliminates illumination-dependent changes in the number of dopaminergic VSC neurons. **a**, Binocular eye enucleation abolishes white-adaptation- and background-illumination-dependent changes in numbers of TH neurons. **b**, Implanted beads delivering activity blockers prevent appearance of TH/NPY neurons in larvae exposed to 2 h light relative to untreated or diffusion marker controls. Transverse sections through the VSC (dashed circles) show NPY/TH colocalization (yellow, arrows) in white-adapted control (WA control) and calcein-bead-implanted larvae that is absent in TTX and BAPTA-bead-implanted larvae. Scale bar: 40 μm. **c**, TH and NPY identify VSC neurons in the WA control and black-adapted (BA) control and after suppression of activity. **a**, **c**, $N \geq 6$ stage-42 larvae raised in the dark. Error bars, s.e.m.; $^{**}P<0.001$.

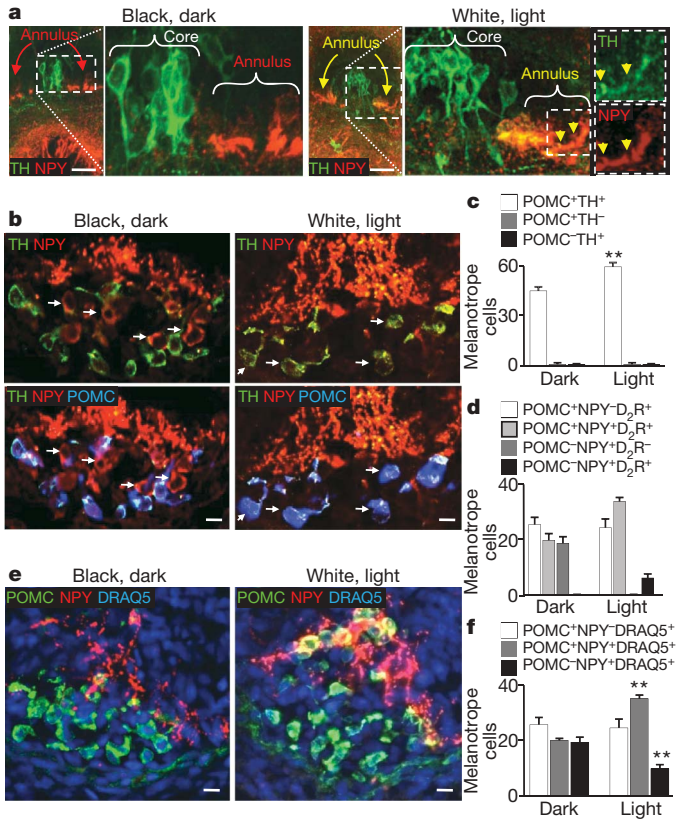


Figure 5 | NPY neurons projecting to melanotrope cells express TH after illumination. **a**, Left, TH is absent in NPY annular neurons (red arrows) in transverse sections of 2-h-dark-adapted larvae. Right, TH is present in somata (yellow arrows) and axons (yellow arrowheads) of NPY annular neurons in 2-h-white- and light-adapted larvae; far right, colour separation of green TH and red NPY (insets). **b**, Left, TH nerve terminals invest POMC-stained melanotrope cells in 2-h-dark-adapted larvae; TH-negative NPY terminals project to POMC-negative melanotrope cells (arrows). Right, TH/NPY terminals project to POMC melanotrope cells (arrows) in 2-h-white-adapted larvae. **c**, Quantification of **b**. **d**, D_2 receptors are expressed in melanotrope cells of 2-h-light-adapted larvae that newly express or may be acquiring POMC (light grey and black bars). **e**, DRAQ5-labelled nuclei identify 2-h-dark/light-dependent changes in melanotrope POMC expression. **f**, Quantification of **e**. **a**–**f**, Larvae raised in the dark. **c**, **d**, **f**, $N \geq 6$ stage-42 larvae. Scale bars: **a**, 100 μm; **b**, **e**, 40 μm. Error bars, s.e.m.; $^{**}P<0.001$.

achieve the same starting level of pigmentation before being tested. Larvae raised in the light, which have more core TH neurons as well as TH/NPY annular neurons, white-adapted more rapidly and to a greater extent than larvae raised in the dark, which have only core TH neurons (Supplementary Fig. 15). Prolonged white adaptation caused increases in the number of both core and annular TH neurons, with a slower time course than the initial changes in pigmentation (Supplementary Fig. 16). These results suggest that transmitter respecification provides a behavioural advantage. In contrast, dark adaptation of larvae with DA annular and core neurons placed on a dark background for 30 min occurred at the same rate as in larvae with only DA core neurons (Supplementary Fig. 17). This result is consistent with dark adaptation in adult frogs, which occurs in the absence of TH VSC inhibition of the melanotrope cells and by means of a different circuit involving the activity of serotonergic raphe neurons[14,18]. These dynamics contribute to plasticity at the neuroendocrine–melanotrope interface[29,30].

To investigate the independent function of these newly dopaminergic neurons we selectively eliminated the core DA neurons. By testing a range of concentrations of the dopaminergic toxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP)[31,32], we found that 700 μM MPTP for 8 h kills the core VSC neurons while sparing other DA nuclei (Fig. 6a). Because the number of TH neurons in the VSC remains constant for the first 40 min of light adaptation (Supplementary Fig. 16), we evaluated dark-raised, MPTP-treated larvae after 30 min exposure to light, and found that this treatment abolishes changes in pigmentation (Fig. 6b, c, left).

To determine the behavioural role of newly dopaminergic neurons, we exposed animals treated with MPTP and lacking VSC core neurons to illumination in the absence of MPTP and the presence of deprenyl to block production of activated MPTP. This protocol again led to induction of TH/NPY annular neurons in the VSC (Fig. 6b), consistent with recruitment from the NPY population. Notably, these neurons alone drive changes in pigmentation in response to light. A high level of ambient illumination led to aggregation of melanin granules in melanocytes, and a low level of illumination led to their dispersion, as in controls (Fig. 6c). These changes in pigmentation were blocked by exposure to sulpiride, identifying the contribution of DA neurons.

Discussion

Our results indicate that physiological levels of environmental illumination dynamically regulate the number of DA VSC neurons innervating hypothalamic melanotrope cells that control pigmentation. Appropriate postsynaptic receptors are regulated in parallel, as observed during transmitter respecification in the peripheral nervous system[8,9]. The recruitment of additional DA neurons is activity- and calcium-dependent and derived from a subpopulation of neurons that display a characteristic molecular signature and already project to a relevant target. When the retinohypothalamic projection is selectively activated, DA synthesis and transport machinery is acquired by annular NPY neurons that surround the TH core of the VSC nucleus. As previously observed in GABA- and TH-expressing neurons of tottering mutant mice[11], respecification of the TH phenotype occurs without influencing expression of the default original transmitter (NPY) in the annular neurons, and seems to follow a homeostatic rule. The number of inhibitory dopaminergic neurons increases after enhanced circuit activation due to higher light exposure, increasing inhibitory input to melanotrope cells and decreasing pigmentation. Conversely, dark-adapted animals respond to lower light-induced activation of this circuit by reducing the number of DA neurons; decreased inhibitory input to the melanotrope cells then boosts pigmentation. This form of plasticity evoked by sensory stimulation parallels use-dependent changes in neuropeptide expression in the hypothalamus[33–35] and activity-dependent alterations in cortical receptive fields in response to alteration of sensory experience[36].

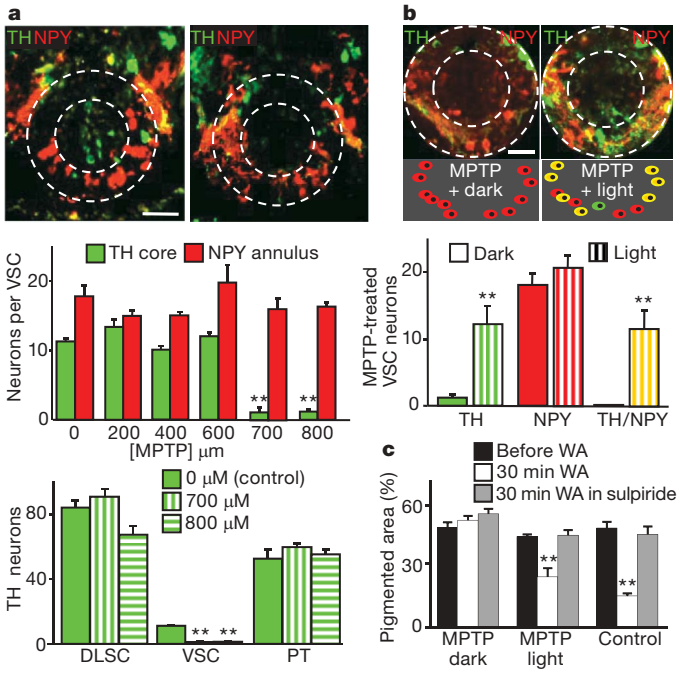


Figure 6 | Newly dopaminergic neurons regulate pigmentation. **a**, MPTP selectively eliminates core VSC neurons. Top, TH/NPY in transverse sections before (left) and after (right) 8 h incubation of larvae in 700 μM MPTP in the dark. Middle, the effect of different MPTP concentrations on numbers of TH/NPY VSC neurons. Bottom, 700 μM MPTP eliminates VSC but not dorsolateral suprachiasmatic nucleus (DLSC) and posterior tuberculum (PT) DA nuclei. **b**, Top, after MPTP treatment TH is largely absent from NPY annular neurons after dark adaptation (left), and is abundant after 2 h light adaptation (right). Middle, scheme illustrating transmitter expression in annular neurons in these conditions. Bottom, quantification of induction of TH; **a**, **b**, $N > 6$ stage-42 larvae raised in the dark. Scale bars: 50 μm. **c**, Illumination-dependent changes in skin pigmentation are rescued when MPTP treatment is followed by 2 h exposure to light. Pigmentation of three groups of larvae raised in the dark. Pigmentation was measured before (black bars) or after (white bars) 30 min of white adaptation (WA), or after WA in the presence of sulpiride (grey bars). MPTP-dark, MPTP followed by 2 h incubation in the dark abolishes the change in pigmentation in response to 30 min white-background adaptation. MPTP-light, MPTP followed by 2 h incubation in the light recovers most of the change in pigmentation in response to a 30-min exposure to white background, relative to control (three right bars). Sulpiride (10 nM) during 30-min exposure to white background blocks the reduction in pigmented area (grey bars, centre and right groups). $N \geq 10$ stage-42 larvae. Error bars, s.e.m. $^{**}P<0.001$.

We propose a model in which the neurons innervating a class of target neurons express different transmitters but belong to a constellation of cells in which the transmitter phenotype can be expanded (by co-expression) or reduced (by elimination), depending on the physiological requirements of the regulated behaviour. Global changes in activity caused by ion channel overexpression throughout the brain drive ectopic TH expression, but this includes neurons that do not express NPY and LIM1/2 and do not innervate melanotrope cells. These findings indicate a general role of activity in homeostatic specification of monoamines expressed by CNS neurons. Our results demonstrate that ectopic neurotransmitter expression can be functionally significant when activity is manipulated by activation of the neural circuit controlling a specific behaviour.

Bright light therapy has been used for more than 20 years to treat seasonal affective disorder (SAD), a form of major depression characterized by recurrent seasonal episodes[37]. Typically a fall/winter onset is followed by full remission of symptoms or a switch into mania during spring/summer. Exposure to bright artificial light, termed phototherapy, a treatment currently recommended in clinical guidelines for SAD[38], is as effective as antidepressant medications.

Dysfunction of dopaminergic signalling seems to be an important element of SAD, because depressed symptoms are induced in SAD patients during summer remission by pharmacologically inhibiting TH[39]. Although it has been shown in primates, as in amphibians, that retinal ganglion cells project monosynaptically to the suprachiasmatic nucleus and DA neurons in the hypothalamus[16], the mechanism through which light therapy exerts its effect is not known. Our results link light exposure to neurotransmitter respecification and neuroendocrine function.

Our findings demonstrate that physiological stimuli can respecify neurotransmitter expression and functional output of a specific neuronal network by harnessing circuit activity: the nervous system then identifies the right molecules to express, and delivers them at the right time, to the right place and in the right dose. Signalling molecules potentially implicated in this form of plasticity include α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)/kainate ionotropic glutamate receptors involved in the light-dependent excitatory response of DA suprachiasmatic neurons[40] and brain-derived growth factor, fibroblast growth factor and microRNAs that have roles in establishing the proper number of DA neurons[41–44]. This approach could guide future research to discover new brain stimulation methods tuned to activate selected neuronal circuits, aimed at preventing or slowing the progression of cognitive and neurodegenerative disorders before discovery of the activity-dependent molecular mechanisms involved.

METHODS SUMMARY

Illumination experiments. We determined the optimal condition for altering skin pigmentation and numbers of TH neurons by raising larvae under two illumination conditions. In the first, larvae were placed on black, grey or white backgrounds (corresponding to N 2.25, N 7.25 or N 9.5 on the Munsell neutral value scale of a monochromatic colour wheel) with constant illumination (331 lux (lx)). In the second, larvae on a white, grey or black background were exposed to constant illumination (6,560 lx, measured with a light meter (Sper Scientific)—equivalent to shady illumination on a sunny day), or to constant dark or to 12 h light/12 h dark. Light on white background proved more effective than light on a dark background. Subsequently, when testing the effect of light on pigmentation or the numbers of TH neurons, animals were usually raised in the dark (stages 35–42) and exposed (typically for 2 h at stage 42) to these optimal conditions. For some tests of the functional significance of newly TH neurons, animals were raised under different light/background conditions and a conditioning step of 30 min light/white background exposure was imposed to achieve similar levels of skin pigmentation before comparing light or dark adaptation between these groups.

Neuropharmacology. *In vivo* drug application was achieved with agarose beads[3], implanted 100–200 μm into the brain of stage-41 larvae. For selective ablation of VSC DA neurons, MPTP was bath-applied in the dark to larvae for 8 h at stage 41. 100 μM deprenyl (Sigma) was bath-applied to inhibit monoamine oxidase.

Immunocytochemistry. The number of neurons in the VSC was quantified by counting the cells within the core and the surrounding annular region in sections of the diencephalon. Details are described in Methods.

Calcium imaging. Procedures are described in Methods.

***In situ* hybridization.** *TH* mRNA was detected by locked nucleic-acid-based *in situ* hybridization[45] optimized for cryostat sections with the tyramide-fluorescein amplification system (Perkin Elmer).

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

Received 28 July; accepted 21 October 2008.

1. Walicke, P. A. & Patterson, P. H. On the role of Ca^{2+} in the transmitter choice made by cultured sympathetic neurons. *J. Neurosci.* **1,** 343–350 (1981).
2. Brosenitsch, T. A. & Katz, D. M. Expression of Phox2 transcription factors and induction of the dopaminergic phenotype in primary sensory neurons. *Mol. Cell. Neurosci.* **20,** 447–457 (2002).
3. Borodinsky, L. N. *et al.* Activity-dependent homeostatic specification of transmitter expression in embryonic neurons. *Nature* **429,** 523–530 (2004).
4. Gomez-Lira, G., Lamas, M., Romo-Parra, H. & Gutierrez, R. Programmed and induced phenotype of the hippocampal granule cells. *J. Neurosci.* **25,** 6939–6946 (2005).
5. Catalano, S. M., Chang, C. K. & Shatz, C. J. Activity-dependent regulation of NMDAR1 immunoreactivity in the developing visual cortex. *J. Neurosci.* **17,** 8376–8390 (1997).
6. Kidd, F. L. & Isaac, J. T. R. Developmental and activity-dependent regulation of kainate receptors at thalamocortical synapses. *Nature* **400,** 569–573 (1999).
7. Shi, J., Townsend, M. & Constantine-Paton, M. Activity-dependent induction of tonic calcineurin activity mediates a rapid developmental downregulation of NMDA receptor currents. *Neuron* **28,** 103–114 (2000).
8. Brunelli, G. *et al.* Glutamatergic reinnervation through peripheral nerve graft dictates assembly of glutamatergic synapses at rat skeletal muscle. *Proc. Natl Acad. Sci. USA* **102,** 8752–8757 (2005).
9. Borodinsky, L. N. & Spitzer, N. C. Activity-dependent neurotransmitter-receptor matching at the neuromuscular junction. *Proc. Natl Acad. Sci. USA* **104,** 335–340 (2007).
10. Fletcher, C. F. *et al.* Absence epilepsy in tottering mutant mice is associated with calcium channel defects. *Cell* **87,** 607–617 (1996).
11. Hess, E. J. & Wilson, M. C. Tottering and leaner mutations perturb transient developmental expression of tyrosine-hydroxylase in embryologically distinct Purkinje-cells. *Neuron* **6,** 123–132 (1991).
12. Ubink, R., Tuinhof, R. & Roubos, E. W. Identification of suprachiasmatic melanotrope-inhibiting neurons in *Xenopus laevis*: A confocal laser-scanning microscopy study. *J. Comp. Neurol.* **397,** 60–68 (1998).
13. Tuinhof, R. *et al.* Involvement of retinohypothalamic input, suprachiasmatic nucleus, magnocellular nucleus and locus-coeruleus in control of melanotrope cells of *Xenopus-laevis* — a retrograde and anterograde tracing study. *Neuroscience* **61,** 411–420 (1994).
14. Kramer, B. M. R. *et al.* Dynamics and plasticity of peptidergic control centres in the retino–brain–pituitary system of *Xenopus laevis. Microsc. Res. Tech.* **54,** 188–199 (2001).
15. Kolk, S. M., Berghs, C., Vaudry, H., Verhage, M. & Roubos, E. W. Physiological control of Xunc18 expression in neuroendocrine melanotrope cells of *Xenopus laevis. Endocrinology* **142,** 1950–1957 (2001).
16. Abizaid, A., Horvath, B., Keefe, D. L., Leranth, C. & Horvath, T. L. Direct visual and circadian pathways target neuroendocrine cells in primates. *Eur. J. Neurosci.* **20,** 2767–2776 (2004).
17. Logan, D. W., Burn, S. F. & Jackson, I. J. Regulation of pigmentation in zebrafish melanophores. *Pigment Cell Res.* **19,** 206–213 (2006).
18. Roubos, E. W., Scheenen, W. & Jenks, B. G. in *Trends in Comparative Endocrinology and Neurobiology* 172–183 (2005).
19. Nordland, J. J., *et al. The Pigmentary System: Physiology and Pathophysiology* (Oxford Univ. Press, 2006).
20. Tonosaki, Y., Nishiyama, K., Honda, T., Ozaki, N. & Sugiura, Y. D-2-like dopamine-receptor mediates dopaminergic or gamma-aminobutyric acidergic inhibition of melanotropin-releasing hormone release from the pars intermedia in frogs (*Rana-nigromaculata*). *Endocrinology* **136,** 5260–5265 (1995).
21. Akopian, A. & Witkovsky, P. D2 dopamine receptor-mediated inhibition of a hyperpolarization-activated current in rod photoreceptors. *J. Neurophysiol.* **76,** 1828–1835 (1996).
22. Wang, Y., Harsanyi, K. & Mangel, S. C. Endogenous activation of dopamine D2 receptors regulates dopamine release in the fish retina. *J. Neurophysiol.* **78,** 439–449 (1997).
23. Li, L. & Dowling, J. E. Effects of dopamine depletion on visual sensitivity of zebrafish. *J. Neurosci.* **20,** 1893–1903 (2000).
24. Green, C. B., Liang, M. Y., Steenhard, B. M. & Besharse, J. C. Ontogeny of circadian and light regulation of melatonin release in *Xenopus laevis* embryos. *Dev. Brain Res.* **117,** 109–116 (1999).
25. Mastick, G. S. & Andrews, G. L. Pax6 regulates the identity of embryonic diencephalic neurons. *Mol. Cell. Neurosci.* **17,** 190–207 (2001).
26. Wullimann, M. F. & Rink, E. Detailed immunohistology of Pax6 protein and tyrosine hydroxylase in the early zebrafish brain suggests role of *Pax6* gene in development of dopaminergic diencephalic neurons. *Dev. Brain Res.* **131,** 173–191 (2001).
27. Vazquez-Martinez, R. *et al.* Melanotrope cell plasticity: a key mechanism for the physiological adaptation to background color changes. *Endocrinology* **142,** 3060–3067 (2001).
28. Berghs, C., Tanaka, S., VanStrien, F. J. C., Kurabuchi, S. & Roubos, E. W. The secretory granule and pro-opiomelanocortin processing in *Xenopus* melanotrope cells during background adaptation. *J. Histochem. Cytochem.* **45,** 1673–1682 (1997).
29. Zhang, H. *et al.* Calcium channel kinetics of melanotrope cells in *Xenopus laevis* depend on environmental stimulation. *Gen. Comp. Endocrinol.* **156,** 104–112 (2008).
30. Jenks, B. G., Kidane, A. H., Scheenen, W. & Roubos, E. W. Plasticity in the melanotrope neuroendocrine interface of *Xenopus laevis. Neuroendocrinology* **85,** 177–185 (2007).
31. Lam, C. S., Korzh, V. & Strahle, U. Zebrafish embryos are susceptible to the dopaminergic neurotoxin MPTP. *Eur. J. Neurosci.* **21,** 1758–1762 (2005).
32. McKinley, E. T. *et al.* Neuroprotection of MPTP-induced toxicity in zebrafish dopaminergic neurons. *Brain Res. Mol. Brain Res.* **141,** 128–137 (2005).
33. Olive, S., Rougon, G., Pierre, K. & Theodosis, D. T. Expression of a glycosyl phosphatidylinositol-anchored adhesion molecule, the glycoprotein F3, in the adult rat hypothalamoneurohypophyseal system. *Brain Res.* **689,** 271–280 (1995).
34. El Majdoubi, M., Poulain, D. A. & Theodosis, D. T. Activity-dependent morphological synaptic plasticity in an adult neurosecretory system: magnocellular oxytocin neurons of the hypothalamus. *Biochem. Cell Biol.* **78,** 317–327 (2000).

35. Mueller, N. K., Di, S., Paden, C. M. & Herman, J. P. Activity-dependent modulation of neurotransmitter innervation to vasopressin neurons of the supraoptic nucleus. *Endocrinology* **146,** 348–354 (2005).
36. Froemke, R. C., Merzenich, M. M. & Schreiner, C. E. A synaptic memory trace for cortical receptive field plasticity. *Nature* **450,** 425–429 (2007).
37. Lam, R. W. & Levitan, R. D. Pathophysiology of seasonal affective disorder: a review. *J. Psychiatry Neurosci.* **25,** 469–480 (2000).
38. Lam, R. W. & Levitt, A. J. *Canadian Consensus Guidelines for the Treatment of Seasonal Affective Disorder* (Clinical and Academic Publishing, 1999).
39. Lam, R. W., Tam, E. M., Grewal, A. & Yatham, L. N. Effects of α-methyl-para-tyrosine-induced catecholamine depletion in patients with seasonal affective disorder in summer remission. *Neuropsychopharmacology* **25,** S97–S101 (2001).
40. Michel, S., Itri, J. & Colwell, C. S. Excitatory mechanisms in the suprachiasmatic nucleus: the role of AMPA/KA glutamate receptors. *J. Neurophysiol.* **88,** 817–828 (2002).
41. Baquet, Z. C., Bickford, P. C. & Jones, K. R. Brain-derived neurotrophic factor is required for the establishment of the proper number of dopaminergic neurons in the substantia nigra pars compActa. *J. Neurosci.* **25,** 6251–6259 (2005).
42. McFarlane, S., McNeill, L. & Holt, C. E. FGF signaling and target recognition in the developing *Xenopus* visual system. *Neuron* **15,** 1017–1028 (1995).
43. Kim, J. *et al.* A microRNA feedback circuit in midbrain dopamine neurons. *Science* **317,** 1220–1224 (2007).
44. Goridis, C. & Rohrer, H. Specification of catecholaminergic and serotonergic neurons. *Nature Rev. Neurosci.* **3,** 531–541 (2002).
45. Obernosterer, G., Martinez, J. & Alenius, M. Locked nucleic acid-based *in situ* detection of microRNAs in mouse tissue sections. *Nature Protocols* **2,** 1508–1514 (2007).

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

Acknowledgements We thank D. Berg, L. Borodinsky and R. Levine for critical comments on the manuscript and I-T. Hsieh and D. Boassa for technical support. This work was supported by a grant to N.C.S. from the National Institutes of Health.

Author Contributions D.D. and N.C.S. planned the project, D.D. designed and carried out the experiments and performed data analysis, and D.D. and N.C.S. wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.D. (ddulcis@biomail.ucsd.edu).

METHODS

Animals. For experiments that did not test illumination-dependence, animals were raised on a 12 h light/12 h dark cycle on a grey or a black background. For experiments testing the effects of illumination, animals were raised as described in the Methods Summary. We quantified skin pigmentation by scoring pixel intensities of lateral views of 200 μm × 600 μm rectangular areas immediately posterior to the gut of stage-42 larvae, thresholding, and expressing the result as the percentage of the black area (Image J; threshold set to 95). The temperature of the water next to the larvae, assessed with a thermistor (Fisher Scientific), was stable across all conditions and was between 22 °C and 23 °C.

Eye enucleation. Experiments were performed to confirm that retinal input controls skin pigmentation. The eye primordium was removed unilaterally or bilaterally at stage 35, leaving the pineal gland intact, and the effect of ambient illumination on pigmentation was tested 24 h later. Sham surgery consisted of creating a small opening in the skin covering the eye and rubbing it with dull forceps.

Neuropharmacology. Bath application of 10 nM sulpiride (Sigma) in 10% Marc's modified ringers (MMR) was used to achieve D_2 DA receptor (D_2R) blockade. To test whether activation of D_2Rs in the retina contributes to background adaptation we exposed stage-42 larvae raised in the dark on a black background (DBB) to a D_2R agonist (1 μM quinpirole hydrochloride, Sigma) in the dark for 30 min. The effect of quinpirole on pigmentation was compared to eye-enucleated stage-42 larvae also exposed for 30 min to quinpirole in the dark. To test whether D_1Rs can contribute to the larval dark/white adaptation we exposed stage-42 larvae raised for 28 h in the light on black background (LBB) to either 1 μM D_1R agonist (SKF-38393; Sigma) or 1 μM D_1R antagonist (SCH23390; Sigma) during 30 min of white (LWB) or dark (DBB) adaptation. Pigmentation displayed by dark- and white-adapted treated larvae was compared to controls. Beads used to deliver drugs to the VSC were washed in 2 mM calcium saline for 2–4 h at 22–23 °C and loaded for 1 h in a bi-directional rotator (Barnstead International) with a solution containing 10 μM calcein-AM (Invitrogen), 10 μM BAPTA-AM (Invitrogen) or 10 μM TTX (Sigma). Larvae were anesthetized with tricaine, recovered after bead implantation, and developed for 24 h before fixation. To quantify diffusion of fluorescent calcein, larvae were first fixed for 40–60 min in a solution of 0.1× PBS containing 40 mg ml^{-1} EDAC (Sigma) and then in 4% paraformaldehyde (PFA) for 1.5 h. Stocks of 10× MMR physiological saline (pH 7.8) were prepared by dissolving NaCl (1 M), KCl (20 mM), $MgSO_4$ (10 mM), HEPES (50 mM), EDTA (1 mM) and $CaCl_2$ (20 mM) in Fisher water and autoclaving for 20 min.

Immunocytochemistry. Larvae were fixed in 4% PFA in phosphate-buffered saline (PBS, pH 7.4) for 2 h at 4 °C, soaked in 30% sucrose for 2.5 h or overnight at 4 °C, and embedded in OCT (Fisher Scientific). For DA immunocytochemistry, the fixation solution contained 0.1 M cacodylate, 10 g l^{-1} sodium metabisulphite and 3% glutaraldehyde (all from Sigma). Ten-micrometre cryostat sections were made through the entire brain, and regions of interest were identified anatomically afterwards. For whole-mount experiments, larvae were fixed and incubated in PBS–0.5% Triton for 3 days. The CNS was then dissected and carried through the immunostaining protocol below. Slides and whole mounts were incubated in a blocking solution of 2% goat serum and 1% fish gelatin for 0.5–2 h at 20 °C, followed by overnight incubation with primary antibodies to TH (Imgenex), NPY (Immunostar Inc.), DAT (Advanced Targeting Systems), VMAT2 (Santa Cruz Biotechnology Inc.), DA (Novus Biologicals), BrdU (Calbiochem), POMC (Novus), PAX6 (Covance) and LIM1/2 (DSHB) at 4 °C; specimens were incubated for 2 h with fluorescently tagged secondary antibodies at 20 °C. Immunoreactivity was examined on a Zeiss Axioskop with a 20× air objective using a Xenon arc lamp, attenuated by neutral density filters and the appropriate excitation and emission filters for Alexa 350, 488, 555, 594 and 647, Cy3, DAPI and DRAQ5 fluorophores. Images were acquired and analysed with Image J, Axiovision and Corel Photopaint. The TH antibody was used as a marker for DA neurons of the VSC, given demonstration of TH/DA double labelling[46] and *TH* mRNA/TH protein colocalization. Differences in labelling were considered significant at $P < 0.05$ (Student's *t*-test). The small size of the VSC, with a transverse diameter of ~250 μm and a rostrocaudal extent of 30–50 μm, facilitated counting of the neurons.

Ion channel overexpression. Human Kir2.1 and rat $Na_v2a\alpha\beta$ DNA constructs were gifts from E. Marban and W. Catterall and were subcloned, transcribed and injected as previously described[3]. Control and activity-manipulated larvae were raised in a 12 h light (331 lx)/12 h dark cycle on a grey background.

Calcium imaging. The CNS was dissected from larvae at stages 35–42 using jewellers' forceps, tungsten needles and glass probes, and pinned with fine tungsten wires in a small chamber. Midbrain dopaminergic neurons were exposed by sectioning through the hypothalamus with micro-scalpels made of sharp razor blades. Fluorescence of the calcium indicator Fluo-4AM (Invitrogen) was used to study increases of $[Ca]_i$ in neurons. Brains were incubated for 30 min to 1 h in MMR containing 2–5 μM dye and 0.01% Pluronic F-127, and washed in saline before imaging. For study of calcium spikes, images were acquired at 0.2 Hz for 1 h periods with a BioRad MRC 1024 or Leica SP5 laser confocal system with a 20× water immersion objective. Image stacks were imported into Image J for analysis. Using the line-drawing tools in the program, a region of interest was traced around the site of the calcium transient and a measurement of the average pixel intensity within the region of interest was acquired. Data were exported to Excel (Microsoft) for analysis. To retain Fluo-4AM fluorescence of the imaged cells, preparations were fixed for 30 min at room temperature with 40 mg ml^{-1} EDAC (Sigma) and 1 h 4% PFA before immunocytochemistry processing.

BrdU and TUNEL staining. To understand whether TH/NPY annular neurons represent newly born cells or pre-existing neurons, cells proliferating during white adaptation were identified by BrdU (Sigma) labelling. Stage-41 larvae were exposed to BrdU by immersion in 4 mg ml^{-1} BrdU and 10% MMR for 6 h. Specimens (Stage 42) were fixed and sucrose-cryoprotected as described previously. Ten-micrometre cryostat sections were treated for 20 min in 2 M hydrochloric acid (HCl; Fisher Scientific) for antigen retrieval, and then washed and incubated with mouse anti-BrdU antibody (Calbiochem) overnight. To determine whether white adaptation influences cell death in the hypothalamus, *in situ* cell death detection KIT fluorescein TUNEL staining (Roche) of 2-h-white-adapted larvae was compared to staining of black-adapted larvae. Positive TUNEL controls were obtained by 10 min incubation with DNase I (3 U ml^{-1} in 50 mM Tris-HCl, Ambion) at 25 °C to induce DNA breaks before labelling procedures.

Statistics. Data were collected from 6–9 *X. laevis* larvae from three or more clutches at stages 35, 37, 42 and 46 and are presented as mean ± s.e.m. Significance was assessed with Student's *t*-test. $^{**}P < 0.001$.

46. Gonzalez, A. & Smeets, W. Comparative analysis of dopamine and tyrosine hydroxylase immunoreactivities in the brain of 2 amphibians, the anuran *Rana ridibunda* and the urodele *Pleurodeles waltlii*. *J. Comp. Neurol.* **303**, 457–477 (1991).

ARTICLES

Deconstructing voltage sensor function and pharmacology in sodium channels

Frank Bosmans[1,2], Marie-France Martin-Eauclaire[3] & Kenton J. Swartz[1]

Voltage-activated sodium (Na_v) channels are crucial for the generation and propagation of nerve impulses, and as such are widely targeted by toxins and drugs. The four voltage sensors in Na_v channels have distinct amino acid sequences, raising fundamental questions about their relative contributions to the function and pharmacology of the channel. Here we use four-fold symmetric voltage-activated potassium (K_v) channels as reporters to examine the contributions of individual S3b–S4 paddle motifs within Na_v channel voltage sensors to the kinetics of voltage sensor activation and to forming toxin receptors. Our results uncover binding sites for toxins from tarantula and scorpion venom on each of the four paddle motifs in Na_v channels, and reveal how paddle-specific interactions can be used to reshape Na_v channel activity. One paddle motif is unique in that it slows voltage sensor activation, and toxins selectively targeting this motif impede Na_v channel inactivation. This reporter approach and the principles that emerge will be useful in developing new drugs for treating pain and Na_v channelopathies.

Na_v channels in nerve and muscle cells open and close, or 'gate', in response to changes in membrane voltage[1]. Mutations in Na_v channels can cause a variety of inherited disorders, such as epilepsy and myotonia[2,3]. Furthermore, Na_v channels are strategically positioned within nociceptive signalling pathways, with mutations leading to severe pain disorders[4,5]. Although the development of drugs interacting with Na_v channels is of widespread medical importance[6], progress has been slow because the architecture and gating mechanisms of these ion channels are complex. Their channel-forming α-subunits contain four homologous domains (I–IV), or pseudosubunits, each containing six transmembrane segments (S1–S6) (Fig. 1a). The S5–S6 segments collectively form a central pore for Na^+, with the S1–S4 segments from each domain forming the surrounding voltage sensors[1]. Each of the four voltage sensors activate in response to changes in voltage; however, those in domains I–III are most important for channel opening, whereas the one in domain IV is crucial for fast inactivation[7–10]. Thus, drugs and toxins can interact with multiple regions of Na_v channels to influence their activity, with the four voltage sensors representing rich, yet complex targets. The similarity of the four voltage sensors raises the possibility of multiple binding sites, with occupancy of each having distinct effects on gating. Indeed, seven different receptor sites have been proposed for drugs and toxins that interact with Na_v channels[1,11], but only a few have been molecularly defined. Here we develop an approach for studying the unique contributions of the four voltage sensors in Na_v channels to gating and pharmacology. We identify S3b–S4 paddle motifs that can be transplanted from Na_v channels into K_v channels without disrupting function, while transferring sensitivity to classical Na_v channel toxins. Our results uncover rich interactions between toxins and multiple Na_v channel paddle motifs, and reveal how domain-specific interactions can be used to reshape Na_v channel activity.

Transferring paddle motifs between Na_v and K_v channels

Studies on voltage sensors in K_v channels have identified an S3b–S4 helix–turn–helix motif, the voltage-sensor paddle, which moves at the protein–lipid interface to drive activation of the voltage sensors and opening of the pore[12–16]. The paddle motif is an important pharmacological target in K_v channels, as tarantula toxins that partition into membranes interact with this region to inhibit channel opening[12,17–24]. Our initial goal was to determine whether paddle motifs can be defined in Na_v channels and whether they fulfil similar functions. As paddle motifs are portable modules that are interchangeable between K_v channels and voltage-sensing proteins[12], we wondered whether distinct paddle motifs in each of the four voltage sensors of Na_v channels could be transplanted into a homotetrameric K_v channel, so as to study them in isolation (Fig. 1a). We initially transplanted the paddle motifs from $rNa_v1.2a$ and $rNa_v1.4$ channels into the $K_v2.1$ channel (see Methods for nomenclature definition). Although the sequences of these Na_v and K_v channel paddle motifs vary substantially, constructs containing specific S3b–S4 regions of Na_v channels result in fully functional channels that display robust voltage-activated potassium currents (Fig. 1b, c, Supplementary Figs 1, 2a, Supplementary Table 1). Functional chimaeras were also obtained by transplanting the paddle motif from $rNa_v1.4$ into $K_v1.3$ and Shaker K_v channels (Supplementary Figs 1, 2b, c, Supplementary Table 2), demonstrating that paddle motifs of Na_v and K_v channels are generally interchangeable.

Transferring toxin sensitivity between Na_v and K_v channels

We next asked whether transplanting isolated paddle motifs from Na_v channels into K_v channels faithfully transfers sensitivity to toxins, in which case the K_v channel could be used as a reporter for investigating interactions between toxins and individual Na_v channel paddle motifs. Tarantula toxins related to those targeting paddle motifs in K_v channels[12,17–22,24] (Supplementary Fig. 1d) inhibit Na_v channels by modifying gating[25–27], but their presumed receptors within the four voltage sensors of Na_v channels have yet to be defined. We focused on three tarantula toxins, PaurTx3, ProTx-I and ProTx-II[25–27], and tested whether they interact with the $K_v2.1$ channel containing each of the four paddle motifs from $rNa_v1.2a$. Remarkably,

[1]Molecular Physiology and Biophysics Section, Porter Neuroscience Research Center, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, Maryland 20892, USA. [2]Laboratory of Toxicology, University of Leuven, 3000 Leuven, Belgium. [3]CNRS UMR 6231, CRN2M, Institut Jean Roche, Université de la Méditerranée, Marseille Cedex 20, France.

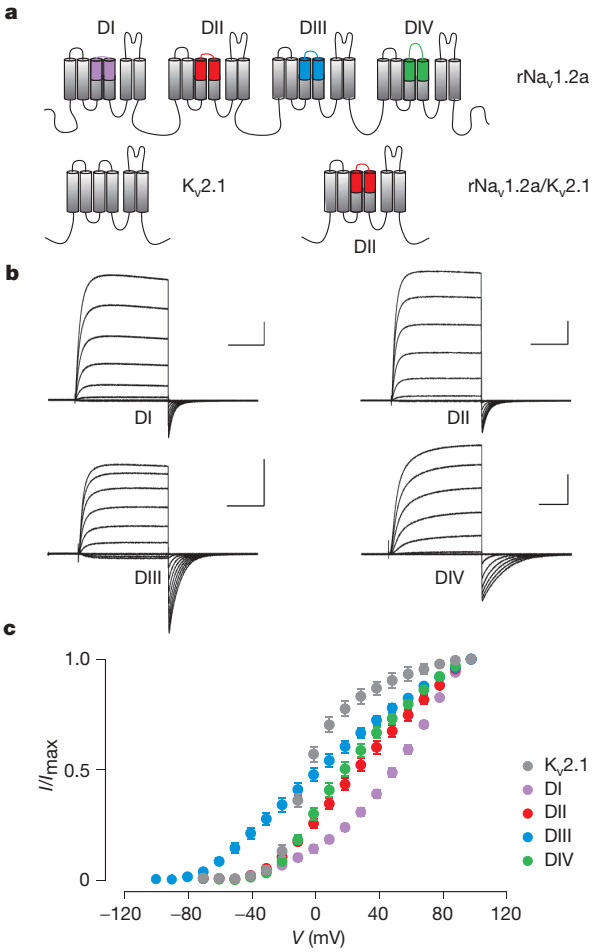


Figure 1 | Transfer of the voltage sensor paddle motifs from rNa$_v$1.2a to K$_v$2.1. a, Cartoon depicting the paddle motif transfer from the Na$_v$ channel S1–S4 voltage sensor of domain II into K$_v$2.1. Purple, domain I paddle (DI); red, domain II paddle (DII); blue, domain III paddle (DIII); green, domain IV paddle (DIV). This colour code is used in all figures. **b**, **c**, Families of potassium currents (**b**) and tail current voltage-activation relationships (**c**) for each chimaera ($n = 18$; error bars, s.e.m.). Holding voltage was −90 mV and the tail voltage was −50 mV (−80 mV for DIII). Scale bars (**b**) are 2 μA (vertical) and 100 ms (horizontal).

sensitivity to these toxins can be transferred to K$_v$2.1 along with specific Na$_v$ channel paddle motifs (Fig. 2a). PaurTx3 has the simplest profile of the three, as this toxin only inhibits the K$_v$ channel containing the paddle motif from domain II (Fig. 2a, Supplementary Fig. 3a, Supplementary Table 3). In contrast, ProTx-I and ProTx-II interact with multiple paddle motifs from rNa$_v$1.2a. ProTx-I binds to K$_v$2.1 containing the paddle motifs from domain II or IV, whereas ProTx-II interacts with those from domains I, II or IV. (ProTx-I also inhibits K$_v$2.1 (ref. 26), but the lack of toxin sensitivity for the domain I and III chimaeras support interaction of the toxin with paddle motifs.) Although it is not straightforward to directly correlate the apparent affinities of K$_v$ channel paddle chimaeras and Na$_v$ channels because the paddles are identical in K$_v$ channels and different in Na$_v$ channels, the concentration-dependence for toxin inhibition of the two channel types are remarkably similar (Supplementary Fig. 4a, b, Supplementary Table 3). These results show that tarantula toxins often interact with multiple paddle motifs in Na$_v$ channels, a scenario that would be difficult to detect using conventional approaches[27,28].

Studies on scorpion venom have established the presence of two classes of toxins, the α- and β-scorpion toxins, that interact with the voltage sensors in domains IV and II of Na$_v$ channels, respectively[1,29–33]. In the conventional view, these toxins interact with solvent-exposed extracellular loops between S3 and S4 to stabilize the voltage sensors in specific states[8,31,34]. Whether these toxins interact with the larger S3b–S4 paddle motif, or can target voltage sensors in domains other than II and IV, is unknown. To explore these possibilities we tested whether we could make K$_v$2.1 sensitive to the classical α- and β-scorpion toxins. As in the case of tarantula toxins, transfer of Na$_v$ channel paddle motifs into K$_v$2.1 renders the channel sensitive to the α-scorpion toxin AaHII; however, in this instance only the paddle motif from domain IV can transfer toxin sensitivity (Fig. 2a, Supplementary Fig. 3a). Although the apparent affinity of AaHII for the domain IV chimaera is considerably less than for rNa$_v$1.2a (Supplementary Fig. 4d, Supplementary Table 3), mutations within S4 can increase the apparent affinity by over 10-fold (Figs 3d, 4c, Supplementary Fig. 4d). These results suggest that α-scorpion toxins interact with the paddle motif from domain IV, but the modest affinity of the paddle chimaera leaves open the possibility that other regions of Na$_v$ channels contribute to the toxin receptor[35].

We also examined whether sensitivity to the β-scorpion toxin TsVII could be transferred with Na$_v$ channel paddle motifs. In contrast to AaHII, transfer of the paddle motifs from domains II, III or IV from rNa$_v$1.2a renders the K$_v$ channel sensitive to TsVII (Fig. 2a, Supplementary Fig. 3a), revealing that β-scorpion toxins can interact

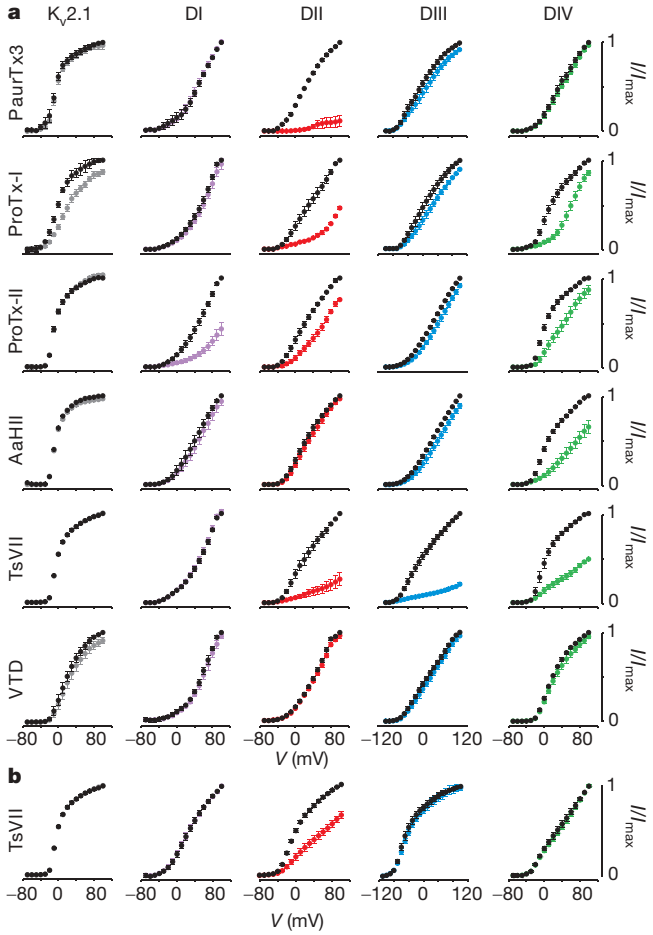


Figure 2 | Sensitivity of rNa$_v$1.2a paddle chimaeras to extracellular toxins. a, Effects of toxins on K$_v$2.1 and chimaeras where paddle motifs were transferred from rNa$_v$1.2a into K$_v$2.1. Normalized tail current voltage-activation relationships are shown, where tail current amplitude (I/I_{max}) is plotted against test voltage before (black) and in the presence of toxins (other colours). Data are grouped per toxin (horizontally; toxin named at left) and per chimaera or wild-type K$_v$2.1 (vertically; see labels at top). Concentrations used are 100 nM (PaurTx3, ProTx-I and ProTx-II); 1 μM (AaHII); 500 nM (TsVII) and 100 μM (veratridine; VTD). The plant alkaloid VTD is used as a negative control. **b**, Effects of TsVII (50 nM) on rNa$_v$1.4 paddle chimaeras. $n = 3$–5; error bars, s.e.m. The holding voltage was −90 mV, the test pulse duration was 300 ms and the tail voltage was −50 mV (−80 mV for DIII).

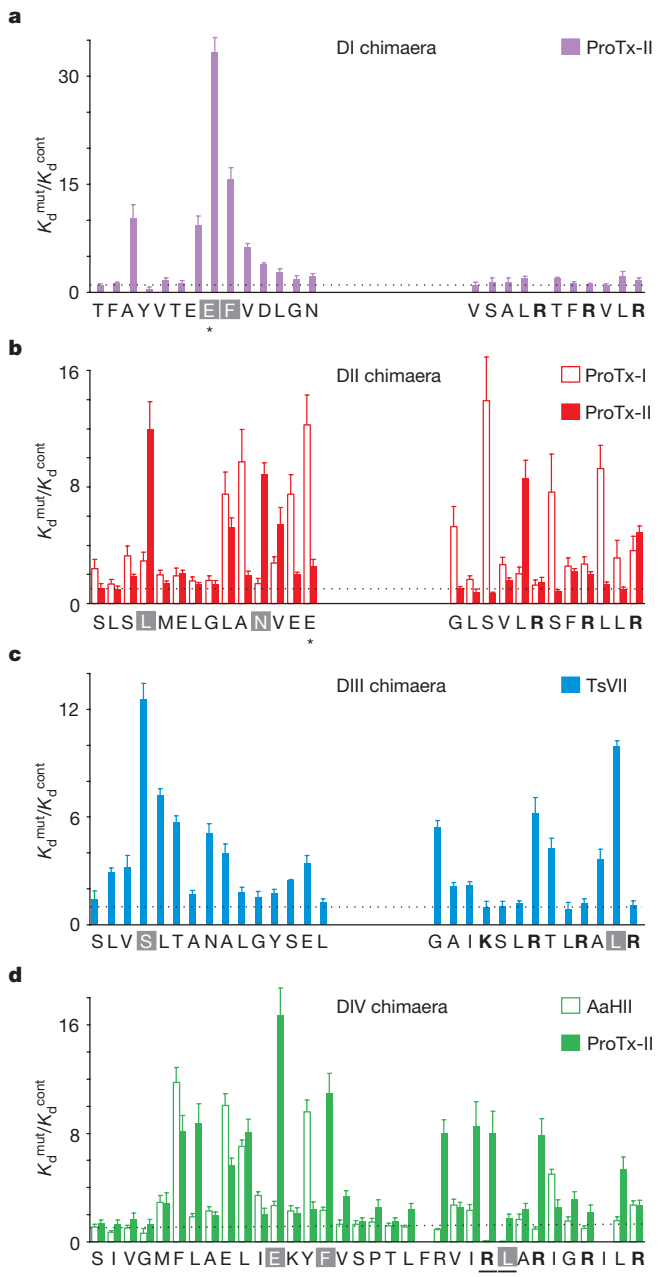


Figure 3 | Scanning mutagenesis of Na_v channel paddle motifs. **a–d**, Ala-scan (see below) of the separate $rNa_v1.2a$ paddle motifs in the DI–DIV chimaeras; changes in apparent toxin affinity (K_d^{mut}/K_d^{cont}) are plotted for individual mutants. See Methods, Supplementary Figs 4, 5c, and Supplementary Table 3 for information on K_d measurements. Most of the residues within the $rNa_v1.2a$ paddle in the four chimaera constructs were individually mutated to Ala (except for native Ala residues, which were mutated to Val). The dashed line marks a value of 1. Each mutant was initially examined using a concentration near the K_d value determined for the control chimaera (Supplementary Fig. 4). Mutants with a K_d^{mut}/K_d^{cont} value greater than five were further examined using a wider range of concentrations. Glu residues marked with asterisks were also mutated to a Lys. Bar diagrams are approximately aligned according to the sequence alignment of the different paddles (Supplementary Fig. 1). Mutants without a corresponding bar did not result in functional channels. Residues with a grey background were used in subsequent tests (Fig. 4). Mutation of two underlined residues in **d** results in an increase in AaHII affinity (K_d = 235 ± 24 nM for R1629A, 205 ± 23 nM for L1630A and 1,902 ± 102 nM for the DIV chimera). Basic residues in S4 were used to align the sequences and are indicated in bold. n = 3–5 for each toxin concentration; error bars, s.e.m.

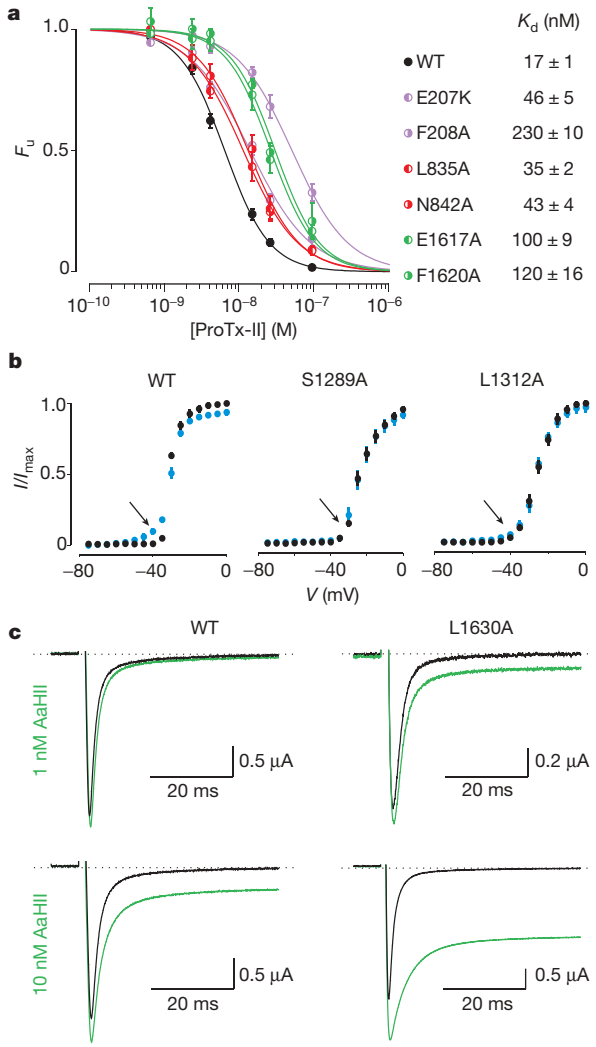


Figure 4 | Reconstitution of paddle mutants into $rNa_v1.2a$ and their effects on toxin–channel interactions. **a**, Concentration dependence for ProTx-II inhibition of $rNa_v1.2a$ (WT) and selected mutants plotted as fraction unbound (F_u) measured at negative voltages. Solid lines are fits of the Hill equation to the data with apparent K_d values shown. See Supplementary Fig. 5b and Supplementary Table 3 for further information. **b**, Normalized conductance–voltage relationships for $rNa_v1.2a$ and two DIII mutants before and after addition of 50 nM TsVII. Arrow indicates toxin effect or lack thereof. Holding voltage was −90 mV. **c**, $rNa_v1.2a$ mutation L1630A increases affinity for AaHII. Sodium currents were elicited by a depolarization to −15 mV from a holding voltage of −90 mV. Green trace is after AaHII addition. Error bars (**a**, **b**), s.e.m.

with multiple Na_v channel paddle motifs. This result is surprising, because studies on $rNa_v1.4$ suggest that TsVII only binds to the voltage sensor in domain II[36,37]. To explore whether scorpion toxin interactions vary between Na_v channel subtypes, we tested the sensitivity of the $rNa_v1.4$ paddle chimaeras to TsVII. By contrast with what is observed with the $rNa_v1.2a$ paddles, TsVII can interact selectively with the domain II paddle from $rNa_v1.4$ (Fig. 2b, Supplementary Fig. 3b). Although the apparent affinities of TsVII for the $rNa_v1.2a$ paddle chimaeras are significantly lower than the $rNa_v1.2a$ channel, the $rNa_v1.4$ domain II paddle chimaera is comparable (Supplementary Fig. 4c, Supplementary Table 3). An interesting feature of TsVII is that the β-scorpion toxin stabilizes a closed state in the paddle chimaeras (Fig. 2), whereas the toxin stabilizes the open state of $rNa_v1.2a$ (Fig. 4b) and $rNa_v1.4$ (ref. 37) channels. β-scorpion toxins can also stabilize a closed state in certain Na_v channels[32], indicating that open state stabilization is not always observed. One explanation is that toxin binding to the paddle

intrinsically stabilizes a closed state, but in certain Na_v channels the toxin can stabilize an open state when it also interacts with other regions[32,33]. Taken together, these data suggest that both tarantula and scorpion toxins interact with paddle motifs, that multiple paddle motifs can be targeted by individual toxins, and that these interactions can differ between Na_v channel subtypes.

Critical residues in the toxin–paddle interaction

To further explore the extent to which toxins interact with S3b–S4 paddle motifs, and to identify mutants for testing whether these toxin–paddle interactions actually occur in Na_v channels, we mutated each residue in the transferred paddle regions and measured changes in toxin affinity (Fig. 3, Supplementary Fig. 4, Supplementary Tables 4–7). One striking feature of the data is that mutations distributed throughout the S3b–S4 paddle motifs cause dramatic perturbations in the apparent affinity for both tarantula and scorpion toxins, suggesting that these toxins target the larger paddle motif. Many of the most influential mutations are of hydrophobic residues, but mutations of polar residues such as Glu residues in S3b and Arg residues in S4 also have pronounced effects. The present mutagenesis of the domain IV paddle motif correctly identifies several mutants previously shown to weaken α-scorpion toxin affinity in Na_v channels[31], but reveals a larger number of influential mutations, many of which have more pronounced effects on toxin affinity. This highlights an advantage of the reporter approach, which allows individual toxin–paddle interactions to be studied in isolation; toxins can interact with up to four different paddle motifs in Na_v channels and the effects of mutations in any one of those can be quite subtle because the others remain unaltered.

Another striking feature of these paddle scans is that mutations perturbing toxin–paddle interactions are unique for each toxin–paddle pair. For example, both ProTx-I and ProTx-II interact with the paddle motif from domain II, but the mutants that weaken toxin affinity only partially overlap. In addition, the mutagenesis results for ProTx-II interacting with paddle motifs from domains I, II and IV identify many influential mutations that differ between the three paddles. This comparison suggests that toxin–paddle interactions do not obey a general lock-and-key mechanism, but that the interfaces vary considerably.

Validating toxin-paddle interactions in $rNa_v1.2a$

To validate that the toxin–paddle interactions identified here actually occur in Na_v channels, we reconstituted the most influential paddle mutants into $rNa_v1.2a$ and measured changes in toxin affinity. For tarantula toxins, we examined mutations influencing ProTx-II affinity, as this toxin interacts with three of the four Na_v channel paddle motifs. Mutations within the paddle motifs of domains I, II and IV weaken the interaction of ProTx-II with $rNa_v1.2$ (Fig. 4a), suggesting that the toxin interacts with all three paddle motifs in the Na_v channel. The mutations tend to have smaller effects in the Na_v channel when compared with the paddle chimaeras (compare Fig. 4a with Fig. 3a, b, d), which makes sense because the mutations in the Na_v channel alter only one of three targeted paddle motifs, whereas in the K_v channel chimaera they alter all four. For the β-scorpion toxin TsVII, we examined mutations within the paddle motif of domain III and found that they diminish the hyperpolarizing shift produced by moderate toxin concentrations (Fig. 4b), confirming that the toxin interacts with domain III (in addition to its canonical interaction with domain II). Finally, in the case of the α-scorpion toxin AaHII, we reconstituted a gain-of-function mutation into $rNa_v1.2a$ and observed that the mutation also increases the apparent affinity of the toxin in the Na_v channel (Fig. 4c). These results demonstrate that the toxin–paddle interactions uncovered using the reporter approach actually occur in Na_v channels.

Unique character of the domain IV paddle motif

The differential coupling of the four voltage sensors in Na_v channels to opening and inactivation[8,10,38–40] is accompanied by differences in the kinetics of voltage sensor activation, with the voltage sensor of domain IV moving slower than the other three[10]. Given that the paddle motif resides in a relatively unconstrained environment in contact with the surrounding lipid membrane, and that it is the region of the voltage sensor that flexes in response to changes in voltage[12–16,41], we wondered whether the paddle motif itself might be responsible for these kinetic differences. To explore this possibility, we measured the kinetics of activation (opening) and deactivation (closing) of the four $rNa_v1.2a$ paddle chimaeras in $K_v2.1$ in response to membrane depolarization and repolarization, respectively (Fig. 5). At most voltages, the time

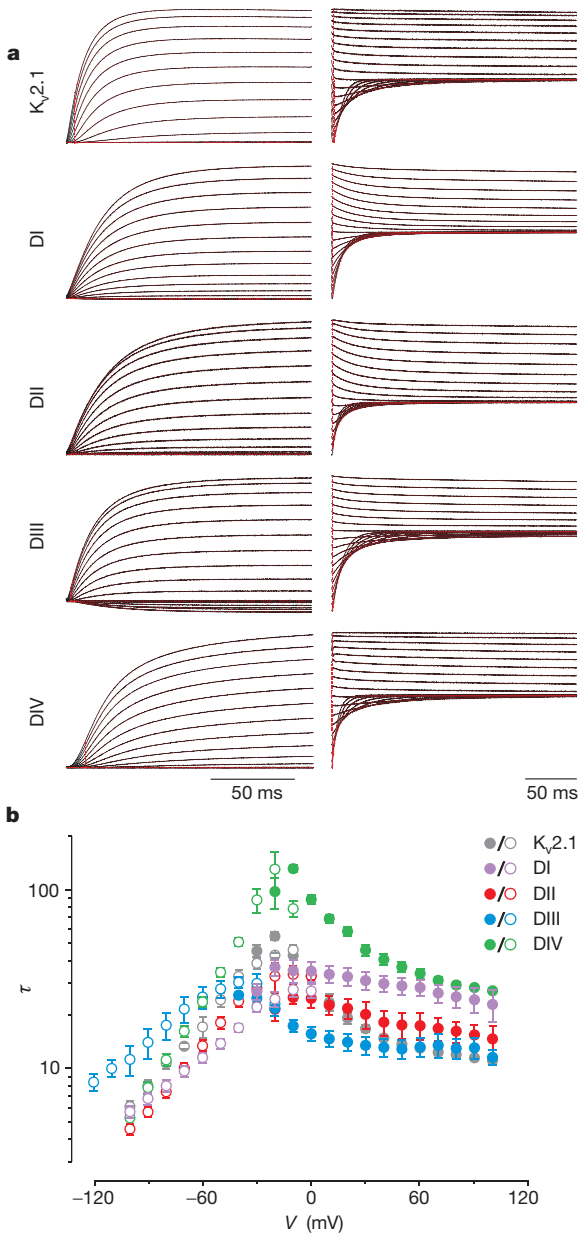


Figure 5 | Kinetics of opening and closing for $rNa_v1.2a/K_v2.1$ chimaeras. **a**, Representative macroscopic currents (black) showing channel activation (left) and channel deactivation (right) using the following voltage protocols: activation, 10 mV incrementing steps to voltages between −40 mV and +100 mV from a holding potential of −90 mV; deactivation, 10 mV incrementing steps to voltages between 0 mV and −100 mV (−120 mV for DIII) from a test voltage of between +80 mV and +100 mV (holding potential is −90 mV). Superimposed red curves are single exponential fits to the current records after initial lags in current activation. **b,** Mean time constants (τ) from single exponential fits to channel activation (filled circles) and deactivation (open circles) plotted as a function of the voltage at which the current was recorded. $n = 4–8$; error bars, s.e.m.

constants obtained by fitting single exponential functions to current relaxations are notably slower for the domain IV paddle chimaera compared with the other chimaeras and to $K_v2.1$. This difference is not unique to chimaeras between the $rNa_v1.2a$ and $K_v2.1$ channel, but is even more pronounced when transplanting paddle motifs from $rNa_v1.4$ into $K_v2.1$, $K_v1.3$ or Shaker K_v channels (Supplementary Fig. 2). These results demonstrate that the paddle motif in domain IV of Na_v channels determines the slower movement of this voltage sensor in response to changes in membrane voltage.

Shaping Na_v channel activity with domain-selective interactions

A particularly intriguing outcome of looking at toxin–paddle interactions in isolation is that patterns emerge between the domains targeted and the effects of the toxin on Na_v channel activity. Inspection of the profiles for toxin–paddle interactions (Fig. 2) reveals that by targeting paddle motifs in domains I, II or III, the overall effect of a toxin is to influence opening of Na_v channels, irrespective of whether domain IV is also targeted. The tarantula toxins PaurTx3, ProTx-I and ProTx-II, as well as the β-scorpion toxin TsVII, are examples of this relationship. In contrast, the requirements for influencing inactivation are more stringent; to do so, a toxin needs to selectively target the paddle motif in domain IV, as is the case with the α-scorpion toxin AaHII. To test whether this pattern is generally applicable across different toxin families, we searched for a tarantula toxin that selectively interacts with the paddle motif in domain IV and asked whether it influences inactivation. Using our reporter constructs, we screened various tarantula toxins, including VSTx1 (ref. 42), GxTx-1E (ref. 43), HaTx (ref. 18) and SGTx1 (ref. 44), all of which interact with paddle motifs in K_v channels[12,17–21,24,44] and are related in sequence to the Na_v channel toxins already studied (Supplementary Fig. 1d). VSTx1 (10 μM) and GxTx-1E (500 nM) are inactive against the Na_v paddle chimaeras (not shown), whereas both HaTx and SGTx1 exhibit robust effects (Fig. 6a). HaTx interacts with paddle motifs from domains I, II and IV, similarly to ProTx-II, whereas SGTx1 interacts selectively with the domain IV paddle chimaera, the profile we sought to find. When tested against $rNa_v1.2a$, HaTx exhibits robust inhibitory effects, similar to ProTx-II, without significantly altering inactivation (Fig. 6b). In contrast, SGTx1 does not inhibit $rNa_v1.2a$ or alter channel opening, but dramatically reduces the extent of inactivation (Fig. 6c), behaving as an α-scorpion toxin. These results suggest that for a toxin to influence inactivation, it must selectively interact with the paddle motif in domain IV; any additional interactions with the other paddles will alter channel opening.

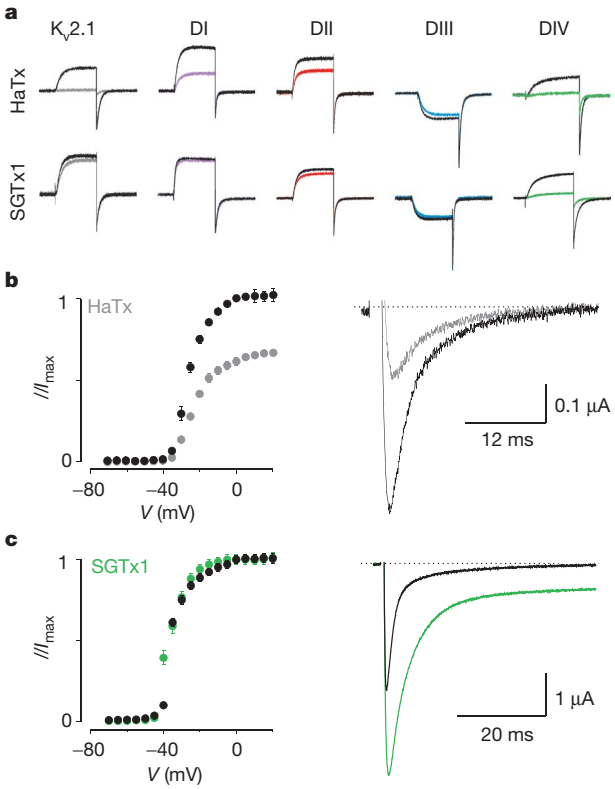


Figure 6 | Identifying a tarantula toxin selective for the paddle motif in domain IV. a, Potassium currents elicited by depolarizations near the foot of the voltage-activation curve for $K_v2.1$ and chimaeras in the absence and presence of 50 nM HaTx or 100 nM SGTx1. The holding voltage was −90 mV, the test pulse duration was 300 ms and the tail voltage was −50 mV (−80 mV for DIII). **b**, Left, conductance–voltage relationships for $rNa_v1.2a$ before and after addition of 50 nM HaTx, normalized to the maximal conductance in control. Right, sodium current elicited by a depolarization to −30 mV before and after addition of 50 nM HaTx. **c**, Left, conductance–voltage relationship of $rNa_v1.2a$ before and after addition of 100 nM SGTx1, individually normalized to the maximal conductance in either control or the presence of toxin. Right, sodium current elicited by a depolarization to −15 mV before and after addition of 100 nM SGTx1. n = 3–5 for each toxin concentration; error bars (**b**, **c**), s.e.m.

Discussion

Our results on isolating Na_v channel paddle motifs using K_v channels as reporters have three fundamental implications for the function and pharmacology of Na_v channels. First, they reveal that paddle motifs can determine the kinetics of voltage sensor activation. For Na_v channels to function properly, inactivation must proceed more slowly than opening, a property that has been attributed to slower activation of the voltage sensor in domain IV[1,8,10,38,45]. Our results show that the paddle motif in domain IV is unique because it systematically slows activation of any voltage sensor into which it is incorporated (Fig. 5, Supplementary Fig. 2). The paddle motif resides in a relatively flexible environment and moves in contact with the surrounding lipid membrane[12–16,41], raising the intriguing possibility that interactions of the domain IV paddle motif with lipids are unique and are responsible for slowing voltage sensor activation. Other aspects of Na_v channel function are unlikely to be determined by the paddle motif (for example, cooperative voltage sensor activation[10,37,46]), but the approaches described here may help to identify them.

Second, our results reveal rich interactions of tarantula and scorpion toxins with paddle motifs, and establish a number of principles for toxin–paddle interactions in Na_v channels where the four paddle motifs are not equivalent. Although α- and β-scorpion toxins are known to interact with domains IV and II, respectively[31,32], the present results demonstrate that each of the paddle motifs are targeted by toxins (Fig. 2). In fact, multiple paddle motifs are often targeted by a toxin. Even the β-scorpion toxins can interact with other domains in addition to the canonical interaction with domain II. One unexpected finding is that the profiles of toxin–paddle interactions can differ between subtypes of Na_v channels, suggesting that such profiles may be a way of distinguishing Na_v channels. Our mutagenesis results with the four paddle motifs in $rNa_v1.2a$ also show that the pattern of mutants altering toxin affinity is specific for each toxin–paddle pair (Fig. 3). The most influential mutants can vary considerably for one toxin interacting with different paddles, or for two related toxins interacting with a single paddle motif.

Last, the profiles of toxin–paddle interactions that emerge from our studies reveal an important relationship between the effect of a toxin on Na_v channel activity and domain-specific interactions. Alterations in channel opening can be achieved with molecules that interact with any of the first three paddle motifs in Na_v channels without regard for whether domain IV is also targeted (for example HaTx, ProTx-II, TsVII), whereas influencing inactivation requires that a molecule interact exclusively with the paddle in domain IV (for example SGTX1, AaHII). These domain-specific interactions have important

implications for designing drugs to reshape Na_v channel activity. Channelopathies like idiopathic ventricular fibrillation[47] and long QT syndrome type 3 (ref. 48) are associated with accelerated Na_v channel inactivation, which could be restored by selectively targeting the domain IV paddle motif. In contrast, the hyperpolarization of Na_v channel opening seen in disorders like generalized epilepsy with febrile seizures type 2 (ref. 49) and hyperkalaemic periodic paralysis[50] could be managed with drugs targeting any paddle motif within the first three domains. Our reporter approach for identifying toxins that specifically target the paddle motif in domain IV (Fig. 6) provides a conceptual example for identifying drugs that interact with specific paddle motifs in Na_v channels.

METHODS SUMMARY

Channel constructs were expressed in *Xenopus* oocytes[21] and studied using two-electrode voltage-clamp recording techniques. Na_v channels were co-expressed with the β_1 subunit. For K_v channel experiments, the external recording solution contained 50 mM KCl, 50 mM NaCl, 5 mM HEPES, 1 mM $MgCl_2$ and 0.3 mM $CaCl_2$, pH 7.6 with NaOH. KCl was replaced by RbCl for Shaker experiments. For Na_v channel experiments, the external recording solution contained 96 mM NaCl, 2 mM KCl, 5 mM HEPES, 1 mM $MgCl_2$ and 1.8 mM $CaCl_2$, pH 7.6 with NaOH.

Voltage–activation relationships were obtained by measuring tail currents for K_v channels or steady-state currents and calculating conductance for Na_v channels, and a single Boltzmann function was fitted to the data. Occupancy of closed or resting channels by toxins was examined using negative holding voltages where open probability was low, and the fraction of unbound channels (F_u) was estimated using depolarizations that are too weak to open toxin-bound channels, as described previously[17–21,24,44]. The apparent equilibrium dissociation constant (K_d) for toxin interaction with K_v channel constructs was calculated assuming four independent toxin-binding sites per channel, with single occupancy being sufficient to inhibit opening in response to weak depolarizations (see Methods, Supplementary Figs 4, 5c and Supplementary Table 3 for information on K_d measurements).

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

Received 22 August; accepted 30 September 2008.

1. Catterall, W. A. From ionic currents to molecular mechanisms: The structure and function of voltage-gated sodium channels. *Neuron* **26,** 13–25 (2000).
2. Cannon, S. C. Pathomechanisms in channelopathies of skeletal muscle and brain. *Annu. Rev. Neurosci.* **29,** 387–415 (2006).
3. George, A. L. Jr. Inherited disorders of voltage-gated sodium channels. *J. Clin. Invest.* **115,** 1990–1999 (2005).
4. Cox, J. J. *et al.* An SCN9A channelopathy causes congenital inability to experience pain. *Nature* **444,** 894–898 (2006).
5. Fertleman, C. R. *et al.* SCN9A mutations in paroxysmal extreme pain disorder: Allelic variants underlie distinct channel defects and phenotypes. *Neuron* **52,** 767–774 (2006).
6. Kaczorowski, G. J., McManus, O. B., Priest, B. T. & Garcia, M. L. Ion channels as drug targets: The next GPCRs. *J. Gen. Physiol.* **131,** 399–405 (2008).
7. Horn, R., Ding, S. & Gruber, H. J. Immobilizing the moving parts of voltage-gated ion channels. *J. Gen. Physiol.* **116,** 461–476 (2000).
8. Sheets, M. F., Kyle, J. W., Kallen, R. G. & Hanck, D. A. The Na channel voltage sensor associated with inactivation is localized to the external charged residues of domain IV, S4. *Biophys. J.* **77,** 747–757 (1999).
9. Yang, N. & Horn, R. Evidence for voltage-dependent S4 movement in sodium channels. *Neuron* **15,** 213–218 (1995).
10. Chanda, B. & Bezanilla, F. Tracking voltage-dependent conformational changes in skeletal muscle sodium channel during activation. *J. Gen. Physiol.* **120,** 629–645 (2002).
11. Catterall, W. A. *et al.* Voltage-gated ion channels and gating modifier toxins. *Toxicon* **49,** 124–141 (2007).
12. Alabi, A. A., Bahamonde, M. I., Jung, H. J., Kim, J. I. & Swartz, K. J. Portability of paddle motif function and pharmacology in voltage sensors. *Nature* **450,** 370–375 (2007).
13. Chakrapani, S., Cuello, L. G., Cortes, D. M. & Perozo, E. Structural dynamics of an isolated voltage-sensor domain in a lipid bilayer. *Structure* **16,** 398–409 (2008).
14. Jiang, Y. *et al.* X-ray structure of a voltage-dependent K^+ channel. *Nature* **423,** 33–41 (2003).
15. Jiang, Y., Ruta, V., Chen, J., Lee, A. & MacKinnon, R. The principle of gating charge movement in a voltage-dependent K^+ channel. *Nature* **423,** 42–48 (2003).
16. Ruta, V., Chen, J. & MacKinnon, R. Calibrated measurement of gating-charge arginine displacement in the KvAP voltage-dependent K^+ channel. *Cell* **123,** 463–475 (2005).
17. Swartz, K. J. & MacKinnon, R. Mapping the receptor site for hanatoxin, a gating modifier of voltage-dependent K^+ channels. *Neuron* **18,** 675–682 (1997).
18. Swartz, K. J. & MacKinnon, R. Hanatoxin modifies the gating of a voltage-dependent K^+ channel through multiple binding sites. *Neuron* **18,** 665–673 (1997).
19. Li-Smerin, Y. & Swartz, K. J. Gating modifier toxins reveal a conserved structural motif in voltage-gated Ca^{2+} and K^+ channels. *Proc. Natl Acad. Sci. USA* **95,** 8585–8589 (1998).
20. Li-Smerin, Y. & Swartz, K. J. Localization and molecular determinants of the Hanatoxin receptors on the voltage-sensing domains of a K(+) channel. *J. Gen. Physiol.* **115,** 673–684 (2000).
21. Lee, H. C., Wang, J. M. & Swartz, K. J. Interaction between extracellular Hanatoxin and the resting conformation of the voltage-sensor paddle in Kv channels. *Neuron* **40,** 527–536 (2003).
22. Lee, S. Y. & MacKinnon, R. A membrane-access mechanism of ion channel inhibition by voltage sensor toxins from spider venom. *Nature* **430,** 232–235 (2004).
23. Milescu, M. *et al.* Tarantula toxins interact with voltage sensors within lipid membranes. *J. Gen. Physiol.* **130,** 497–511 (2007).
24. Phillips, L. R. *et al.* Voltage-sensor activation with a tarantula toxin as cargo. *Nature* **436,** 857–860 (2005).
25. Bosmans, F. *et al.* Four novel tarantula toxins as selective modulators of voltage-gated sodium channel subtypes. *Mol. Pharmacol.* **69,** 419–429 (2006).
26. Middleton, R. E. *et al.* Two tarantula peptides inhibit activation of multiple sodium channels. *Biochemistry* **41,** 14734–14747 (2002).
27. Smith, J. J., Cummins, T. R., Alphy, S. & Blumenthal, K. M. Molecular interactions of the gating modifier toxin ProTx-II with NaV 1.5: Implied existence of a novel toxin binding site coupled to activation. *J. Biol. Chem.* **282,** 12687–12697 (2007).
28. Sokolov, S., Kraus, R. L., Scheuer, T. & Catterall, W. A. Inhibition of sodium channel gating by trapping the domain II voltage sensor with protoxin II. *Mol. Pharmacol.* **73,** 1020–1028 (2008).
29. Cahalan, M. D. Modification of sodium channel gating in frog myelinated nerve fibres by *Centruroides sculpturatus* scorpion venom. *J. Physiol. (Lond.)* **244,** 511–534 (1975).
30. Koppenhofer, E. & Schmidt, H. Effect of scorpion venom on ionic currents of the node of Ranvier. II. Incomplete sodium inactivation. *Pflugers Arch.* **303,** 150–161 (1968).
31. Rogers, J. C., Qu, Y., Tanada, T. N., Scheuer, T. & Catterall, W. A. Molecular determinants of high affinity binding of alpha-scorpion toxin and sea anemone toxin in the S3-S4 extracellular loop in domain IV of the Na^+ channel alpha subunit. *J. Biol. Chem.* **271,** 15950–15962 (1996).
32. Cestele, S. *et al.* Voltage sensor-trapping: Enhanced activation of sodium channels by beta-scorpion toxin bound to the S3-S4 loop in domain II. *Neuron* **21,** 919–931 (1998).
33. Cestele, S. *et al.* Structure and function of the voltage sensor of sodium channels probed by a beta-scorpion toxin. *J. Biol. Chem.* **281,** 21332–21344 (2006).
34. Cohen, L. *et al.* Direct evidence that receptor site-4 of sodium channel gating modifiers is not dipped in the phospholipid bilayer of neuronal membranes. *J. Biol. Chem.* **281,** 20673–20679 (2006).
35. Tejedor, F. J. & Catterall, W. A. Site of covalent attachment of alpha-scorpion toxin derivatives in domain I of the sodium channel alpha subunit. *Proc. Natl Acad. Sci. USA* **85,** 8742–8746 (1988).
36. Marcotte, P., Chen, L. Q., Kallen, R. G. & Chahine, M. Effects of *Tityus serrulatus* scorpion toxin gamma on voltage-gated Na^+ channels. *Circ. Res.* **80,** 363–369 (1997).
37. Campos, F. V., Chanda, B., Beirao, P. S. & Bezanilla, F. Beta-scorpion toxin modifies gating transitions in all four voltage sensors of the sodium channel. *J. Gen. Physiol.* **130,** 257–268 (2007).
38. Campos, F. V., Chanda, B., Beirao, P. S. & Bezanilla, F. Alpha-scorpion toxin impairs a conformational change that leads to fast inactivation of muscle sodium channels. *J. Gen. Physiol.* **132,** 251–263 (2008).
39. Cha, A., Ruben, P. C., George, A. L. Jr, Fujimoto, E. & Bezanilla, F. Voltage sensors in domains III and IV, but not I and II, are immobilized by Na^+ channel fast inactivation. *Neuron* **22,** 73–87 (1999).
40. Sheets, M. F., Kyle, J. W. & Hanck, D. A. The role of the putative inactivation lid in sodium channel gating current immobilization. *J. Gen. Physiol.* **115,** 609–620 (2000).
41. Banerjee, A. & MacKinnon, R. Inferred motions of the S3a helix during voltage-dependent K^+ channel gating. *J. Mol. Biol.* **381,** 569–580 (2008).
42. Ruta, V., Jiang, Y., Lee, A., Chen, J. & MacKinnon, R. Functional analysis of an archaebacterial voltage-dependent K^+ channel. *Nature* **422,** 180–185 (2003).
43. Herrington, J. *et al.* Blockers of the delayed-rectifier potassium current in pancreatic beta-cells enhance glucose-dependent insulin secretion. *Diabetes* **55,** 1034–1042 (2006).
44. Lee, C. W. *et al.* Solution structure and functional characterization of SGTx1, a modifier of Kv2.1 channel gating. *Biochemistry* **43,** 890–897 (2004).
45. Armstrong, C. M. Na channel inactivation from open and closed states. *Proc. Natl Acad. Sci. USA* **103,** 17991–17996 (2006).
46. Chanda, B., Asamoah, O. K. & Bezanilla, F. Coupling interactions between voltage sensors of the sodium channel as revealed by site-specific measurements. *J. Gen. Physiol.* **123,** 217–230 (2004).

47. Wan, X., Chen, S., Sadeghpour, A., Wang, Q. & Kirsch, G. E. Accelerated inactivation in a mutant Na(+) channel associated with idiopathic ventricular fibrillation. *Am. J. Physiol. Heart Circ. Physiol.* **280,** H354–H360 (2001).
48. Bennett, P. B., Yazawa, K., Makita, N. & George, A. L. Jr. Molecular mechanism for an inherited cardiac arrhythmia. *Nature* **376,** 683–685 (1995).
49. Spampanato, J., Escayg, A., Meisler, M. H. & Goldin, A. L. Generalized epilepsy with febrile seizures plus type 2 mutation W1204R alters voltage-dependent gating of Na(v)1.1 sodium channels. *Neuroscience* **116,** 37–48 (2003).
50. Bendahhou, S., Cummins, T. R., Tawil, R., Waxman, S. G. & Ptacek, L. J. Activation and inactivation of the voltage-gated sodium channel: Role of segment S5 revealed by a novel hyperkalaemic periodic paralysis mutation. *J. Neurosci.* **19,** 4762–4771 (1999).

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

Acknowledgements We thank J. W. Kyle, D. A. Hanck and A. L. Goldin for the $rNa_v1.2a$, $rNa_v1.4$ and β_1 clones, C. Deutsch for $K_v1.3$, M. M. Smith for GxTx-1E, K. M. Blumenthal and J. B. Herrington for ProTx-II, L. D. Possani for a sample of TsVII, the NINDS DNA sequencing facility for DNA sequencing, and the NINDS protein sequencing facility for mass spectrometry and peptide sequencing. We thank A. A. Alabi for helping with K_v and Na_v channel alignments and T.-H. Chang for assistance with Na_v channel mutants. We also thank A. A. Alabi, M. Holmgren, M. Mayer, M. Milescu, J. Mindell, A. Plested, S. Silberberg and members of the Swartz laboratory for discussions. This work was supported by the Intramural Research Program of the NINDS, NIH (K.J.S.) and by an NIH-FWO postdoctoral fellowship (F.B.).

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.J.S. (swartzk@ninds.nih.gov).

METHODS

Channel and chimaera constructs. Chimaeras and point mutations were generated using sequential polymerase chain reactions (PCR) with $K_v2.1\Delta7$ (refs 19, 51), $K_v1.3$ (ref. 52), Shaker (ref. 53), $rNa_v1.2a$ (neuronal type)[54] or $rNa_v1.4$ (muscle type)[55] as templates. Channel nomenclature consists of a numerical system in which the number following the subscript indicates the gene subfamily, followed by a number indicating the specific channel isoform. Splice variants of each family member are identified by lowercase letters following the numbers. For Nav channels the prefix 'r' indicates the organism (rat) from which the channel was cloned. The $K_v2.1\Delta7$ construct contains seven point mutations in the outer vestibule[19], rendering the channel sensitive to agitoxin-2, a pore-blocking toxin from scorpion venom[56]. The Shaker construct contains a deletion of residues 6–46 to remove inactivation[57]. The DNA sequence of all constructs and mutants was confirmed by automated DNA sequencing, and cRNA was synthesized using T7/SP6 polymerase (Message Machine kit, Ambion) after linearizing the DNA with appropriate restriction enzymes.

Spider and scorpion toxin purification. PaurTx3 (ref. 25) was purified from the venom of *Phrixotrichus auratus* (Spider Pharm) using a new two-step HPLC protocol (Supplementary Fig. 5a). Identity and purity were determined with mass spectrometry and automated peptide sequencing. Hanatoxin was purified from *Grammostola spatulata* venom (Spider Pharm) as described previously[58]. SGTx1 and VSTx1 were synthesized using solid-phase chemical methods, folded *in vitro* and purified as described previously[44,59]. AaHII from *Androctonus australis hector* venom (animals collected near Tozeur, Tunisia), and TsVII from *Tityus serrulatus* venom (gift from C. Diniz) were purified as described previously[60,61]. Synthetic ProTx-I was acquired from Peptides International. ProTx-II was provided by K. M. Blumenthal (SUNY) and J. B. Herrington (Merck Research Labs, Rahway). GxTx-1E was provided by M. M. Smith (Merck Research Labs, Rahway). The plant alkaloid veratridine (VTD; Sigma) is used as a negative control in Fig. 2 as the binding site of this lipid-soluble compound is thought to consist of the S6 transmembrane segments[62]. Toxins were kept at −20 °C. Before experiments, toxin aliquots were dissolved in appropriate solutions containing 0.1% BSA (or 1% BSA in the case of TsVII).

Two-electrode voltage-clamp recording from *Xenopus* oocytes. Channel constructs were expressed in *Xenopus* oocytes[21] and studied following 1–2 days incubation after cRNA injection (incubated at 17 °C in 96 mM NaCl, 2 mM KCl, 5 mM HEPES, 1 mM $MgCl_2$ and 1.8 mM $CaCl_2$, 50 μg ml^{-1} gentamycin, pH 7.6 with NaOH) using two-electrode voltage-clamp recording techniques (OC-725C, Warner Instruments) with a 150 μl recording chamber. Data were filtered at 1 kHz and digitized at 10 kHz using pClamp software (Axon). Microelectrode resistances were 0.1–1 MΩ when filled with 3 M KCl. For most K_v channel experiments, the external recording solution contained 50 mM KCl, 50 mM NaCl, 5 mM HEPES, 1 mM $MgCl_2$ and 0.3 mM $CaCl_2$, pH 7.6 with NaOH. KCl was replaced by RbCl for Shaker experiments. For Na_v channel experiments, the external recording solution contained 96 mM NaCl, 2 mM KCl, 5 mM HEPES, 1 mM $MgCl_2$ and 1.8 mM $CaCl_2$, pH 7.6 with NaOH. For Na_v channel experiments, oocytes were co-injected with the β_1 subunit in a 1:5 molar ratio. All experiments were performed at room temperature (~22 °C). Leak and background conductances, identified by blocking the channel with agitoxin-2, have been subtracted for all of the K_v channel currents shown[21]. Tetrodotoxin subtraction (TTX) or an online P/-4 prepulse protocol was used to subtract linear leak and membrane capacitive currents to isolate Na_v channel currents.

Analysis of channel activity and toxin–channel interactions. Voltage–activation relationships were obtained by measuring tail currents for K_v channels or steady-state currents and calculating conductance for Na_v channels, and a single Boltzmann function was fitted to the data according to $I/I_{max} = [1 + \exp(-zF(V - V_{1/2})/RT)]^{-1}$, where I/I_{max} is the normalized tail-current amplitude, z is the equivalent charge, $V_{1/2}$ is the half-activation voltage, F is Faraday's constant, R is the gas constant and T is temperature in kelvin.

Occupancy of closed or resting channels by toxins was examined using negative holding voltages where open probability was low, and the fraction of unbound channels (F_u) was estimated using depolarizations that are too weak to open toxin-bound channels, as described previously[17–21,24,44]. After addition of the toxin to the recording chamber, the equilibration between the toxin and the channel was monitored using weak depolarizations elicited at 5–10 s intervals. For all channels, we recorded voltage–activation relationships in the absence and presence of different concentrations of toxin. The ratio of currents (I/I_0) recorded in the presence (I) and absence (I_0) of toxin was calculated for various strength depolarizations, typically –70 mV to +10 mV. The value of I/I_0 measured in the plateau phase at voltages where toxin-bound channels do not open was taken as F_u (Supplementary Fig. 5b, c). The apparent equilibrium dissociation constant (K_d) for K_v channels was calculated according to $K_d = ((1/(1 - Fu^{1/4})) - 1)[\text{toxin}]$ assuming four independent toxin-binding sites per channel, with single occupancy being sufficient to inhibit opening in response to weak depolarizations. For all chimaeras and mutants, voltage protocols were adjusted appropriately so that the plateau phase in the I/I_0–voltage relationship was well defined. Example traces showing the inhibitory activity of tarantula toxins were taken for relatively weak depolarizations within the plateau phase for that particular channel construct. Off-line data analysis was performed using Clampfit (Axon), Origin 7.5 (Originlab) and Microsoft Solver (Microsoft Excel).

51. Frech, G. C., VanDongen, A. M., Schuster, G., Brown, A. M. & Joho, R. H. A novel potassium channel with delayed rectifier properties isolated from rat brain by expression cloning. *Nature* **340**, 642–645 (1989).
52. Stuhmer, W. *et al.* Molecular basis of functional diversity of voltage-gated potassium channels in mammalian brain. *EMBO J.* **8**, 3235–3244 (1989).
53. Tempel, B. L., Papazian, D. M., Schwarz, T. L., Jan, Y. N. & Jan, L. Y. Sequence of a probable potassium channel component encoded at Shaker locus of Drosophila. *Science* **237**, 770–775 (1987).
54. Auld, V. J. *et al.* A rat brain Na^+ channel alpha subunit with novel gating properties. *Neuron* **1**, 449–461 (1988).
55. Trimmer, J. S. *et al.* Primary structure and functional expression of a mammalian skeletal muscle sodium channel. *Neuron* **3**, 33–49 (1989).
56. Garcia, M. L., Garcia-Calvo, M., Hidalgo, P., Lee, A. & MacKinnon, R. Purification and characterization of three inhibitors of voltage-dependent K^+ channels from *Leiurus quinquestriatus* var. *hebraeus* venom. *Biochemistry* **33**, 6834–6839 (1994).
57. Hoshi, T., Zagotta, W. N. & Aldrich, R. W. Biophysical and molecular mechanisms of Shaker potassium channel inactivation. *Science* **250**, 533–538 (1990).
58. Swartz, K. J. & MacKinnon, R. An inhibitor of the Kv2.1 potassium channel isolated from the venom of a Chilean tarantula. *Neuron* **15**, 941–949 (1995).
59. Jung, H. J. *et al.* Solution structure and lipid membrane partitioning of VSTx1, an inhibitor of the KvAP potassium channel. *Biochemistry* **44**, 6015–6023 (2005).
60. Ceard, B., De Lima, M. E., Bougis, P. E. & Martin-Eauclaire, M. F. Purification of the main beta-toxin from *Tityus serrulatus* scorpion venom using high-performance liquid chromatography. *Toxicon* **30**, 105–110 (1992).
61. Martin, M. F., Rochat, H., Marchot, P. & Bougis, P. E. Use of high performance liquid chromatography to demonstrate quantitative variation in components of venom from the scorpion *Androctonus australis hector*. *Toxicon* **25**, 569–573 (1987).
62. Wang, G. K., Quan, C., Seaver, M. & Wang, S. Y. Modification of wild-type and batrachotoxin-resistant muscle mu1 Na^+ channels by veratridine. *Pflugers Arch.* **439**, 705–713 (2000).

ARTICLES

Structure of the guide-strand-containing argonaute silencing complex

Yanli Wang[1], Gang Sheng[1], Stefan Juranek[2], Thomas Tuschl[2] & Dinshaw J. Patel[1]

The slicer activity of the RNA-induced silencing complex is associated with argonaute, the RNase H-like PIWI domain of which catalyses guide-strand-mediated sequence-specific cleavage of target messenger RNA. Here we report on the crystal structure of *Thermus thermophilus* argonaute bound to a 5′-phosphorylated 21-base DNA guide strand, thereby identifying the nucleic-acid-binding channel positioned between the PAZ- and PIWI-containing lobes, as well as the pivot-like conformational changes associated with complex formation. The bound guide strand is anchored at both of its ends, with the solvent-exposed Watson–Crick edges of stacked bases 2 to 6 positioned for nucleation with the mRNA target, whereas two critically positioned arginines lock bases 10 and 11 at the cleavage site into an unanticipated orthogonal alignment. Biochemical studies indicate that key amino acid residues at the active site and those lining the 5′-phosphate-binding pocket made up of the Mid domain are critical for cleavage activity, whereas alterations of residues lining the 2-nucleotide 3′-end-binding pocket made up of the PAZ domain show little effect.

RNA interference is routinely used in biological research to effectively silence genes of interest[1–3] and is being increasingly investigated as a therapeutic modality against diverse human diseases[4,5]. The argonaute (Ago) protein[6–11], as a key catalytic component of the RNA-induced silencing complex, has a central role in the RNA interference pathway by mediating the maturation of small interfering RNA (siRNA) through initial degradation of the passenger strand, followed by guide-strand-mediated sequence-specific cleavage of target mRNA.

X-ray crystal structures of archaeal *Pyrococcus furiosus* Ago in its free state[12,13] and eubacterial *Aquifex aeolicus* Ago in its free state[14] and when externally bound to siRNA[15] have defined the folds and relative positioning of protein domains and connecting linker segments (see Supplementary Fig. 1 for sequence alignments of prokaryotic and eukaryotic Agos). A mechanistic understanding of guide-strand-mediated mRNA recognition and cleavage chemistry mediated by Ago requires additional structural information on the binary complex of Ago with bound guide strand substrate and, in the longer term, on the ternary complex with added mRNA strand.

Given that binding and cleavage studies have identified eubacterial *A. aeolicus* Ago to be a site-specific DNA-guided endoRNase[14], we attempted to crystallize the binary complex of an eubacterial *T. thermophilus* Ago with bound 5′-phosphorylated DNA guide strands ranging in length from 8–12 to 14, 16, 18–24, 26 and 28 bases. Diffraction quality crystals were obtained solely for binary complexes with 21-base and 10-base DNA guide strands. No crystals were obtained for the complex with the corresponding 5′-phosphorylated 21-base RNA guide strand.

Structure of the 21-base binary complex

The 3.0 Å crystal structure of *T. thermophilus* Ago complexed with a 5′-phosphorylated (5′-phos-$T_1GAGG_5TAGTA_{10}GGTTG_{15}TATAG_{20}T$) 21-base DNA guide strand is shown in Fig. 1a, with crystallographic statistics listed in Supplementary Table 1. The protein in a ribbon representation is colour-coded by domains (N, PAZ, Mid and PIWI) and linkers (L1 and L2). The bound DNA is shown in red, with the chain traceable for nucleotides 1–11 and 18–21, including two phosphates within the remaining disordered segment. The corresponding view to that in Fig. 1a, with the protein in a surface representation, colour-coded according to its electrostatic potential, is shown in Fig. 1b. The bound DNA guide strand threads its way within a central basic channel between the PAZ-containing (N, L1 and PAZ) and PIWI-containing (Mid and PIWI) lobes of the bilobal scaffold of Ago, thereby contacting all domains and linkers and defining the nucleic-acid-binding channel within the protein (Fig. 1a, b).

Base stacking within the seed segment

An alternate view of the complex, with the protein in a ribbon representation, is shown in stereo in Fig. 1c. A rotating view of Fig. 1c can be accessed as Movie 3DLH in the Supplementary Information. Within its traceable 1 to 11 segment, the bound DNA is continuously stacked from residues 2 to 10, with distinct breaks at the 1–2 and 10–11 steps (Fig. 1c, Supplementary Fig. 2 and $2F_o - F_c$ density map of bound DNA in Supplementary Fig. 3a). The observed integrity of stacked residues in the 5′-half (residues 2 to 10) but not the 3′-half of the 21-base DNA in the complex (Supplementary Fig. 2) is consistent with conclusions of functional experiments that established a toleration for mismatches within the 3′-segment but not the 5′-segment of the guide strand in its complex with mRNA[16,17].

The corresponding expanded view of the complex with the protein in an electrostatic surface representation is shown in Supplementary Fig. 4. The Watson–Crick edges of bases 2 to 6 are exposed, identifying this segment of the guide strand as the nucleation site for pairing with mRNA. Our structural results are consistent with previous functional studies that identified a 'seed' segment (residues 2 to 8) within the guide strand as critical for mRNA pairing[18,19]. In contrast, bases 7 to 11 are threaded inside the Ago scaffold in the binary complex (Supplementary Fig. 4).

Anchoring both ends of bound guide strand

The 21-base DNA is anchored at both its 5′ and 3′ ends in the complex, thereby constraining the trajectory of the bound guide

[1]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA. [2]Howard Hughes Medical Institute, Laboratory of RNA Molecular Biology, The Rockefeller University, New York, New York 10065, USA.

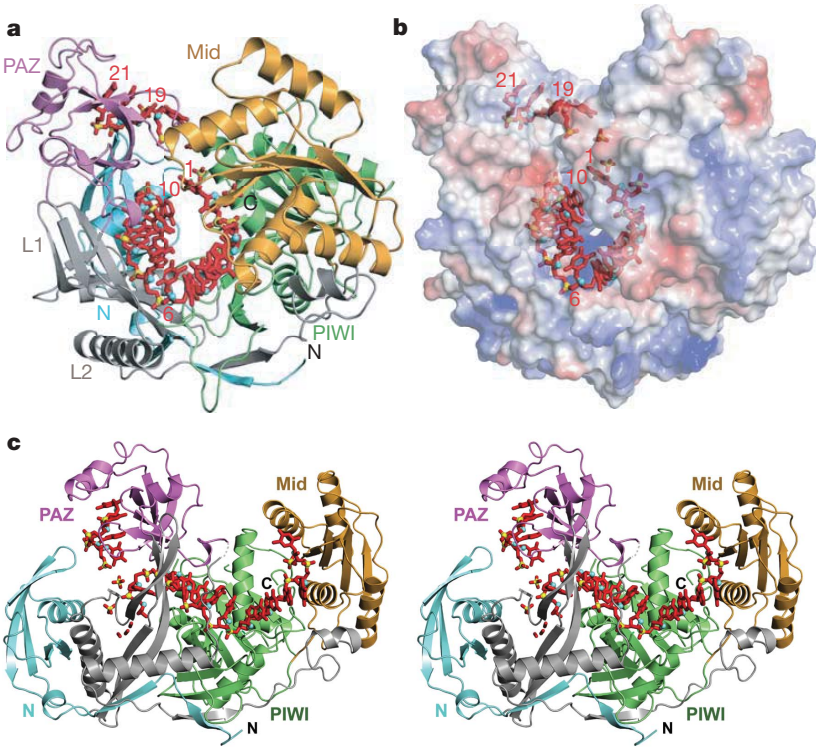


Figure 1 | Crystal structure of *T. thermophilus* argonaute bound to a 5′-phosphorylated 21-base DNA guide strand. **a**, Stick (DNA) and ribbon (Ago) view of the 3.0 Å structure of the complex. Individual domains and linkers are colour-coded. The bound 21-base DNA guide strand is coloured red (with phosphorus atoms in yellow), and can be traced for nucleotides 1 to 11 and 18 to 21, together with phosphates at positions 16 and 17. **b**, Corresponding stick (DNA) and electrostatically colour-coded surface (Ago) view of the complex. **c**, An alternate stereo view of the complex from that shown in **a**. For rotation of **c**, see Supplementary Movie 2DLH.

strand. The 5′-phosphate is anchored within the binding pocket in the Mid domain, with its oxygens hydrogen-bonded to the side chains of highly conserved residues (R418, K422, S432, Q433, K457) and a bound magnesium ion (Fig. 2a) as previously observed in the *A. fulgidus* Piwi protein–siRNA complex[20,21]. The magnesium is coordinated to the first and third phosphates from the 5′ end, as well as the carboxy-terminal carboxylate end (V685) of the PIWI domain (Fig. 2a). This result reinforces earlier functional conclusions that the integrity of the 5′-phosphate-binding pocket within the Mid domain is critical for slicing fidelity[21].

The bases are splayed apart at the 1–2 step, with base 1 stacked over the side chain of an arginine (R418, Fig. 2a). This distortion at the 5′ end of the guide strand reinforces an earlier structural observation in the *A. fulgidus* Piwi–siRNA complex[20,21], as well as functional experiments[16–18], which indicated that non-canonical pairing with mRNA is preferred at the 1 position in the guide strand.

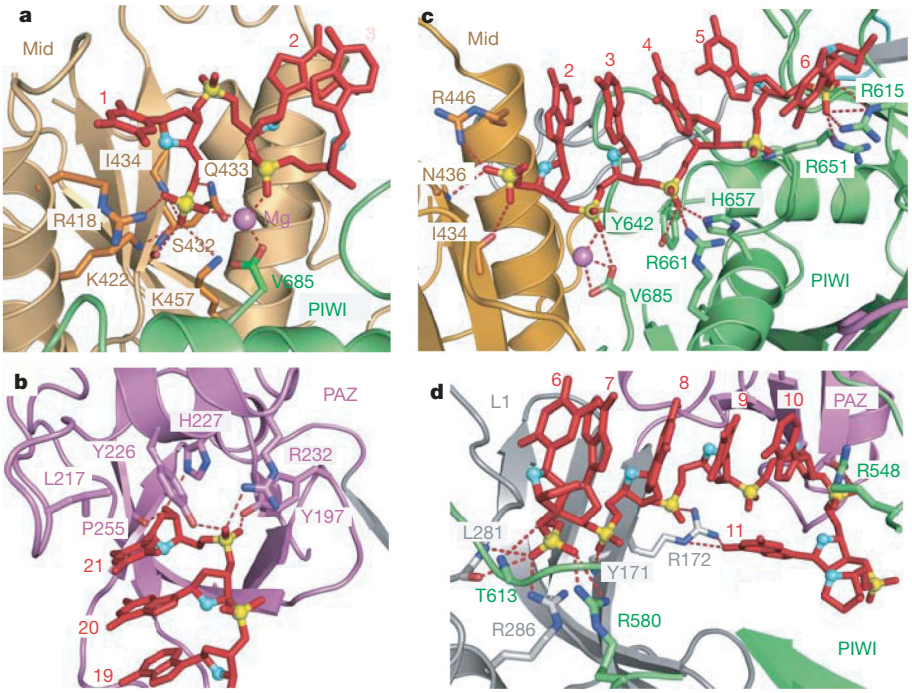


Figure 2 | Intermolecular hydrogen-bonding alignments in the *T. thermophilus* Ago bound to a 21-base DNA guide strand. **a**, Insertion of the 5′-phosphate of the DNA guide strand into the binding pocket in the Mid domain. **b**, Positioning of stacked 3′-end residues 20 and 21 of the DNA guide strand into the binding pocket in the PAZ domain. **c**, Positioning of stacked residues 2 to 6 of the DNA guide strand, with emphasis on intermolecular interactions. **d**, Positioning of stacked residues 6 to 10 and 11 of the DNA guide strand, with emphasis on intermolecular interactions.

Residues 20 and 21 at the 3′ end of the complex are anchored within the binding pocket in the PAZ domain (Fig. 2b and $2F_o - F_c$ density map of bound DNA in Supplementary Fig. 3b), with the oxygens of the phosphate linking residues 20 and 21 and the 3′-OH of residue 21 hydrogen-bonded to aromatic (Y197, Y226, H227) and basic (R232) side chains, as previously observed in PAZ–siRNA[22] and PAZ–single-stranded RNA[23] complexes.

The positioning of the RNase H-type catalytic residues D546, D478 (on adjacent β-strands) and D660 of the PIWI domain of *T. thermophilus* Ago relative to segment 1–11 of the bound 21-base DNA guide strand is outlined in Supplementary Fig. 2.

There is extensive hydrogen bonding and salt bridge formation between the backbone phosphates of the DNA guide strand and basic Arg side chains spanning the various domains in the protein (Fig. 2c, d). These extensive intermolecular contacts are consistent with related interactions located towards the 5′ end of the guide strand in the earlier *A. fulgidus* Piwi–siRNA structure[20,21] and proposed models of the catalytic cycle[2,7,14], where the anchored guide strand is retained while the cleaved message dissociates during each cycle of the multiple turnover catalytic process[10,11].

The 10–11 cleavage site step

The guanidinium group of R548 (conserved amongst the thermophilic Agos, whereas Val occupies this position in human AGOs) stacks on the base at position 10, with the resulting block enforcing an unanticipated orthogonal alignment of bases 10 and 11 (Fig. 2d). In addition, the guanidinium group of R172 (also observed in *A. aeolicus* Ago, but otherwise not conserved) has a bridging role through hydrogen bonding both to the phosphate group of residue 9 and to the base edge of residue 11 (Fig. 2d and $2F_o - F_c$ density map in Supplementary Fig. 3c). We anticipate that Arg residues R172 and R548, which enforce the pronounced kink at the 10–11 step in the Ago–guide-strand binary complex (Fig. 2d), will probably undergo a conformational switch on mRNA pairing during the propagation step of ternary complex formation, so as to generate a catalytically competent undistorted helical conformation at the 10–11 step associated with the mRNA cleavage site[17,24,25].

The helical trajectory of the stacked 2 to 10 segment of the bound 21-base DNA guide strand in the binary Ago complex (in red, Supplementary Fig. 5a, b) deviates from the helical trajectory of both A-form (in cyan, Supplementary Fig. 5a) and B-form (in green, Supplementary Fig. 5b) helices, following best-fit superposition of residues 2 to 6. This probably implies that either anchoring of both ends (Figs 1c and 2a, b) and/or intermolecular interactions (Fig. 2c, d) perturb the helical trajectory of the bound DNA guide strand in the complex. Nevertheless, it is significant that the 5′ end of the bound DNA is ordered in the absence of pairing with message.

Conformational transitions

We have been unable to grow crystals of *T. thermophilus* Ago in the free state, but have succeeded in solving the 2.7 Å structure with a bound 5′-phosphorylated 10-base DNA guide strand (5′-phos-$T_1GAGG_5TAGTA_{10}$, Fig. 3a), with crystallographic statistics listed in Supplementary Table 1. We can trace 3′-end residues 6 to 10 in the bound 10-base DNA sequence (Supplementary Fig. 6), with residues 9 and 10 inserted into the PAZ pocket (Fig. 3a). The C terminus of the *T. thermophilus* Ago protein is disordered in the 10-base DNA guide strand complex, and hence the 5′-phosphate-binding pocket in the Mid domain must be disrupted in the absence of an inserted C-terminal carboxylate group. The relative positioning of the 3′ ends of the bound 21-base (residues 18 to 21, in magenta) and 10-base (residues 6 to 10, in green) DNAs in their respective complexes after superposition of their PAZ domains are shown in Fig. 3b. The two 3′-end nucleotides that are inserted within their PAZ pockets superposition quite well, but the remaining nucleotides (18 and 19 in the 21-base and 6 to 8 in the 10-base complexes) follow different trajectories (Fig. 3b). This reflects differences in the relative orientations of the N and PAZ domains between the two complexes (Supplementary Fig. 7a), as well as intermolecular contacts (Supplementary Fig. 7b, c), suggestive of at least two nucleic-acid-binding channels leading from the PAZ pocket.

The relative domain alignments of Ago in complex with bound 10-base and 21-base DNA guide strands, after superposition of their PAZ domains (boxed region), is shown in Fig. 3c and in stereo in Supplementary Fig. 8. There are significant conformational changes

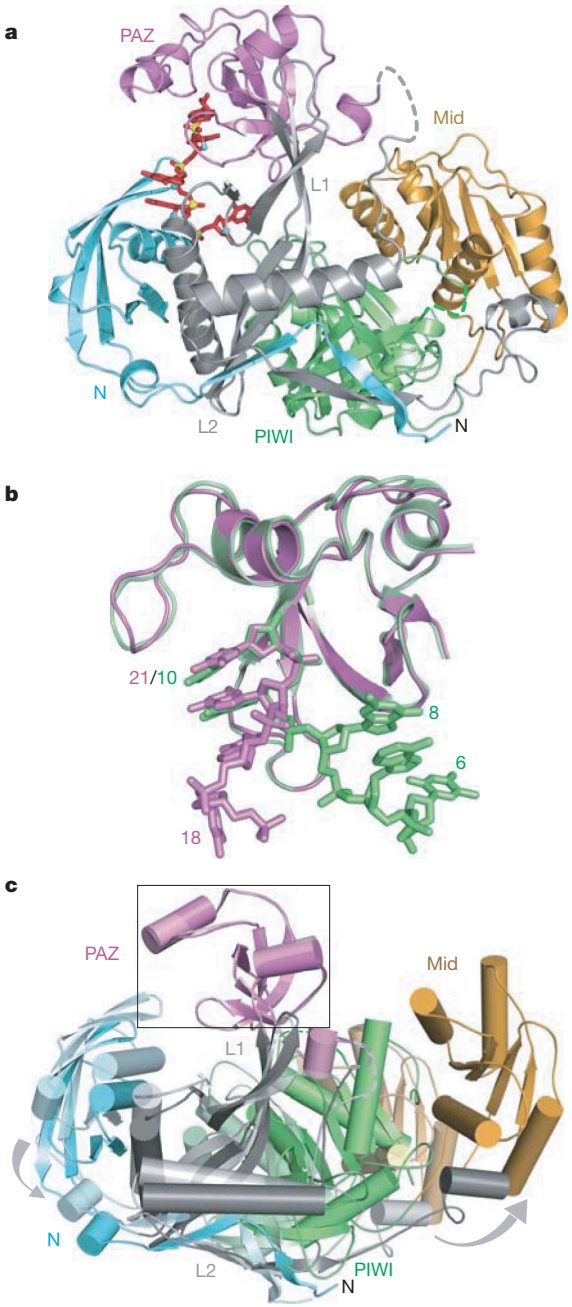


Figure 3 | Conformational changes in *T. thermophilus* Ago on formation of the 21-base DNA guide strand complex. **a**, The 2.7 Å structure of *T. thermophilus* Ago bound to a 5′-phosphorylated 10-base DNA guide strand. The bound 10-base DNA can be traced for nucleotides 6 to 10. **b**, Superpositioning of the PAZ domains together with observable 3′ ends in DNA 21-base (in magenta) and 10-base (in green) guide strand complexes. **c**, Alignment of Agos in complexes with bound 10-base (in lighter shades) and 21-base (in darker shades) DNA guide strands, after superpositioning of their PAZ domains (boxed segment). The grey arrows indicate the magnitude of the conformational changes on proceeding from the 10-base to the 21-base complexes. For transition between **a** and Fig. 1c, see Supplementary Movie 3DLB–3DLH.

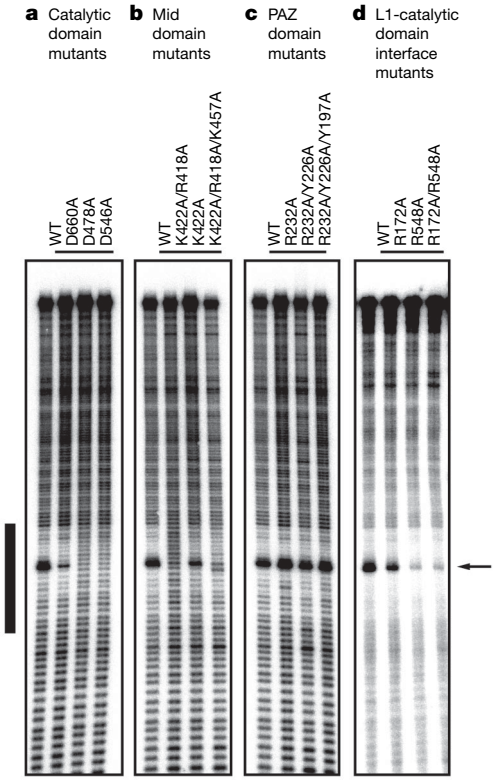


Figure 4 | DNA-guide-dependent RNA cleavage activity of wild-type and mutant *T. thermophilus* Ago. A 5′-phosphorylated 21-nucleotide DNA oligonucleotide sequence was pre-incubated with the indicated Ago proteins followed by the addition of 5′-radiolabelled RNA cleavage substrate. Cleavage products were resolved using denaturing polyacrylamide gel electrophoresis gels and visualized by phosphoimaging. The black bar indicates the region of the cleavage substrate covered by the guide DNA, with the cleavage site shown by an arrow. **a**, Catalytic Asp mutants (left panel); **b**, Mid domain mutants (central left panel); **c**, PAZ domain mutants (central right panel); and **d**, the L1-catalytic domain interface mutants (right panel).

on proceeding from the 10-base complex with an anchored 3′ end (lighter colours, Fig. 3c) to the 21-base complex with both 5′- and 3′-anchored ends (darker colours in Fig. 3c), with grey arrows highlighting the magnitude of the transitions. Specifically, the Mid domain rotates by 22° (Supplementary Fig. 9a) and the PAZ-containing lobe rotates by 25° (Supplementary Fig. 9b), both with respect to the PIWI domain. These movements can be tracked in Movie 3DLB–3DLH provided in the Supplementary Information. The net result of rotating the Mid domain relative to the PIWI domain is to increase the length of the nucleic-acid-binding channel towards the 5′-binding site by 8 Å, thereby accommodating the full length of the bound 21-base DNA guide strand. By contrast, the conformational change in the PAZ-containing lobe relative to the PIWI domain reflects a closing movement that results in the generation of a narrower nucleic acid-binding channel necessary to hold and orient the bound 21-base DNA guide strand.

Cleavage assays

To characterize the catalytic activity of the *T. thermophilus* Ago protein, we performed target RNA cleavage assays (Fig. 4). Similar to the observations made previously for *A. aeolicus* Ago protein[14], the cleavage activity was enhanced by addition of divalent manganese ions and the cleavage position was located 10 nucleotides downstream from the residue paired with the 5′ end of the DNA guide strand (see also Supplementary Fig. 10). Mutation of any one of the Asp residues of the catalytic triad abolished or substantially reduced target RNA cleavage (Fig. 4a). Alteration of K422 in the 5′-phosphate-binding Mid domain pocket also reduced activity, and double or triple mutations including R418 and K457 completely abolished activity (Fig. 4b). In contrast, alterations of up to three conserved residues of the 2-nucleotide 3′-end-binding PAZ domain pocket showed no reduction of cleavage activity (Fig. 4c). We also evaluated the impact on cleavage after mutation of the two arginines whose interactions promote the orthogonal alignment of bases 10 and 11 in the complex. The R172A mutant (linker L1) had a small, but noticeable, impact on cleavage, while significantly reduced cleavage was observed for the R548A mutant (PIWI domain), as well as for the R172A/R548A dual mutant (Fig. 4d).

Our long-term goal is to capture Ago–nucleic acid complexes during distinct assembly and functional steps of the RNA interference catalytic cycle[26,27]. The earlier structures of the *A. fulgidus* Piwi protein[28] bound to siRNA duplexes[20,21] together with functional implications[10,11] and our current structure of the *T. thermophilus* Ago bound to a 21-base DNA guide strand represent initial but critical contributions to the success of this endeavour. We expect that the structural insights thus derived will be portable to eukaryotic AGOs that function as RNA-guide-strand-mediated endoRNases, thereby providing the molecular underpinnings defining guide-strand-mediated recognition and processing of mRNA[29,30], with eventual impact on design and delivery strategies for silencing human disease using RNA interference[4,5].

METHODS SUMMARY

Wild-type and mutant *T. thermophilus* Agos were overexpressed from *Escherichia coli* and purified by column chromatography. The structure of Ago complexed with the 5′-phosphorylated 10-base DNA was determined by multi-wavelength anomalous diffraction on the selenomethionine (SeMet)-modified protein. The structure of Ago complexed with 5′-phosphorylated 21-base DNA was determined by molecular replacement using the domains of the Ago–10-base DNA complex structure as search models. Details of all biochemical and crystallographic procedures are listed in Methods.

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

Received 19 April; accepted 6 August 2008.
Published online 27 August 2008.

1. Baulcombe, D. RNA silencing in plants. *Nature* **431**, 356–363 (2004).
2. Filipowicz, W. The nuts and bolts of the RISC machine. *Cell* **122**, 17–20 (2005).
3. Rana, T. M. Illuminating the silence: understanding the structure and function of small RNAs. *Nature Rev. Mol. Cell Biol.* **8**, 23–36 (2007).
4. Kim, D. H. & Rossi, J. J. Strategies for silencing human disease using RNA interference. *Nature Rev. Genet.* **8**, 173–184 (2007).
5. De Fougerolles, A., Vornlocher, H.-P., Maraganore, J. & Lieberman, J. Interfering with disease: a progress report on siRNA-based therapeutics. *Nature Rev. Drug. Discov.* **6**, 443–453 (2007).
6. Hall, T. M. Structure and function of Argonaute proteins. *Structure* **13**, 1403–1408 (2005).
7. Tomari, Y. & Zamore, P. D. Perspective: machines for RNAi. *Genes Dev.* **19**, 517–529 (2005).
8. Tolia, N. H. & Joshua-Tor, L. Slicer and the argonautes. *Nature Chem. Biol.* **3**, 36–43 (2007).
9. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nature Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
10. Parker, J. S. & Barford, D. Argonaute: a scaffold for the function of short regulatory RNAs. *Trends Biochem. Sci.* **31**, 622–630 (2006).
11. Patel, D. J. *et al.* Structural biology of RNA silencing and its functional implications. *Cold Spring Harb. Symp. Quant. Biol.* **71**, 81–93 (2006).
12. Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. Crystal structure of Argonaute and its implications for RISC slicer activity. *Science* **305**, 1434–1437 (2004).
13. Rivas, F. V. *et al.* Purified Ago2 and an siRNA form recombinant human RISC. *Nature Struct. Mol. Biol.* **12**, 340–349 (2005).
14. Yuan, Y. R. *et al.* Crystal structure of *A. aeolicus* argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol. Cell* **19**, 405–419 (2005).
15. Yuan, Y. R. *et al.* A potential protein–RNA recognition event along the RISC loading pathway from the structure of *A. aeolicus* Ago with externally bound siRNA. *Structure* **14**, 1557–1565 (2006).
16. Doench, J. G. & Sharp, P. A. Specificity of miRNA target selection in translational repression. *Genes Dev.* **18**, 504–511 (2004).
17. Haley, B. & Zamore, P. D. Kinetic analysis of the RNAi enzyme complex. *Nature Struct. Mol. Biol.* **11**, 599–606 (2004).

18. Lewis, B. P. *et al.* Prediction of mammalian microRNA targets. *Cell* **115,** 787–798 (2003).
19. Stark, A., Brennecke, J., Russel, R. B. & Cohen, S. M. Identification of *Drosophila* microRNA targets. *PLoS Biol.* **1,** 397–409 (2003).
20. Parker, J. S., Roe, S. M. & Barford, D. Structural insights into mRNA recognition from a PIWI domain–siRNA guide complex. *Nature* **434,** 663–666 (2005).
21. Ma, J. B. *et al.* Structural basis for 5′-end-specific recognition of guide RNA by the *A. fulgidus* Piwi protein. *Nature* **434,** 666–670 (2005).
22. Ma, J., Ye, K. & Patel, D. J. Structural basis for overhang-specific small interfering RNA recognition by the Paz domain. *Nature* **429,** 318–322 (2004).
23. Lingel, A., Simon, B., Izaurralde, E. & Sattler, M. Nucleic acid 3′-end recognition by the Argonaute2 Paz domain. *Nature Struct. Mol. Biol.* **11,** 576–577 (2004).
24. Nowotny, M., Gaidamakov, S. A., Crouch, R. J. & Yang, W. Crystal structures of RNase H bound to an RNA/DNA hybrid: substrate specificity and metal-dependent catalysis. *Cell* **121,** 1005–1016 (2005).
25. Chiu, Y. L. & Rana, T. M. RNAi in human cells: Basic structural and functional features of small interfering RNA. *Mol. Cell* **10,** 549–561 (2002).
26. Elbashir, S. M., Lendeckel, W. & Tuschl, T. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* **15,** 188–200 (2001).
27. Martinez, J. & Tuschl, T. RISC is a 5′-phosphomonoester-producing RNA endonuclease. *Genes Dev.* **18,** 975–980 (2004).
28. Parker, J. S., Roe, S. & Barford, D. Crystal structure of a PIWI protein suggests mechanisms for siRNA recognition and slicer activity. *EMBO J.* **23,** 4727–4737 (2004).
29. Liu, J. *et al.* Argonaute2 is the catalytic engine of RNAi. *Science* **305,** 1437–1441 (2004).
30. Meister, G. *et al.* Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15,** 185–197 (2004).

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

Acknowledgements The research was supported by funds from the NIH to D.J.P. and T.T. We thank the staff of NE-CAT beam line at the Advanced Photon Source, Argonne National Laboratory, supported by the US Department of Energy, for assistance with data collection.

Author Contributions Y.W. and G.S. expressed and purified *T. thermophilus* Ago and its mutants, and grew crystals of the complex. Y.W. collected X-ray diffraction data and solved the structure of the complex. The structural studies were undertaken under the supervision of D.J.P. S.J. was responsible for the cleavage assays on the wild-type and mutant Agos under the supervision of T.T. All authors read and approved the submitted manuscript.

Author Information The structures of *T. thermophilus* Ago bound to 5′-phosphorylated 21-base and 10-base DNAs have been deposited to the Protein Data Bank under accession codes 3DLH and 3DLB, respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.J.P. (pateld@mskcc.org).

METHODS

Protein expression and purification. An expression vector (PET-SUMO, Invitrogen) was constructed to produce full-length *T. thermophilus* Ago. The *E. coli* cells were collected for 8 h after incubation with 0.1 mM IPTG at 20 °C. The cells were lysed in 20 mM Tris-HCl buffer, 1 M NaCl, 2 mM $MgCl_2$, pH 7.5, at 25 °C. The protein was first purified using His-select nickel agarose resin (GE Healthcare). The protein was digested by ubiquitin-like protein 1 (Ulp1) SUMO protease (Invitrogen) and dialysed against 20 mM Tris-HCl, 0.5 M NaCl, 2 mM $MgCl_2$, pH 7.5, for 16 h at room temperature. The protein was then heated at 55 °C for 15 min and next purified by His-select nickel agarose and gel filtration (Superdex 200, Amersham) chromatography. Finally, it was concentrated to 30 mg ml^{-1} in 20 mM Tris-HCl, 0.5 M NaCl and 2 mM $MgCl_2$.

The Ago mutants were made with the Quik-Change kit (Stratagene) and verified by sequencing. We expressed and purified the mutant proteins following the same protocol as that used for the wild-type protein.

Crystallization and data collection. DNA oligonucleotides were purchased from Invitrogen. All crystals were grown by the hanging-drop vapour diffusion method. Native Ago or selenomethionine (SeMet)-labelled Ago was mixed with 5′-phosphorylated 10-base DNA at 1:2 molar ratio. The SeMet–Ago–10-base DNA crystals were grown from 3–5% (v/v) polyethylene glycerol 550 monomethyl ether (PEGMME550), 50 mM Tris-HCl, pH 7.5, 0.1 M KCl and 10 mM $MgCl_2$ at 35 °C. Crystals were flash-frozen in cryoprotecting solutions containing 3% PEGMME550, 50 mM Tris-HCl, pH 7.5, 0.1 M KCl, 10 mM $MgCl_2$ and 25% (v/v) glycerol.

The Ago–21-base DNA complex was prepared for crystallization by mixing the purified Ago with 1:1 molar ratio of 5′-phosphorylated 21-base DNA. The crystals were grown from 18% (v/v) polyethylene glycerol (PEG)8000, 50 mM MES, pH 6.5, 200 mM KCl and 80 mM $MgAc_2$ at 40 °C. They were flash-frozen in cryoprotecting buffer containing 16% PEG8000, 50 mM MES, pH 6.5, 200 mM KCl, 80 mM $MgAc_2$ and 25% (v/v) glycerol.

Diffraction data were collected at 100 K at beam line NE-CAT 24-IDC, the Advanced Photon Source (APS), Argonne National Laboratory. All data were processed with the HKL2000 suite[31] and data processing statistics are summarized in Supplementary Table 1. The Ago–10-base DNA complex belonged to the $P2_1$ space group and diffracted to 2.7 Å. The Ago–21-base DNA complex belonged to the $P2_12_12_1$ space group and diffracted to 3.0 Å.

Structure determination and refinement. The structure of the Ago–10-base DNA complex was solved by the multi-wavelength anomalous diffraction method using the SeMet-labelled complex. Phases were calculated with programs SOLVE and RESOLVE[32–34]. The model was built with program COOT[35] and was improved by several cycles of manual rebuilding and refinement with CNS[36] and REFMAC[37]. There are two molecules in the asymmetric unit. For the Ago–10-base DNA complex, 43 residues at the C terminus could not be traced, nor could we trace DNA residues 1 to 5 in the complex.

The structure of the Ago–21-base DNA complex was solved by molecular replacement with the program PHASER[38], using the domains of the Ago–10-base DNA complex structure as a search model. The final model of the Ago–21-base DNA complex contains two molecules in the asymmetric unit. Almost all protein residues and nucleotides 1–11 and 18–21 of the bound DNA in one molecule and nucleotides 1–10 and 19–21 of the other molecule in the asymmetric unit could be traced in the complex.

Cleavage activity assay of *T. thermophilus* Ago. 5′-phosphorylated 21-nucleotide guide oligodeoxynucleotides were prepared by solid-phase synthesis using standard DNA amidites (Sigma-Proligo) and chemical phosphorylating reagent (Glen Research) on an ABI 3400 DNA synthesizer. The 177-nucleotide cleavage substrate was prepared by *in vitro* transcription from a PCR template as described previously[39]. The transcript was dephosphorylated using alkaline phosphatase (Roche) and then 5′-radiolabelled using T4 polynucleotide kinase (Fermentas) and γ-^{32}P-ATP.

Recombinant *T. thermophilus* Ago (0.5 μM) was incubated with a reaction mixture containing 10 mM HEPES-KOH, pH 7.5, 100 mM NaCl, 5 mM $MnCl_2$, 8 U of RNasin and 0.1 μM guide strand for 90 min at 55 °C in a final volume of 15 μl. Next, ATP and GTP were added to a final concentration of 1 mM and 0.2 mM, respectively, together with 0.1 μM 5′-^{32}P-labelled RNA substrate at a final concentration of 0.5 μM. The incubation was continued for 60 min. The reaction was stopped by addition of 185 μl proteinase K solution (1 mg ml^{-1} proteinase K, 20 mM HEPES-KOH, pH 7.5, 1.5 mM EDTA, 100 mM NaCl, 1.5 mM $CaCl_2$, 1.5% SDS), and incubated at 55 °C for 10 min, followed by phenol-chloroform extraction and ethanol precipitation. The cleavage products were resolved by 8% denaturing PAGE, and radioactivity was detected by phosphorimaging.

31. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
32. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallogr. D* **55**, 849–861 (1999).
33. Terwilliger, T. C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59**, 38–44 (2003).
34. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
35. Emsley, P. and Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
36. Brunger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
37. The CCCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
38. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
39. Martinez, J., Patkaniowska, A., Urlaub, H., Lührmann, R. & Tuschl, T. Single-stranded anti-sense siRNAs guide target RNA cleavage in RNAi. *Cell* **110**, 563–574 (2002).

LETTERS

Complex structure within Saturn's infrared aurora

Tom Stallard[1], Steve Miller[2], Makenzie Lystrup[2], Nicholas Achilleos[2], Emma J. Bunce[1], Christopher S. Arridge[3], Michele K. Dougherty[4], Stan W. H. Cowley[1], Sarah V. Badman[1], Dean L. Talboys[1], Robert H. Brown[5], Kevin H. Baines[6], Bonnie J. Buratti[6], Roger N. Clark[7], Christophe Sotin[6], Phil D. Nicholson[8] & Pierre Drossart[9]

The majority of planetary aurorae are produced by electrical currents flowing between the ionosphere and the magnetosphere which accelerate energetic charged particles that hit the upper atmosphere. At Saturn, these processes collisionally excite hydrogen, causing ultraviolet emission[1–8], and ionize the hydrogen, leading to H_3^+ infrared emission[9–15]. Although the morphology of these aurorae is affected by changes in the solar wind[6,11], the source of the currents which produce them is a matter of debate[16,17]. Recent models predict only weak emission away from the main auroral oval[18]. Here we report images that show emission both poleward and equatorward of the main oval (separated by a region of low emission). The extensive polar emission is highly variable with time, and disappears when the main oval has a spiral morphology; this suggests that although the polar emission may be associated with minor increases in the dynamic pressure from the solar wind, it is not directly linked to strong magnetospheric compressions. This aurora appears to be unique to Saturn and cannot be explained using our current understanding of Saturn's magnetosphere. The equatorward arc of emission exists only on the nightside of the planet, and arises from internal magnetospheric processes that are currently unknown.

Saturn's aurorae have previously been examined using observations of ultraviolet emission by NASA's Voyager 2 spacecraft[1] and Hubble Space Telescope (HST)[2–8] and observations of infrared emission using ground-based telescopes[9–15]. Observations from Earth are significantly limited in that Saturn can only be viewed from the dayside, and because Saturn's southern hemisphere has been in summer for the past fifteen years, generally only the southern pole has been viewed. Preliminary results from both the Ultraviolet Imaging Spectrograph (UVIS)[19] and the Visual and Infrared Mapping Spectrometer (VIMS)[20] instruments on NASA's Cassini spacecraft provide new observations from orbit around Saturn; for a more detailed description of these observations, see the Supplementary Information. This paper concerns VIMS images, made between October 2006 and June 2007, that form a significant part of a larger, ongoing set of observations that began while Cassini travelled towards Saturn. These images are much higher in quality than any previous infrared images made of Saturn's aurora[15], and as such can be directly compared with the HST ultraviolet images, surpassing the latter by imaging the aurora from multiple viewing angles.

HST observations have previously shown that Saturn's ultraviolet auroral region is dominated by a single main auroral oval. When the solar wind is rarefied, the main oval is almost circular, typically of relatively low intensity and located at latitude ~75° S (ref. 7). During periods of raised dynamic pressure in the solar wind, the main oval becomes more active, contracting poleward with brighter emission on the dawn side. Under higher dynamic pressures, the auroral oval takes on a spiral morphology[5] produced by compression-induced tail reconnection[21], and during the highest pressure solar wind events the ultraviolet emission infills the entire dawn half of the polar cap and is almost absent on the dusk side[6]. Ground-based infrared observations, by comparison, have a significantly reduced spatial resolution owing to the distorting effect of the Earth's atmosphere, but have shown that the infrared aurora appears to have a similar morphology[11].

Two leading theories have been put forward to explain the origin of this main auroral oval. The first proposes that the location of these aurorae maps to the magnetospheric boundary between closed field lines and those open to the solar wind. The resultant shear in rotational flow produces a ring of upward-directed current that results in electrons being accelerated into the ionosphere[16]. The second proposes that variations in the dynamic pressure of the solar wind result in centrifugal instabilities in the outer magnetosphere, and that the resultant variation in angular velocity drives currents into the ionosphere[17]. Recent analysis has shown that there is a strong correspondence between the location of auroral emission on the planet observed by HST and the boundary between open and closed field lines measured *in situ* by Cassini[8], suggesting that the first of these models may better explain the formation of the main auroral oval.

Ultraviolet emission both poleward and equatorward of the main auroral oval is very low in intensity, not rising significantly above the level of noise[5]. Ground-based infrared observations have shown that there are likely to be raised levels of emission poleward of the main oval, as well as an equatorward secondary oval formed by internal magnetospheric processes in a similar way to the main auroral oval at Jupiter[14]. Models of the aurorae, however, predict that this secondary oval is very weak in comparison with the main auroral oval and that no emission occurs within the polar cap.

The images shown here, whose production is described in detail in the Supplementary Information, suggest that the main auroral oval observed in the infrared is comparable to that observed in the ultraviolet, with examples of morphology similar to that of the ultraviolet oval under both rarefied (Fig. 1) and compressed (Fig. 2) solar wind conditions. However, unlike in previous ultraviolet observations, auroral emission can also be seen in these images both poleward and equatorward of the main auroral oval, at a level significantly brighter than the body of the planet. The region poleward of the main oval, which contains no discernable emission in the HST observations, is shown here to have a significant but highly variable level of infrared emission. Previous ground-based infrared spectroscopic

[1]Department of Physics and Astronomy, University of Leicester, Leicester LE1 7RH, UK. [2]Atmospheric Physics Laboratory, Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK. [3]Department of Space and Climate Physics, Mullard Space Science Laboratory, University College London, Holmbury St Mary, Dorking, Surrey RH5 6NT, UK. [4]Space and Atmospheric Physics Group, Department of Physics, Imperial College of Science, Technology and Medicine, South Kensington Campus, London SW7 2AZ, UK. [5]Lunar and Planetary Laboratory and Steward Observatory, University of Arizona, Tucson, Arizona 85721-0092, USA. [6]Jet Propulsion Laboratory, California Institute of Technology, M/S 183-601, 4800 Oak Grove Drive, Pasadena, California 91109-8099, USA. [7]US Geological Survey, Denver, Colorado 80225, USA. [8]Cornell University, Astronomy Department, Ithaca, New York 14853, USA. [9]Observatoire de Paris, Meudon 92195, France.

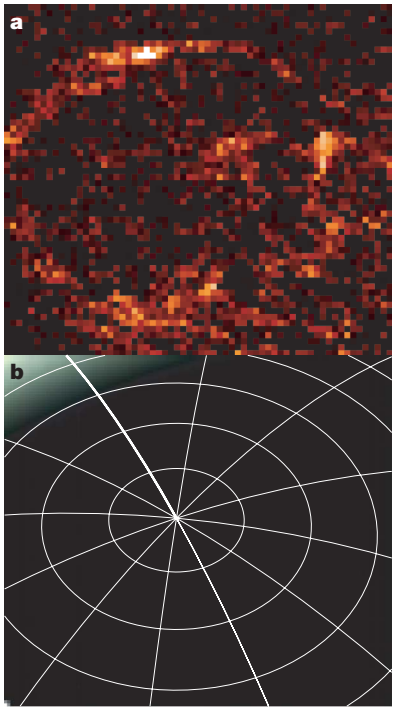


Figure 1 | Quiescent auroral conditions. **a**, H_3^+ emission imaged by the VIMS instrument on Cassini. **b**, Three-colour visible image of Saturn, providing context for the image in **a**. A latitudinal grid in 5° steps and a longitudinal grid in 30° steps has been overlaid on **b**, with noon highlighted in bold. This image was made at 18:21 UT on 10 November 2006, at a time when Cassini's inclined orbit allowed the entire auroral oval to be viewed from an effective planetary inclination of 52° N, providing a view of the entire northern aurora at a distance of $17.6R_S$ (R_S = 60,268 km) from the dusk–midnight sector. The main auroral oval is somewhat fragmented in appearance, consisting of a weak oval located at a latitude of ~75° N, interspersed with several brighter spots along its length, with the brightest of these in the dawn sector approaching noon. The dusk sector emissions are often relatively weak and irregular. This appears to be analogous to auroral emission in the ultraviolet during periods of rarefaction of the solar wind. The main ultraviolet oval also consists of several arcs of different widths and brightnesses, rotating at 65% of the full planetary rotation[18]; such features in the infrared have also been inferred from ground-based spectroscopic observations[14]. The oval is more fragmented in the midnight sector, although this is partly due to poor instrument response in the lower-right part of the array in this particular image. Inside the main oval, there is a bright spot of polar emission in the dawn sector, as intense as some parts of the main auroral oval. The rest of the polar cap appears to contain only low levels of emission.

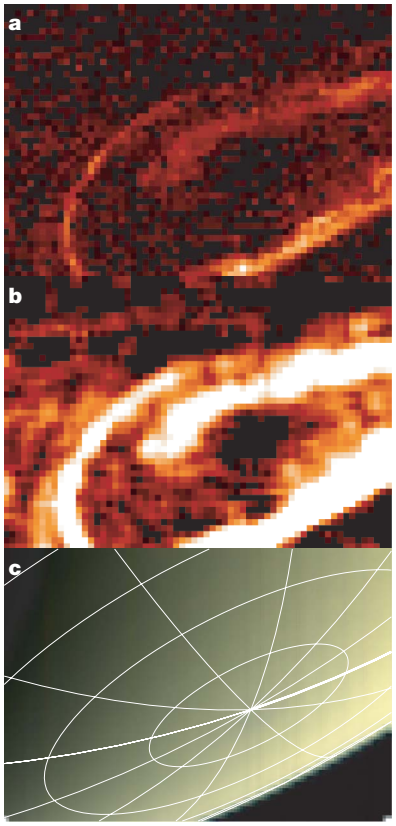


Figure 2 | Spiral auroral morphology. **a**, The southern auroral oval viewed at 02:38 UT, 11 October 2006, from the dawn–midnight sector at a distance of $11.14R_S$. The main oval is shifted significantly poleward, lying at ~80° S, and has a spiral morphology similar to that observed in ultraviolet images, which is characteristic of periods of moderate dynamic pressure in the solar wind[6]. This image provides a clear view of the overlap of the spiral on the dawn side of the planet, with the inner arm at ~85° S and the outer arm at ~81° S. The dusk side of the oval is relatively bright, at least in part because of line-of-sight brightening near the planetary limb. **b**, The same image as in **a**, processed to show low-level emission by smoothing with a three-pixel boxcar window to improve the signal-to-noise ratio and by magnifying the colour scale to show the intensity range from 0–20% of the peak auroral brightness. This reveals low-level emission both equatorward and poleward of the main auroral oval that is significantly brighter than the body of the planet. A second arc of emission, ~25% as bright as the main auroral oval, can be clearly seen ~1° equatorward of the main auroral oval near midnight (on the far left of the image), and extends around the oval as far as the dayside. There are also moderate levels of emission poleward of the main oval, except in a small region close to the pole, which are strongest near noon on the dayside. No distinction can be made between day and night sides of the planet, implying that any H_3^+ created on the dayside by solar extreme ultraviolet emission is not bright enough to be detected. This means that all the emission seen in these observations is auroral in origin, making the relative effect of extreme ultraviolet ionization, in comparison with auroral ionisation, significantly weaker on Saturn than on either Jupiter or Uranus. **c**, Three-colour visible image, overlaid with a grid as in Fig. 1b.

observations have hinted at the presence of such emission[11,15], and preliminary results from the UVIS instrument also see some poleward aurora[19]. However, the relative strength and dynamics of this newly observed aurora is unprecedented.

There are periods when the polar aurora is relatively quiet, with the emission taking the form of localized bright spots, low levels of patchy emission across the pole or extended arcs of emission. However, beyond this general poleward activity, the aurora in this region sometimes become as bright, or brighter, than the main auroral oval (Fig. 3), filling the region between ~82° N and the pole with emission, unlike anything that has been observed in previous observations. This emission is also highly variable with time, completely disappearing on timescales as short as 45 min. It occurs when the morphology of the main auroral oval is relatively circular, although the main oval appears to brighten at the same time as the polar emission, with some poleward emission in the dawn sector also occurring, within a time frame of a few hours. This, combined with the lack of such bright polar emission when the main oval has a spiral morphology (Fig. 2), suggests that although these features may be associated with minor increases in the dynamic pressure of the solar wind, they are not directly linked to strong magnetospheric compressions.

This polar cap emission is much stronger, relative to the brightness of the respective main auroral ovals, than that usually seen at Jupiter[22,23]. The differences in emission strength between the ultraviolet and infrared observations are important and may point towards an explanation of why this emission occurs; the less extreme differences seen at Jupiter have previously been explained by the infrared aurora being formed by comparatively low-energy particle precipitation[24]. Unlike Jupiter, Saturn's ultraviolet and infrared emissions are predominantly produced by similar distributions of particle energies[1]. However, the brightening may be a result of a

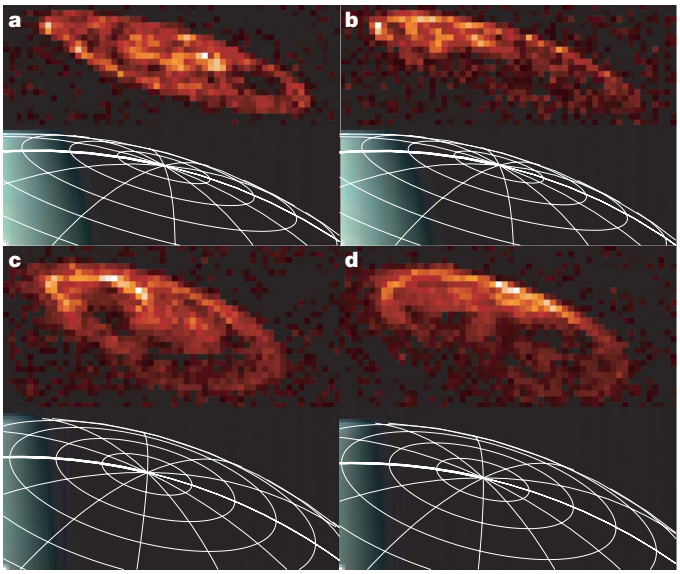


Figure 3 | Bright polar auroral emission. a–d, Images, made on 9 June (**a**, **b**) and 24 May (**c**, **d**) 2007, viewing the northern dusk–noon sector at a distance between 24.4R_S and 26.4R_S. Below each aurora plot is a three-colour visible image overlaid with a grid as in Fig. 1b. The images are separated by 3 h, 18 min (**a**, **b**) and 4 h, 6 min (**c**, **d**). In all the images, the main auroral oval can be seen to extend all the way around the planet, although the emission is typically weaker in the dusk sector. However, it is the emission across the polar cap that is particularly noteworthy. In **a** and **c**, there are large regions of strong polar auroral activity inside the main auroral oval. These are brightest at a latitude $>82^\circ$ N, and extend from noon to midnight and well inside both the dawn and dusk sectors. There is a band of relatively low emission that separates this emission from the main auroral oval. The brightness of the main auroral oval appears to increase when this polar cap emission occurs. The strength and location of this aurora is unlike any emission previously seen in the ultraviolet, or within the aurorae of other planets. In **b** and **d**, significant poleward emission in the dawn–noon sector can also be seen. Given the viewing angle, it is difficult to determine if this emission is separate from the main auroral oval or if it constitutes a poleward extension of the main oval. This emission is more clearly observed than dawn brightening events in the ultraviolet, but could be associated with features previously seen in the ultraviolet emission[9]. There is also some evidence of emission equatorward of the main auroral oval at and around midnight in all these images.

raised temperature across the pole, which would only affect the infrared auroral brightness. Such an increase in temperature could be due to increased heating, or may indicate lower-energy particles precipitating higher into the atmosphere.

That this bright polar aurora appears to be separated from the main auroral oval by a region of low emission is of particular interest. At Earth, bright aurorae do occur across the polar cap region, but only in the form of 'theta' aurorae, namely arcs of auroral emission that extend from a point close to noon to one near midnight (local times) across the polar cap, not the high-latitude disks of emission seen here. Theta aurorae at Earth are believed to be associated with large-scale reconfigurations of magnetospheric open flux that occur over several hours during specific orientations of the. However, because the timescale for reconfigurations of Saturn's open flux (a few days) is generally much longer than the timescale for variations of the interplanetary magnetic field (a few hours), this phenomenon is unlikely to occur in Saturn's magnetosphere[21].

Thus, the bright polar aurora appears to be unique to Saturn and is unexpected on the basis of existing discussions of solar-wind-driven magnetosphere–auroral dynamics. However, it may be associated with varying ion velocities recently measured across the polar region using ground-based spectroscopy[11]. These observations show that some regions poleward of the main auroral oval return to co-rotation with the planet, a phenomena that is also currently unexplained by models of ionospheric–magnetospheric interaction. This produces a velocity shear within the polar region that might be related to the formation of this bright polar emission.

Equatorward of the main auroral oval, emission generally falls off relatively quickly, except for there being a consistent arc of emission just equatorward of the main oval on the nightside (Fig. 2). This arc can extend as far as the dayside, and sometimes appears to form an outer oval extending over all local times. This correlates with brief, limb-brightened nightside ultraviolet emission seen by HST[5] and in preliminary UVIS observations[19].

Given its location equatorward of the main auroral oval, the arc is likely to be caused by an internal magnetospheric process. It could be indicative of energetic particle injections into the inner magnetosphere, which at Earth are often positioned at a local time near midnight[25]. However, because these injections are relatively uncommon and the equatorward arc is seen to re-occur across the ultraviolet and infrared data sets, this explanation seems unlikely. Thus, the equatorward arc at Saturn has no known analogy at another planet; it has also never been predicted by theoretical modelling. Thus, the presence of such an aurora reveals important internal magnetospheric dynamics driving a significant, yet previously undiscussed, current system. The location of the aurora strongly suggests that this current system is associated with dynamics within the magnetotail.

This arc could be the equatorward 'jovian-like' auroral emission recently detected by ground-based infrared spectroscopy[14,26]. The significant brightening on the nightside suggests a non-axisymmetric breakdown in co-rotation in the plasma caused by sub-co-rotating flux tubes expanding outwards on the nightside, owing to lack of confinement by the solar wind. However, it is not clear whether these two separately detected auroral features are co-located, and they may constitute two separate current systems forming auroral emission at Saturn. Along with the presence of large regions of auroral emission across the polar cap, this suggests that although recent observations have finally begun to explain the origin of the main auroral oval at Saturn, our understanding of the currents flowing between the ionosphere and the magnetosphere has barely begun to explain the auroral emission seen from Saturn.

Received 8 April; accepted 12 September 2008.

1. Sandel, B. R. *et al.* Extreme ultraviolet observations from the Voyager 2 encounter with Saturn. *Science* **215**, 548–553 (2005).
2. Trauger, J. T. *et al.* Saturn's hydrogen aurora: Wide field and planetary camera 2 imaging from the Hubble Space Telescope. *J. Geophys. Res.* **103**, 20237–20244 (1998).
3. Gérard, J.-C. *et al.* Characteristics of Saturn's FUV aurora observed with the Space Telescope Imaging Spectrograph. *J. Geophys. Res.* **109**, A09207 (2004).
4. Clarke, J. T. *et al.* Morphological differences between Saturn's ultraviolet aurorae and those of Earth and Jupiter. *Nature* **433**, 717–719 (2005).
5. Grodent, D., Gérard, J.-C., Cowley, S. W. H., Bunce, E. J. & Clarke, J. T. Variable morphology of Saturn's southern ultraviolet aurora. *J. Geophys. Res.* **110**, A07215 (2005).
6. Bunce, E. J. *et al.* Cassini observations of the interplanetary medium upstream of Saturn and their relation to Hubble Space Telescope auroral data. *Adv. Space Res.* **38**, 806–814 (2006).
7. Badman, S. V., Cowley, S. W. H., Gérard, J.-C. & Grodent, D. A statistical analysis of the location and width of Saturn's southern auroras. *Ann. Geophys.* **24**, 3533–3545 (2006).
8. Bunce, E. J. *et al.* Origins of Saturn's aurora: Simultaneous observations by Cassini and the Hubble Space Telescope. *J. Geophys. Res.* **113**, A09209 (2008).
9. Stallard, T. *et al.* The H_3^+ latitudinal profile of Saturn. *Astrophys. J.* **521**, L149–L152 (1999).
10. Stallard, T., Miller, S., Trafton, L. M., Geballe, T. R. & Joseph, R. D. Ion winds in Saturn's southern auroral/polar region. *Icarus* **167**, 204–211 (2004).
11. Stallard, T. *et al.* Saturn's auroral/polar H_3^+ infrared emission I: General morphology and ion velocity structure. *Icarus* **189**, 1–13 (2007).
12. Stallard, T. *et al.* Saturn's auroral/polar H_3^+ infrared emission II: A comparison with plasma flow models. *Icarus* **191**, 678–690 (2007).
13. Melin, H., Miller, S., Stallard, T., Trafton, L. M. & Geballe, T. R. Variability in the H_3^+ emission of Saturn: Consequences for ionisation rates and temperature. *Icarus* **186**, 234–241 (2007).
14. Stallard, T. *et al.* Jovian-like aurorae on Saturn. *Nature* **453**, 1083–1085 (2008).
15. Stallard, T., Lystrup, M. & Miller, S. Emission-line imaging of Saturn's H_3^+ aurora. *Astrophys. J.* **675**, L117–L120 (2008).

16. Cowley, S. W. H., Bunce, E. J. & O'Rourke, J. M. A simple quantitative model of plasma flows and currents in Saturn's polar ionosphere. *J. Geophys. Res.* **109,** A05212 (2004).
17. Sittler, E. C., Blanc, M. F. & Richardson, J. D. Proposed model for Saturn's auroral response to the solar wind: Centrifugal instability model. *J. Geophys. Res.* **111,** A06208 (2006).
18. Cowley, S. W. H. & Bunce, E. J. Corotation-driven magnetosphere-ionosphere coupling currents in Saturn's magnetosphere and their relation to the auroras. *Ann. Geophys.* **21,** 1691–1707 (2003).
19. Pryor, W. R. *et al. Auroral movies and spectroscopy from Cassini UVIS.* Abstract no. P31A–0187 (Fall Meeting, American Geophysical Union, 2007).
20. Drossart, P. *et al.* Mapping the H_3^+ aurora on Saturn from Cassini/VIMS observations. *Bull. Am. Astron. Soc.* **37,** 657 (2005).
21. Cowley, S. W. H. *et al.* Reconnection in a rotation-dominated magnetosphere and its relation to Saturn's auroral dynamics. *J. Geophys. Res.* **110,** A02201 (2005).
22. Grodent, D. *et al.* Jupiter's polar auroral emissions. *J. Geophys. Res.* **108,** 1366–1374 (2003).
23. Stallard, T., Miller, S., Millward, G. & Joseph, R. D. On the dynamics of the Jovian ionosphere and thermosphere II: The measurement of H_3^+ vibrational temperature, column density, and total emission. *Icarus* **156,** 498–514 (2002).
24. Millward, G., Miller, S., Stallard, T., Aylward, A. & Achilleos, N. On the dynamics of the Jovian ionosphere and thermosphere III: The modelling of auroral conductivity. *Icarus* **160,** 95–104 (2002).
25. Baker, D. N., Higbie, P. R., Hones, E. W. & Belian, R. D. High resolution energetic particle measurements at $6.6R_E$ 3. Low energy electron anisotropies and short-term substorm predictions. *J. Geophys. Res.* **83,** 4863–4868 (1978).
26. Hill, T. W. The Jovian auroral oval. *J. Geophys. Res.* **106,** 8101–8108 (2001).

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

Acknowledgements This work was supported by a Research Councils UK Fellowship (T.S.) and by the UK Science and Technology Facilities Council (N.A., S.V.B., D.L.T., C.S.A., E.J.B., M.K.D.). The European authors are part of the Europlanet European Planetology Network, supported by the European Union's Sixth Framework Programme.

Author Contributions T.S. analysed the data and wrote the paper; S.M., M.L. and N.A. aided data analysis; E.J.B., C.S.A., M.K.D., S.W.H.C., S.V.B. and D.L.T. provided discussion as members of the UK MAG-VIMS collaboration team; and R.H.B., K.H.B., B.J.B., R.N.C., C.S., P.D.N. and P.D. provided the reduced data and discussion as members of the Cassini VIMS team. All authors commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.S. (tss@ion.le.ac.uk).

LETTERS

Complete quantum control of a single quantum dot spin using ultrafast optical pulses

David Press[1], Thaddeus D. Ladd[1,2], Bingyang Zhang[1] & Yoshihisa Yamamoto[1,2]

A basic requirement for quantum information processing systems is the ability to completely control the state of a single qubit[1–6]. For qubits based on electron spin, a universal single-qubit gate is realized by a rotation of the spin by any angle about an arbitrary axis. Driven, coherent Rabi oscillations between two spin states can be used to demonstrate control of the rotation angle. Ramsey interference, produced by two coherent spin rotations separated by a variable time delay, demonstrates control over the axis of rotation. Full quantum control of an electron spin in a quantum dot has previously been demonstrated using resonant radio-frequency pulses that require many spin precession periods[7–10]. However, optical manipulation of the spin allows quantum control on a picosecond or femtosecond timescale[11–18], permitting an arbitrary rotation to be completed within one spin precession period[6]. Recent work in optical single-spin control has demonstrated the initialization of a spin state in a quantum dot[19–22], as well as the ultrafast manipulation of coherence in a largely unpolarized single-spin state[17]. Here we demonstrate complete coherent control over an initialized electron spin state in a quantum dot using picosecond optical pulses. First we vary the intensity of a single optical pulse to observe over six Rabi oscillations between the two spin states; then we apply two sequential pulses to observe high-contrast Ramsey interference. Such a two-pulse sequence realizes an arbitrary single-qubit gate completed on a picosecond timescale. Along with the spin initialization and final projective measurement of the spin state, these results demonstrate a complete set of all-optical single-qubit operations.

Coherent control of a single qubit is often accomplished by driving the qubit at its resonant frequency. For a qubit composed of a single electron spin in a magnetic field, resonant coherent control requires the use of radio-frequency pulses of at least nanosecond duration[7–10]. One way to reduce this timescale is to construct qubits of multiple, coupled particles and to rapidly manipulate their coupling potential, as exemplified by the several-hundred-picosecond gate times of exchange-coupled electron pairs in electrically controlled quantum dots[23]. For an isolated, optically controlled quantum dot spin such as the one studied here, even shorter operation times may be achieved using ultrafast optical pulses[11–18]. Using such optical pulses, the axis of rotation of the qubit is determined by the arrival time of the pulse with respect to the qubit oscillation period[6]. A single-qubit gate consisting of an arbitrary rotation about any axis may thereby be completed in a single Larmor period. For electron spin qubits, a large magnetic field is therefore necessary to increase the speed of a single-qubit gate, and quantum information processing with clock speeds in excess of 10 GHz may be possible[6].

Our scheme to rotate a single electron spin using a picosecond pulse is shown in Fig. 1a. A single electron is confined in the quantum dot. The electron spin states $|\downarrow\rangle$ and $|\uparrow\rangle$ are split by an externally applied magnetic field $B_{\mathrm{ext}} = 7$ T, aligned parallel to the z axis (Voigt geometry; see Fig. 1d), to provide a large Larmor precession frequency of $\delta_{\mathrm{e}}/2\pi = 26.3$ GHz. The lowest energy interband transitions are to the two trion states consisting of a pair of electrons in a spin singlet and an unpaired heavy hole[24], denoted $|\uparrow\downarrow, \Downarrow\rangle$ and $|\uparrow\downarrow, \Uparrow\rangle$, which are split by a frequency δ_{h}. Each trion state forms an independent Λ system with the two metastable states $|\downarrow\rangle$ and $|\uparrow\rangle$. Optical selection rules dictate that the vertical and cross transitions in Fig. 1a couple to orthogonal linear polarizations of light, denoted H and V, and are $\pi/2$ out of phase with each other. The exact orientations of H and V are determined by the shape and strain of the quantum dot[25]. Each transition has a Rabi frequency $\Omega_{\mathrm{H}} = \mu E_{\mathrm{H}}/\hbar$ or $\Omega_{\mathrm{V}} = \mu E_{\mathrm{V}}/\hbar$, where μ is the transition's dipole strength and E_{H} and E_{V} are the complex electric field amplitudes of the rotation pulse in the corresponding polarization basis. A circularly polarized rotation pulse ensures that the probability amplitudes from the two Λ systems add constructively, and a large detuning Δ minimizes undesired population in the excited states. Hence, a single broadband rotation pulse will coherently change the spin from $|\downarrow\rangle$ to $|\uparrow\rangle$ and back through a stimulated Raman transition. The dynamics may be qualitatively described by the condition that $\Omega_{\mathrm{H}} \ll \Delta$ and $\Omega_{\mathrm{V}} \ll \Delta$, under which the upper levels can be adiabatically eliminated. Doing so, we expect to find two-state Rabi oscillations with an effective Rabi frequency $\Omega_{\mathrm{eff}} \approx \Omega_{\mathrm{H}}\Omega_{\mathrm{V}}/\Delta$ between states $|\downarrow\rangle$ and $|\uparrow\rangle$. The spin rotation may alternatively be described in terms of an optical Stark shift[17].

In addition to rotations, a complete set of single-qubit operations also requires initialization and measurement. We perform both of these tasks by optical pumping (Fig. 1b). A narrowband, continuous-wave laser optically drives the $|\downarrow\rangle \leftrightarrow |\uparrow\downarrow, \Downarrow\rangle$ transition with rate Ω_{P}. The optical pumping laser has negligible effect on the spin rotation because $\Omega_{\mathrm{P}} \ll \Omega_{\mathrm{eff}}$. Spontaneous decay into the two spin states at half the trion's total spontaneous emission rate, denoted Γ, quickly initializes the electron into the $|\uparrow\rangle$ state. After spin rotation, the population in the $|\downarrow\rangle$ state is measured using the same optical pumping process. If the spin is rotated to $|\downarrow\rangle$, the quantum dot will emit a single photon from the $|\uparrow\downarrow, \Downarrow\rangle \rightarrow |\uparrow\rangle$ transition, which can be detected using a single-photon counter.

Our single-spin measurement technique has been proposed for use in quantum computation[1], and offers the experimental convenience of including measurement and initialization in the same step. However, the fidelity of a single-shot readout is limited by the photon collection efficiency. An optical microcavity would boost the measurement scheme's efficiency, and could also enable coherent conversion of spin qubits into photon qubits for quantum networking[26]. Resonant absorption measurements[19–21] offer similar advantages, but also require a microcavity-enhanced absorption cross-section to enable single-shot readout. Quantum non-demolition measurements based on dispersive Kerr rotation[27], Faraday rotation[28] or a recycling transition[29] use many photons to measure the spin and are therefore more robust to photon loss, but they require a separate initialization step.

[1]E. L. Ginzton Laboratory, Stanford University, Stanford, California 94305, USA. [2]National Institute of Informatics, Hitotsubashi 2-1-2, Chiyoda-ku, Tokyo 101-8403, Japan.

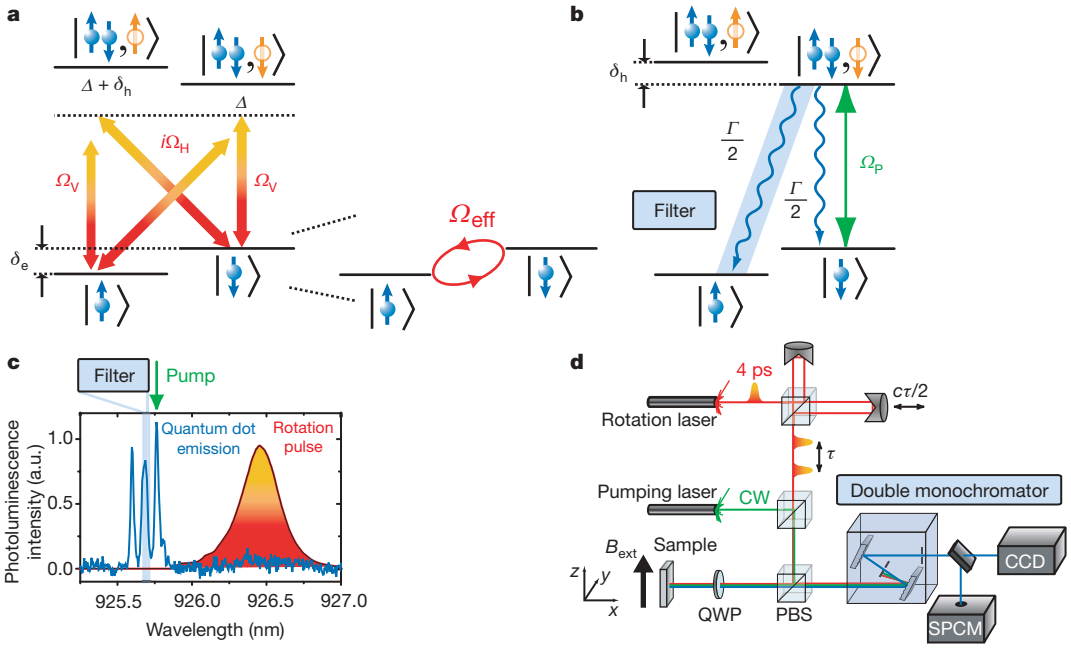


Figure 1 | Experimental methods to initialize, control and measure a single electron spin. **a**, The spin rotation scheme involves a stimulated Raman transition through two independent Λ systems. The four-level system is effectively reduced to a two-level system if the Rabi frequencies, Ω_H and Ω_V, are much smaller than the detuning, Δ. **b**, The spin initialization and measurement scheme performed by optical pumping. **c**, Measured photoluminescence spectrum of the charged quantum dot excited by an above-bandgap, 785-nm laser. The rotation pulse is detuned by $\Delta/2\pi = 290$ GHz below the lowest transition frequency. **d**, Experimental set-up. During each experimental cycle, one or two rotation pulses may be sent to the sample to observe Rabi oscillations or Ramsey interference, respectively. The time delay, τ, between pairs of pulses is controlled by a retroreflector mounted on a computer-controlled translation stage. CW, continuous wave; QWP, quarter-wave plate; PBS, polarizing beam splitter; SPCM, single-photon counting module; CCD, charge-coupled device; c, speed of light.

The optical initialization is calibrated by measuring the single-photon count rate as a function of optical pumping power, P_{OP}, following a fixed rotation through angle $\Theta = \pi$ (Fig. 2a). The signal saturates at around $P_{OP} \approx 15\,\mu W$ as the population in the $|\downarrow\rangle$ state becomes almost completely initialized to $|\uparrow\rangle$. In all of our remaining experiments, P_{OP} is fixed just above the saturation value of the optical pumping curve. To quantify the initialization fidelity, we make a time-resolved measurement of photon count rate following a rotation of $\Theta = \pi$ (Fig. 2b). The count rate is proportional to the instantaneous population in $|\downarrow\rangle$. Immediately following the rotation pulse, the population in $|\downarrow\rangle$ is near unity and the signal is maximized. The signal drops as the spin is pumped back to $|\uparrow\rangle$ in a characteristic time of 3.4 ns, orders of magnitude faster[21] than optical pumping schemes involving a dipole-forbidden transition[19]. The minimum count rate, which is measured just before the next rotation pulse, corresponds to the remnant population in $|\downarrow\rangle$ due to imperfect initialization. By comparing the count rates immediately before and after the rotation pulse, we estimate (see Supplementary Information) the spin initialization fidelity to be $F_0 = 92 \pm 7\%$.

Rabi oscillations between the two spin states are evident in the photon count rate as the rotation pulse power P_{RP} is varied (Fig. 3a). By contrast with the adiabatic-elimination model discussed earlier, in which $\Theta \propto P_{RP}$, we empirically determine that $\Theta \propto P_{RP}^{0.68}$ in the range $\pi \leq \Theta \leq 13\pi$ (Fig. 3b). This sublinear dependency is a consequence of the breakdown of the adiabatic approximation ($\Omega_H \ll \Delta$, $\Omega_V \ll \Delta$), as non-negligible virtual population is present

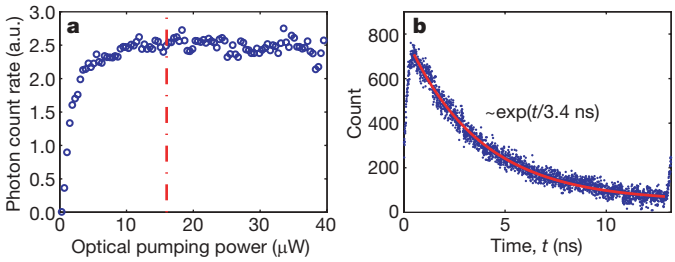


Figure 2 | Initialization by optical pumping. **a**, Saturation of the spin initialization process by means of optical pumping, showing single-photon signal as a function of optical pumping power for a fixed rotation angle $\Theta = \pi$. The operating power, P_{OP}, for the optical pump in all subsequent experiments is indicated by the dash–dot line. **b**, Time-resolved measurement of optical pumping following a fixed rotation by $\Theta = \pi$. The count rate is fitted by an exponential decay with a 3.4-ns time constant. a.u., arbitrary units.

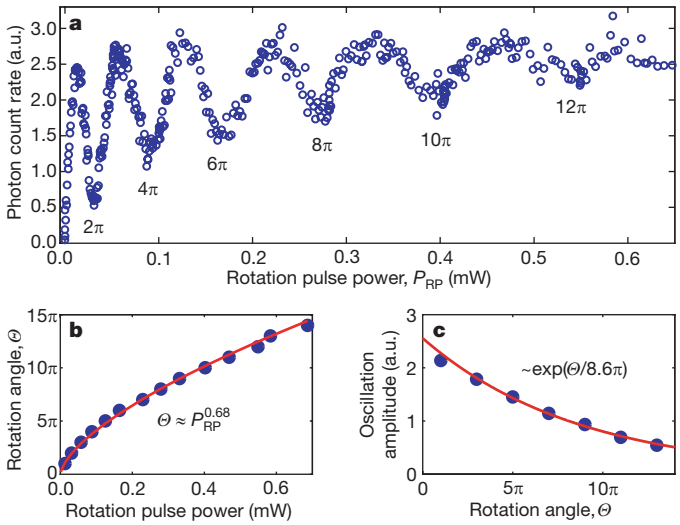


Figure 3 | Experimental demonstration of Rabi oscillations. **a**, Rabi oscillations between the spin states are evident in the oscillating photon signal as rotation pulse power, P_{RP}, is increased. **b**, The rotation angle as a function of rotation pulse power, showing an empirical fit to a power-law dependence. **c**, Amplitude of measured Rabi oscillations as a function of rotation angle, with an empirical exponential fit.

in the excited states during the rotation pulse. A four-level master-equation simulation of stimulated Raman scattering with no adjustable parameters and no decoherence processes (see Supplementary Information) produces oscillations that are well fit by $\Theta \propto P_{\mathrm{RP}}^{0.65}$, in reasonable agreement with experiment.

The amplitude of the Rabi oscillations shrinks owing to incoherent processes such as trion dephasing. This may be understood as a decrease of the length of the Bloch vector of the two-state system as Θ increases. This decreasing length is well fit (excluding the first data point) by an empirical exponential decay proportional to $\exp(-\Theta/8.6\pi)$, as shown in Fig. 3c. These incoherent processes transform the virtual population in the excited states during the rotation pulse into real population, which contributes to the photon count rate as background noise. The increasing background is responsible for the overall upwards slope of the data in Fig. 3a.

The experimentally determined trajectory of the Bloch vector as it undergoes Rabi oscillations is parametrically plotted in Fig. 4 as a function of rotation pulse power. The methods we used to generate this trajectory are described in the Supplementary Information. For small rotation angles $\Theta \lesssim \pi$, the vector rotates about a tilted axis because the Larmor precession frequency, δ_e, is non-negligible in comparison with the effective Rabi frequency, Ω_{eff}. This tilted axis of rotation causes the reduction in height of the first peak in Fig. 3a and the lowering of the first oscillation amplitude in Fig. 3c. For larger rotation angles, $\Omega_{\mathrm{eff}} \gg \delta_e$ and the rotation is very nearly about the x axis.

Rabi oscillations demonstrate the rotation of a qubit by an arbitrary angle about a single axis, that is, U(1) control. Full control over the Bloch sphere (SU(2) control) requires rotation about a second axis. The natural Larmor precession of the spin about the z axis accomplishes this rotation, and can be investigated using Ramsey interferometry.

In a Ramsey interferometer, the spin population is measured following a pair of $\pi/2$ rotations about the x axis separated by a variable free precession of time τ about the z axis. Ramsey fringes are shown in Fig. 5a. The fringe amplitude decays with a time constant $T_2^* = 185$ ps. This small time constant is a consequence of the optical pumping laser remaining on between the two rotation pulses, and could be increased by switching the optical pump off between pulses using a fast electro-optic modulator. We determine the electron g-factor magnitude to be $|g_e| = 0.267$ from the Larmor frequency $\delta_e/2\pi = 26.3$ GHz. To estimate the fidelity of each $\pi/2$ pulse, we assume that the Bloch vector initially has length $L_0 = 0.83$ (determined from our initialization fidelity of 92%) and is directed towards the south pole of the Bloch sphere, and that it shrinks in length by a factor of $D_{\pi/2}$ with each pulse. The Bloch vector length after two pulses is thus $L_0 D_{\pi/2}^2$, and the population in state $|\downarrow\rangle$ oscillates between $(1 + L_0 D_{\pi/2}^2)/2$ and $(1 - L_0 D_{\pi/2}^2)/2$ with a Larmor period of $2\pi/\delta_e$. The fidelity of each rotation is then given by $(1 + D_{\pi/2})/2$. We estimate the fidelity by considering the Ramsey fringe amplitude at the shortest measured delay time (see Supplementary Information for details). This gives a $\pi/2$ pulse fidelity of $F_{\pi/2} = 94\%$.

To investigate the quality of our π pulses, we perform a similar experiment with two π pulses separated by a variable time delay, as shown in Fig. 5b. Ideally, the signal would remain constant at $L_0(1 - D_\pi^2)/2$ with no oscillations. The signal shows an overall upwards slope, again due to the optical pump remaining on between the two π pulses and pumping population from the $|\downarrow\rangle$ state into $|\uparrow\downarrow, \Downarrow\rangle$ where it is later detected. Small oscillations remain in the signal because our π pulse is not exactly around the x axis, as discussed earlier. We estimate from the phase of these remaining fringes that our π pulse rotates the spin about a vector tilted 0.17 rad from the x axis. If we simply model the rotation pulse as a rectangular pulse with constant frequency Ω_{eff} applied over 4 ps, we would expect to rotate around an axis tilted by roughly $\delta_e/\Omega_{\mathrm{eff}} = 0.21$ rad, in reasonable agreement with experiment. By comparing the length and orientation of the Bloch vector after our π pulse with a unit vector in the direction of the north pole of the Bloch sphere, we estimate our π pulse fidelity to be roughly $F_\pi = 91\%$.

To construct a general SU(2) single-qubit gate, we may adjust the intensities of the first and second rotation pulses and the precession duration, τ, thus applying three rotations through Euler angles about the x, z and x axes. In Fig. 5c we explore the entire surface of the Bloch sphere by varying the rotation angle of both rotation pulses as well as the delay time τ. The fringe amplitude is shown as a function of rotation angle in Fig. 5d. High-contrast Ramsey fringes are visible when each rotation angle is a half-integer multiple of π, and the fringes vanish when each rotation angle is an integer multiple of π.

In conclusion, we have demonstrated the optical initialization, rotation by arbitrary angle, and projective measurement of an electron spin in a quantum dot. This forms a complete set of all-optical single-qubit operations. A single-qubit gate, consisting of three independent rotations about different axes, is accomplished in less than one Larmor period of 38 ps. Coherence times of $T_2 = 3.0$ μs have

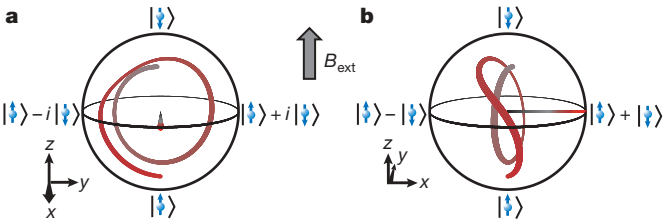


Figure 4 | Reconstructed evolution of the Bloch vector. The curves trace out the tip of the Bloch vector in the one-pulse (Rabi oscillation) experiment over the range of rotation angles $0 \leq \Theta \leq 3\pi$. The colour scale indicates the length of the Bloch vector, which shrinks exponentially with Θ. Views are from the perspective of the x axis (**a**) and the $-y$ axis (**b**) of the Bloch sphere. The rotation angle and the length of the Bloch vector are extracted from the extrema of the Rabi oscillation data shown in Fig. 3, and the azimuthal position of the Bloch vector is determined from the phase of the Ramsey fringes shown in Fig. 5.

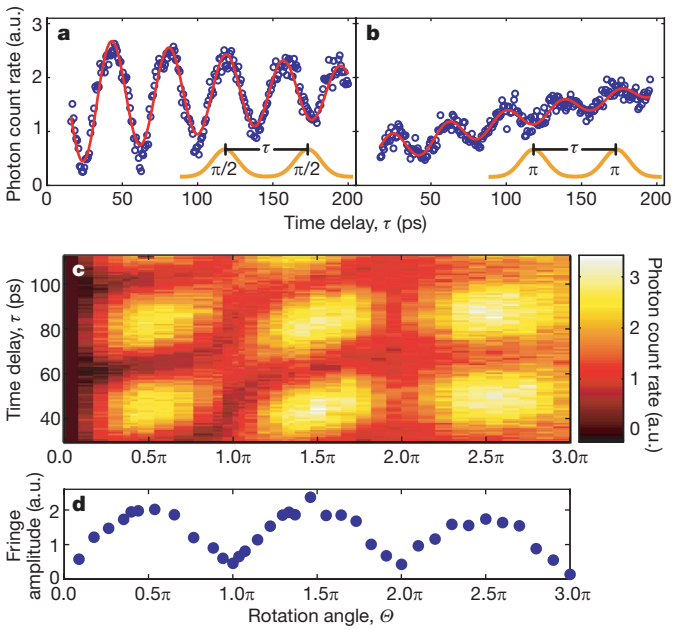


Figure 5 | Experimental demonstration of Ramsey fringes. **a**, Ramsey interference for a pair of $\pi/2$ pulses, showing photon count rate as a function of the time delay between pulses. **b**, Ramsey fringes for a pair of π pulses. The data in **a** and **b** are fitted to an exponentially decaying sinusoid with a linear offset (see Supplementary Information for details). **c**, Photon count rate is colour-mapped as a function of rotation angle, Θ, and delay time between pulses, τ. **d**, The amplitude of Ramsey fringes for various rotation angles. Fringe amplitudes are determined by fitting the data shown in **c** with decaying sinusoids.

been reported for quantum dot electron spins[13], so nearly 10^5 gate operations may be possible within the qubit's coherence time. The rotation pulses are of sufficient fidelity to be applied to a simple spin-based quantum information processing system. Our results are also readily applicable to a spin–photon interface for quantum networks, and suggest methods of greatly increasing the speed and number of operations in general spin-based quantum information processing schemes.

METHODS SUMMARY

The sample contained about $5 \times 10^9\ \mathrm{cm}^{-2}$ self-assembled InGaAs quantum dots grown by the Stranski–Krastanow method on a GaAs substrate. A δ-doping layer of Si donors with a surface density of roughly $10^{10}\ \mathrm{cm}^{-2}$ was grown 20 nm below the layer of quantum dots to probabilistically dope them. Roughly half of the quantum dots were charged, and could be identified by their splitting into a symmetrical quadruplet at high magnetic field (Fig. 1c). The sample was etched into 600-nm-diameter mesa structures containing roughly ten quantum dots that could be individually spectroscopically probed.

The experimental setup is shown in Fig. 1d. The sample was cooled to 1.5 K in a superconducting magnetic cryostat. The optical pumping and rotation pulse lasers were focused onto the sample using an aspheric objective lens (numerical aperture of 0.68) placed inside the cryostat. The sample was positioned relative to the objective using piezoelectric 'slip-stick' positioners. Single-photon photoluminescence was collected through the same lens and directed onto a single-photon counter or charge-coupled device. Collection efficiency was further enhanced using a hemispherical solid immersion lens. The quantum dot emission was spectrally dispersed and filtered using a double monochromator with 0.02-nm resolution. Scattered laser light was further rejected by double-passing through a quarter-wave plate and polarizer. The optical pumping laser was tuned to be resonant with the lowest energy excitonic transition, and the rotation pulse laser was detuned by $\Delta/2\pi = 290$ GHz below the lowest transition frequency. The rotation pulse laser was modulated at 500 Hz using an optical chopper, and the signal from the single-photon counter was detected using a lock-in amplifier. For the time-resolved measurement (Fig. 2b), a timing histogram was generated using a time-interval analyser without lock-in amplification. Instead, the background from scattered optical pump photons and detector dark counts was independently measured and subtracted.

Received 1 July; accepted 6 October 2008.

1. Imamoglu, A. *et al.* Quantum information processing using quantum dot spins and cavity QED. *Phys. Rev. Lett.* **83,** 4204–4207 (1999).
2. Combescot, M. & Betbeder-Matibet, O. Theory of spin precession monitored by laser pulse. *Solid State Commun.* **132,** 129–134 (2004).
3. Chen, P., Piermarocchi, C., Sham, L. J., Gammon, D. & Steel, D. G. Theory of quantum optical control of a single spin in a quantum dot. *Phys. Rev. B* **69,** 075320 (2004).
4. Pryor, C. E. & Flatté, M. E. Predicted ultrafast single-qubit operations in semiconductor quantum dots. *Appl. Phys. Lett.* **88,** 233108 (2006).
5. Economou, S. E., Sham, L. J., Wu, Y. & Steel, D. G. Proposal for optical U(1) rotations of electron spin trapped in a quantum dot. *Phys. Rev. B* **74,** 205415 (2006).
6. Clark, S. M., Fu, K.-M. C., Ladd, T. D. & Yamamoto, Y. Quantum computers based on electron spins controlled by ultrafast off-resonant single optical pulses. *Phys. Rev. Lett.* **99,** 040501 (2007).
7. Koppens, F. H. L. *et al.* Driven coherent oscillations of a single electron spin in a quantum dot. *Nature* **442,** 766–771 (2006).
8. Tokura, Y., van der Wiel, W., Obata, T. & Tarucha, S. Coherent single electron spin control in a slanting Zeeman field. *Phys. Rev. Lett.* **96,** 047202 (2006).
9. Nowack, K. C., Koppens, F. H. L., Nazarov, Y. V. & Vandersypen, L. M. K. Coherent control of a single electron spin with electric fields. *Science* **318,** 1430–1433 (2007).
10. Koppens, F. H. L., Nowack, K. C. & Vandersypen, L. M. K. Spin echo of a single electron spin in a quantum dot. *Phys. Rev. Lett.* **100,** 236802 (2008).
11. Gupta, J. A., Knobel, R., Samarth, N. & Awschalom, D. D. Ultrafast manipulation of electron spin coherence. *Science* **292,** 2458–2461 (2001).
12. Dutt, M. V. G. *et al.* Ultrafast optical control of electron spin coherence in charged GaAs quantum dots. *Phys. Rev. B* **74,** 125306 (2006).
13. Greilich, A. *et al.* Mode locking of electron spin coherences in singly charged quantum dots. *Science* **313,** 341–345 (2006).
14. Wu, Y. *et al.* Selective optical control of electron spin coherence in singly charged GaAs-$Al_{0.3}Ga_{0.7}As$ quantum dots. *Phys. Rev. Lett.* **99,** 097402 (2007).
15. Greilich, A. *et al.* Robust manipulation of electron spin coherence in an ensemble of singly charged quantum dots. *Phys. Rev. B* **75,** 233301 (2007).
16. Carter, S. G., Chen, Z. & Cundiff, S. T. Ultrafast below-resonance Raman rotation of electron spins in GaAs quantum wells. *Phys. Rev. B* **76,** 201308(R) (2007).
17. Berezovsky, J., Mikkelson, M. H., Stoltz, N. G., Coldren, L. A. & Awschalom, D. D. Picosecond coherent optical manipulation of a single electron spin in a quantum dot. *Science* **320,** 349–352 (2008).
18. Fu, K.-M. C. *et al.* Ultrafast control of donor-bound electron spins with single detuned optical pulses. *Nature Phys.* **4,** 780–784 (2008).
19. Atatüre, M. *et al.* Quantum-dot spin-state preparation with near-unity fidelity. *Science* **312,** 551–553 (2006).
20. Gerardot, B. D. *et al.* Optical pumping of a single hole spin in a quantum dot. *Nature* **451,** 441–443 (2008).
21. Xu, X. *et al.* Fast spin state initialization in a singly charged InAs-GaAs quantum dot by optical cooling. *Phys. Rev. Lett.* **99,** 097401 (2007).
22. Ramsay, A. J. *et al.* Fast optical preparation, control, and readout of a single quantum dot spin. *Phys. Rev. Lett.* **100,** 197401 (2008).
23. Petta, J. R. *et al.* Coherent manipulation of coupled electron spins in semiconductor quantum dots. *Science* **309,** 2180–2184 (2005).
24. Bayer, M. *et al.* Fine structure of neutral and charged excitons in self-assembled In(Ga)As/(Al)GaAs quantum dots. *Phys. Rev. B* **65,** 195315 (2002).
25. Krizhanovskii, D. N. *et al.* Individual neutral and charged $In_xGa_{1-x}As$-GaAs quantum dots with strong in-plane optical anisotropy. *Phys. Rev. B* **72,** 161312(R) (2005).
26. Cirac, J. I., Zoller, P., Kimble, H. J. & Mabuchi, H. Quantum state transfer and entanglement distribution among distant nodes in a quantum network. *Phys. Rev. Lett.* **78,** 3221–3224 (1997).
27. Berezovsky, J. *et al.* Nondestructive optical measurements of a single electron spin in a quantum dot. *Science* **314,** 1916–1920 (2006).
28. Atatüre, M., Dreiser, J., Badolato, A. & Imamoglu, A. Observation of Faraday rotation from a single confined spin. *Nature Phys.* **3,** 101–105 (2007).
29. Shabaev, A., Efros, A. L., Gammon, D. & Merkulov, I. A. Optical readout and initialization of an electron spin in a single quantum dot. *Phys. Rev. B* **68,** 201305(R) (2003).

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

Acknowledgements This work was supported by JST/SORST, NICT, MEXT, MURI (ARMY, DAAD 19-03-1-0199) and Special Coordination Funds for Promoting Science and Technology. We thank S. Koseki, S. Götzinger, C. Santori and Q. Zhang for their assistance.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.P. (dlpress@stanford.edu).

LETTERS

Nanoscale chemical imaging of a working catalyst by scanning transmission X-ray microscopy

Emiel de Smit[1], Ingmar Swart[1], J. Fredrik Creemer[2], Gerard H. Hoveling[3], Mary K. Gilles[4], Tolek Tyliszczak[4], Patricia J. Kooyman[5], Henny W. Zandbergen[6], Cynthia Morin[1], Bert M. Weckhuysen[1] & Frank M. F. de Groot[1]

The modern chemical industry uses heterogeneous catalysts in almost every production process[1]. They commonly consist of nanometre-size active components (typically metals or metal oxides) dispersed on a high-surface-area solid support, with performance depending on the catalysts' nanometre-size features and on interactions involving the active components, the support and the reactant and product molecules. To gain insight into the mechanisms of heterogeneous catalysts, which could guide the design of improved or novel catalysts, it is thus necessary to have a detailed characterization of the physicochemical composition of heterogeneous catalysts in their working state at the nanometre scale[1,2]. Scanning probe microscopy methods have been used to study inorganic catalyst phases at subnanometre resolution[3–6], but detailed chemical information of the materials in their working state is often difficult to obtain[5–7]. By contrast, optical microspectroscopic approaches offer much flexibility for *in situ* chemical characterization; however, this comes at the expense of limited spatial resolution[8–11]. A recent development promising high spatial resolution and chemical characterization capabilities is scanning transmission X-ray microscopy[4,12,13], which has been used in a proof-of-principle study to characterize a solid catalyst[14]. Here we show that when adapting a nanoreactor specially designed for high-resolution electron microscopy[7], scanning transmission X-ray microscopy can be used at atmospheric pressure and up to 350 °C to monitor *in situ* phase changes in a complex iron-based Fisher–Tropsch catalyst and the nature and location of carbon species produced. We expect that our system, which is capable of operating up to 500 °C, will open new opportunities for nanometre-resolution imaging of a range of important chemical processes taking place on solids in gaseous or liquid environments.

In situ solid-state microscopic catalyst characterization studies that use scanning transmission electron microscopy/electron energy-loss spectroscopy[3,4], scanning tunnelling microscopy[5,6] or optical microspectroscopic techniques[8–11] focus either on identifying the inorganic catalyst phase responsible for the catalytic activity, or on characterizing organic reactants, intermediates and products interacting with the catalyst surface. Scanning transmission X-ray microscopy (STXM) is a particularly promising recent addition to the field of chemical microspectroscopy[4,12,13] in that the use of soft X-rays (200–2,000 eV) as a probe makes it possible to image both the active phase of the catalyst (through absorption edges of the inorganic species) and the organic reactant phase of the catalyst (through the absorption edges of carbon, oxygen and nitrogen) with a spatial resolution of ~15 nm. In addition, the X-ray probe induces less radiation damage in comparison with the related scanning transmission electron microscopy/electron energy-loss spectroscopy technique[4].

The main experimental challenge to overcome when applying the technique *in situ* is the strong attenuation of soft X-rays by matter. As a result, stringent restrictions apply to the thickness of the sample and the X-ray path length through the medium. For example, the transmission of 700-eV X-rays through 1-bar CO decreases from 60% after 50 μm to ~0.1% after 250 μm. This difficulty meant that the first study[14] of a catalytic solid using STXM under *in situ* conditions was restricted to sample treatments in diluted gases at temperatures below 260 °C, which precluded the study of many catalytic systems. For the present work, we adapt a nanoreactor originally designed for *in situ* transmission electron microscopy studies[7] for use in STXM and are thus able to expose our sample to a reactant gas atmosphere (typically 1.2 bar) at temperatures up to 500 °C. In addition, as gas-phase attenuation for STXM is lower than in transmission electron microscopy experiments[4], the nanoreactor allows the use of CO in the *in situ* STXM experiments.

The nanoreactor, manufactured as a microelectromechanical system, consists of a reactor chamber connected by micrometre-size gas-flow channels (see Fig. 1 for details of the set-up). Two amorphous 1.2-μm-thick SiN_x windows separate the reactor from the outer environment. The windows are etched down to a thickness of 10 nm in certain areas to maximize the transmission of the soft X-rays. The height of the reactor is about 50 μm and ensures minimum attenuation of the X-rays by gas-phase molecules. Heating is provided by a platinum resistive heater spiral embedded in one of the windows[7]. The reactor is supported on an adaptor which is mounted on an interferometrically controlled, piezoelectric stage that can translate the sample in the X-ray beam with nanometre precision. The adaptor is designed to hold up to two separate nanoreactors and connects the reactors to the external gas supply and electronics.

We use the system for an *in situ* STXM study of a fully promoted iron-based Fischer–Tropsch catalyst. In Fischer–Tropsch synthesis (FTS), synthesis gas (a mixture of CO and H_2) is converted into hydrocarbon chains through a surface polymerization reaction[15–17]. This reaction enables the production of high-purity chemicals and transportation fuels from sources other than conventional crude oil, most notably natural gas, coal and biomass. The iron-based catalyst consists of an iron oxide phase dispersed on silicon oxide (SiO_2), with copper oxide and potassium oxide promoters added to improve its selectivity, activity and stability. During FTS, iron oxide and metallic iron (α-Fe) usually coexist, with the iron phases largely converted into iron carbides (essentially α-Fe species with carbon dissolved in their interstitial vacancies). Owing to the complexity of the iron–oxygen–carbon system, the identity of the active phase(s) has long been controversial. In fact, α-Fe, bulk and surface iron carbides and Fe_3O_4 have all been suggested as active phases for FTS[17].

[1]Inorganic Chemistry and Catalysis, Debye Institute for Nanomaterials Science, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands. [2]DIMES-ECTM, Delft University of Technology, PO Box 5053, 2600 GB Delft, The Netherlands. [3]DEMO, Delft University of Technology, PO Box 5031, 2600 GA Delft, The Netherlands. [4]Advanced Light Source, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. [5]DelftChemTech and National Centre for High Resolution Electron Microscopy, Delft University of Technology, Julianalaan 136, 2628 BL Delft, The Netherlands. [6]Kavli Institute of NanoScience, National Centre for High Resolution Electron Microscopy, Delft University of Technology, PO Box 5046, 2600 GA Delft, The Netherlands.

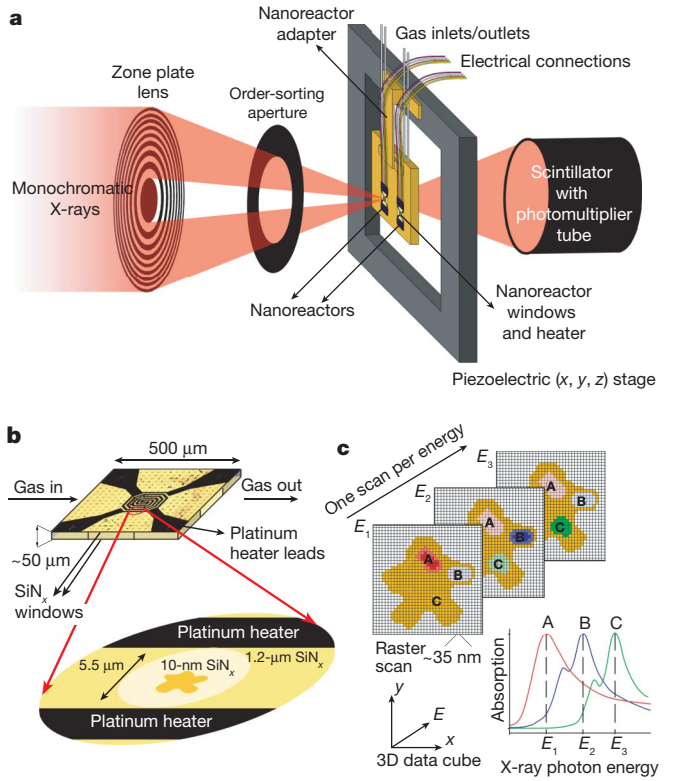


Figure 1 | Experimental set-up and data acquisition method. a, Diagram of the *in situ* STXM technique. Soft X-ray light is focused on the sample using a Fresnel-type zone plate lens. An order-sorting aperture filters out higher-order diffraction orders. The nanoreactor containing the sample is placed in an adaptor that holds up to two nanoreactors at the same time. The adaptor can be translated with nanometre precision by an interferometrically controlled (x, y, z) piezoelectric stage, allowing the acquisition of raster scans. **b**, Close-up of the nanoreactor, showing the windows and the embedded heater spiral[7]. The reactor dimensions are $\sim 500 \times 500 \times 50\ \mu m^3$. The platinum heater has four electronic connections for simultaneous power supply and resistive temperature measurement. The sample itself is supported on the SiN_x windows. The measurements are performed in the circular areas (5.5-µm diameter) where the 1.2-µm-thick SiN_x windows are etched down to a thickness of ~10 nm. **c**, Diagram of a typical STXM data acquisition method. By acquiring images at different X-ray photon energies (for example E_1, E_2 and E_3), a three-dimensional (3D) data cube with full spectral information at every pixel is obtained. This data can be used to image and distinguish between specific chemical species (for example species A, B and C).

In our experiment, we image the reduction of supported iron oxide particles upon heating to 350 °C in H_2 (1 bar) and, subsequently, the working catalyst during the FTS reaction at 250 °C in synthesis gas (1 bar).

We use the carbon K edge (284.2 eV), the oxygen K edge (543.1 eV) and the iron L_2 and L_3 edges (706.8 eV and 719.9 eV, respectively) to image and characterize the sample. The iron L_2 and L_3 edges are used to monitor the valence and coordination[18,19] of iron species, and the oxygen K edge is used to distinguish between different oxygen-containing species. The pre-edge in the oxygen K-edge spectrum is due to the mixing of the oxygen $2p$ states with transition metal $3d$ states and can be used as a measure for the unoccupied iron $3d$ states[20]. In addition, the oxygen K edge of SiO_2 is sufficiently distinct from the iron oxide spectra to specifically map the location of the silicon and iron oxide species. The carbon K edge is measured to image the type and location of carbon species present in the catalyst during reaction.

Before the *in situ* experiments, we analyse the material at room temperature (25 °C) in a helium atmosphere by recording images at the iron L_2 and L_3 edges (Fig. 2d) and the oxygen K edge (Fig. 2e). The iron phase is mainly present as α-Fe_2O_3, as is made evident by X-ray absorption spectra of the iron L_2 and L_3 edges and the oxygen K edge[18–20]. Quantitative analysis by linear-combination fitting of the oxygen K-edge spectrum shows that the α-Fe_2O_3 phase contributes 25% of the spectrum, whereas SiO_2 contributes 75%. Distinct regions are found in which there is no absorption at the iron L_2 and L_3 edges and strong absorption at the oxygen K edge. The oxygen K edge of these regions is characteristic of the SiO_2 phase[21,22]. No significant contribution of carbon species is observed in the starting material. By fitting the experimental spectrum to a linear combination of the spectra of the pure phases, we construct composite contour maps, as shown in Fig. 2a for SiO_2 and Fe_2O_3.

After a 2-h exposure to H_2 at 350 °C, the material shows significant changes. Figure 2b is a composite contour map of the region in Fig. 2a. Analysis of the iron L_2 and L_3 edges (Fig. 2f) and the oxygen K edge (Fig. 2g) of the two regions indicated in Fig. 2b shows that the Fe_2O_3 is converted into a mixture of iron oxides and metallic iron. The iron L_3 edges show an enhanced intensity at 707.7 eV, which is attributed to the occurrence of Fe^{2+} species in octahedral coordination[18,19]. These observations suggest the formation of Fe_3O_4 and an octahedral Fe^{2+} system, probably Fe_2SiO_4. Because the intensity ratio of the Fe^{2+} peak to the Fe^{3+} peak is higher than would be expected for Fe_3O_4 (ref. 18), the feature is assigned to the occurrence of Fe_2SiO_4. From the spectra alone, it is not possible to distinguish between Fe_2SiO_4 and other octahedral Fe^{2+} oxides, such as FeO. However, it has been reported that a reaction between FeO and the SiO_2 support can cause the loss of iron species into the support during the activation treatment and FTS[23,24]. In addition, under our reaction conditions FeO is expected to disproportionate into Fe_3O_4 and Fe^0 species[25].

Apart from the spectral contributions of the Fe_3O_4 and Fe_2SiO_4 phases, there is also a small contribution observed at 706.8 eV, corresponding to the occurrence of Fe^0 species. The iron L_2- and L_3-edge spectra are analysed quantitatively by least-squares fitting with reference spectra. The calculated phase composition of the selected particle from linear combination fitting is 25% Fe^0, 25% Fe_2SiO_4 and 50% Fe_3O_4 for region one and 33% Fe^0, 50% Fe_2SiO_4 and 17% Fe_3O_4 for region two (these regions are shown in Fig. 2b, c). The pre-edge intensity of the oxygen K edge (Fig. 2g) of both regions decreased significantly, which is due to (1) the decreased overlap between iron and oxygen states with lower valences and (2) the reduction to Fe^0, implying less oxidic iron. The total contribution of iron oxide phases to the oxygen K-edge spectrum decreases from 20% to 10%. As can be seen in Fig. 2a, b, the size and shape of the particle change significantly, indicating morphological changes in the sample as a result of the transition from Fe_2O_3 to Fe_3O_4, Fe_2SiO_4 and Fe^0.

During the subsequent FTS reaction in synthesis gas at 250 °C, the Fe_3O_4 phase is further converted to Fe^0 and Fe_2SiO_4. Figure 2c shows the chemical composite contour map of the selected region after 4 h of FTS. The iron L_3 edge of the selected regions (Fig. 2h) shows an increase in the intensity of the feature at 707.7 eV. The pre-edge in the oxygen spectrum (Fig. 2i) is almost absent in these regions. The spectrum mainly shows a contribution from SiO_2, plus a shoulder at the edge (as indicated by the arrows) indicating strong interaction between small iron particles and the SiO_2 support[26]. Quantitative analysis of the iron L_2 and L_3 edges acquired from region one show a phase composition of 20% Fe^0, 75% Fe_2SiO_4 and 5% Fe_3O_4. Region two has a composition of approximately 40% Fe^0, 50% Fe_2SiO_4 and 10% Fe_3O_4. Analysis of the carbon K edge shows that carbon is preferentially present in iron-rich regions (Fig. 3). Furthermore, regions of higher Fe^0 intensity have carbon K-edge spectra different than those of regions of lower Fe^0 intensity. The high contribution of the absorption band at 285 eV in the carbon K-edge spectrum (Fig. 3a), corresponding to the $1s \rightarrow \pi^*$ transition, indicates the presence of sp^2 hybridized carbon species in the iron-rich region[27]. This observation, in combination with the almost featureless absorption band due to the $1s \rightarrow \sigma^*$ transition at ~295 eV, suggests that the iron phase is converted to iron carbide[28]. In the regions where less iron is

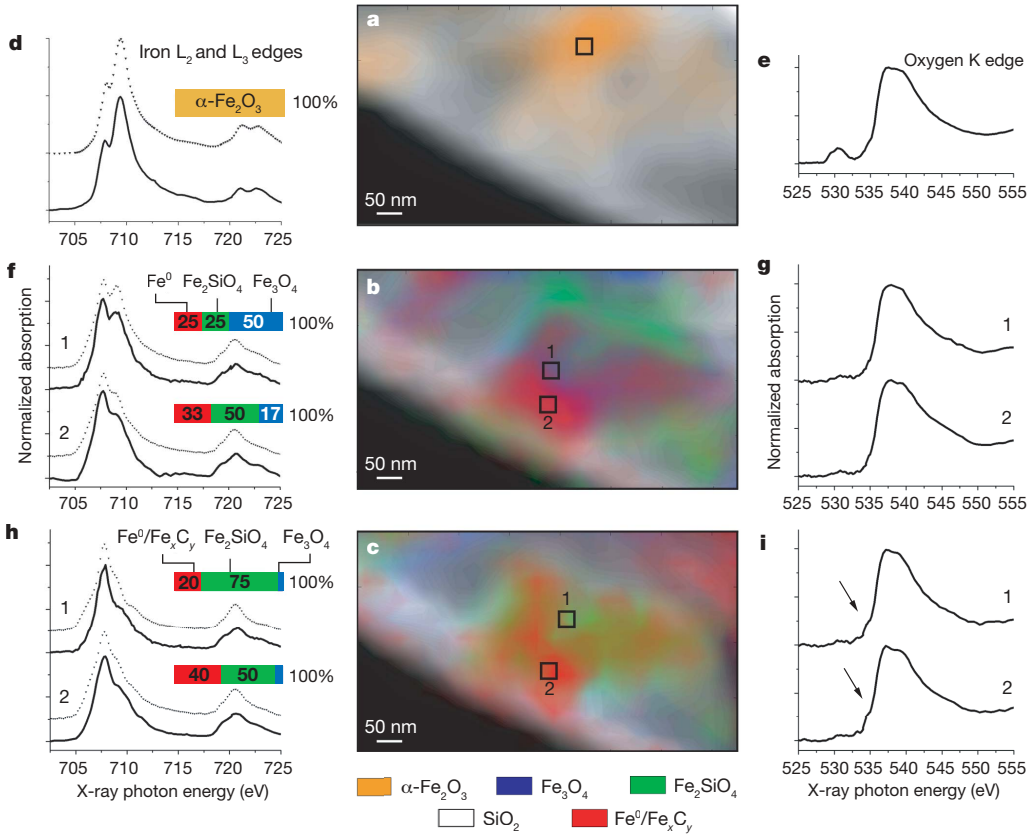


Figure 2 | Chemical contour maps and X-ray absorption spectra of the catalyst material during the different stages of reaction. Chemical contour maps (**a**, **b**, **c**) of a 400 nm × 750 nm region and corresponding iron L_2- and L_3-edge (**d**, **f**, **h**) and oxygen K-edge (**e**, **g**, **i**) X-ray absorption spectra, all normalized to a maximum absorption of one. The contour lines have been removed for clarity. **a**, Before treatment, at room temperature (25 °C) in helium; **b**, after 2 h in H_2 at 350 °C; **c**, after 4 h in synthesis gas at 250 °C. Specific sampling regions and the corresponding X-ray absorption spectra are indicated in the figures. The dotted lines in **d**, **f** and **h** indicate the spectra fitted by a linear combination of reference spectra. The bar graphs represent the calculated relative percentage contributions of the different iron phases at the sampling points. The arrows in **i** indicate the observed shoulder feature.

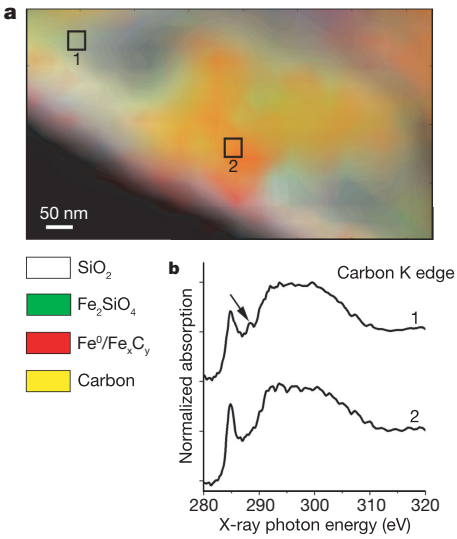


Figure 3 | Location and nature of carbon species on the catalyst material. **a**, Chemical contour map of the 400 nm × 750 nm region of the catalyst material represented in Fig. 2, showing the distribution of carbon species after 4 h in synthesis gas at 250 °C, overlaid on the iron species map. **b**, Carbon K-edge spectra corresponding to the indicated sample regions. The arrow indicates the additional peak at 288 eV.

present, the $1s \rightarrow \pi^*$ absorption peak is lower in intensity and less sharp, suggesting a more sp^3-like character[27] (Fig. 3b), which can be attributed to the FTS reaction products. Furthermore, an additional peak at 288 eV, characteristic of carboxylic species[27], is observed. The presence of reactant carbon species in iron-deficient areas indicates that the support might have some role in the spillover of (hydro)-carbon species from the metal to the support, thereby preventing blocking of the active sites of the catalyst.

These observations demonstrate that STXM can image a catalytic system as complex as an iron-based Fischer–Tropsch catalyst under catalytically relevant reaction conditions, and deliver detailed information on the morphology and composition of the catalyst material. We anticipate that improvements in the optics and imaging methods used with the X-ray microscope will result in a higher achievable spatial resolution[29], and that further development of the detection techniques will improve the count rates to the degree that it might be possible to carry out time-resolved experiments. Moreover, improved designs of dedicated reactor cells may enable studies at higher pressures and temperatures and/or X-ray tomography applications[4,30]. In short, *in situ* STXM is a promising technique that we expect to develop and deliver new insights into many complex chemical problems.

METHODS SUMMARY

Materials. The promoted iron-based $Fe_2O_3/CuO/K_2O/SiO_2$ Fischer–Tropsch catalyst was synthesized following a previously reported procedure[15]. The catalyst material was ground, suspended in ethanol, treated in an ultrasonic bath and

loaded into the nanoreactor by flowing the suspension through the reactor. Remaining ethanol was removed from the reactor by drying at room temperature.

Chemical imaging. All experiments were performed on the interferometrically controlled STXM microscope at beamline 11.0.2 (ref. 13) of the Advanced Light Source at the Lawrence Berkeley National Laboratory, USA. As a compromise between spatial resolution and the longer working distance required, a 35-nm zone plate lens was used with a spatial resolution of ~40 nm. The X-ray absorption spectra and the images were measured with a 35 nm × 35 nm step size of the piezoelectric sample stage.

Data analysis. STXM data files were analysed using the aXis2000 software package (http://unicorn.mcmaster.ca/aXis2000.html), which allowed detailed interactive processing of the images and least-squares linear-combination fitting of the X-ray absorption spectra.

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

Received 29 June; accepted 24 September 2008.

1. Bell, A. T. The impact of nanoscience on heterogeneous catalysis. *Science* **299**, 1688–1691 (2003).
2. Weckhuysen, B. M. Chemistry: Catalysts live and up close. *Nature* **439**, 548 (2006).
3. Hansen, T. W. *et al.* Atomic-resolution in situ transmission electron microscopy of a promoter of a heterogeneous catalyst. *Science* **294**, 1508–1510 (2001).
4. Hitchcock, A. P., Dynes, J. J., Johansson, G., Wang, J. & Botton, G. Comparison of NEXAFS microscopy and TEM-EELS for studies of soft matter. *Micron* **39**, 311–319 (2008).
5. Zambelli, T., Barth, J. V., Wintterlin, J. & Ertl, G. Complex pathways in dissociative adsorption of oxygen on platinum. *Nature* **390**, 495–497 (1997).
6. Frenken, J. & Hendriksen, B. The reactor-STM: A real-space probe for operando nanocatalysis. *MRS Bull.* **32**, 1015–1021 (2007).
7. Creemer, J. F. *et al.* Atomic-scale electron microscopy at ambient pressure. *Ultramicroscopy* **108**, 993–998 (2008).
8. Roeffaers, M. B. J. *et al.* Spatially resolved observation of crystal-face-dependent catalysis by single turnover counting. *Nature* **439**, 572–575 (2006).
9. Zurner, A., Kirstein, J., Doblinger, M., Brauchle, C. & Bein, T. Visualizing single-molecule diffusion in mesoporous materials. *Nature* **450**, 705–708 (2007).
10. Kox, M. H. F., Stavitski, E. & Weckhuysen, B. M. Nonuniform catalytic behavior of zeolite crystals as revealed by in situ optical microspectroscopy. *Angew. Chem. Int. Ed.* **46**, 3652–3655 (2007).
11. Stavitski, E., Kox, M. H. F., Swart, I., de Groot, F. M. F. & Weckhuysen, B. M. In situ synchrotron-based IR microspectroscopy to study catalytic reactions in zeolite crystals. *Angew. Chem. Int. Ed.* **47**, 3543–3547 (2008).
12. Warwick, T. *et al.* A scanning transmission X-ray microscope for materials science spectromicroscopy at the Advanced Light Source. *Rev. Sci. Instrum.* **69**, 2964–2973 (1998).
13. Kilcoyne, A. L. D. *et al.* Interferometer-controlled scanning transmission X-ray microscopes at the Advanced Light Source. *J. Synchrotron Radiat.* **10**, 125–136 (2003).
14. Drake, I. J. *et al.* An in situ cell for characterization of solids by soft X-ray absorption. *Rev. Sci. Instrum.* **75**, 3242–3247 (2004).
15. Dry, M. E. in *Catalysis - Science and Technology* (eds Anderson, J. R. & Boudart, M.) 160–255 (Springer, 1981).
16. Van der Laan, G. P. & Beenackers, A. Kinetics and selectivity of the Fischer-Tropsch synthesis: A literature review. *Catal. Rev. Sci. Eng.* **41**, 255–318 (1999).
17. de Smit, E. & Weckhuysen, B. M. The renaissance of iron-based Fischer-Tropsch synthesis: On the multifaceted catalyst deactivation behaviour. *Chem. Soc. Rev.* doi:10.1039/B805427D (in the press).
18. Heijboer, W. M. *et al.* In-situ soft X-ray absorption of over-exchanged Fe/ZSM5. *J. Phys. Chem. B* **107**, 13069–13075 (2003).
19. de Groot, F. & Kotani, A. *Core Level Spectroscopy of Solids* (Taylor & Francis, 2008).
20. de Groot, F. M. F. *et al.* Oxygen 1s x-ray-absorption edges of transition-metal oxides. *Phys. Rev. B* **40**, 5715–5723 (1989).
21. Sharp, T. *et al.* Distinction between six- and fourfold coordinated silicon in SiO_2 polymorphs via electron loss near edge structure (ELNES) spectroscopy. *Phys. Chem. Miner.* **23**, 17–24 (1996).
22. Davoli, I. *et al.* Structure of densified vitreous silica: Silicon and oxygen XANES spectra and multiple scattering calculations. *Phys. Chem. Miner.* **19**, 171–175 (1992).
23. Wielers, A. F. H., Kock, A. J. H. M., Hop, C. E. C. A., Geus, J. W. & van der Kraan, A. M. The reduction behavior of silica-supported and alumina-supported iron catalysts: A Mössbauer and infrared spectroscopic study. *J. Catal.* **117**, 1–18 (1989).
24. Zhang, C. H., Wan, H. J., Yang, Y., Xiang, H. W. & Li, Y. W. Study on the iron-silica interaction of a co-precipitated Fe/SiO_2 Fischer-Tropsch synthesis catalyst. *Catal. Commun.* **7**, 733–738 (2006).
25. Jozwiak, W. K., Kaczmarek, E., Maniecki, T. P., Ignaczak, W. & Maniukiewicz, W. Reduction behavior of iron oxides in hydrogen and carbon monoxide atmospheres. *Appl. Catal. A Gen.* **326**, 17–27 (2007).
26. Wang, F., Malac, M. & Egerton, R. F. Alternative methods of identifying the oxidation of metallic nanoparticles embedded in a matrix. *Micron* **38**, 371–376 (2007).
27. Braun, A. *et al.* Advantages of soft X-ray absorption over TEM-EELS for solid carbon studies - a comparative study on diesel soot with EELS and NEXAFS. *Carbon* **43**, 117–124 (2005).
28. Urbonaite, S. *et al.* EELS studies of carbide derived carbons. *Carbon* **45**, 2047–2053 (2007).
29. Thibault, P. *et al.* High-resolution scanning X-ray diffraction microscopy. *Science* **321**, 379–382 (2008).
30. Le Gros, M. A., McDermott, G. & Larabell, C. A. X-ray tomography of whole cells. *Curr. Opin. Struct. Biol.* **15**, 593–600 (2005).

Acknowledgements We acknowledge financial support for this research work from the Dutch National Science Foundation in the form of two VICI grants (to F.M.F.d.G. and B.M.W.), a grant from the Netherlands Research School Combination on Catalysis (to B.M.W. and F.M.F.d.G.) and a grant from Shell Global Solutions (to B.M.W.). The Advanced Light Source is supported by the Director, Office of Science, Office of Basic Energy Sciences, US Department of Energy. The nanoreactors were fabricated with the assistance of the DIMES ICP-group and the Nanofacility of TU Delft.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.M.F.d.G. (f.m.f.degroot@uu.nl) or B.M.W. (b.m.weckhuysen@uu.nl).

METHODS

Catalyst material preparation. The $Fe_2O_3/CuO/K_2O/SiO_2$ catalyst was prepared by precipitation from a ferric nitrate solution using a basic sodium carbonate solution[15]. We dissolved 25 g $Fe(NO_3)_3 \cdot 9H_2O$ (98+% ACS reagent, Acros) and 1.2 g $Cu(NO_3)_2 \cdot 3H_2O$ (p.a. 99.5%, Merck) in 100 ml distilled water. The solution was heated to near its boiling point, after which it was slowly (over ~1 min) added to a vigorously stirred, near-boiling solution of 25 g Na_2CO_3 in 100 ml distilled water. As sodium has been reported to negatively influence the catalytic performance of the catalysts, the resulting precipitate was filtered and reslurried in about 1 l of near-boiling distilled water to remove any residual sodium. This process was repeated about four times, until the pH of the solution was about 7.

The washed precipitate was reslurried in 200 ml distilled water. About 8 g of potassium water-glass solution ($K_2O:SiO_2$ (1:2.15), Akzo-PQ) was added to the slurry under vigorous stirring and 1.5 ml of concentrated HNO_3 was added to precipitate the SiO_2 and lower the total potassium content. After this, the precipitate was dried for ~6 h at 60 °C and subsequently for 24 h at 120 °C. Finally, the material was calcined in a flow of air at 300 °C for 5 h using a heating ramp of 5 °C.

The final catalyst compositions were confirmed by X-ray fluorescence analysis using a Goffin Meyvis Spectro X-lab 2000. The relative molar composition of the catalyst was Fe:Cu:K:Si 100:7.5:5.9:15.6.

Nanoreactor and adaptor design. The design and assembly of the nanoreactor itself is described in more detail elsewhere[7]. The nanoreactor adaptor was fabricated from a solid slab of brass. Holes accommodating the 1/16-inch gas tubing were drilled into the brass slab and connected to the nanoreactors from their rear side. Each nanoreactor sits on top of a gas channel inlet and outlet, which were sealed using Viton O rings. The nanoreactors are held in place by a cover plate. Electronic connections, four for each reactor, were connected to the nanoreactor by means of the wire bonding method, using aluminium wires.

STXM data analysis. Chemical contour maps and the relative iron phase compositions were acquired by using the aXis2000 software package (http://unicorn.mcmaster.ca/aXis2000.html). The software uses linear regression to fit the absorbance (A) spectrum of each ($35 \times 35\ nm^2$) pixel to a linear combination of reference spectra. Absorbance is defined as $A = -\log_{10}(I/I_0)$, where I and I_0 respectively denote the transmitted and incident intensities. The fit procedure minimizes $(A^{measured}_{x,y,E} - A^{calculated}_{x,y,E})^2$ over an energy range, where $A^{measured}_{x,y,E}$ is the absorbance measured at point (x, y) and energy E and $A^{calculated}_{x,y,E}$ is defined as

$$A^{calculated}_{x,y,E} = \mathrm{const}_E + a_{1,E}A_{1,E} + a_{2,E}A_{2,E} + \cdots + a_{i,E}A_{i,E}$$

Here const_E is a constant, $A_{i,E}$ is the absorbance of compound i at energy E and $a_{i,E}$ is the expansion coefficient of compound i at that energy. The coefficients are varied to fit the measured spectrum.

LETTERS

Transient nature of late Pleistocene climate variability

Thomas J. Crowley[1] & William T. Hyde[2]

Climate in the early Pleistocene[1] varied with a period of 41 kyr and was related to variations in Earth's obliquity. About 900 kyr ago, variability increased and oscillated primarily at a period of ~100 kyr, suggesting that the link was then with the eccentricity of Earth's orbit. This transition has often[2–5] been attributed to a nonlinear response to small changes in external boundary conditions. Here we propose that increasing variablility within the past million years may indicate that the climate system was approaching a second climate bifurcation point, after which it would transition again to a new stable state characterized by permanent mid-latitude Northern Hemisphere glaciation. From this perspective the past million years can be viewed as a transient interval in the evolution of Earth's climate. We support our hypothesis using a coupled energy-balance/ice-sheet model, which furthermore predicts that the future transition would involve a large expansion of the Eurasian ice sheet. The process responsible for the abrupt change seems to be the albedo discontinuity at the snow–ice edge. The best-fit model run, which explains almost 60% of the variance in global ice volume[6] during the past 400 kyr, predicts a rapid transition in the geologically near future to the proposed glacial state. Should it be attained, this state would be more 'symmetric' than the present climate, with comparable areas of ice/sea-ice cover in each hemisphere, and would represent the culmination of 50 million years of evolution from bipolar nonglacial climates to bipolar glacial climates.

A recent composite[1] of benthic $\delta^{18}O_{cal}$ (that is, calcite) records summarizes the general pattern of climate back to 3.0 Myr ago (Fig. 1a). A 500-kyr moving boxcar window (see ref. 5) highlights long-term changes in $\delta^{18}O_{cal}$ variability (Fig. 1b). This window length was chosen to mute glacial–interglacial variations and any possible influence of the 413-kyr eccentricity cycle. There is an increase in variability after the Pliocene/Pleistocene boundary (~1.8 Myr ago), with clear evidence for steps ~665 kyr and ~885 kyr ago. The trend is similar (Fig. 1b) to strontium isotope changes[7] that are sometimes interpreted to reflect variations in continental weathering.

Histograms (Fig. 1c) of equal-sized, 450-kyr sections of the latest Pleistocene indicate continued evolution and bimodality, with an interglacial mode being more common in the earlier section. Despite the presence of more extreme interglacial states in the later section (Fig. 1a), its mean $\delta^{18}O_{cal}$ value is isotopically heavier (about 0.1‰, equivalent to ~10 m in sea-level change), and the glacial mode is more common. Two of the four most extreme $\delta^{18}O_{cal}$ maxima occurred during the last two glacial periods (Fig. 1a). The above changes, plus a smaller step ~1.12 Myr ago, correspond to Pleistocene 'megacycle' boundaries identified[8] in major European glacial advances and European and Chinese loess sequences.

Trends in variability may also apply on millennial scales. A core[9] from the subpolar North Atlantic (Fig. 1d) suggests generally enhanced variability from glacial stage 12 (~430 kyr ago) to stages 2–4 of the last glacial cycle. However, this interpretation requires further testing.

Because prior energy-balance modelling (EBM) work (see fig. 4 of ref. 10) suggests that enhanced variability may indicate a system near a climate bifurcation point, one interpretation of the observed Pleistocene trend is that the system was poised to flip to a different climate state. We test this hypothesis by employing, without significant modification, a published version of an energy-balance/ice-sheet model[11]. The coupled model consists of four submodels that predict ice flow, mass balance, temperature and bedrock sinking. Ice is assumed to flow subject to a temperature-independent rheology. Ice sheets are driven by a uniform precipitation of 0.6 m yr^{-1} that decreases with height and temperature. This value is characteristic of current Northern Hemisphere mid-latitude land areas. Precipitation is computed according to statistical models, which themselves take as input monthly temperatures from a nonlinear, two-dimensional (latitude–longitude grid), diffusive seasonal EBM[12]. Bedrock sinking is assumed to occur with a time constant of 4 kyr. The ice-sheet model has a resolution of $0.5° \times 0.5°$. The time step for the EBM is 40 yr, whereas for the ice-sheet model it is never more than 20 yr and, during times of rapid change, is as small as 2.5 yr. The coupled model has been previously applied to topics ranging from millennial variability to Neoproterozoic glaciation[13,14].

The coupled model is forced with orbital insolation variations as derived in ref. 15. To isolate as much as possible the inherent response of the ice sheet, we do not consider glacial–interglacial changes in CO_2 or North Atlantic heat sources.

Plotting operating curves (Fig. 2a) is a standard method of depicting model behaviour, with solar constant changes used as a proxy for other changes in boundary conditions. As radiative forcing decreases, there is an approximately linear increase in ice volume until the model reaches a critical point. Over an interval of 1.4 W m^{-2} (less than the ice age change in radiative forcing from CO_2), variability increases substantially. The solution then stabilizes on a second branch, after which ice volume again changes nearly linearly with forcing. Simulated ice sheets (Fig. 3a–c) indicate that the North American and Eurasian ice sheets are broadly similar in location to, but only about half the volume of, those observed for the last glacial maximum. Glacial–interglacial CO_2 variations are a plausible explanation for the principal difference, with the lack of a variable North Atlantic heat source perhaps contributing additionally to Eurasian ice-volume changes.

As suggested by theory, variability reaches a maximum in the transition zone (Fig. 2b) and is virtually the same for the two stable states, despite their having very different ice-volume levels. In the middle of the rapid growth phase of the North American ice sheet, a very large Eurasian ice sheet (Fig. 2a) develops from a perturbation of only 0.07 W m^{-2} (three times smaller than the radiative forcing from the 11-yr solar cycle). This ice sheet (Fig. 3d) is more than twice the

[1]School of Geosciences, The University of Edinburgh, Edinburgh EH9 3JW, UK. [2]Department of Physics, The University of Toronto, Toronto, Ontario M5S 1A7, Canada.

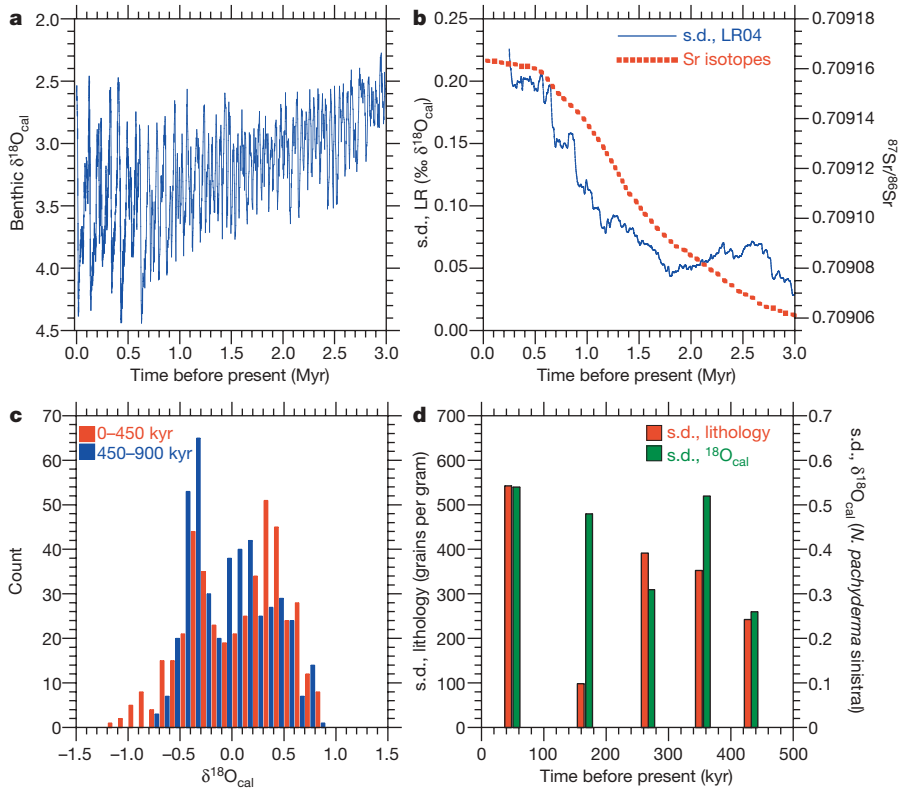


Figure 1 | Evidence for trends in climate variability in the Plio-Pleistocene. **a**, Plio-Pleistocene composite $\delta^{18}O_{cal}$ record from ref. 1; more positive values indicate colder climates. ($\delta^{18}O = (^{18}O/^{16}O)_{sample}/(^{18}O/^{16}O)_{standard} - 1$ is a measure of the difference in a sample, with respect to a standard, in the stable isotope ratios of ^{18}O to ^{16}O, usually expressed in parts per thousand.) **b**, 500-kyr moving-boxcar-windowed standard deviation of the record in **a** (termed LR in label) compared with the $^{87}Sr/^{86}Sr$ isotope record from ref. 7; the trend in $\delta^{18}O_{cal}$ variability is relatively robust to changes in window width. **c**, Histograms of $\delta^{18}O_{cal}$ composite (with mean removed) for the past 900 kyr (see text), divided into two equal subsamples; more negative values are interpreted as having less ice volume. **d**, Standard deviations of two indices of variability within the five glacial stages of the last 0.5 Myr in densely sampled parts of a subpolar North Atlantic core[9], showing a general increase in within-stage variations (towards the present) of lithic (ice-rafted) grains and the $\delta^{18}O_{cal}$ of cold-water planktonic foraminifer *Neogloboquadrina pachyderma* sinistral (left-coiling). This general 'increase towards the present' is similar to that found in orbital-scale variations (see text).

volume of the modelled North American ice sheet. Because our ice-sheet model does not allow for multi-doming, we probably overestimate[16] its volume by ~20%. Nevertheless, the general success of model simulations for the Pleistocene suggests that the prediction of a very large Eurasian ice sheet is not implausible; such a response has already been found in another model[17].

To determine whether patterns in the operating curves are within the range of inferred Plio-Pleistocene changes, we simulate the last 3.0 Myr using only orbital forcing plus long-term CO_2 changes. Although the strontium isotope curve (Fig. 1b) suggests a sigmoidal variation in weathering (and perhaps CO_2) since 3.0 Myr ago, for the sake of simplicity we mainly use linear ramp forcing (Fig. 4a; in one run, CO_2 stabilized at 240 p.p.m. 1.0 Myr ago). The initial Pliocene CO_2 level is constrained by leaf stomata data[18] at about 360 p.p.m., and the level at 0 Myr is set at the mean CO_2 level of the last 20 kyr (~240 p.p.m.). Because the original CO_2 reference level of the coupled ice-sheet model was 360 p.p.m., we adjust the outgoing radiative forcing term in the energy balance model to account for the anthropogenic effect. As the model may have some biases, relative changes in radiative forcing are probably a more reliable indicator of model behaviour than the absolute value of CO_2 at any particular time.

There are a number of interesting features of the model response to one small (~1.0–2.0 W m^{-2}), linear change in boundary conditions (Fig. 4). All runs extended beyond the present day (Fig. 4b) transition to the postulated second equilibrium climate state within a strikingly narrow mean CO_2 value of about 240 ± 5 p.p.m. The run with the smallest change (295–240 p.p.m.) has the highest correlation ($r = 0.70$) with the Plio-Pleistocene composite of ref. 1 (see Fig. 1a). Separation of the Northern Hemisphere trend into individual ice-sheet responses (Fig. 4c) indicates that, as in the operating curves, the North American ice sheet first crossed the bifurcation threshold

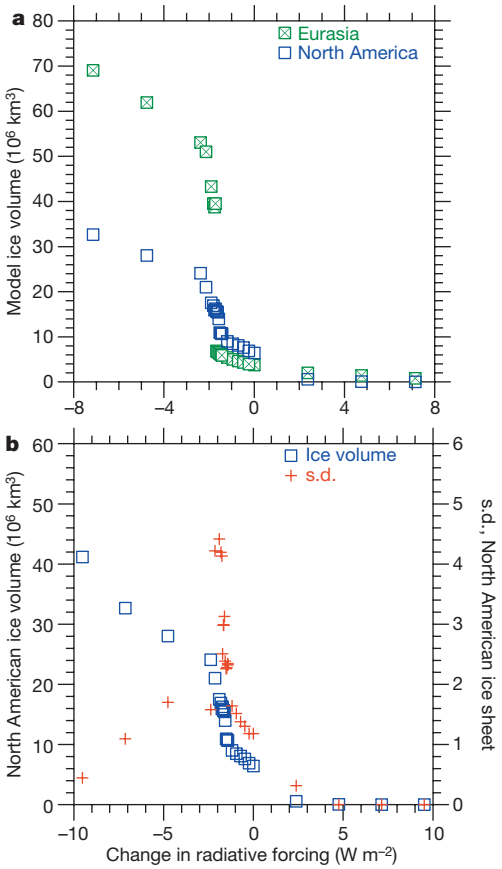


Figure 2 | Operating curves for climate/ice-sheet model as a function of the solar constant (abscissa). The standard value of the solar constant for the present is 1.0 (average for Earth, ~340 W m^{-2}). These units can be translated into equivalent CO_2 changes: a 1.5% change in the solar constant is approximately equivalent to a doubling of CO_2 (~3.7 W m^{-2}), after factoring in the 30% albedo of the Earth. **a**, Ice-volume changes for North America and Eurasia; (b) North American ice volume versus standard deviation for the 150–200-kyr interval of each of 36 model runs. Note that variability increases sharply near the bifurcation.

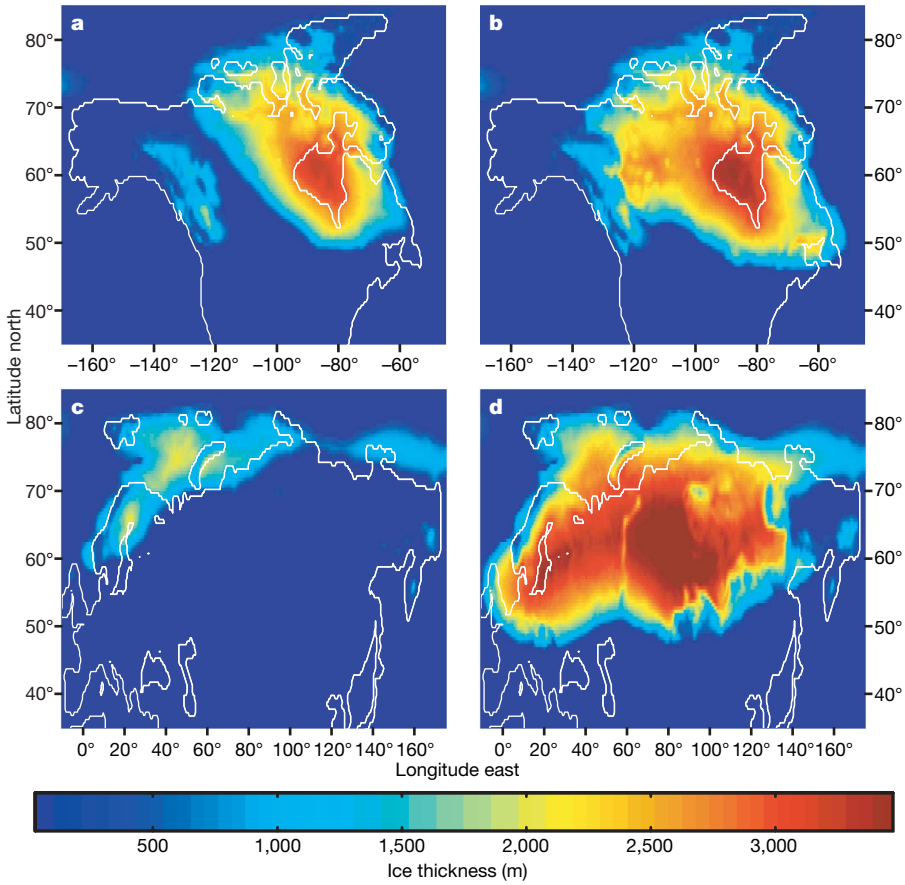


Figure 3 | Simulated ice extent for different stages of North American and Eurasian ice-sheet growth. **a**, Small North American ice sheet; **b**, large North American ice sheet; **c**, small Eurasian ice sheet; **d**, large Eurasian ice sheet that is the postulated second stable state for Northern Hemisphere glaciation.

~1.14 Myr ago, during a low-eccentricity interval. This coincides almost exactly with the beginning of one of the small steps (see above) in the variability curve (Fig. 1b). These responses are consistent with one explanation[12] for glacial inception: because North America is smaller than Eurasia, it will not warm up in the summer as much as Eurasia and, therefore, will respond earlier than Eurasia at a glaciation threshold.

In the last 0.6–0.7 Myr (Fig. 5a) of the best-fit run, lower frequency oscillations occur with ~100-kyr periods (as in observations). Although low-frequency oscillations are present in other runs, the flatter CO_2 ramp for the 295–240-p.p.m. run allows more time for ~100-kyr variability before transitioning to the large Eurasian ice sheet. Deglaciations are similar to 'terminations' in the marine $\delta^{18}O_{cal}$ record. In our model, rapid melting can be attributed to continued southward flow of the ice sheet into the isostatic depression it created[19]. An initial attempt to superimpose observed CO_2 changes[20] during the last 650 kyr triggered the Eurasian ice-sheet instability, most probably because of the model constraints addressed above.

There is agreement in both the time ($r = 0.77$) and the frequency domains (Fig. 5) of the scaled best-fit run with the $\delta^{18}O_{sw}$ best-guess[6] seawater reconstruction of global ice volume, with both records having substantially less ~100-kyr-ago power than the $\delta^{18}O_{cal}$ record (Fig. 5b). This difference reflects the deep-water temperature overprint on observed $\delta^{18}O_{sw}$. Model variability in the last 410 kyr is about 25% greater than in the preceding segment (Fig. 5a) and is primarily manifested in the ~23-kyr precession and 41-kyr obliquity bands (Fig. 5b). Model ~100-kyr-ago power is greater than a 'background' level from earlier in the record and is most strongly present in the North American ice sheet (Supplementary Fig. 1), first emerging in the 820–410-kyr segment and then increasing in the Eurasian ice sheet over the last 410-kyr segment. Again, it is necessary to recall that some frequency-domain differences between model and observations may be partly due to lack of glacial–interglacial CO_2 fluctuations in our runs, plus no attempt to simulate variable North Atlantic heat sources. Regardless, our results clearly indicate that ~100-kyr-ago power is embedded in the physics of the coupled climate/ice-sheet system and is not just a consequence of, for example, glacial–interglacial CO_2 variations[6].

For the best-fit run, transition to the large Eurasian ice sheet occurs shortly after the present (Fig. 5a). Our results therefore suggest that the actual climate system may have been geologically close (10^4–10^5 yr) to the final phase of a 50-Myr evolution from bipolar warm climates to permanent bipolar glaciation. (Presumably, future society could prevent this transition indefinitely with very modest adjustments to the atmospheric CO_2 level.)

Because our model has no ocean dynamics, rapid change requires a different explanation than the oft-discussed thermohaline instability. In our simulations, the instability resembles the well-known[21] albedo discontinuity at the snow–ice boundary. To persist, an ice cap must be stable against warming. This stability is conferred by the balance between the energy sink provided by the ice cap and the energy source (transport) onto the ice cap from warmer regions. The energy sink is a function of the ice cap's higher albedo (ameliorated by reduced infrared output from the higher, colder surface) and the size of the ice cap.

For radiative adjustment and transport values typical of the Earth, it has been analytically determined[22] that at a bifurcation point snow cover expands on a scale of ~1500–2000 km. This length scale can be found in mid/high-latitude instrumental records[23], some modelling studies[24–26] and the latitudinal shift associated with the large Eurasian ice sheet (see Fig. 3c, d). Our model also displays hysteresis behaviour characteristic of such systems[27].

In addition to the tests outlined in this paper, more complex ice-sheet models could be used. Even though these models are more 'realistic', we are not convinced that they would show a fundamentally different result. There is already a huge difference in complexity between the original EBM[21] and the model we use, yet to first order the bifurcation diagrams are the same. We hypothesize that the future behaviour of more complicated coupled models will be influenced as much by the albedo discontinuity as by any added features of ice-sheet complexity.

Observational tests involve the ~200 kyr before 400 kyr ago (Fig. 5a). If our agreement with the $\delta^{18}O_{sw}$ record is not a coincidence, then we predict that sea level/ice volume differences from present during the precession peaks ~480 and 500 kyr ago would be significantly less (on the order of 0 ± 10 m) than the ~20–25-m sea level drop suggested by the $\delta^{18}O_{cal}$ record (Fig. 1a). Because the $\delta^{18}O$ of atmospheric oxygen ($\delta^{18}O_{atm}$) is a good approximation to the $\delta^{18}O_{sw}$ estimate ($r = 0.92$), publication of the extended Dome C (Antarctica) $\delta^{18}O_{atm}$ record will be one test of our hypothesis.

Older geological intervals can also be used to test whether the entire transition to an alternative stable state may have occurred

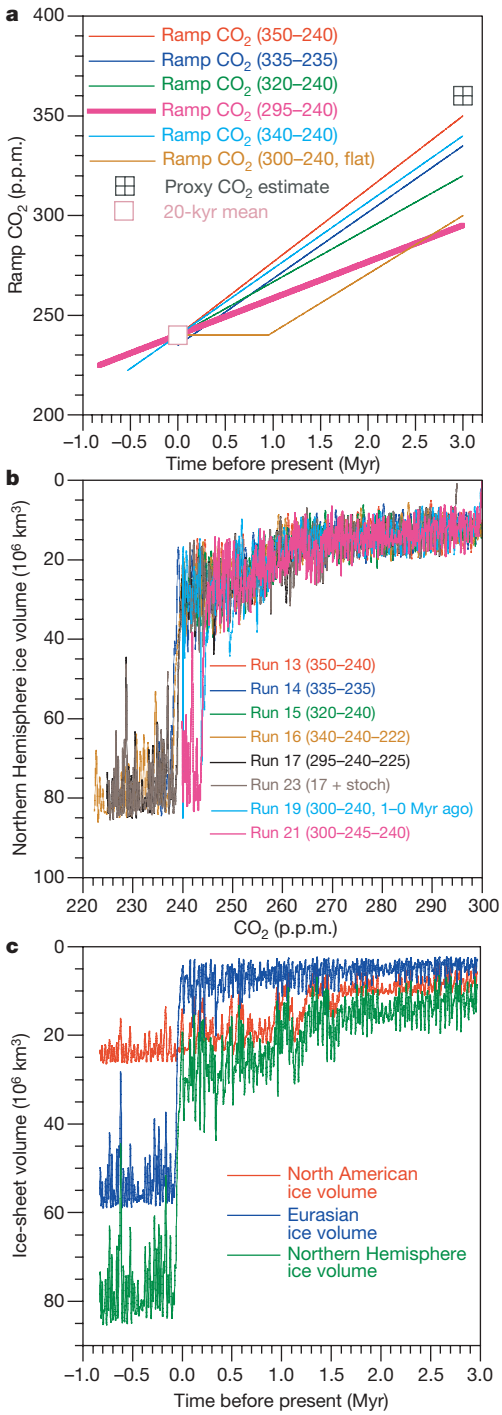


Figure 4 | Principal results of the ramp model runs. a, Plio-Pleistocene CO_2 scenarios for runs (see text); **b**, plots of Northern Hemisphere ice volume as a function of CO_2 for the different ramp scenarios. Labels refer to CO_2 values (p.p.m.) used for each run. In each label, the first value is the starting point, 3.0 Myr ago, and the second value is the end point, 0.0 Myr ago (except for run 19, in which the ramp stopped 1.0 Myr ago and constant CO_2 levels were used for the remaining 1.0 Myr). Where there is a third value, it refers to the future projection using the trend line for the corresponding run. In run 23, the best-fit run (run 17) was modified with a stochastic forcing of ~0.4 W m^{-2} to mimic decadal-scale oscillations. We note that, regardless of forcing scenario, runs generally transition to the large Eurasian ice sheet at a background CO_2 level of ~240 p.p.m. **c**, Details of the best-fit run (see Fig. 5) with respect to total ice volume and contribution from each ice sheet (see Supplementary Data).

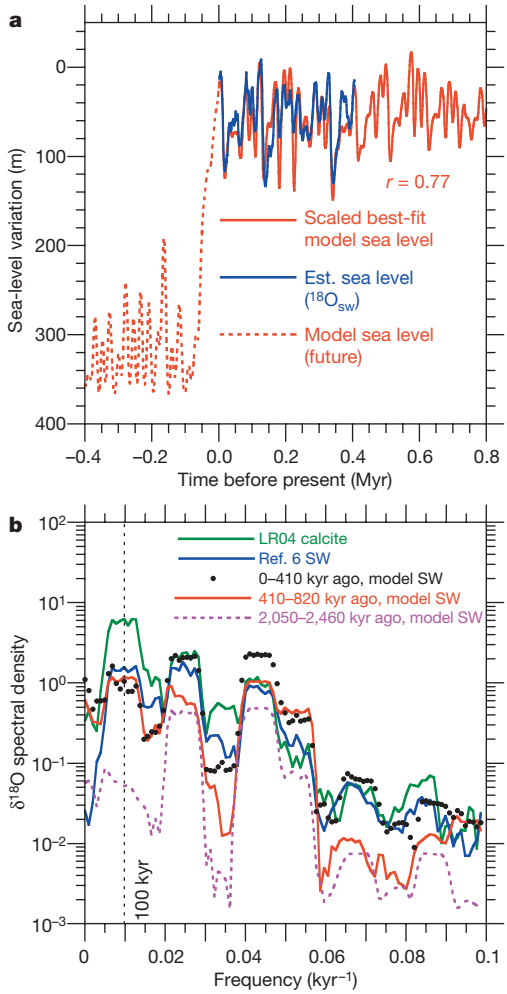


Figure 5 | Further analyses of best-fit run. a, Comparison with best-guess seawater ($\delta^{18}O_{sw}$) reconstruction of ref. 6, and projection into the future. **b**, Multi-taper-method frequency-domain comparison of the best-fit model with the results of ref. 6 and the $\delta^{18}O_{cal}$ records, and two earlier intervals of the best-fit model run. The ice-volume record for the best-fit run was converted to $\delta^{18}O_{sw}$ by setting the last interglacial best-guess value to −0.04‰ and matching the model to the preceding maximum in $\delta^{18}O_{sw}$ (see Supplementary Data). The mean value of the $\delta^{18}O_{sw}$ record of ref. 6 was also adjusted to set the last interglacial value to −0.04‰. All analyses were done for identical intervals of 410 kyr. Solid lines indicate a run was significant at the 95% level for 100-kyr-ago power (higher frequency intervals at 23 kyr and 41 kyr were always significant).

before. The advance of the Antarctic ice sheet to the edge of the continent near the Eocene/Oligocene boundary (~33.6 Myr ago) represents one possible case[28]. A composite benthic $\delta^{18}O_{cal}$ reconstruction[29] indicates that $\delta^{18}O_{cal}$ standard deviations of the ~400-kyr time intervals before and after the transition are nearly ~0.14‰, whereas during the transition (see Supplementary Fig. 2) the standard deviation is 80% higher (20% detrended). This response is consistent with our interpretation of multiple stable states. There were also numerous late Cenozoic oscillations of a wet-based East Antarctic ice sheet[30] before it transitioned to a cold-based glacier ~14 Myr ago, following which at least parts of it have apparently remained stable.

Finally, our hypothesis may be relevant to interpretation of millennial-scale variability in the late Pleistocene. The tentative link in Fig. 1b, d between orbital- and millennial-scale variability suggests that the many studies and interpretations of the latter may be missing a fundamental perspective, unless they are interpreted within the framework of the overall higher and (we suggest) transient variability on longer timescales.

Received 21 March; accepted 28 August 2008.

1. Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records. *Paleoceanography* **20**, doi:10.1029/2004PA001071 (2005).
2. Maasch, K. & Saltzman, B. A low-order dynamical model of global climatic variability over the full Pleistocene. *J. Geophys. Res.* **95**, 1955–1963 (1990).
3. Berger, A., Li, X. S. & Loutre, M.-F. Modeling northern hemisphere ice volume over the last three million years. *Quat. Sci. Rev.* **18**, 1–11 (1999).
4. Huybers, P. Glacial variability over the last two million years: An extended depth-derived age model, continuous obliquity forcing, and the Pleistocene progression. *Quat. Sci. Rev.* **26**, 37–55 (2007).
5. Lisiecki, L. & Raymo, M. E. Plio-Pleistocene climate evolution: Trends and transitions in glacial cycle dynamics. *Quat. Sci. Rev.* **26**, 56–69 (2007).
6. Shackleton, N. J. The 100,000 year ice-age cycle identified and found to lag temperature, carbon dioxide, and orbital eccentricity. *Science* **289**, 1897–1902 (2000).
7. McArthur, J. M., Howarth, R. J. & Bailey, T. R. Strontium isotope stratigraphy: LOWESS version 3: Best fit to the marine Sr-isotope curve for 0–509 Ma and accompanying look-up table for deriving numerical age. *J. Geol.* **109**, 155–170 (2001).
8. Kukla, G. & Cilek, V. Plio-Pleistocene megacycles: Record of climate and tectonics. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **120**, 171–194 (1996).
9. McManus, J. F., Oppo, D. W. & Cullen, J. L. A 0.5 million year record of millennial scale climate variability in the North Atlantic. *Science* **283**, 971–975 (1999).
10. Crowley, T. J. & North, G. R. Abrupt climate change and extinction events in Earth history. *Science* **240**, 996–1002 (1988).
11. Tarasov, L. & Peltier, W. R. Terminating the 100 kyr ice age cycle. *J. Geophys. Res.* **18**, 21665–21693 (1997).
12. North, G. R., Mengel, J. G. & Short, D. A. Simple energy balance model resolving the seasons and the continents: Application to the astronomical theory of glaciation. *J. Geophys. Res.* **88**, 6576–6586 (1983).
13. Hyde, W. T. & Crowley, T. J. Stochastic forcing of Pleistocene ice sheets: Implications for the origin of millennial-scale climate oscillations. *Paleoceanography* 17, 10.1029/2001PA000669 (2002).
14. Hyde, W. T., Crowley, T. J., Baum, S. K. & Peltier, W. R. Neoproterozoic 'Snowball Earth' simulations with a coupled climate/ice sheet model. *Nature* **405**, 425–429 (2000).
15. Berger, A. Long term variations of daily insolation and Quaternary climate changes. *J. Atmos. Sci.* **35**, 2362–2367 (1978).
16. Crowley, T. J. & Baum, S. K. Estimating Carboniferous sea-level fluctuations from Gondwanan ice extent. *Geology* **19**, 975–977 (1991).
17. Calov, R. & Ganopolski, A. Multistability and hysteresis in the climate-cryosphere system under orbital forcing. *Geophys. Res. Lett.* **32**, doi:10.1029/2005GL024518 (2005).
18. van der Burgh, J., Visscher, H., Dilcher, D. L. & Kurschner, W. M. Paleoatmospheric signatures in Neogene fossil leaves. *Science* **260**, 1788–1790 (1993).
19. Hyde, W. T. & Peltier, W. R. Sensitivity experiments with a model of the ice age cycle: The response to harmonic forcing. *J. Atmos. Sci.* **42**, 2170–2188 (1985).
20. Lüthi, D. *et al.* High-resolution carbon dioxide concentration record 650,000–800,000 years before present. *Nature* **453**, 379–382 (2008).
21. Budyko, M. I. The effect of solar radiation changes on the climate of the Earth. *Tellus* **21**, 611–619 (1969).
22. North, G. R. The small ice cap instability in diffusive climate models. *J. Atmos. Sci.* **41**, 3390–3395 (1984).
23. Hansen, J. E. & Lebedeff, S. Global trends of measured surface air temperature. *J. Geophys. Res.* **92**, 13345–13372 (1987).
24. Manabe, S. & Broccoli, A. J. The influence of continental ice sheets on the climate of an ice age. *J. Geophys. Res.* **90**, 2167–2190 (1985).
25. Crowley, T. J., Yip, K.-Y. J. & Baum, S. K. Snowline instability in a general circulation model: Application to Carboniferous glaciation. *Clim. Dyn.* **10**, 363–374 (1994).
26. Baum, S. K. & Crowley, T. J. The snow/ice instability as a mechanism for rapid climate change: A Neoproterozoic Snowball Earth model example. *Geophys. Res. Lett.* **30**, doi:10.1029/2003GL017333 (2003).
27. Crowley, T. J., Hyde, W. T. & Baum, S. K. CO_2 levels required for deglaciation of a "Near-Snowball" Earth. *Geophys. Res. Lett.* **28**, 283–286 (2001).
28. DeConto, R. M. & Pollard, D. Rapid Cenozoic glaciation of Antarctica induced by declining atmospheric CO_2. *Nature* **421**, 245–249 (2003).
29. Zachos, J., Pagani, M., Sloan, L., Thomas, E. & Billups, K. Trends, rhythms, and aberrations in global climate, 65 Ma to Present. *Science* **292**, 686–693 (2001).
30. Jamieson, S. S. R. & Sugden, D. E. in *Antarctica: A Keystone in a Changing World* (eds Cooper, A. K. & Raymond, C.) 39–54 (Proc. 10th Intl Symp. Antarctic Earth Sci., National Academies, 2008).

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

Acknowledgements This research received partial support the Scottish Alliance for Geoscience, the Environment, and Society (SAGES) and from the US National Science Foundation. We thank G. North for numerous discussions over the years and S. Obrochta for valuable assistance.

Author Contributions W.T.H. had primary responsibility for model simulations and T.J.C. for analysis of model output, comparison with geological data and write-up of results.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.J.C. (thomas.crowley@ed.ac.uk).

LETTERS

Radiative conductivity in the Earth's lower mantle

Alexander F. Goncharov[1], Benjamin D. Haugen[1,2], Viktor V. Struzhkin[1], Pierre Beck[1] & Steven D. Jacobsen[3]

Iron in crustal and mantle minerals adopts several possible oxidation states: this has implications for biogeochemical processes[1], oxygenation of the atmosphere[2] and the oxidation state of the mantle[3,4]. In the deep Earth, iron in silicate perovskite, $(Mg_{0.9}Fe_{0.1})SiO_3$, and ferropericlase, $(Mg_{0.85}Fe_{0.15})O$, influences the thermal conductivity of the lower mantle and therefore heat flux from the core. Little is known, however, about the effect of iron oxidation states on transport properties. Here we show that the radiative component of thermal conductivity in the dominant silicate perovskite material of Earth's lower mantle is controlled by the amount of ferric iron, Fe^{3+}. We obtained the optical absorption spectra of silicate perovskite and ferropericlase at pressures up to 133 GPa, corresponding to pressures at the core–mantle boundary. Absorption spectra of ferropericlase up to 800 K and 60 GPa exhibit minimal temperature dependence. The results on silicate perovskite show that optical absorption in the visible and near-infrared spectral range is dominated by O–Fe^{3+} charge transfer and Fe^{3+}–Fe^{2+} intervalence transitions, whereas a contribution from the Fe^{2+} crystal-field transitions is substantially smaller. The estimated pressure-dependent radiative conductivity, k_{rad}, from these data is 2–5 times lower than previously inferred from model extrapolations, with implications for the evolution of the mantle, such as generation and stability of thermo-chemical plumes in the lower mantle.

The lower mantle extends from the 660-km seismic discontinuity to the core–mantle boundary at 2,900 km depth and constitutes roughly half of the Earth's mass. Dominated by just two silicate and oxide minerals—silicate perovskite and ferropericlase—the lower mantle controls the flow of latent and radiogenic heat out of the core, which contributes an estimated 25–30% of the total heat flux measured at the surface[5]. Thermal conductivity of the lower-mantle mineral assemblage therefore provides a constraint on the ability of the mantle to transfer heat non-convectively, but also influences convection[6]. The importance of temperature- and pressure-dependent thermal conductivity[7,8] has been recognized in recent geodynamic models of mantle convection, which incorporate variable thermal conductivity[9] and variable viscosity[6,10].

Thermal conductivity in metals is dominated by electron transport, whereas heat conduction in insulators (dielectrics) such as silicates and oxides is essentially dominated by vibrational (phonon) transport. At high temperature, thermal conductivity of mantle materials has a competing second contribution—radiative conductivity[7,11,12]. The radiative component of total conduction in defect-free material is determined by the material's optical properties, namely absorptivity. In silicates and oxides, direct photon transfer depends critically on the concentration and valence state of *d*-block elements, especially iron, whereas diffusive radiative transfer also depends on macroscopic properties, such as grain size[8]. Because the influence of transition metal elements such as iron on direct photon transfer is related to electronic structure, the spin state of iron is also an important factor in a material's infrared and visible spectral ranges where crystal-field transitions occur[13–15]. Application of laboratory-derived spectral data to conductivity in the lower mantle not only requires information on each component of heat transfer, but also their respective pressure and temperature dependencies[8].

Previous measurements of thermal diffusivity in mantle minerals at high pressures have been limited to the 5–10 GPa pressure range (that is, 150–300 km depth) at room temperature[16], with a few optical absorption studies of the direct radiative component of conductivity in olivine or silicate spinels at 20–30 GPa (refs 17, 18). At high temperatures, thermal diffusivity in olivine has been measured up to about 1,500 °C at ambient pressure[19]. At conditions deeper in the mantle and beyond previous experimental studies, theoretical models have been used to estimate pressure–temperature variation of thermal conductivity, but the validity of model predictions require experimental verification at conditions relevant to the lower mantle.

Optical absorption spectra of ferropericlase were recently reported to about 80 GPa at room temperature[20,21], and showed a strong dependence of absorptivity on iron concentration and pressure, and possibly on spin state. In ferropericlase, interpretation of optical spectra are complicated by the defect structure of (Mg,Fe)O, resulting from variable amounts of Fe^{3+} in samples annealed at different conditions of oxygen fugacity and pressure[21]. Absorption spectra for Fe-bearing silicate perovskite have been reported[22,23], but so far only at ambient experimental pressure–temperature conditions.

The effect of pressure on optical absorption of Fe-bearing silicate perovskite is shown in Fig. 1. Observation of some bands requires thicker samples (>50 μm), which are difficult to accommodate in the diamond anvil cell at very high pressures (Supplementary Fig. 1). The crystal-field transition near 8,000 cm^{-1} could not be detected with samples thinner than 15 μm. However, using an oil medium, we were able to study the sample of approximately 30 μm initial thickness to 46 GPa (Fig. 1a, Supplementary Fig. 1). Increasing pressure shifts the crystal-field band at 8,000 cm^{-1} to higher energy, accompanied by a gradual decrease in absorption above about 36 GPa. Experiments in Ar media show that the crystal-field band becomes unobservable in the 43–84 GPa pressure range. However, the weakness of this band relative to the background in our spectra at high pressure prevents us from drawing definitive conclusions about whether this band disappears or falls below detection level. During the highest-pressure experiments, the crystal-field band was barely detectable (Fig. 1c) because thinner samples were used. Also, non-hydrostatic stresses, which develop within 'hard' pressure media (such as Ar at high pressure) or without a transmitting medium, are observed to broaden the crystal-field band, at least partly because of the increase in splitting between the individual components of the ^5E_g–$^5T_{2g}$ transitions.

In agreement with the previous ambient pressure study[22], all three sets of measurements on silicate perovskite reveal a broad band near 15,000 cm^{-1}, with an ultraviolet absorption edge above 25,000 cm^{-1} that has a broad absorption tail extending through the visible and near-infrared. The 15,000 cm^{-1} band has previously been assigned to intervalence charge transfer[13,22,24]. Our measurements show a blue

[1]Geophysical Laboratory, Carnegie Institution of Washington, 5251 Broad Branch Road NW, Washington DC 20015, USA. [2]Department of Geological Sciences, University of Colorado, Boulder, Colorado 80309, USA. [3]Department of Earth and Planetary Sciences, Northwestern University, Evanston, Illinois 60208, USA.

shift of this band with pressure, similar to iron-bearing ringwoodite[18]. In perovskite, the intervalence band becomes sharper at approximately 40 GPa (Fig. 1a) and is almost invariant in frequency with increasing pressure above ~80 GPa. The ultraviolet absorption edge, which was assigned previously[22] to O^{2-}–Fe^{3+} charge transfer, shows a moderate red shift (or increase of intensity) with pressure (Fig. 1). It has been argued[8] that this 'background' ultraviolet feature is due to grain boundary scattering. However, we also observe this feature in our single-crystal samples. Moreover, an Al-bearing perovskite crystal with 4% Fe also shows the ultraviolet absorption edge (Supplementary Fig. 1), which minimizes the likelihood that this band results from charged defects. We therefore tentatively assign the ultraviolet band to O^{2-}–Fe^{3+} charge transfer, on the basis of the systematics of charge-transfer gap values in oxides[25]. However, we cannot exclude the possibility of a small contribution from the charge-transfer transition O^{2-}–Fe^{2+}. Ultraviolet–visible absorption is much smaller for samples with lower iron concentration (see Supplementary Information for detailed results), which confirms the proposed band assignments.

Relatively high radiative conductivity of the lower mantle has been inferred as a consequence of the high-spin to low-spin electronic transition of Fe^{2+} in silicate perovskite and ferropericlase[14,15,26]. This conjecture draws from predictions of blue-shifting crystal-field transitions, which were believed to be the main mechanism of the absorption in the near-infrared and visible spectral ranges. As the thermal radiation corresponding to the Earth's lower mantle (2,200–3,000 K) has a maximum in the near-infrared, a blue shift would allow thermal radiation to propagate longer distances, thus increasing the radiative component of thermal conductivity. Our results show that this effect does occur in perovskite, but only over the pressure range up to ~50 GPa. The decrease in intensity of the ^5E_g–$^5T_{2g}$ band at 7,000 cm^{-1} (Fig. 1a, b) at 45–60 GPa is consistent with the pressure range of spin-pairing of Fe^{2+} (refs 15, 27).

Because crystal-field transitions of Fe^{3+} are much weaker (they are forbidden by spin conservation[13]) and because the intervalence Fe^{2+}–Fe^{3+} optical transitions also change in the pressure range where the ^5E_g–$^5T_{2g}$ transition diminishes (Fig. 1), we interpret the spectra with a spin transition of Fe^{2+} in the dodecahedral (A) site in the 45–60 GPa range[15]. No other qualitative change in absorption is observed in our spectra (Fig. 1) above 60 GPa (compare ref. 15). The change in optical absorption near 50 GPa is rather subtle, being overwhelmed by other absorption mechanisms at higher pressures. Similarly to our previous results in ferropericlase[20] (see also Supplementary Fig. 2), at higher pressures we observe an increase in the absorption related to a red shift of the ultraviolet absorption edge, which causes an overall absorption increase that compensates or exceeds the effects of pressure on the crystal-field transitions. Unlike ferropericlase with high Fe concentration (>11%, which corresponds to the percolation limit), this increase of the absorption edge is rather moderate in silicate perovskite, so we assign it to the Fe–O charge transfer transitions (see also ref. 22). We observed a similar subtle change of the optical absorption in our 6% Fe (Mg,Fe)O samples[20].

The current results from optical absorption of silicate perovskite and ferropericlase (see also Supplementary Fig. 2) with iron concentrations relevant to Earth's lower mantle provide a basis for determining the direct radiative heat transfer of minerals over a wide depth range in the mantle. Although the effect of temperature on the absorption spectra of some minerals may be substantial[12], we do not expect this to be the case for silicate perovskite and ferropericlase, as the main absorption mechanisms are different to that in olivine studied in ref. 12. Our absorption measurements on 15% Fe (Mg,Fe)O samples to 60 GPa and 800 K did not show any substantial temperature dependence, which would affect the calculations of the radiative conductivity presented below (see Supplementary Fig. 3).

The radiative part of the thermal conductivity, $k_{rad}(T)$, can be calculated as follows[8]:

$$k_{rad}(T) = \frac{4}{3}\int_0^\infty \frac{\Lambda}{[1+\ln 10\Lambda\alpha(\nu,T)]} \times \frac{1}{\nu^2}\frac{\partial[n(\nu,T)^2(1-\exp(-\ln 10\Lambda\alpha(\nu,T)))I(\nu,T)]}{\partial T}\,d\nu \quad (1)$$

where $I(\nu,T)$ is the Planck function (T is temperature), ν is wavenumber, $n(\nu,T)$ is the refraction index, Λ is the grain size (scattering length), and α is the measured absorption coefficient. Assuming the grain size is much larger than the typical absorption length (~$1/\alpha$), we have calculated the radiative contribution to thermal conductivity in the lower mantle (Fig. 2), using our experimental absorption spectra (Fig. 1, Supplementary Fig. 2).

Our estimated radiative conductivity is 2–5 times smaller than the previous report[7,8] based on extrapolations of mineral optical properties to higher pressures. Substantial contribution from the

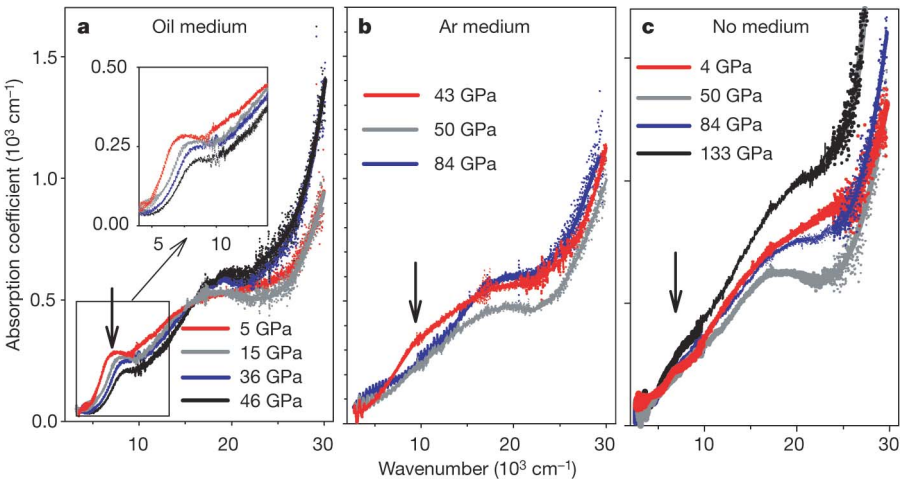


Figure 1 | Optical absorption spectra of silicate perovskite (10 mol% Fe) up to 133 GPa at room temperature in various pressure media. a, Oil; **b**, Ar, **c**, no medium. The absorption coefficient has been calculated using $\alpha = A\ln 10/d$, where $A = \log_{10}(I_0/I)$, I_0 and I are the reference and sample signals respectively, and d is the sample thickness measured from white-light interferometry at ambient pressure and calculated at high pressure using an isothermal equation of state. Although substantial care has been taken to measure I_0 and I in the same conditions (for example, pressure), sample imperfections and limited sample dimensions affected the accuracy of the baseline, which has up to 50 cm^{-1} uncertainty. Ripples in the spectra at low pressures are interference fringes resulting from multiple reflections between the parallel polished sample surfaces. The interference patterns and noisy spectral areas are smoothed (thick solid lines) for clarity. Vertical arrows designate the Fe^{2+} crystal-field transitions, which are characteristic of the high-spin phase. Inset in **a** shows an enlarged view of the near-infrared spectral range (boxed).

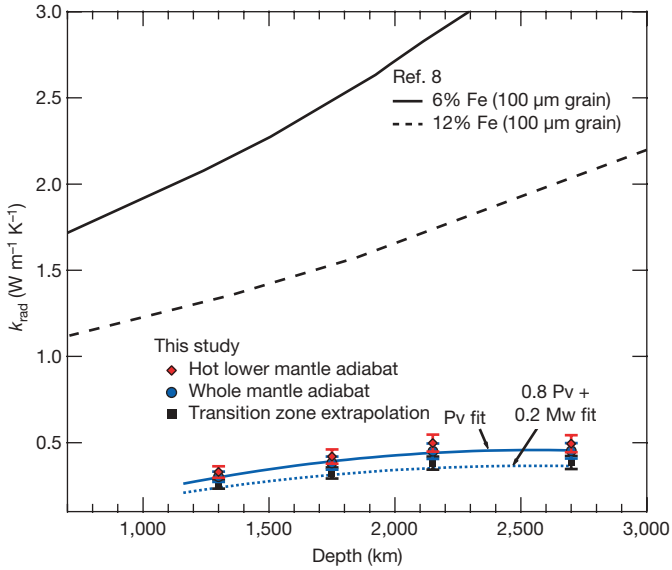


Figure 2 | Radiative part of the Earth's lower-mantle thermal conductivity as a function of depth. Three different geotherms were used, following ref. 8: a hot lower-mantle adiabat (red filled diamonds), a whole mantle adiabat (blue filled circles) and the extrapolated transition zone adiabat (black filled squares). Solid and dashed black lines present previous estimates of k_{rad} for 6% and 12% iron concentrations, respectively[8]. The absorption coefficient was assumed to be temperature independent, which could have caused an overestimation of k_{rad}. We also suppose that the product $\alpha\Lambda$ is large ($\Lambda > 0.5$ mm). The radiative conductivity was calculated using equation (1) in two limiting cases: pure perovskite mantle (ferropericlase has a much larger absorption coefficient and therefore will not contribute to the diffusive radiative heat transfer) and a mixture of 80% of silicate perovskite, $(Mg_{0.9}Fe_{0.1})SiO_3$ and 20% of ferropericlase with 15% Fe composition (see Supplementary Information for the results in ferropericlase), which contribute proportionally (which would be valid in the case of the layered mantle). The results are shown as solid and dotted blue lines labelled 'Pv fit' and '0.8 Pv + 0.2 Mw fit', respectively. Error bars are projected uncertainties of the absorption coefficient measurements and their temperature corrections.

ultraviolet–visible absorption tails to the radiative conductivity, which could only be inferred in previous models[8], is responsible for the discrepancy. As absorption from the ultraviolet–visible tail is mostly due to the O–Fe^{3+} charge transfer (see Supplementary Fig. 1), and the contribution from the Fe^{2+}–Fe^{3+} intervalence charge transfer is substantial, the amount of the ferric iron in the mantle (that is, redox state) essentially controls the radiative heat conductivity (Fig. 2). According to our estimates, the radiative contribution does not exceed 0.54 W m^{-1} K^{-1} at the top of the D″ layer. The value of the thermal conductivity, k, in the thermal boundary layer strongly influences its thinness and stability: low values of k typically imply a less stable and thinner boundary layer[28]. Moreover, a decrease in k also implies a decrease in plume temperature[29]. These results should be taken into account when carrying out model calculations of the mantle dynamics[9,10,30]. Substantial efforts are still required to estimate the lattice contribution to thermal diffusivity at relevant pressures and temperatures.

METHODS SUMMARY

Sample preparation and characterization. Single crystals of silicate perovskite were synthesized from synthetic ^{57}Fe-enriched enstatite starting material (En_{90}) at 23 GPa and 1,800 °C. Electron microprobe analysis of the recovered samples confirmed the approximate chemical formula $(Mg_{0.9}Fe_{0.1})SiO_3$. Mössbauer spectroscopy indicates that 11 ± 3% of the iron is ferric (Fe^{3+}). Single-crystal structure refinements indicate that iron occupies the dodecahedral (A) site, without detectable iron in the octahedral (B) site (~2 mol% or less). For the optical absorption measurements at high pressures, samples were polished to 15–25 μm thickness and selected on the basis of their visible clarity. Silicate perovskite crystals exhibited sharp optical extinction in cross-polarized light, confirming their crystalline state.

Optical absorption measurements. Spectra were obtained in the 2,000–40,000 cm^{-1} range using a custom infrared microscope and all-reflecting relay optics coupled to a Fourier-transform infrared spectrometer and a near-infrared-to-ultraviolet single-stage grating spectrograph with a CCD array detector. At each pressure, a spectrum through the sample and a spectrum through the pressure medium for background correction were obtained. A double confocal geometry was used to minimize scattered and spurious light, so the spectra of highly absorbing samples could be recorded.

High-pressure techniques. High-pressure measurements were carried out in a symmetric diamond anvil cell with various pressure media (Ar, Ne, silicone oil) to 87 GPa or without pressure medium (to 133 GPa).

Received 5 July 2007; accepted 2 September 2008.

1. Johnson, C. M. & Beard, B. L. Biogeochemical cycling of iron isotopes. *Science* **309**, 1025–1027 (2005).
2. Rouxel, O. J., Bekker, A. & Edwards, K. J. Iron isotope constraints on the Archean and Paleoproterozoic ocean redox state. *Science* **307**, 1088–1091 (2005).
3. McCammon, C. Perovskite as a possible sink for ferric iron in the lower mantle. *Nature* **387**, 694–696 (1997).
4. Frost, D. J. *et al.* Experimental evidence for the existence of iron-rich metal in the Earth's lower mantle. *Nature* **428**, 409–412 (2004).
5. Zhong, S. Constraints on thermochemical convection of the mantle from plume heat flux, plume excess temperature, and upper mantle temperature. *J. Geophys. Res.* **111**, B04409, doi:10.1029/2005JB003972 (2006).
6. Yanagawa, T. K. B., Nakada, M. & Yuen, D. A. Influence of lattice thermal conductivity on thermal convection with strongly temperature-dependent viscosity. *Earth Planets Space* **57**, 15–28 (2005).
7. Hofmeister, A. M. Mantle values of thermal conductivity and the geotherm from phonon lifetimes. *Science* **283**, 1699–1706 (1999).
8. Hofmeister, A. M. & Yuen, D. A. Critical phenomena in thermal conductivity: Implications for lower mantle dynamics. *J. Geodyn.* **44**, 186–199 (2007).
9. Dubuffet, F., Yuen, D. A. & Rainey, E. S. G. Controlling thermal chaos in the mantle by positive feedback from radiative thermal conductivity. *Nonlin. Process. Geophys.* **9**, 1–13 (2002).
10. van den Berg, A. P., Rainey, E. S. G. & Yuen, D. A. The combined influences of variable thermal conductivity, temperature- and pressure-dependent viscosity and core–mantle coupling on thermal evolution. *Phys. Earth Planet. Inter.* **149**, 259–278 (2005).
11. Clark, S. P. Jr. Radiative transfer in the Earth's mantle. *Trans. Am. Geophys. Union* **38**, 931–938 (1957).
12. Shankland, T. J., Nitsan, U. & Duba, A. G. Optical absorption and radiative heat transport in olivine at high temperature. *J. Geophys. Res.* **84**, 1603–1610 (1979).
13. Burns, R. G. *Mineralogical Applications of Crystal Field Theory* 2nd edn (Cambridge Univ. Press, 1993).
14. Sherman, D. M. The high-pressure electronic structure of magnesiowüstite (Mg,Fe)O: Applications to the physics and chemistry of the lower mantle. *J. Geophys. Res.* **96**, 14299–14312 (1991).
15. Badro, J. *et al.* Transitions in perovskite: Possible nonconvecting layers in the lower mantle. *Science* **305**, 383–386 (2004).
16. Ross, R. G., Andersson, P., Sundqvist, B. & Bäckström, G. Thermal conductivity of solids and liquids under pressure. *Rep. Prog. Phys.* **47**, 1347–1402 (1984).
17. Mao, H. K. & Bell, P. M. Electrical conductivity and the red shift of absorption in olivine and spinel at high pressure. *Science* **176**, 403–406 (1972).
18. Keppler, H. & Smyth, J. R. Optical and near infrared spectra of ringwoodite to 21.5 GPa: Implications for radiative heat transport in the mantle. *Am. Mineral.* **90**, 1209–1212 (2005).
19. Pertermann, M. & Hofmeister, A. M. Thermal diffusivity of olivine-group minerals at high temperature. *Am. Mineral.* **91**, 1747–1760 (2006).
20. Goncharov, A. F., Struzhkin, V. V. & Jacobsen, S. D. Reduced radiative conductivity of low-spin (Mg,Fe)O in the lower mantle. *Science* **312**, 1205–1208 (2006).
21. Keppler, H., Kantor, I. & Dubrovinsky, L. S. Optical absorption spectra of ferropericlase to 84 GPa. *Am. Mineral.* **92**, 433–436 (2007).
22. Keppler, H., McCammon, C. A. & Rubie, D. C. Crystal-field and charge-transfer spectra of (Mg,Fe)SiO_3 perovskite. *Am. Mineral.* **79**, 1215–1218 (1994).
23. Shen, G., Fei, Y., Hålenius, U. & Wang, Y. Optical absorption spectra of (Mg,Fe)SiO_3 silicate perovskites. *Phys. Chem. Miner.* **20**, 478–482 (1994).
24. Mattson, S. M. & Rossman, G. R. Identifying characteristics of charge transfer transitions in minerals. *Phys. Chem. Miner.* **14**, 94–99 (1987).
25. Bocquet, A. E. *et al.* Electronic structure of early 3*d*-transition-metal oxides by analysis of the 2*p* core-level photoemission spectra. *Phys. Rev. B* **53**, 1161–1170 (1996).
26. Badro, J. *et al.* Iron partitioning in Earth's mantle: Toward a deep lower mantle discontinuity. *Science* **300**, 789–791 (2003).
27. Li, J. *et al.* Pressure effect on the electronic structure of iron in (Mg,Fe)(Si,Al)O_3 perovskite: A combined synchrotron Mössbauer and X-ray emission spectroscopy study up to 100 GPa. *Phys. Chem. Miner.* **33**, 575–585 (2006).
28. Turcotte, D. L. & Schubert, G. *Geodynamics* (Cambridge Univ. Press, 2002).

29. Montague, N. L., Kellogg, L. H. & Manga, M. High-Rayleigh number thermo-chemical models of a dense boundary layer in D. *Geophys. Res. Lett.* **25,** 2345–2348 (1998).
30. Naliboff, J. B. & Kellogg, L. H. Can large increases in viscosity and thermal conductivity preserve large-scale heterogeneity in the mantle? *Phys. Earth Planet. Inter.* **161,** 86–102 (2007).

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

Acknowledgements We acknowledge support from NSF/EAR, DOE/BES, DOE/NNSA (CDAC) and the W. M. Keck Foundation. S.D.J. thanks D. J. Frost, S. J. Mackwell and D. P. Dobson for help with sample synthesis, C. A. McCammon for Mössbauer spectroscopy, J. R. Smyth for single-crystal X-ray diffraction, H. Watson for electron microprobe analysis of the silicate perovskite material and the NSF and Bayerishes Geoinstitut Visitor's Program for support. B.D.H. was supported by the NSF Research Experience for Undergraduates (REU) Program at the Carnegie Institution of Washington. P.B. was partially supported by the Balzan Foundation.

Author Contributions A.F.G., V.V.S. and S.D.J. designed the research programme; S.D.J. synthesized and polished the single crystals; A.F.G. and B.D.H. performed high-pressure experiments at room temperature; A.F.G., V.V.S. and P.B. performed high-temperature experiments; A.F.G., V.V.S., B.D.H. and P.B. analysed the data; P.B. developed a thermal conductivity model; V.V.S, A.F.G. and S.D.J interpreted the results; A.F.G, S.D.J., P.B. and V.V.S. wrote the paper. All authors discussed the results and commented on the manuscripts.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.F.G. (goncharov@gl.ciw.edu).

LETTERS

The interaction between predation and competition

Peter Chesson[1] & Jessica J. Kuang[1]

Competition and predation are the most heavily investigated species interactions in ecology, dominating studies of species diversity maintenance. However, these two interactions are most commonly viewed highly asymmetrically. Competition for resources is seen as the primary interaction limiting diversity, with predation modifying what competition does[1], although theoretical models have long supported diverse views[1–5]. Here we show, using a comprehensive three-trophic-level model, that competition and predation should be viewed symmetrically: these two interactions are equally able to either limit or promote diversity. Diversity maintenance requires within-species density feedback loops to be stronger than between-species feedback loops. We quantify the contributions of predation and competition to these loops in a simple, interpretable form, showing their equivalent potential to strengthen or weaken diversity maintenance. Moreover, we show that competition and predation can undermine each other, with the tendency of the stronger interaction to promote or limit diversity prevailing. The past failure to appreciate the symmetrical effects and interactions of competition and predation has unduly restricted diversity maintenance studies. A multitrophic perspective should be adopted to examine a greater variety of possible effects of predation than generally considered in the past. Conservation and management strategies need to be much more concerned with the implications of changes in the strengths of trophic interactions.

We focus on the middle trophic level in a three-trophic-level system (Fig. 1), and address how both competition for resources (the trophic level below) and predation (the trophic level above) affect species coexistence in the middle trophic level. Several decades ago, MacArthur[6] formulated the definitive model for resource competition in the Lotka–Volterra form. This model leads to a measure of niche overlap, ρ, between any pair of species[7], and also measures κ_j (originally k_j)[8] defining the fitness of any species, j. Coexistence occurs in two-species Lotka–Volterra competition if the competitive effect that a species has on the other species (interspecific competition, α_{ij}) is less than the competitive effect that it has on itself (intraspecific competition, α_{jj})[8]. Notably, the ratio of these competitive effects can be expressed in terms of fitnesses and niche overlap[8]:

$$\frac{\alpha_{ij}}{\alpha_{jj}} = \frac{\kappa_j}{\kappa_i}\rho \qquad (1)$$

Species j dominates over species i, and excludes it from the system, if expression (1) is greater than one. When niche overlap is complete, ρ equals one and the species with the larger fitness excludes the other. Otherwise, ρ is less than one and the relative fitness (κ_j/κ_i) must be discounted by ρ (how much the species interact) to see if exclusion occurs (that is, to see if interspecific competition exceeds intraspecific competition). Neither species can exclude the other when expression (1) is less than one for both $(i,j) = (1,2)$ and $(i,j) = (2,1)$, a condition equivalent to[9]

$$\rho < \frac{\kappa_1}{\kappa_2} < \frac{1}{\rho} \qquad (2)$$

When condition (2) holds, the species coexist. Thus, niche overlap, ρ, constrains the fitness differences compatible with coexistence. Low overlap (ρ near to zero) means that the species can differ greatly in fitness and still coexist with each other, whereas large overlap (ρ near to one) means fitnesses must be nearly equal for coexistence to occur (Fig. 2).

The new finding with a three-trophic-level Lotka–Volterra system is that these same conditions continue to hold, including predation in the same terms as resource competition (Box 1). To achieve this outcome, however, a new assumption is necessary: the focal species in the middle trophic level must not be the sole food source for the predators. Prey outside the focal group prevent the predators from becoming extinct when the focal species are at zero density. Although this is not the usual assumption made in models, it is not an unreasonable case to consider: often the focal group is not the entirety of a predator's prey, and predators often range more widely than their prey so that the predator is not solely supported by the region in which the focal group resides[10]. When this assumption is removed, the main conclusions here are retained (see Supplementary Information). We make a similar assumption with respect to resource competition: focal species must not drive their resources to extinction. The key conclusions are retained when this assumption is violated (Supplementary Information), but ρ is no longer a constant, complicating coexistence conditions[11].

In the original MacArthur model, fitness (κ_j) is proportional to the net excess resource intake of a species over its maintenance requirements[9]. With three trophic levels, subtracted from this net excess is the mortality due to predation when predators are at their equilibrium densities in the absence of the focal species. These new fitnesses, κ_j, are maximal quantities representing the abilities of focal species to gather resources and avoid predation. These quantities have the

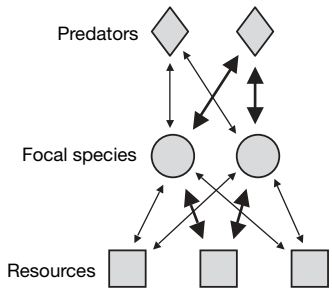


Figure 1 | Simplified three-trophic-level food web. The heavy lines highlight linkages between focal species through a shared resource and a shared predator. Double-headed arrows indicate that linkages are bidirectional, creating feedback loops. For example, high focal density of a species increases predator density, which then feeds back to greater predation on both the same focal species and the other focal species (apparent competition). Similarly, feedback loops through resources create resource competition. Each bidirectional linkage by itself is an intraspecific feedback loop for a focal species. Linkages between focal species through a shared predator or shared resource are interspecific feedback loops.

[1]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721, USA.

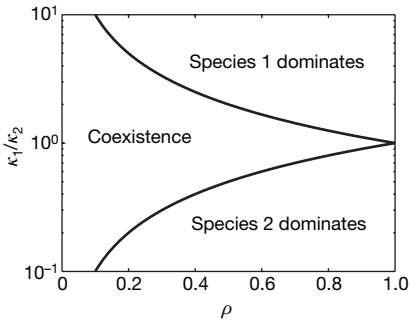


Figure 2 | Coexistence and exclusion regions. Two species coexist when the niche overlap, ρ, and fitness ratio, κ_1/κ_2, lie within the central wedge, in which condition (2) is satisfied. Exclusion occurs outside this wedge. The log scale for the κ_1/κ_2-axis preserves symmetry.

essential property of predicting the winning species in cases in which there is no resource or predator partitioning (that is, in situations in which there is no possibility of coexistence[8,12]).

The quantities α_{jj} and α_{ij} now represent the total strengths of intraspecific and interspecific density dependence, combining both competition and predation. Thus, α_{jj} measures the combined strengths of the feedback loops from species j to species j through both resources and predators, whereas α_{ij} measures the combined strengths of all such loops from species j to species i (Fig. 1). The fact that feedback loops through predators lead to mutually negative indirect interactions between prey, analogous to competition, is the important insight of Holt[2,13], who coined the term "apparent competition" for this outcome (Fig. 1). The simple idea that competitive coexistence requires intraspecific competition to exceed interspecific competition is now generalized to the idea that intraspecific density dependence must exceed interspecific density dependence. The critical ratio of interspecific density dependence to intraspecific density dependence is again given by expression (1), showing that the ability of a species to exclude another depends simply and intuitively on its relative fitness, κ_j/κ_i, discounted by niche overlap, ρ. Niches now involve how the species relate to their predators in addition to how they relate to their resources (Fig. 3). Niche overlap once again determines the breadth of the opportunities for coexistence according to condition (2), illustrated in Fig. 2. This condition is derived from the requirement that $(\kappa_j/\kappa_i)\rho$ must always be less than one for coexistence (that is, α_{jj} should always be greater than α_{ij}).

The measures α_{jj}, α_{ij} and ρ depend on each feedback loop according to its strength. This fact is intuitive but of profound consequence: competition and predation can each undermine the predicted effects of the other (either coexistence or exclusion) depending on which is stronger. Niche overlap jointly represents the overlap between species in their patterns of resource dependence and their patterns of predator susceptibility (Fig. 3), but the dependence of ρ on predators and resources reflects the tendencies of these trophic levels to dominate focal species interactions. If resources strongly dominate, ρ approaches the limiting value ρ^R based on resource overlap alone. If predators strongly dominate, ρ approaches the predator overlap value ρ^P. Which of these dominates depends on the relative strengths of the density-dependent feedback loops through resources and through predators (that is, on which of these more strongly regulates the densities of the focal species). A complex of factors determine which feedback loops are strongest, but, simply put, resource loops are strong if resources regenerate slowly, and predation loops are strong if predators are primarily controlled by prey in the focal group (Box 1).

Whether coexistence or exclusion is promoted is determined by whether partitioning of the dominant interaction occurs—be that competition or predation. As the relative intensity of predation and competition is changed, niche overlap, ρ, changes as depicted in Fig. 4. Cases in which there is resource partitioning, but no predator partitioning (Fig. 3b), make ρ an increasing function of relative predation intensity, having a low value when competition is dominant, increasing to a value of one when predation dominates (Fig. 4, curve b). Thus, broad opportunities for coexistence in terms of potentially broad differences in κ values are permitted when competition dominates, but not when predation dominates. Notably, the opposite pattern of predator partitioning without resource partitioning (Fig. 3c) provides the strongest opportunities for coexistence (lowest ρ) under dominance by predation (Fig. 4, curve c). When there is no partitioning at either level, coexistence is still possible if there is a trade-off across species between resource sensitivity and predation sensitivity (Fig. 3d). In this case, opportunities for coexistence arise for a broad region of intermediate values of relative competition and predation intensities (Fig. 4, curve d). However, in the absence of the trade-off between competition and predation, ρ is instead one for all predation and competition intensities (Fig. 4, curve e).

Box 1 | Model and analysis

Lotka–Volterra equations for three trophic levels (focal species, N_j, resources, R_l, and predators, P_m) are

$$\frac{1}{N_j}\frac{dN_j}{dt} = \sum_l c_{jl}v_l R_l - \sum_m a_{jm}P_m - \mu_j$$
$$\frac{1}{R_l}\frac{dR_l}{dt} = r_l^R\left(1-\alpha_l^R R_l\right) - \sum_j N_j c_{jl} \qquad (3)$$
$$\frac{1}{P_m}\frac{dP_m}{dt} = r_m^P\left(1-\alpha_m^P P_m\right) + \sum_j wN_j a_{jm}$$

with parameters c_{jl} (consumption of resource l by focal species j), a_{jm} (attack rate of focal species j by predator m), r_l^R and r_m^P (predator and resource intrinsic rates of increase), α_l^R and α_m^P(resource and predator intraspecific competition—reciprocals of carrying capacities), v_l (unit value of resource l), μ_j (resource maintenance requirement of focal species j), and w (value of a unit of prey to a predator).

For any pair, j and k, of focal species, methods previously described[7] give the overlap measure

$$\rho = \frac{\sum_l \frac{c_{jl}v_l c_{kl}}{r_l^R\alpha_l^R} + \sum_m \frac{a_{jm}wa_{km}}{r_m^P\alpha_m^P}}{\sqrt{\left(\sum_l \frac{c_{jl}^2 v_l}{r_l^R\alpha_l^R} + \sum_m \frac{a_{jm}^2 w}{r_m^P\alpha_m^P}\right)\left(\sum_l \frac{c_{kl}^2 v_l}{r_l^R\alpha_l^R} + \sum_m \frac{a_{km}^2 w}{r_m^P\alpha_m^P}\right)}} \qquad (4)$$

To obtain ρ^R and ρ^P, the predator and resource terms, respectively, are set to zero.

Joint sensitivity to predation and competition is measured as

$$s_j = \sqrt{\left(\sum_l \frac{c_{jl}^2 v_l}{r_l^R\alpha_l^R} + \sum_m \frac{a_{jm}^2 w}{r_m^P\alpha_m^P}\right)} \qquad (5)$$

following Appendix D of ref. 9.

Fitness measures,

$$\kappa_j = \frac{1}{s_j}\left(\sum_l \frac{c_{jl}v_l}{\alpha_l^R} - \sum_m \frac{a_{jm}}{\alpha_m^P} - \mu_j\right) \qquad (6)$$

are focal species per capita growth rates at zero densities of all focal species, divided by s_j (Appendix D of ref. 9).

Intraspecific and interspecific coefficients of density dependence are

$$\alpha_{jj} = s_j/\kappa_j \text{ and } \alpha_{ij} = \rho s_j/\kappa_i \qquad (7)$$

as explained in Supplementary Information. The invasibility criterion for coexistence of two species[8] leads to condition (2). See Supplementary Information for details.

Competition is strong if resources regenerate slowly (that is, if r_l^R is small). Density dependence due to predation is strong if predators depend only weakly on prey outside the focal group (that is, if r_m^P is small). The importance of predation or competition is thus inverse to r_l^R or r_m^P. Figure 4 represents a common linear increase from left to right in each $1/r_m^P$ with a corresponding linear decrease in each $1/r_l^R$.

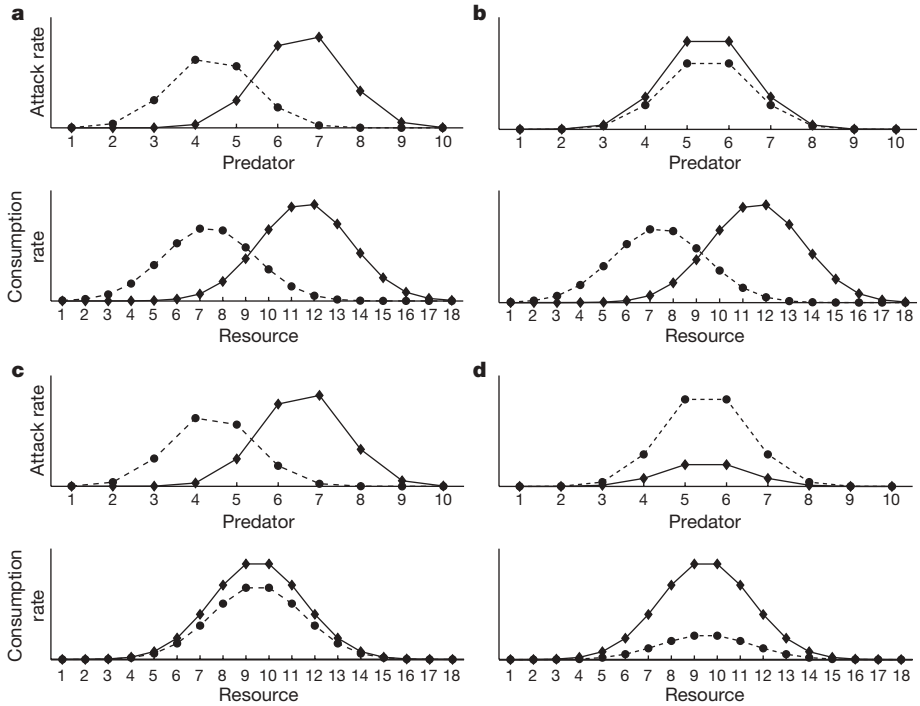


Figure 3 | Niches of two focal species in terms of resources and predators. The x axis labels identify predator and resource species. The filled circles and filled diamonds define focal species. The y axes are the rates a_{jm} and c_{jl} of Box 1. **a**, Niches separated by resource consumption ($\rho^R = 0.33$) and predation ($\rho^P = 0.31$). **b**, Niches separated by resources ($\rho^R = 0.33$), but not predators ($\rho^P = 1$). **c**, Niches separated by predators ($\rho^P = 0.31$), but not resources ($\rho^R = 1$). **d**, No separate predator or resource niche differences ($\rho^R = \rho^P = 1$). In **a**–**c**, overall niche overlap, ρ, is intermediate between ρ^R and ρ^P. In **d**, ρ is less than 1 whenever neither predators nor resources dominate focal species interactions.

These results show that predation and resource competition have the potential to affect diversity in the same way, and each may either promote coexistence or promote exclusion. Each promotes exclusion when it does not differentiate between species; conversely, each promotes coexistence when it does differentiate between species. Moreover, predation and competition interact with each other. If one is much stronger than the other, the predictions of the stronger prevail. When present together, both differentiating between species, coexistence is promoted more strongly by their joint action. When only one process differentiates between species, the joint outcome is less coexistence than when only the differentiating process is present. When of similar strength, predation and competition can jointly promote coexistence through a trade-off between competition and predation. In this last way, the predation–competition interaction differentiates between species, and creates an effective mechanism where previously there was none.

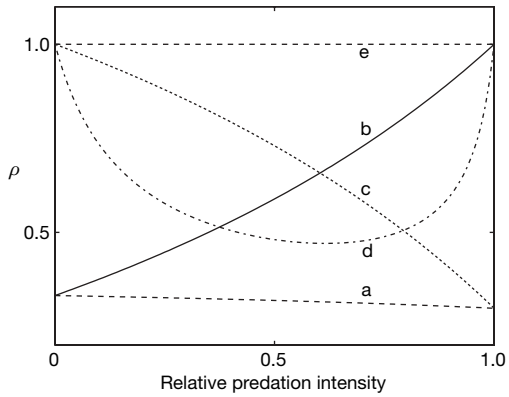


Figure 4 | Niche overlap as a function of relative predation intensity. Changes in niche overlap as the intensity of density dependence owing to predation is varied relative to the intensity of competition. Curves correspond to panels in Fig. 3, with the extra curve e for the case in which there is no partitioning of resources or predators and no trade-off between susceptibility to predation and sensitivity to resources (like Fig. 3d but with the symbols switched in the attack-rate graph to match the consumption-rate graph).

These outcomes are not restricted to the two-focal species illustrations given here, as multispecies findings for MacArthur's model[14] allow their extension. This new understanding suggests a classification of coexistence mechanisms as competition-based (when intraspecific competition exceeds interspecific competition), predation-based (when intraspecific density feedback through predators is stronger than interspecific density feedback through predators) or jointly based on competition and predation[4,15]. In his classic work on diversity maintenance, Hutchinson[16] focused coexistence studies within trophic levels, spawning an industry in the study of interspecific competition. The fact that predation may function in a very similar way in the maintenance of diversity within trophic levels means that studies of predation–competition interactions should move beyond the notion that competition is the primary interaction limiting diversity, with predation modifying what competition does. We echo the call by Holt a quarter of a century ago to treat predation and competition even-handedly[2]. The absence of an accurate sense of theoretical predictions for the role of predation in diversity maintenance has led to much confusion at the interface between theoretical and empirical work[1]. Better targeted empirical investigations should be possible with the new understanding from this and related, although more complex, work for non-Lotka–Volterra systems[1,15,17].

Maintaining predators is of increasing concern in conservation and management strategies[18–21]. Humans have had major effects on the trophic structure of terrestrial, freshwater and marine ecosystems, with implications that are yet to be fully realized[22–24]. The individual species approach to conservation often enshrined in law, such as the US Endangered Species Act, focuses on minimal populations for a species' own conservation without considering its role in an ecosystem. Our findings emphasize the profound effects that one trophic level can have on diversity maintenance in other trophic

levels. They argue for conservation strategies that seek to maintain trophic structure and the strengths of trophic links. In particular, they reinforce concerns that the decimation of carnivores that is occurring in most environments on the planet may have major impacts on diversity maintenance in lower trophic levels[18,23,24].

Received 23 May; accepted 11 July 2008.

1. Chase, J. M. *et al.* The interaction between predation and competition: a review and synthesis. *Ecol. Lett.* **5,** 302–315 (2002).
2. Holt, R. D. Spatial heterogeneity, indirect interactions, and the coexistence of prey species. *Am. Nat.* **124,** 377–406 (1984).
3. Kotler, B. P. & Holt, R. D. Predation and competition: the interaction of two types of species interactions. *Oikos* **54,** 256–260 (1989).
4. Grover, J. P. & Holt, R. D. Disentangling resource and apparent competition: realistic models for plant-herbivore communities. *J. Theor. Biol.* **191,** 353–376 (1998).
5. Krivan, V. Competitive co-existence by adaptive predators. *Evol. Ecol. Res.* **5,** 1163–1182 (2003).
6. MacArthur, R. Species packing and competitive equilibrium for many species. *Theor. Popul. Biol.* **1,** 1–11 (1970).
7. Chesson, P. MacArthur's consumer–resource model. *Theor. Popul. Biol.* **37,** 26–38 (1990).
8. Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Syst.* **31,** 343–366 (2000).
9. Chesson, P. & Huntly, N. The roles of harsh and fluctuating conditions in the dynamics of ecological communities. *Am. Nat.* **150,** 519–553 (1997).
10. Holt, R. D., Grover, J. & Tilman, D. Simple rules for interspecific dominance in systems with exploitative and apparent competition. *Am. Nat.* **144,** 741–771 (1994).
11. Abrams, P. A. High competition with low similarity and low competition with high similarity: exploitative and apparent competition in consumer-resource systems. *Am. Nat.* **152,** 114–128 (1998).
12. Chesson, P. in *Unity in Diversity: Reflections on Ecology after the Legacy of Ramon Margalef* (eds Valladares, F. *et al.*) 119–164 (Fundacion BBVA, 2008).
13. Holt, R. D. Predation, apparent competition, and the structure of prey communities. *Theor. Popul. Biol.* **12,** 197–229 (1977).
14. Haygood, R. Coexistence in MacArthur-style consumer–resource models. *Theor. Popul. Biol.* **61,** 215–223 (2002).
15. Kuang, J. J. & Chesson, P. Predation–competition interactions for seasonally recruiting species. *Am. Nat.* **171,** E119–E133 (2008).
16. Hutchinson, G. E. Homage to Santa Rosalia or why are there so many kinds of animals? *Am. Nat.* **93,** 145–159 (1959).
17. Kuang, J. J. & Chesson, P. Coexistence of annual plants: generalist seed predation weakens the storage effect. *Ecology*. (in the press).
18. Soule, M. E. & Terborgh, J. Conserving nature at regional and continental scales—a scientific program for North America. *Bioscience* **49,** 809–817 (1999).
19. Sala, E. Top predators provide insurance against climate change. *Trends Ecol. Evol.* **21,** 479–480 (2006).
20. Schmitz, O. J. Predators have large effects on ecosystem properties by changing plant diversity, not plant biomass. *Ecology* **87,** 1432–1437 (2006).
21. Johnson, C. N., Isaac, J. L. & Fisher, D. O. Rarity of a top predator triggers continent-wide collapse of mammal prey: dingoes and marsupials in Australia. *Proc. R. Soc. B* **274,** 341–346 (2007).
22. Borrvall, C. & Ebenman, B. Early onset of secondary extinctions in ecological communities following the loss of top predators. *Ecol. Lett.* **9,** 435–442 (2006).
23. Myers, R. A., Baum, J. K., Shepherd, T. D., Powers, S. P. & Peterson, C. H. Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* **315,** 1846–1850 (2007).
24. Heithaus, M. R., Frid, A., Wirsing, A. J. & Worm, B. Predicting ecological consequences of marine top predator declines. *Trends Ecol. Evol.* **23,** 202–210 (2008).

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

Acknowledgements This work was supported by National Science Foundation grants DEB-0542991 and DEB-0717222.

Author Contributions P.C. and J.J.K. jointly developed the model. P.C. derived the coexistence conditions and wrote the first draft of the manuscript. P.C. and J.J.K. jointly prepared the figures and all revisions of the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.C. (pchesson@u.arizona.edu) or J.J.K. (jjkuang@u.arizona.edu).

LETTERS

The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes

Chris Bowler[1,2], Andrew E. Allen[1,3], Jonathan H. Badger[3], Jane Grimwood[4], Kamel Jabbari[1], Alan Kuo[5], Uma Maheswari[1], Cindy Martens[6], Florian Maumus[1], Robert P. Otillar[5], Edda Rayko[1], Asaf Salamov[5], Klaas Vandepoele[6], Bank Beszteri[7], Ansgar Gruber[8], Marc Heijde[1], Michael Katinka[9], Thomas Mock[10]†, Klaus Valentin[7], Fréderic Verret[11], John A. Berges[12], Colin Brownlee[11], Jean-Paul Cadoret[13], Anthony Chiovitti[14], Chang Jae Choi[12], Sacha Coesel[2]†, Alessandra De Martino[1], J. Chris Detter[5], Colleen Durkin[10], Angela Falciatore[2], Jérome Fournet[15], Miyoshi Haruta[16], Marie J. J. Huysman[6,17], Bethany D. Jenkins[18], Katerina Jiroutova[19], Richard E. Jorgensen[20], Yolaine Joubert[15], Aaron Kaplan[21], Nils Kröger[22], Peter G. Kroth[8], Julie La Roche[23], Erica Lindquist[5], Markus Lommer[23], Véronique Martin–Jézéquel[15], Pascal J. Lopez[1], Susan Lucas[5], Manuela Mangogna[2], Karen McGinnis[20], Linda K. Medlin[7,11], Anton Montsant[1,2], Marie-Pierre Oudot–Le Secq[24], Carolyn Napoli[20], Miroslav Obornik[19], Micaela Schnitzler Parker[10], Jean-Louis Petit[9], Betina M. Porcel[9], Nicole Poulsen[25], Matthew Robison[16], Leszek Rychlewski[26], Tatiana A. Rynearson[27], Jeremy Schmutz[4], Harris Shapiro[5], Magali Siaut[2]†, Michele Stanley[28], Michael R. Sussman[16], Alison R. Taylor[11,29], Assaf Vardi[1,30], Peter von Dassow[31], Wim Vyverman[17], Anusuya Willis[14], Lucjan S. Wyrwicz[26], Daniel S. Rokhsar[5], Jean Weissenbach[9], E. Virginia Armbrust[10], Beverley R. Green[24], Yves Van de Peer[6] & Igor V. Grigoriev[5]

Diatoms are photosynthetic secondary endosymbionts found throughout marine and freshwater environments, and are believed to be responsible for around one-fifth of the primary productivity on Earth[1,2]. The genome sequence of the marine centric diatom *Thalassiosira pseudonana* was recently reported, revealing a wealth of information about diatom biology[3–5]. Here we report the complete genome sequence of the pennate diatom *Phaeodactylum tricornutum* and compare it with that of *T. pseudonana* to clarify evolutionary origins, functional significance and ubiquity of these features throughout diatoms. In spite of the fact that the pennate and centric lineages have only been diverging for 90 million years, their genome structures are dramatically different and a substantial fraction of genes (~40%) are not shared by these representatives of the two lineages. Analysis of molecular divergence compared with yeasts and metazoans reveals rapid rates of gene diversification in diatoms. Contributing factors include selective gene family expansions, differential losses and gains of genes and introns, and differential mobilization of transposable elements. Most significantly, we document the presence of hundreds of genes from bacteria. More than 300 of these gene transfers are found in both diatoms, attesting to their ancient origins, and many are likely to provide novel possibilities for metabolite management and for perception of environmental signals. These findings go a long way towards explaining the incredible diversity and success of the diatoms in contemporary oceans.

The sequenced diatoms represent two of the major classes of diatoms—the bi/multipolar centrics (Mediophyceae), to which *T. pseudonana* belongs, and the pennates (Bacillariophyceae), to which *P. tricornutum* belongs (Supplementary Fig. 1). The earliest fossil deposit from centrics is 180 million years (Myr) old and that from pennates is 90 Myr old[6,7]. Although being the youngest, the pennates are by far the most diversified, and they are major components of both

[1]CNRS UMR8186, Department of Biology, Ecole Normale Supérieure, 46 rue d'Ulm, 75005 Paris, France. [2]Stazione Zoologica 'Anton Dohrn', Villa Comunale, I-80121 Naples, Italy. [3]J. Craig Venter Institute, San Diego, California 92121, USA. [4]Joint Genome Institute-Stanford, Stanford Human Genome Center, 975 California Avenue, Palo Alto, California 94304, USA. [5]Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, California 94598, USA. [6]VIB Department of Plant Systems Biology, Ghent University, Technologiepark 927, B-9052 Ghent, Belgium. [7]Alfred Wegener Institute for Polar and Marine Research, Am Handelshafen 12, 27570 Bremerhaven, Germany. [8]Fachbereich Biologie, University of Konstanz, 78457 Konstanz, Germany. [9]Genoscope, CEA-Institut de Génomique, UMR CNRS no. 8030, 2 rue Gaston Crémieux, 91057 Evry Cedex, France. [10]School of Oceanography, University of Washington, Seattle, Washington 98195, USA. [11]Marine Biological Association of the UK, The Laboratory, Citadel Hill, Plymouth PL1 2PB, UK. [12]Department of Biological Sciences, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA. [13]PBA, IFREMER, BP 21105, 44311 Nantes Cedex 03, France. [14]School of Botany, The University of Melbourne, Victoria 3010, Australia. [15]EA 2160, Laboratoire 'Mer, Molécule, Santé', Faculté des Sciences et Techniques, Université de Nantes, 2 rue de la Houssinière, 44322, BP 92208, 44322 Nantes Cedex 3, France. [16]University of Wisconsin Biotechnology Center, 425 Henry Mall, Madison, Wisconsin 53706, USA. [17]Laboratory of Protistology and Aquatic Ecology, Ghent University, Krijgslaan 281-S8, B-9000 Ghent, Belgium. [18]Department of Cell and Molecular Biology and Graduate School of Oceanography, University of Rhode Island, 316 Morrill Hall, 45 Lower College Road, Kingston, Rhode Island 02881, USA. [19]Biology Centre ASCR, Institute of Parasitology and University of South Bohemia, Faculty of Science, Branisovska 31, 370 05 Ceske Budejovice, Czech Republic. [20]Bio5 Institute and Department of Plant Sciences, University of Arizona, Tucson, Arizona 85719, USA. [21]Department of Plant and Environmental Sciences, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel. [22]School of Chemistry and Biochemistry, School of Materials Science and Engineering, School of Biology, Georgia Institute of Technology, 901 Atlantic Drive NW, Atlanta, Georgia 30332-0400, USA. [23]Leibniz-Institut für Meereswissenschaften, 24105 Kiel, Germany. [24]Department of Botany, University of British Columbia, 3529-6270 University Boulevard, Vancouver, British Columbia V6T 1Z4, Canada. [25]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400, USA. [26]BioInfoBank Institute, Limanowskiego 24A/16, 60-744 Poznan, Poland. [27]Graduate School of Oceanography, University of Rhode Island, South Ferry Road, Narragansett, Rhode Island 02882-1197, USA. [28]Microbial & Molecular Biology, Scottish Association for Marine Science, Dunstaffnage Marine Laboratory, Oban, Argyll PA37 1QA, UK. [29]Department of Biology and Marine Biology, The University of North Carolina Wilmington, 601 South College Road, Wilmington, North Carolina 28403, USA. [30]Environmental Biophysics and Molecular Ecology Group, Institute of Marine and Coastal Sciences, Rutgers University, 71 Dudley Road, New Brunswick, New Jersey 08901, USA. [31]CNRS UMR7144, Station Biologique de Roscoff, Place George Teissier BP74, 29682 Roscoff Cedex, France. †Present addresses: University of East Anglia, School of Environmental Sciences, Norwich NR4 7TJ, UK (T.M.); Institute for Systems Biology, 1441 North 34th Street, Seattle, Washington 98103, USA (S.C.); CEA, DSV, IBEB, SBVME, UMR 6191 CNRS/CEA/Université Aix-Marseille, Laboratoire de Bioénergétique et Biotechnologie des Bactéries et Microalgues, Cadarache, Saint-Paul-lez-Durance F-13108, France (M.S.).

Table 1 | Major features of the *P. tricornutum* and *T. pseudonana* genomes

	P. tricornutum	*T. pseudonana*
Genome size	27.4 Mb	32.4 Mb
Predicted genes	10,402	11,776
Core genes*	3,523	4,332
Diatom-specific genes*	1,328	1,407
Unique genes*	4,366	3,912
Introns	8,169	17,880
Introns per gene	0.79	1.52
Long-terminal-repeat retrotransposon content	5.8%	1.1%

* Different classes of genes were assigned by comparing the *P. tricornutum* and *T. pseudonana* predicted proteomes with those from two plants, three green algae, one red alga, three metazoans, two fungi and ten other chromalveolates (see Supplementary Information) by all-against-all BLASTP using an expected cut-off value of 10^{-5}. Core genes were defined as being present in representatives from all these eukaryotic groups, diatom-specific genes were present in both of the diatoms but not elsewhere, and unique genes were only found in one of the two diatoms. The different numbers of diatom-specific genes in the two diatoms is a consequence of species-specific gene duplication events.

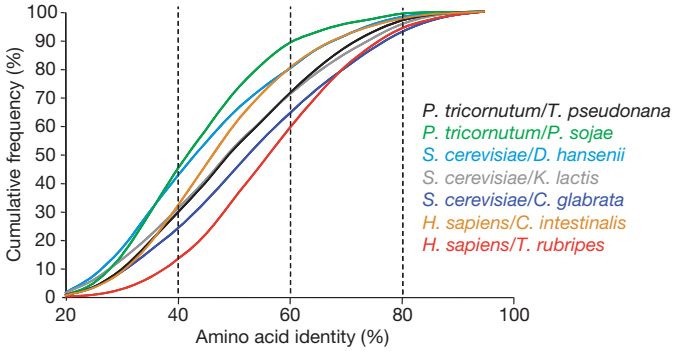


Figure 1 | Molecular divergence between *P. tricornutum* and *T. pseudonana*. Analysis of molecular divergence between the diatoms and other heterokonts, and comparison with selected hemiascomycetes and chordates. The diatom–oomycete pair displays the lowest amino acid identity (43.3%), in agreement with their proposed ancient separation, around 700 Myr ago[10]. The divergence between the pennate and centric diatom is similar to the fish–mammal divergence, which probably occurred in the Proterozoic era (550 Myr ago)[11]. The centric–pennate divergence, on the other hand, has been dated to at least 90 Myr ago[7]. In the figure, we represent the cumulative frequencies of amino acid identity across each set of potential orthologous pairs shown in Table 2.

pelagic and benthic habitats[7]. They display a range of features, including their bilateral symmetry, that distinguish them from centric species. For example, they have amoeboid isogametes, by contrast with the motile sperm and oogamy observed in centric species; they are major biofoulers; they include toxic species; and they generally respond most strongly to mesoscale iron fertilization[7,8]. Furthermore, members of the raphid pennate clade can glide actively along surfaces.

The completed *P. tricornutum* genome is approximately 27.4 megabases (Mb) in size, which is slightly smaller than *T. pseudonana* (32.4 Mb), and *P. tricornutum* is predicted to contain fewer genes (10,402 as opposed to 11,776; Table 1, Supplementary Fig. 2). Gene identification and functional analysis was facilitated by the availability of more than 130,000 expressed sequence tags (ESTs) generated from cells grown under 16 different conditions. In total, 86% of gene predictions had EST support (Supplementary Table 1).

P. tricornutum shares 57% of its genes with *T. pseudonana* (see Supplementary Information for criteria used), of which 1,328 are absent from other sequenced eukaryotes (Table 1). The molecular divergence between the two diatoms was assessed by examining the percentage amino acid identity of 4,267 orthologous gene pairs (Table 2, Fig. 1). We found an average identity of 54.9% between diatom orthologues, in comparison with approximately 43% between the diatoms and a more distantly related heterokont, the non-photosynthetic oomycete *Phytophthora sojae*. This agrees with the predicted ancient separation (around 700 Myr ago) of these lineages[9,10]. The divergence between the two diatoms is similar to what is observed between *Saccharomyces cerevisiae* and the related yeast *Kluyveromyces lactis*, and about halfway between the *Homo sapiens*/*Takifugu rubripes* (pufferfish) divergence and the *H. sapiens*/*Ciona intestinalis* (sea squirt) divergence (Table 2, Fig. 1). The more rapid evolutionary rates of diatoms compared with other organismal groups (for example, the fish–mammal divergence probably occurred in the Proterozoic era earlier than 550 Myr ago[11]) is consistent with previous observations[6,7]. As has been found in the two yeasts[12], no major conservation of gene order (synteny) could be detected between the two diatom genomes other than in a few examples of microclusters of up to eight genes (Supplementary Fig. 3). Furthermore, approximately two-thirds of intron positions are unique to each species (Supplementary Information). The widespread intron gain that has been reported in *T. pseudonana*[13] was not found in *P. tricornutum* (Table 1), suggesting that it may be a recent event in the centric diatom.

Large-scale within-genome duplication events do not appear to have played a major role in driving the generation of diatom diversity (Supplementary Information), by contrast with what has been found in yeasts and metazoans[14,15]. The observed high levels of diatom species diversity must therefore have been generated by other mechanisms. Whereas intron gain may be one factor in centric diatoms, the dramatic expansion of diatom-specific copia-retrotransposable elements may have contributed in the *P. tricornutum* genome (Table 1, Supplementary Figs 2, 4). These elements also appear to have expanded in other pennate diatoms (Supplementary Information), so they may have been a significant driving force in the generation of pennate diatom diversity through transpositional duplications and subsequent genome fragmentation.

Diatoms, and heterokonts in general, are believed to be derived from a secondary endosymbiotic process that took place around one billion years ago between a red alga and a heterotrophic eukaryote[16]. Diatom chloroplast genomes have fewer genes than red algal chloroplast genomes, indicating that a number of chloroplast genes were transferred to the nucleus after secondary endosymbiosis, and a few more genes appear to be in the process of transfer in one diatom species or the other[5]. It is generally thought that the diatom mitochondrion originated in the host, and the mitochondrial gene complement is almost identical to that of haptophytes and cryptophytes, which are other algal phyla that may have originated from the same secondary endosymbiotic event. We used a phylogenomic approach to search for genes of red algal origin in the two diatoms and the two sequenced oomycetes, *Phytophthora ramorum* and *Phytophthora sojae*[9], using *Cyanidioschyzon merolae* as reference red algal genome[17]. We classified 171 genes as being of red algal origin, on

Table 2 | Molecular divergence between different organism pairs

Pairwise comparison	Mean identity (%)	Number of compared pairs
Phaeodactylum tricornutum/*Thalassiosira pseudonana*	54.9	4,267
Phaeodactylum tricornutum/*Phytophthora sojae*	43.3	2,952
Saccharomyces cerevisiae/*Debaryomyces hansenii*	50.1	2,694
Saccharomyces cerevisiae/*Kluyveromyces lactis*	54.8	4,246
Saccharomyces cerevisiae/*Candida glabrata*	58.2	4,484
Homo sapiens/*Ciona intestinalis*	52.6	5,208
Homo sapiens/*Takifugu rubripes*	61.4	10,225

Summary of numbers of orthologous pairs (reciprocal best hits with an expected cut-off value of 10^{-10}) for each organism comparison and their mean percentage identities.

the basis of strong (>85%) bootstrap support for the red-alga-plus-heterokont clade (Supplementary Table 2). Of the 171 high-scoring genes, 108 were shared between the two diatoms and 74 (43%) were predicted to be plastid targeted. In addition, 11 of these genes were also present in oomycetes, as expected if the common ancestor of diatoms and oomycetes had a red algal plastid that was subsequently lost in the oomycetes[9]. The results of this survey support there being a red algal origin for the diatom plastid and many gene transfers from the red algal nucleus to the host nucleus before the former was lost.

A remarkably high number of *P. tricornutum* predicted genes appear to have been transferred between diatoms and bacteria (784; 7.5% of gene models). Specifically, by searching for orthologous genes in 739 prokaryotic genomes, followed by automated phylogenetic tree construction and manual curation, we confirmed that 587 putative *P. tricornutum* genes clustered with bacteria-only clades or formed a sister group to clades that included only bacterial genes (with or without other heterokonts). This finding indicates that horizontal gene transfer between bacteria and diatoms is pervasive and is much higher than has been found in other sequenced eukaryotes[18,19]. Of the 587 identified sequences, 42% are only found in *P. tricornutum* whereas 56% are present in both diatoms (Fig. 2a), attesting to their ancient origin. Only 73 sequences are shared between *P. tricornutum* and *Phytophthora* spp. (Fig. 2a, Supplementary Table 3), 59 of which are also present in *T. pseudonana*, suggesting that the vast majority of gene transfers occurred after the divergence of photosynthetic heterokonts and oomycetes.

Many of the genes shared between diatoms and bacteria encode components that are likely to provide novel metabolic capacities, for example for organic carbon and nitrogen utilization[20](xylanases and glucanases, prismane, carbon-nitrogen hydrolase, amidohydrolase), functioning of the diatom urea cycle[3] (carbamoyl transferase, carbamate kinase, ornithine cyclodeaminase) and polyamine metabolism related to diatom cell wall silicification[21] (*S*-adenosylmethionine (SAM)-dependent decarboxylases and methyltransferases). Others are likely to encode novel cell wall components, and to provide unorthodox mechanisms of DNA replication, repair and recombination for a eukaryotic cell (Supplementary Table 3).

Bacterial genes in diatoms do not appear to be derived from any one specific source, but from a range of origins including proteobacteria, cyanobacteria and archaea (Fig. 2a, b, Supplementary Table 3). Heterotrophic bacteria and cyanobacteria, especially diazotrophs and planctomycete bacteria, have been found in various close associations with diatoms[22–24], which may explain the unprecedented levels of horizontal gene transfer events that appear to have occurred. In *P. tricornutum*, bacterial genes are distributed throughout the genome, although several clusters, as well as regions devoid of bacterial genes, can be observed (Supplementary Fig. 5). Some of these genes in diatoms share bacterial-specific gene fusions that support phylogenetic associations, such as assimilatory nitrite reductase B and D subunits; these are apparently of planctomycete origin (Fig. 2c).

Bacterial histidine-kinase-based phosphorelay two-component systems, which are involved in environmental sensing, also appear to be highly developed in diatoms. For example, *P. tricornutum* contains a wide range of two-component signalling proteins sometimes organized in novel domain associations (Fig. 3). One of these proteins bears the classical features of bacterial phytochrome photoreceptors,

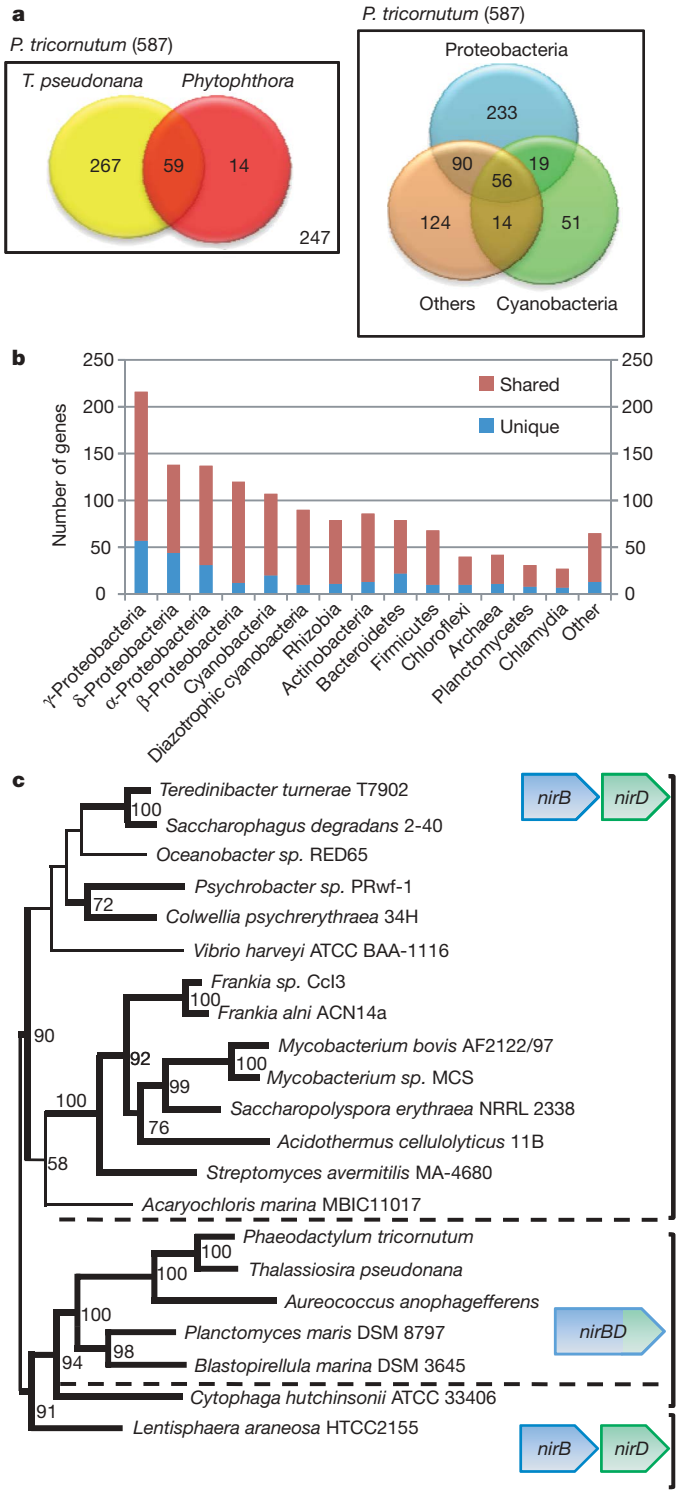


Figure 2 | Bacterial genes in diatoms. a, Venn diagrams showing how many of the bacterial genes identified in *P. tricornutum* are also found in other heterokonts (left panel), and which bacterial classes are most related phylogenetically (right panel). In each case, the Venn diagrams indicate the number of trees in which the designated taxa occur within the same clade or in a sister clade of *P. tricornutum*. **b**, Breakdown of different bacterial groups that occur in the same clade or in a sister clade of *P. tricornutum*. 'Unique' denotes a gene found only in a particular bacterial class; 'shared' denotes a gene that is most similar to a gene of that specific bacterial class but that is also present in other bacterial groups. **c**, PhyML maximum likelihood tree (log likelihood ratio, −22,358.321320) as inferred from the amino acid sequences of the large subunit of NAD(P)H assimilatory nitrite reductase (NirB). The choice of protein evolution model was WAG with gamma-distributed rates ($\alpha = 0.80$), as suggested by a ProtTest analysis of the alignment (see Supplementary Information for methods). Numbers above selected branches indicate maximum likelihood bootstrap support (100 replicates). Gene fusions and distinct open reading frames are indicated adjacent to the appropriate clades. In most cases, the large (NirB) and small (NirD) subunits of NAD(P)H assimilatory nitrite reductase are encoded by distinct open reading frames, but in diatoms and planctomycetes the *nirD* and *nirB* open reading frames have been fused to encode a single gene product. A total of 587 trees show evidence for prokaryotic origins of diatom genes and are available in Supplementary Information.

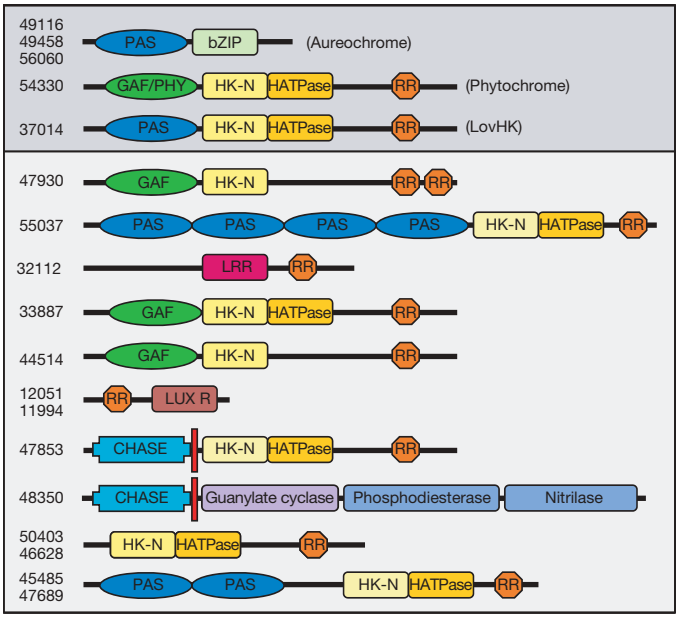


Figure 3 | Domain structures of two-component systems found in *P. tricornutum*. Domains likely to be involved in signalling are illustrated schematically and *P. tricornutum* protein identifiers are indicated on the left. Proteins corresponding to putative photoreceptors (aureochrome, phytochrome and LovHK) are indicated above (in grey, above the horizontal line). For further information about two-component systems, see Supplementary Information. Domain abbreviations are PAS, Per–Arnt–Sim; bZIP, basic region/leucine zipper; GAF, cGMP phosphodiesterase/adenylyl cyclase/FhlA; PHY, phytochrome; HK-N, histidine kinase N-terminal domain; RR, response regulator; LRR, leucine-rich repeat; LUX R, LuxR transcriptional activator; CHASE, cyclases/histidine kinases associated sensory extracellular.

as previously noted in *T. pseudonana*[3,4]. Another domain combination present in both diatoms resembles aureochrome blue-light photoreceptors[25], and *P. tricornutum* contains orthologues of LovHK and other light-dependent histidine kinases reported in bacteria[26,27].

To identify additional novel features of the diatom gene repertoire, we compared the gene family content of the two diatoms with other eukaryotes (Fig. 4, Supplementary Figs 6, 7). Diatoms contain many species-specific multicopy gene families, as well as large numbers of species-specific single-copy genes (denoted orphans in Fig. 4a). The higher number of species-specific gene families in *P. tricornutum* may suggest that the more recent pennate diatoms possess more specialized functions, perhaps related to the heterogeneity of the benthic environments that they commonly inhabit. The centric diatom, by contrast, has retained more features found in other eukaryotes (Fig. 4b, Table 1), such as the flagellar apparatus[28]. We found a similar number of diatom-specific gene families (1,011) and eukaryotic gene families not found in diatoms (1,062), revealing that the rates of gene gain and gene loss are very similar and consistent with the high diversification rates observed in diatoms. We also found that diatom-specific genes are evolving faster than other genes in diatom genomes (Fig. 4c), providing a further explanation for the rapid diatom divergence rates[6,7].

Of the gene families found in the diatoms, some contain higher numbers of genes in comparison with other eukaryotes (Supplementary Table 4, Supplementary Fig. 7); for example, genes involved in polyamine metabolism are over-represented. The expansion of polyamine-related components is of interest in consideration of the role of long-chain polyamines in silica nanofabrication[21]. Of the eight predicted spermine/spermidine synthase-like genes in *P. tricornutum*, three encode potentially bifunctional enzymes bearing both an aminopropyltransferase domain and a SAM decarboxylase domain. Interestingly, the only other organisms containing such bifunctional proteins are *T. pseudonana* (four copies) and the bacteria *Bdellovibrio bacteriovorus* and *Delftia acidovorans*. Silaffins and silacidins are proteins/peptides believed to be involved in diatom silica formation[21,29]. *P. tricornutum* contains only one silaffin-like protein, and no homologues of silacidin. Frustulin genes, encoding proteins that form organic constituents of the biosilica cell wall but

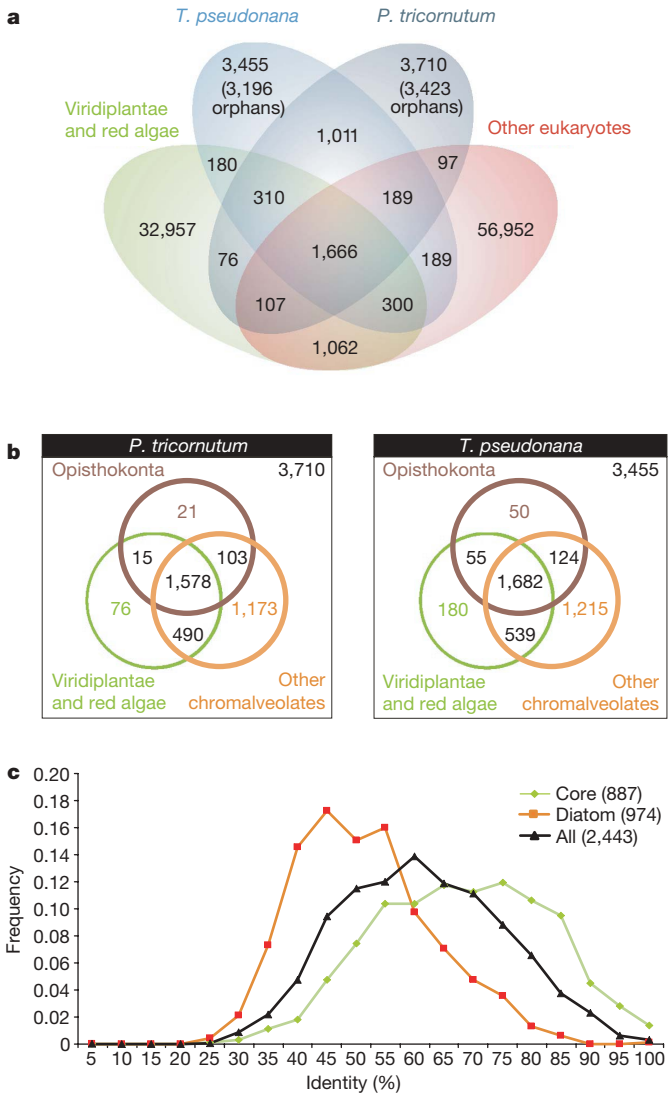


Figure 4 | Shared and unique gene families. **a**, Venn diagram representation of shared/unique gene families in *P. tricornutum*, *T. pseudonana*, Viridiplantae (plants and green algae) and red algae, and other eukaryotes (that is, other chromalveolates and Opisthokonta (fungi and metazoa)). In addition to the total number of gene families specific to *P. tricornutum* and *T. pseudonana*, the number of families consisting of a single gene (denoted 'orphans') is also indicated. For example, of the 3,710 gene families that are only found in *P. tricornutum*, 3,423 consist of single-copy genes whereas 287 gene families have at least two members. **b**, Venn diagram of the distribution of *P. tricornutum* (left panel) and *T. pseudonana* (right panel) gene families with homology to proteins from the Viridiplantae & red algae, Opisthokonta and other chromalveolates (including the other diatom). The numbers outside the circles indicate the number of *P. tricornutum* (left panel) or *T. pseudonana* (right panel) gene families with no homology to the examined proteomes. **c**, Percentage amino acid identity plot of orthologues (based on reciprocal best hits) of different classes of diatom genes identified in **a**. Numbers in parentheses indicate the numbers of orthologues per class. 'Diatom' corresponds to genes only found in *P. tricornutum* and *T. pseudonana* (members of the 1,011 gene families shown in **a**), 'core' corresponds to genes present in all eukaryotic groups (members of the 1,666 gene families shown in **a**), and 'all' corresponds to all orthologous gene pairs in *P. tricornutum* and *T. pseudonana*.

are not involved in silica formation, are present in large numbers and are highly expressed. Both diatoms contain a similar number of silicic acid transporters.

Other noteworthy diatom-specific expansions include histidine kinases (see above and Fig. 3), cyclins and heat-shock transcription factors. Cyclins are major regulators of the cell cycle in eukaryotes. In addition to members of each of the canonical families of cyclins, we found 10 and 42 diatom-specific cyclin genes in *P. tricornutum* and *T. pseudonana*, respectively. The dramatic expansion of this gene family may reflect the unusual characteristics of diatom life cycles due to the rigid nature of their cell walls, such as the control of cell size reduction, the activation of sexual reproduction at a critical size threshold, and life in rapidly changing and unpredictable environments[7]. Conversely, it may be significant that genes encoding RCC1 proteins (RCC, regulator of chromosome condensation), which are also involved in cell cycle control, have been expanded in both diatom genomes (Supplementary Table 4). For the putative heat-shock transcription factors, we found 69 copies in *P. tricornutum* and 89 copies in *T. pseudonana*[4]. These numbers represent almost 50% of the total number of transcription factors in the two sequenced diatoms. The significance of this expansion is unclear, but EST data indicates that the majority are expressed and that some are induced specifically in response to certain growth conditions (Supplementary Fig. 8).

In conclusion, through our comparative analyses we have revealed diverse origins of diatom genes. Diatom-specific genes may have arisen from genome rearrangements and subsequent domain recombinations due to the action of diatom-specific transposable elements, from selective gene family expansions/contractions and from intron gain/loss. It was previously shown that diatoms have retained genes from both partners of the secondary endosymbiosis[3], thus bringing together primary metabolic processes such as photosynthetic carbon fixation and organic nitrogen production by means of the urea cycle in a single organism[30]. Our studies now suggest that genes acquired after secondary endosymbiosis by gene transfer from bacteria are pervasive in diatoms and represent at least 5% of their gene repertoires. This level of horizontal gene transfer is around one order of magnitude higher than has been found in other free-living eukaryotes, and is similar to the rates found between bacteria[19]. Although our analyses may be biased by the currently poor taxon sampling of whole genome sequences in eukaryotes (relative to that for prokaryotes), they are nonetheless supported by molecular phylogenies. We therefore propose that gene transfer from bacteria to diatoms, and perhaps vice versa, has been a common event in marine environments and has been a major driving force during diatom evolution. It has also brought together highly unorthodox combinations of genes permitting non-canonical management of carbon and nitrogen in primary metabolism and the sensing of external stimuli adapted to aquatic environments. The combination of mechanisms reported here may underlie the rapid diversification rates observed in diatoms and may explain why they have come to dominate contemporary marine ecosystems in a relatively short period of time.

METHODS SUMMARY

High-molecular-weight DNA was extracted from axenic cultures of *P. tricornutum* accession Pt1 8.6 (deposited as CCMP2561 in the Provasoli–Guillard National Center for Culture of Marine Phytoplankton) and used to construct replicate libraries containing inserts of 2–3 kb, 6–8 kb and 35–40 kb. Using the Joint Genome Institute (JGI) JAZZ assembler, approximately 556,000 reads involving 564 Mb of sequence were trimmed, filtered for short reads and assembled. All low-quality areas and gaps were identified and converted into targets for manual finishing. The draft genome sequence of *T. pseudonana*[3] was finished in a similar way. Both diatom genomes were annotated using the JGI annotation pipeline, which combines several gene prediction, annotation and analysis tools. Complementary DNA libraries were constructed from messenger RNA extracted from *P. tricornutum* cultures grown under 16 different conditions. More than 130,000 ESTs were generated. Full information about all methods used for the analyses reported here is available in Supplementary Information.

Received 20 June; accepted 8 September 2008.
Published online 15 October 2008.

1. Falkowski, P. G., Barber, R. T. & Smetacek, V. Biogeochemical controls and feedbacks on ocean primary production. *Science* **281,** 200–206 (1998).
2. Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281,** 237–240 (1998).
3. Armbrust, E. V. *et al.* The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. *Science* **306,** 79–86 (2004).
4. Montsant, A. *et al.* Identification and comparative genomic analysis of signaling and regulatory components in the diatom *Thalassiosira pseudonana*. *J. Phycol.* **43,** 585–603 (2007).
5. Oudot-Le Secq, M.-P. *et al.* Chloroplast genomes of the diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*: comparison with other plastid genomes of the red lineage. *Mol. Gen. Genom.* **277,** 427–439 (2007).
6. Sims, P. A., Mann, D. G. & Medlin, L. K. Evolution of the diatoms: insights from fossil, biological and molecular data. *Phycologia* **45,** 361–402 (2006).
7. Kooistra, W. H. C. F., Gersonde, R., Medlin, L. K. & Mann, D. G. in *Evolution of Primary Producers in the Sea* (eds Falkowski, P. G. & Knoll, A. H.) 207–249 (Academic Press, 2007).
8. de Baar, H. J. W. *et al.* Synthesis of iron fertilization experiments: From the Iron Age in the Age of Enlightenment. *J. Geophys. Res.* **110,** doi:10.1029/2004JC002601 (2005).
9. Tyler, B. M. *et al. Phytophthora* genome sequences uncover evolutionary origins and mechanisms of pathogenesis. *Science* **313,** 1261–1266 (2006).
10. Yoon, H. S., Hackett, J. D., Ciniglia, C., Pinto, G. & Bhattacharya, D. A molecular timeline for the origin of photosynthetic eukaryotes. *Mol. Biol. Evol.* **21,** 809–818 (2004).
11. Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* **392,** 917–920 (1998).
12. Dujon, B. Yeasts illustrate the molecular mechanisms of eukaryotic genome evolution. *Trends Genet.* **22,** 375–387 (2006).
13. Roy, S. W. & Penny, D. A very high fraction of unique intron positions in the intron-rich diatom *Thalassiosira pseudonana* indicates widespread intron gain. *Mol. Biol. Evol.* **24,** 1447–1457 (2007).
14. Scannell, D. R., Byrne, K. P., Gordon, J. L., Wong, S. & Wolfe, K. H. Multiple rounds of speciation associated with reciprocal gene loss in polyploid yeasts. *Nature* **440,** 341–345 (2006).
15. Semon, M. & Wolfe, K. H. Reciprocal gene loss between Tetraodon and zebrafish after whole genome duplication in their ancestor. *Trends Genet.* **23,** 108–112 (2007).
16. Bhattacharya, D., Archibald, J. M., Weber, A. P. & Reyes-Prieto, A. How do endosymbionts become organelles? Understanding early events in plastid evolution. *Bioessays* **29,** 1239–1246 (2007).
17. Matsuzaki, M. *et al.* Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **428,** 653–657 (2004).
18. Martens, C., Vandepoele, K. & Van de Peer, Y. Whole-genome analysis reveals molecular innovations and evolutionary transitions in chromalveolate species. *Proc. Natl Acad. Sci. USA* **105,** 3427–3432 (2008).
19. Keeling, P. J. & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nature Rev. Genet.* **9,** 605–618 (2008).
20. Kroth, P. G. *et al.* A model for carbohydrate metabolism in the diatom Phaeodactylum tricornutum deduced from whole genome analysis and comparative genomic analyses with Thalassiosira pseudonana and other photoautotrophs. *PLoS One* **3,** e1426 (2008).
21. Kröger, N. Prescribing diatom morphology: toward genetic engineering of biological nanomaterials. *Curr. Opin. Chem. Biol.* **11,** 662–669 (2007).
22. Carpenter, E. J. & Janson, S. Intracellular cyanobacterial symbionts in the marine diatom *Climacodium frauenfeldianum* (Bacillariophyceae). *J. Phycol.* **36,** 540–544 (2000).
23. Schmid, A.-M. M. Endobacteria in the diatom *Pinnularia* (Bacillariophyceae). I. ''Scattered ct-nucleoids'' explained: DAPI-DNA complexes stem from exoplastidial bacteria boring into the chloroplasts. *J. Phycol.* **39,** 122–138 (2003).
24. Zehr, J. P., Carpenter, E. J. & Villareal, T. A. New perspectives on nitrogen-fixing microorganisms in tropical and subtropical oceans. *Trends Microbiol.* **8,** 68–73 (2000).
25. Takahashi, F. *et al.* AUREOCHROME, a photoreceptor required for photomorphogenesis in stramenopiles. *Proc. Natl Acad. Sci. USA* **104,** 19625–19630 (2007).
26. Purcell, E. B., Siegal-Gaskins, D., Rawling, D. C., Fiebig, A. & Crosson, S. A photosensory two-component system regulates bacterial cell attachment. *Proc. Natl Acad. Sci. USA* **104,** 18241–18246 (2007).
27. Swartz, T. E. *et al.* Blue-light-activated histidine kinases: two-component sensors in bacteria. *Science* **317,** 1090–1093 (2007).
28. Merchant, S. S. *et al.* The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318,** 245–250 (2007).
29. Sumper, M. & Brunner, E. Silica biomineralization in diatoms: the model organism *Thalassiosira pseudonana*. *ChemBioChem* **9,** 1187–1194 (2008).
30. Allen, A. E., Vardi, A. & Bowler, C. An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Curr. Opin. Plant Biol.* **9,** 264–273 (2006).

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

Acknowledgements Diatom genome sequencing at the JGI (USA) was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Berkeley National Laboratory, under contract no. DE-AC02-05CH11231, Lawrence Livermore National Laboratory under contract no. DE-AC52-07NA27344 and Los Alamos National Laboratory under contract no. DE-AC02-06NA25396. *P. tricornutum* ESTs were generated at Genoscope (France). Funding for this work was also obtained from the EU-funded FP6 Diatomics project (LSHG-CT-2004-512035), the EU-FP6 Marine Genomics Network of Excellence (GOCE-CT-2004-505403), an ATIP 'Blanche' grant from the CNRS (France) and the Agence Nationale de la Recherche (France). We are grateful to M. Muffato and H.-R. Crollius for the analysis reported in Supplementary Fig. 3a.

Author Contributions C.B. coordinated *Phaeodactylum* genome annotation and manuscript preparation. E.V.A. coordinated *Thalassiosira* genome annotation. D.S.R. and I.V.G. coordinated diatom genome sequencing and analysis at JGI. J.W. coordinated EST sequencing at Genoscope. A.E.A., J.H.B., J.G., K.J., A.K., U.M., C.M., F.M., R.P.O., E.R., A.S. and K.V. made equivalent and substantial contributions to the data presented, and should be considered joint second authors. B.B., A.G., M.H., M.K., T.M., K.V. and F.V. also made significant contributions. A.E.A., E.V.A., B.R.G., Y.V.d.P. and I.V.G. assisted in data interpretation and manuscript preparation. Other authors contributed as members of the *Phaeodactylum* genome sequencing consortium.

Author Information Assemblies and annotations of the *P. tricornutum* and *T. pseudonana* genomes are available through the JGI Genome Portal at http://www.jgi.doe.gov/phaeodactylum and http://www.jgi.doe.gov/thalassiosira. Genome assemblies together with predicted gene models and annotations have been deposited at DDBJ, EMBL and GenBank under the project accessions ABQD00000000 and AAFD00000000, respectively. The versions described in this paper are the first version, ABQD01000000, for *P. tricornutum*, which includes complete chromosomes 3 and 11 (CP001142 and CP001141), and the second version, AAFD02000000, for *T. pseudonana*, also including complete chromosomes 7 and 18 (CP001160 and CP001159). *P. tricornutum* EST expression profiles can be found at http://www.biologie.ens.fr/diatomics/EST3, which also provides links to gene models on the JGI genome browser. ESTs have been deposited at NCBI dbEST with GenBank accession numbers CD374840-CD384835, BI306757-BI307753, CT868744-CT950687 and CU695349-CU740080. Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. Correspondence and requests for materials should be addressed to C.B. (cbowler@biologie.ens.fr).

LETTERS

Associative learning of social value

Timothy E. J. Behrens[1,2]*, Laurence T. Hunt[1,2]*, Mark W. Woolrich[1] & Matthew F. S. Rushworth[1,2]

Our decisions are guided by information learnt from our environment. This information may come via personal experiences of reward, but also from the behaviour of social partners[1,2]. Social learning is widely held to be distinct from other forms of learning in its mechanism and neural implementation; it is often assumed to compete with simpler mechanisms, such as reward-based associative learning, to drive behaviour[3]. Recently, neural signals have been observed during social exchange reminiscent of signals seen in studies of associative learning[4]. Here we demonstrate that social information may be acquired using the same associative processes assumed to underlie reward-based learning. We find that key computational variables for learning in the social and reward domains are processed in a similar fashion, but in parallel neural processing streams. Two neighbouring divisions of the anterior cingulate cortex were central to learning about social and reward-based information, and for determining the extent to which each source of information guides behaviour. When making a decision, however, the information learnt using these parallel streams was combined within ventromedial prefrontal cortex. These findings suggest that human social valuation can be realized by means of the same associative processes previously established for learning other, simpler, features of the environment.

To compare learning strategies for social and reward-based information, we constructed a task in which each outcome revealed information both about likely future outcomes (reward-based information) and about the trust that should be assigned to future advice from a confederate (social information).

Twenty-four subjects performed a decision-making task requiring the combination of information from three sources (Fig. 1, Methods and Supplementary Information): (1) the reward magnitude of each option (generated randomly at each trial); (2) the likely correct response (blue or green) based on their own experience of rewards on each option; and (3) the confederate's advice, and how trustworthy the confederate currently was. When a new outcome was witnessed, subjects could use this single outcome to learn in parallel about the likely correct action, and the trustworthiness of the confederate.

The investigation resembles previous experiments that have compared animate and inanimate conditions in different trials or experiments[5,6]. Here, however, both sources of information were present on each trial outcome but the relevance of each was manipulated continuously allowing determination of both the functional magnetic resonance imaging (fMRI) signal and the behavioural influence associated with each source of information.

Optimal behaviour in this task requires the subject to track the probability of the correct action and the probability of correct advice independently, and to combine these two probabilities into an overall probability of the correct response (Supplementary Information). Computational models of reinforcement learning (RL) have had considerable success in predicting how such probabilities are tracked in learning tasks outside the social domain[7]. The simplest RL models integrate information over trials by maintaining and updating the expected value of each option. When new information is observed this value is updated by the product of the prediction error and the learning rate[7]. In our task, there are two dissociable prediction errors: the reward prediction error (actual reward − expected value), for learning about the correct option, and the confederate prediction error (actual − expected fidelity), for learning about the trustworthiness of the confederate. The optimal learning rate depends on the volatility of the underlying information source[8–10]. In volatile conditions, subjects should give more weight to recent information, using a fast learning rate. In stable conditions, subjects should weigh recent and historical information almost equally, using a slow learning rate. By ensuring that the correct option and the confederate's advice became volatile at different times, we ensured that the learning rate for these two sources of information varied independently. We used a Bayesian RL[8] model (Supplementary Information) to generate the optimal estimates of prediction error, volatility and outcome probability separately for each source of information (Fig. 1b–d).

We first sought to establish whether human behaviour matched predictions from the RL model. We used logistic regression to determine the degree to which subject choices were influenced by the optimally tracked confederate and outcome probabilities, and by the difference in reward magnitudes between options. Parameter estimates for all three information sources were significantly greater than zero, and there was no significant difference in the degree to which subjects used reward and social information to determine their behaviour (Fig. 1e). Furthermore there was no significant effect either of subjects blindly following confederate advice without learning its value, or of subjects assuming that the confederate would behave in the same way as the previous trial (Fig. 1e). Hence subjects were able to integrate the fidelity of the confederate over many trials in an RL-like fashion.

We then investigated whether the fMRI signal reflected the model's estimates of prediction error and volatility, for both social and reward information, when subjects witnessed new outcomes. In the reward domain, neural responses have been identified that encode these key parameters[8,11–16]. Dopamine neurons in the ventral tegmental area (VTA) code reward prediction errors[12,13,17]. Similar signals are reported in the dopaminoceptive striatum[11,18] and even in the VTA itself, when specialized strategies are used in human fMRI studies[19]. fMRI correlates of the learning rate in the reward domain have been reported in anterior cingulate sulcus (ACCs)[8]. If humans can learn from social information in a similar fashion, it should be possible to detect signals that co-vary with the same computational parameters, but in the social domain.

We observed blood-oxygen-level-dependent (BOLD) correlates of the confederate prediction error in dorsomedial prefrontal cortex (DMPFC) in the vicinity of the paracingulate sulcus, right middle temporal gyrus (MTG), and in the right superior temporal sulcus at the temporoparietal junction (STS/TPJ) (Fig. 2a). Equivalent signals

[1]FMRIB Centre, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, UK. [2]Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.
*These authors contributed equally to this work.

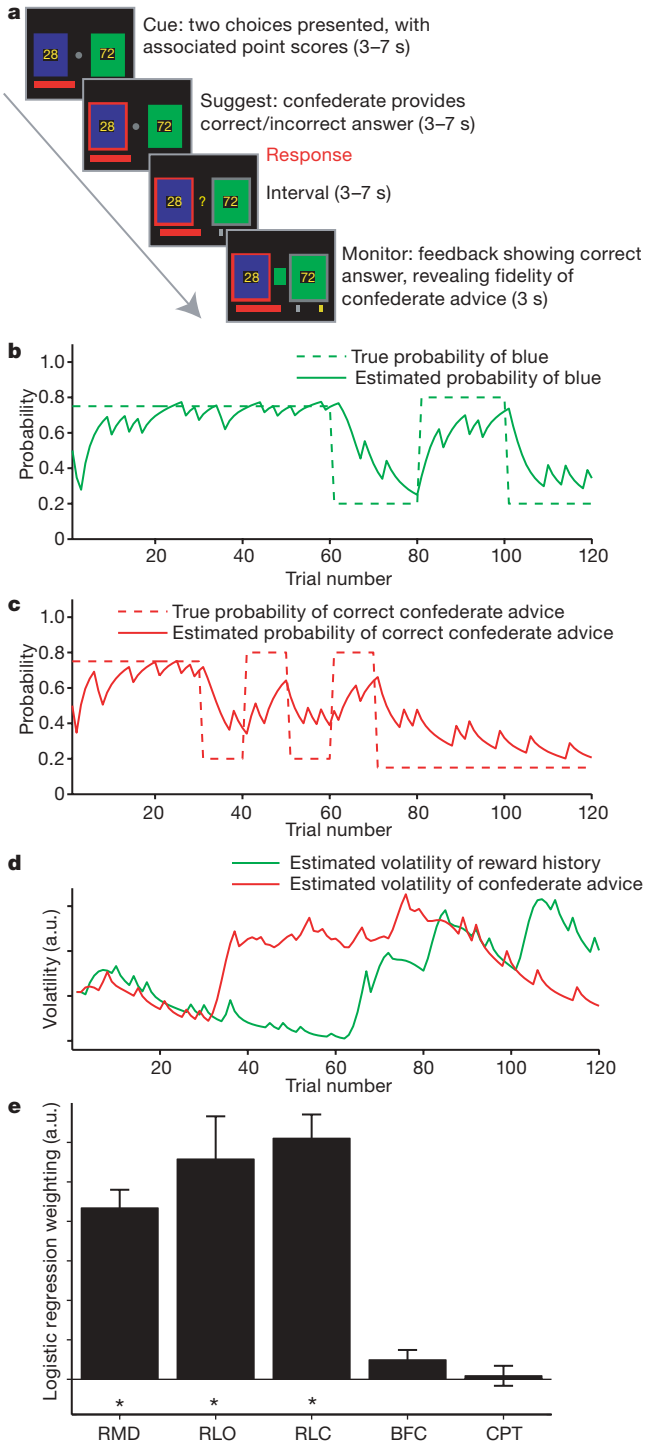


Figure 1 | Experimental task and behavioural findings. a, Experimental task (see Methods and Supplementary Information). Each trial consists of four phases. Subjects are presented with a decision (Cue), receive the advice (red square) of the confederate (Suggest) and respond using a button press (grey square). An 'Interval' period follows, before the correct outcome is revealed (Monitor). If the subject chooses correctly the red bar is incrementally increased by the number of points on the chosen option. **b**, **c**, Reward schedules for reward (**b**) and social (**c**) information. Dashed lines show the true probability of blue being correct (**b**) and the true probability of correct confederate advice (**c**). Each schedule underwent periods of stability and volatility. Solid lines show the model's estimate of the probabilities. **d**, Optimal model estimates of the volatility of reward (green) and social (red) information. **e**, Logistic regression on subject behaviour. Factors included were the reward magnitude difference between options (RMD); the outcome probability derived from the model using reward outcomes (RLO); the outcome probability derived from the model using confederate advice (RLC); the possibility that the subjects would blindly follow the confederate without learning (BFC); and the possibility that subjects would assume the confederate would behave as in the previous trial (CPT). The logistic regression analysis revealed significant effects only on RMD, RLO and RLC (asterisks). Error bars show s.e.m.; a.u., arbitrary units.

were present in the left hemisphere at the same threshold, but did not pass the cluster extent criterion; similar effects were also found bilaterally in the cerebellum (Supplementary Information). Notably, these regions showed a pattern of activation similar to known dopaminergic activity in reward learning[13], but for social information. Activity correlated with the probability of a confederate lie after the subject decision but before the outcome was revealed (a prediction signal). When the subjects observed the trial outcome, activity correlated negatively with this same probability, but positively with the event of a confederate lie (Fig. 2b). This signal reflects both components of a prediction error signal for social information: the outcome (lie or truth) minus the expectation (Fig. 2b). These signals cannot be influenced by reward prediction errors as the two types of prediction error were decorrelated in the task design. The presence of this prediction error signal in the brain is a prerequisite for any theory of an RL-like strategy for social valuation.

We performed a similar analysis for prediction errors on reward information (reward minus expected reward). We found a significant effect of reward prediction error in the ventral striatum (Fig. 2c), the ventromedial prefrontal cortex, and anterior cingulate sulcus (see Supplementary Information). As in the social domain, we observed significant effects of all three elements of the reward prediction error (Fig. 2d; see Supplementary Information for discussion).

As previously demonstrated[8], the volatility of action–outcome associations predicted BOLD signal in a circumscribed region of the ACCs (Fig. 3a). This effect varied across people such that those whose behaviour relied more on their own experiences (Supplementary Information) showed a greater volatility-related signal in this region (Fig. 3b). The volatility of confederate advice correlated with BOLD signal in a circumscribed region in the adjacent ACC gyrus (ACCg) (Fig. 3a). Subjects whose behaviour relied more on this advice showed greater signal change in this region (Fig. 3c). Notably, this double dissociation (reflected in a three-way interaction between area (ACCs versus ACCg), volatility type (social versus outcome) and degree of reliance on social ($F_{1,20} = 7.145$, $P = 0.015$) or experiential information ($F_{1,20} = 5.379$, $P = 0.031$) in an analysis of covariance) can be understood by reference to a dissociation in macaque monkeys. Selective lesions to ACCs but not ACCg impair reward-guided decision making in the reward domain[20]. In the social domain, male macaques will forego food to acquire information about other individuals[21,22]. Selective lesions to ACCg but not ACCs abolish this effect[23]. We found that BOLD signals in these two regions reflect the respective values of the same outcome for learning about the two different sources of information.

Learning about reward probability from vicarious and personal experiences recruits distinct neural systems, but subjects combine information across both sources when making decisions (Fig. 1e). A ventromedial portion of the prefrontal cortex (VMPFC) has been shown to code such an expected value signal for the chosen action[24,25] during decision making.

We computed two probabilities of reward on the subject's chosen option: one based only on experience and one based only on confederate advice. BOLD signal in the VMPFC was significantly correlated with both probabilities (Fig. 4a and Supplementary Fig. 4). However, there was subject variability in whether the VMPFC signal better reflected the reward probability based on outcome history or on social information. The extent to which the VMPFC data reflected each source of information (at the time of the decision) was predicted by the ACCs/ACCg response to outcome/social volatility (at the time when the outcomes were witnessed) (Fig. 4b, c).

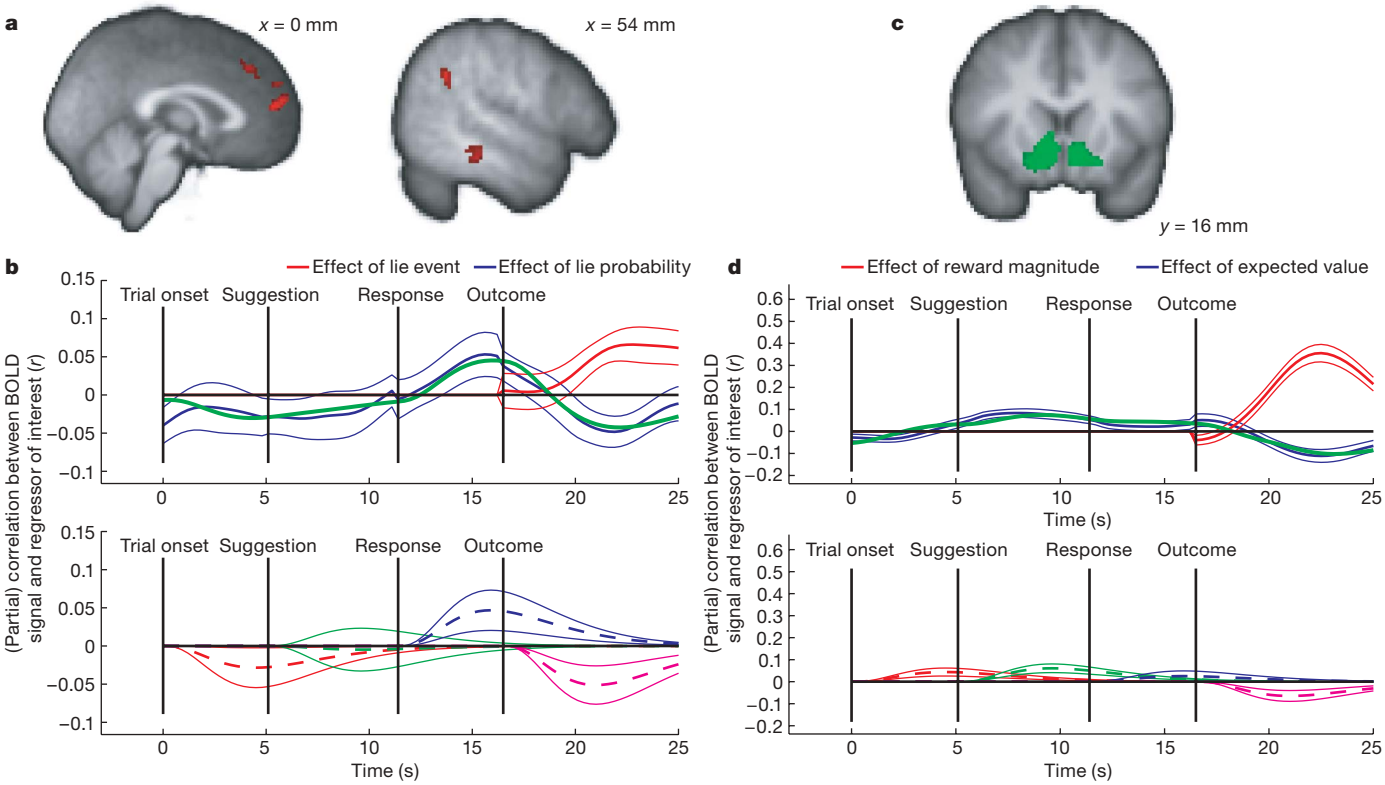


Figure 2 | Predictions and prediction errors in social and non-social domains. Time courses show (partial) correlations ± s.e.m. (See Supplementary Fig. 2.) **a**, Activation in the DMPFC, right TPJ/STS and MTG correlate with the social prediction error at the outcome (threshold set at $Z > 3.1$, cluster size >50 voxels). **b**, Deconstruction of signal change in the DMPFC. Similar results were found in the MTG and TPJ/STS. Top: following the outcome, areas that encode prediction error correlate positively with the outcome and negatively with the predicted probability. Red, effect size of the confederate lie outcome (1 for lie, 0 for truth); blue, effect size of the predicted confederate lie probability. To perform inference, we fit a haemodynamic model in each subject to the time course of this effect (that is, to the blue line). The green line in the top panel shows the mean overall fit of this haemodynamic model (for comparison with the blue line). Bottom: the effect of lie probability (blue line from top panel) is decomposed into a haemodynamic response function at each trial event (corresponding to the four colours in the bottom panel) (see Supplementary Fig. 2). Dashed and solid lines show mean responses ± s.e.m. Each region showed a significant positive effect of predicted confederate lie probability after the decision ($t_{22} = 1.96$ ($P < 0.05$), 1.73 ($P < 0.05$), 1.74 ($P < 0.05$) for DMPFC, MTG and TPJ/STS, respectively). Crucially, each brain region showed a significant negative effect of predicted confederate lie probability after the outcome ($t_{22} = 2.68$ ($P < 0.005$), 2.35 ($P < 0.05$), 3.27 ($P < 0.005$)). **c**, Ventral striatum is taken as an example of a number of regions revealed by the voxel-wise analysis of reward prediction error (threshold set at $Z > 3.1$, cluster size >100 voxels). **d**, Panels are exactly as in **b**, but coded in terms of reward and not in terms of confederate fidelity. The top panel shows the parameter estimate relating to the expected value of the trial (blue line) and, after the outcome, the parameter estimate relating to the magnitude of these rewards (red line). To test for prediction error coding, we again fit a haemodynamic model to the expectation parameter estimate (shown by the green line, for comparison with blue line). Bottom panel: the time course showed a significant positive effect during the time of the decision ($t_{22} = 3.32$ ($P < 0.002$)), and a significant negative effect after the trial outcome ($t_{22} = 2.50$ ($P < 0.05$)). (See Supplementary Information for further discussion.)

Here, we have shown that the weighting assigned to social information is subject to learning and continual update via associative mechanisms. We use techniques that predict behaviour when learning from personal experiences to show that similar mechanisms explain behaviour in a social context. Furthermore, we demonstrate fundamental similarities between the neural encoding of key parameters for reward-based and social learning. Despite using similar mechanisms, distinct anatomical structures code learning parameters in the two domains. However, information from both is combined in ventromedial prefrontal cortex when making a decision.

By comparing the two sources of information, we find that social prediction error signals similar to those reported in dopamine neurons for reward-based learning are coded in the MTG, STS/TPJ and DMPFC. BOLD signal fluctuations in these regions are often seen in social tasks[26,27], and in tasks which involve the attribution of motive to stimuli[28]. Such activations have been thought critical in studies of the theory of mind[28]. That these regions should code quantitative prediction and prediction error signals about a confederate lends more weight to the argument that social evaluation mechanisms are able to rely on simple associative processes.

A second crucial parameter in reinforcement learning models is the learning rate, reflecting the value of each new piece of information. In the context of reward-based learning, this parameter predicts BOLD signal fluctuations in the ACCs at the crucial time for learning[8]—a finding that is replicated here. We further demonstrate that the exact same computational parameter, in the context of social learning, predicts BOLD fluctuations in the neighbouring ACCg. This functional dissociation is mirrored by differences in the regions' anatomical connectivity. In the macaque monkey, connections with motor regions lie predominantly in ACCs[29], giving access to information about the monkey's own actions. Connections with visceral and social regions, including the STS, lie predominantly in ACCg[29], giving access to information about other agents. Nevertheless, that it is the same computational parameter that is represented in ACCs and ACCg suggests that parallel streams of learning occur within ACC for social and non-social information.

It has been suggested that VMPFC activity might represent a common currency in which the value of different types of items might be encoded[25,30]. Here we show that the same portion of the VMPFC represents the expected value of a decision based on the combination

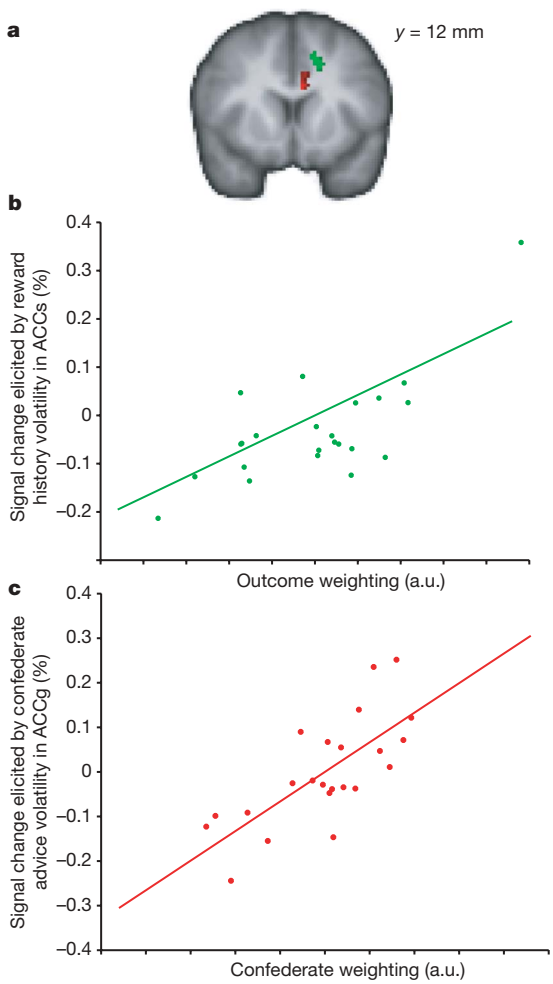


Figure 3 | Agency-specific learning rates dissociate in the ACC. **a**, Regions where the BOLD correlates of reward (green) and confederate (red) volatility predict the influence that each source of information has on subject behaviour ($Z > 3.1$, $P < 0.05$ cluster-corrected for cingulate cortex). **b**, Subjects with high BOLD signal changes in response to reward volatility in the ACCs are guided strongly by reward history information (maximum $Z = 3.7$, correlation $R = 0.7163$, $P < 0.0001$). **c**, Subjects with high BOLD signal changes in response to confederate advice volatility in the ACCg are guided strongly by social information (maximum $Z = 4.1$, correlation $R = 0.7252$, $P < 0.0001$). See Supplementary Information.

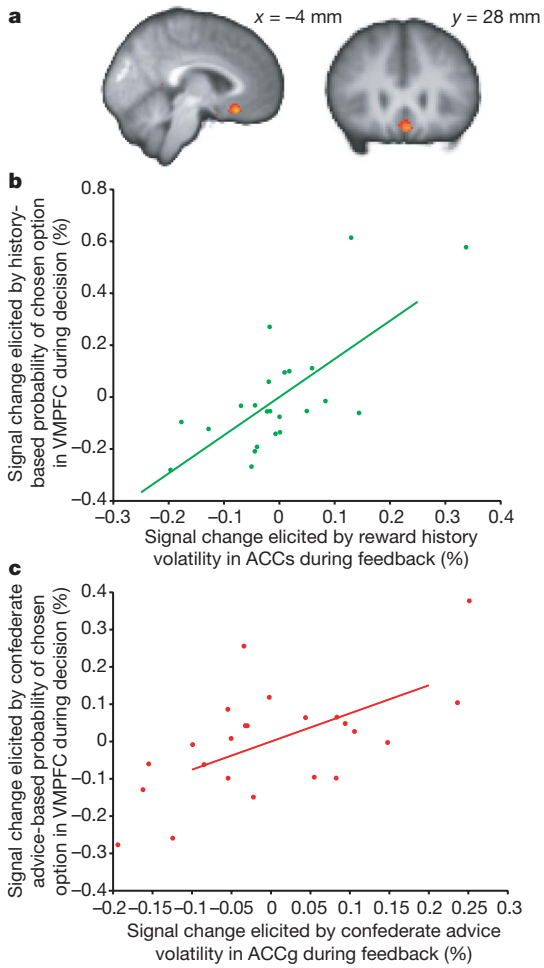


Figure 4 | Combination of expected value of chosen option in VMPFC. **a**, Activation for the combination (mean contrast) of experience-based probability during the Cue and Suggest phases, and advice-based probability during the Suggest phase (threshold set at $Z > 3.1$, $P < 0.005$ cluster-corrected for VMPFC). These phases represent the times at which subjects had these probabilities available to them (see Supplementary Fig. 4). **b**, Correlation between the effect of outcome-based probability in VMPFC during the decision and the effect of outcome volatility in ACCs during the Monitor phase ($R = 0.7113$, $P < 0.0002$). **c**, Correlation between the effect of confederate-based probability in VMPFC during the decision and the effect of confederate volatility in ACCs during the Monitor phase ($R = 0.6119$, $P < 0.002$). See Supplementary Information.

of information from social and experiential sources. However, the extent to which the VMPFC signal reflects each source of information during a decision is predicted by the extent to which the ACCs and ACCg modulate their activity at the point when information is learnt. If, as is suggested, the VMPFC response codes the expected value of a decision, then the ACCs response to each new outcome predicts the extent that this outcome will determine future valuation of an action; the ACCg response predicts the extent to which this outcome will determine future valuation of an individual.

METHODS SUMMARY

Short description of task (Fig. 1a). Subjects performed a decision-making task while undergoing fMRI, repeatedly choosing between blue and green rectangles, each of which had a different reward magnitude available on each trial. The chance of the rewarded colour being blue or green depended on the recent outcome history. Before the experiment, subjects were introduced to a confederate. At each trial, the confederate would choose between supplying the subject with the correct or incorrect option, unaware of the number of points available. The subject's goal was to maximize the number of points gained during the experiment. In contrast, the confederate's goal was to ensure that the eventual score would lie within one of two pre-defined ranges, known to the confederate but not the subject. The confederate might therefore reasonably give consistently helpful or unhelpful advice, but this advice might change as the game progressed (Supplementary Information). During the experiment, the confederate was replaced by a computer that gave correct advice on a prescribed set of trials. Subjects knew that the trial outcomes were determined by an inanimate computer program, but believed that the social advice came from an animate agent's decision.

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

Received 8 September; accepted 14 October 2008.

1. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785–791 (2003).
2. Maynard Smith, J. *Evolution and the Theory of Games* (Cambridge Univ. Press, 1982).
3. Delgado, M. R., Frank, R. H. & Phelps, E. A. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neurosci.* **8**, 1611–1618 (2005).
4. King-Casas, B. *et al.* Getting to know you: reputation and trust in a two-person economic exchange. *Science* **308**, 78–83 (2005).
5. Rilling, J. *et al.* A neural basis for social cooperation. *Neuron* **35**, 395–405 (2002).

6. Gallagher, H. L., Jack, A. I., Roepstorff, A. & Frith, C. D. Imaging the intentional stance in a competitive game. *Neuroimage* **16,** 814–821 (2002).
7. Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (MIT Press, 1998).
8. Behrens, T. E., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. Learning the value of information in an uncertain world. *Nature Neurosci.* **10,** 1214–1221 (2007).
9. Courville, A. C., Daw, N. D. & Touretzky, D. S. Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* **10,** 294–300 (2006).
10. Dayan, P., Kakade, S. & Montague, P. R. Learning and selective attention. *Nature Neurosci.* **3** (Suppl.), 1218–1223 (2000).
11. O'Doherty, J. *et al.* Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304,** 452–454 (2004).
12. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275,** 1593–1599 (1997).
13. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412,** 43–48 (2001).
14. Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neurosci.* **10,** 647–656 (2007).
15. Tanaka, S. C. *et al.* Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nature Neurosci.* **7,** 887–893 (2004).
16. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441,** 876–879 (2006).
17. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47,** 129–141 (2005).
18. Haruno, M. & Kawato, M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J. Neurophysiol.* **95,** 948–959 (2006).
19. D'Ardenne, K., McClure, S. M., Nystrom, L. E. & Cohen, J. D. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* **319,** 1264–1267 (2008).
20. Kennerley, S. W., Walton, M. E., Behrens, T. E., Buckley, M. J. & Rushworth, M. F. Optimal decision making and the anterior cingulate cortex. *Nature Neurosci.* **9,** 940–947 (2006).
21. Deaner, R. O., Khera, A. V. & Platt, M. L. Monkeys pay per view: adaptive valuation of social images by rhesus macaques. *Curr. Biol.* **15,** 543–548 (2005).
22. Shepherd, S. V., Deaner, R. O. & Platt, M. L. Social status gates social attention in monkeys. *Curr. Biol.* **16,** R119–R120 (2006).
23. Rudebeck, P. H., Buckley, M. J., Walton, M. E. & Rushworth, M. F. A role for the macaque anterior cingulate gyrus in social valuation. *Science* **313,** 1310–1312 (2006).
24. O'Doherty, J. P. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* **14,** 769–776 (2004).
25. Kable, J. W. & Glimcher, P. W. The neural correlates of subjective value during intertemporal choice. *Nature Neurosci.* **10,** 1625–1633 (2007).
26. Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nature Rev. Neurosci.* **7,** 268–277 (2006).
27. Allison, T., Puce, A. & McCarthy, G. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* **4,** 267–278 (2000).
28. Castelli, F., Frith, C., Happe, F. & Frith, U. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* **125,** 1839–1849 (2002).
29. Van Hoesen, G. W., Morecraft, R. J. & Vogt, B. A. in *Neurobiology of Cingulate Cortex and Limbic Thalamus* (eds Vogt, B. A. & Gabriel, M.) (Birkhäuser, 1993).
30. Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* **27,** 9984–9988 (2007).

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

Acknowledgments We would like to acknowledge funding from the UK MRC (T.E.J.B., M.F.S.R.), the Wellcome Trust (L.T.H.) and the UK EPSRC (M.W.W.). We thank S. Knight for helping with data acquisition, and K. Watkins for help with figure preparation.

Author contributions All four authors contributed to generating the hypothesis and designing the experiment. Where specific roles can be assigned, L.T.H. collected the data, T.E.J.B. and L.T.H. analysed the data, T.E.J.B. and M.W.W. built the model, and M.F.S.R. supervised the project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.E.J.B. (behrens@fmrib.ox.ac.uk).

METHODS

Detailed analysis of the task, the learning model, the behavioural analysis, the data acquisition and pre-processing, and several further results and discussion can be found in the Supplementary Information. Here, we describe aspects of the fMRI modelling that may be relevant to the interpretation of our results. Further technical details can also be found in the Supplementary Information.

fMRI single-subject modelling. We performed two fMRI GLM analyses using FMRIB's Software library (FSL, ref. 31). The first looked for learning-related activity (Figs 2, 3 and Supplementary Fig. 3), the second for decision-related activity (Fig. 4 and Supplementary Fig. 4). In each case a general linear model was fit in pre-whitened data space (to account for autocorrelation in the fMRI residuals)[32]. Regressors were convolved and filtered according to FSL defaults (see Supplementary Information).

The following regressors (plus their temporal derivatives) were included in the time series model (learning-related activity): four regressors defining the different times during the task (see Fig. 1 and Supplementary Information), namely Cue, Suggest, Interval, Monitor; four regressors defining key learning parameters when the outcomes are presented (see Supplementary Information), namely (Monitor × Reward history volatility), (Monitor × Confederate volatility), (Monitor × Reward prediction error), (Monitor × Confederate prediction error).

The following regressors (plus their temporal derivatives) were included in the time series model (decision-related activity): four regressors defining the different times during the task (see Fig. 1 and Supplementary Information), namely Cue, Suggest, Interval, Monitor; seven regressors defining key decision parameters at the times when they were available during the decision (see Supplementary Information), namely (Cue × Experience-based probability), (Suggest × Experienced-based probability), (Suggest × Confederate-based probability), (Cue × Chosen reward magnitude), (Suggest × Chosen reward magnitude), (Cue × Unchosen reward magnitude), (Suggest × Unchosen reward magnitude). Note that probabilities were log-transformed such that their linear combination in the GLM would approximate the optimal combination for behaviour (see Supplementary Information). Figure 4a was generated using the mean ([1 1 1]) contrast of all probability-related regressors.

fMRI group modelling. fMRI group analyses were carried out using a GLM with three regressors: a group mean, the weight for reward history information based on each subject's behaviour (see Supplementary Information), and the weight for confederate information based on each subject's behaviour (see Supplementary Information).

fMRI region of interest analyses (Fig. 2). The following processing steps are illustrated schematically in Supplementary Fig. 2 and described in more detail in the Supplementary Information. Individual subject data were taken from regions of interest defined by the group clusters. Data from each trial were up-sampled and re-aligned to points in the trial corresponding to the onset of the four trial stages. Data were *Z*-normalized across trials at each time point in the trial. We then performed two general linear models across trials for both reward and confederate prediction errors. This allowed us (1) to test at which points in the trial the data correlated with the prediction of reward, or the prediction of confederate fidelity, and (2) to test at which points after the outcome the data correlated with the trial outcome, or actual confederate fidelity. A prediction error signal should comprise three parts. (1) A positive correlation with the prediction after the decision; (2) a positive correlation with the trial outcome at the time of this outcome; (3) a negative correlation with the prediction at the time of the outcome (as a prediction error is defined as the outcome minus the prediction).

We witnessed all three parts of the confederate prediction error as deflections in BOLD correlations at the relevant times. However, owing to the nature of the haemodynamic response, it is difficult to test significance from just these deflections. We therefore fit a haemodynamic model to these correlation profiles in each subject (see Supplementary Information). The key test was whether the time course of correlations with the prediction could be accounted for by a positive haemodynamic impulse at the time of the decision and a negative haemodynamic impulse at the time of the outcome; and whether the time course of correlations with the outcome could be accounted for by a positive haemodynamic impulse at the time of the outcome. By fitting the haemodynamic model we were able to measure three parameter estimates for each of these three haemodynamic impulses in each subject, and perform random-effects *t*-tests to measure statistical significance of each.

31. Smith, S. M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23** (Suppl. 1), S208–S219 (2004).
32. Woolrich, M. W., Ripley, B. D., Brady, M. & Smith, S. M. Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* **14**, 1370–1386 (2001).

LETTERS

Oligopotent stem cells are distributed throughout the mammalian ocular surface

François Majo[1,3]†, Ariane Rochat[1,3], Michael Nicolas[1,3], Georges Abou Jaoudé[2] & Yann Barrandon[1,3]

The integrity of the cornea, the most anterior part of the eye, is indispensable for vision. Forty-five million individuals worldwide are bilaterally blind and another 135 million have severely impaired vision in both eyes because of loss of corneal transparency[1]; treatments range from local medications to corneal transplants, and more recently to stem cell therapy[2]. The corneal epithelium is a squamous epithelium that is constantly renewing, with a vertical turnover of 7 to 14 days in many mammals[3]. Identification of slow cycling cells (label-retaining cells) in the limbus of the mouse has led to the notion that the limbus is the niche for the stem cells responsible for the long-term renewal of the cornea[4]; hence, the corneal epithelium is supposedly renewed by cells generated at and migrating from the limbus, in marked opposition to other squamous epithelia in which each resident stem cell has in charge a limited area of epithelium[5,6]. Here we show that the corneal epithelium of the mouse can be serially transplanted, is self-maintained and contains oligopotent stem cells with the capacity to generate goblet cells if provided with a conjunctival environment. Furthermore, the entire ocular surface of the pig, including the cornea, contains oligopotent stem cells (holoclones)[7,8] with the capacity to generate individual colonies of corneal and conjunctival cells. Therefore, the limbus is not the only niche for corneal stem cells and corneal renewal is not different from other squamous epithelia. We propose a model that unifies our observations with the literature and explains why the limbal region is enriched in stem cells.

To explore the function of limbal stem cells (Fig. 1a, b), portions (1.5 mm × 0.3 mm) of the limbus of athymic mice were excised and replaced with limbus from β-gal-ROSA26 mice (Fig. 1c). In most cases, epithelial continuity was rapidly restored (Supplementary Fig. 1). Transplanted eyes were usually collected 4 months later and once after 11 months. Notably, β-galactosidase (β-gal)-labelled cells never

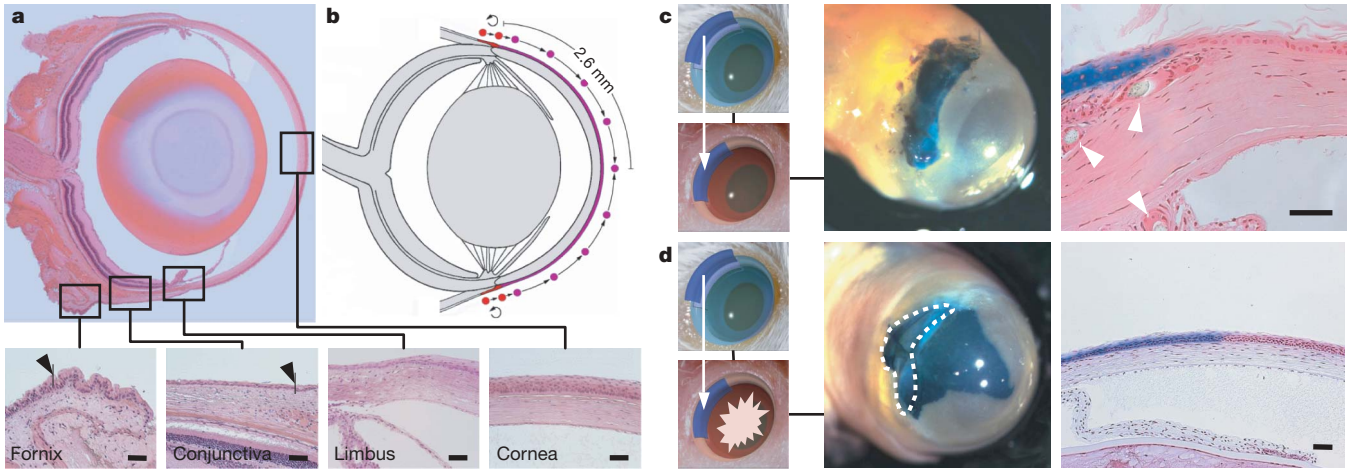


Figure 1 | The limbus does not contribute to steady-state corneal renewal, but to corneal repair. a, Anatomy of a mouse eye. The limbus is the zone at the junction of the cornea and the conjunctiva. Arrows denote mucin-producing cells (goblet cells). **b**, The limbal paradigm[4] is shown. The cornea does not contain stem cells and is renewed by transient amplifying cells (purple) generated by limbal stem cells (red) dividing asymmetrically. Transient amplifying cells then migrate centripetally to the central cornea while renewing the corneal epithelium *en route*. The distance separating the limbus from the corneal midline is 2.6 mm. **c**, The limbus does not contribute to the steady-state renewal of the corneal epithelium. A limbal fragment of an athymic mouse was excised and replaced by an equivalent limbal fragment obtained from a β-gal-ROSA26 mouse (left panel). The transplant was stitched with its corneal side facing the cornea of the recipient mouse (normal orientation). Cells did not migrate out of the transplant 3 months after the transplantation (β-gal staining; middle panel). Similar results were obtained when the limbal transplant was stitched with its conjunctival side facing the cornea of the recipient mouse (reverse orientation) or using SCID mice as recipients. Note the clear boundary between transplanted limbal (stained blue) and corneal (unstained) cells (right panel); white arrows indicate nylon stitches. See also Supplementary Fig. 1a. **d**, The limbus can contribute to corneal repair. An extensive corneal wound challenged the limbal transplant 4 months after transplantation (left panel). Labelled cells migrated out of the transplant and restored the corneal integrity in 7 days (β-gal staining) (middle panel). The migrating cells did not cross the corneal midline (right panel). These experiments indicate that the limbus contributes little to steady renewal of the cornea, but becomes instrumental when the cornea is extensively wounded. Scale bars, all 50 μm.

[1]Laboratory of Stem Cell Dynamics, [2]Laboratory of Informatics and Visualization, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne CH, Switzerland. [3]Department of Experimental Surgery Lausanne University Hospital (CHUV), 1011 Lausanne CH, Switzerland. †Present address: Hôpital Ophtalmique Jules Gonin, Avenue de France 15, 1004 Lausanne CH, Switzerland.

migrated out of the grafts onto the cornea (Fig. 1c), whether the grafts were transplanted with their corneal (normal orientation) or conjunctival (reverse orientation) side facing the cornea of the recipient mouse ($n = 39$ and 28, respectively). Similar results were also obtained using SCID mice as recipients, thus excluding the remote possibility that the migration of the transplanted limbal cells was adversely affected in athymic mice. Next we chemically or physically wounded the cornea of recipient mice[4] in which limbal grafts had been in place for up to 6 months. As expected, labelled cells migrated out of the grafts along with unlabelled limbal cells of the recipient mice resulting in a mosaicism of the healed corneal epithelium (Fig. 1d), whether the graft was in a normal ($n = 39$) or a reverse ($n = 32$) orientation (data not shown). The healed cornea was usually transparent but corneal opacity was observed when the stroma was altered ($n = 13$). Cells migrating from the limbus usually covered the distance to the central part of cornea (2.6 mm) in a week and had a typical corneal phenotype (Fig. 1d). Collectively, these experiments indicated that the absence of a contribution of limbal stem cells to corneal renewal did not result from an incapacity of the transplanted cells to migrate, and confirmed that limbal stem cells could efficiently heal an extreme corneal wound[2,4].

In humans, limbal deficiency is a condition thought to result from a complete loss of limbal stem cells[9]. As a consequence, the corneal epithelium is replaced by a conjunctival-like epithelium containing mucus-producing goblet cells, and vision is impaired. To mimic limbal deficiency in the mouse, we cauterized the entire circumference of the limbus of wild-type mice (12 eyes). The burn entirely destroyed the limbal epithelium and the underlying matrix, reaching the ciliary bodies in many places, making it unlikely that any limbal stem cells had escaped destruction (Supplementary Fig. 2a). The cornea of burned eyes remained transparent for the duration of the experiment (a minimum of 4 months), and neither corneal ulcers nor stromal blood vessels were observed (Supplementary Fig. 2b), strongly suggesting self-maintenance of the corneal epithelium.

To evaluate further the ability of the cornea to self-renew, full-thickness grafts including the corneal epithelium and its stroma were obtained from the central part of the cornea of β-gal-ROSA26 donor mice older than 6 months (up to 1 year), using a 1.5 mm diameter ophthalmic trephine. The grafts were then transplanted at the limbal region of athymic mice as described earlier. Most of the grafts ($n = 11$) were rapidly incorporated into the limbal region of the recipient mice, and remained viable for months with no signs of necrosis (Fig. 2a). We next challenged the grafts at the time of transplantation or several weeks later by extensive corneal wounds (Fig. 2b). In most cases (47 out of 51) cells quickly migrated out of the grafts onto the denuded corneal stroma and corneal integrity was restored within 7 days. Notably, healed corneas were transparent and the regenerated epithelium remained β-gal-labelled for the entire duration of the experiments, up to 9 months (Fig. 2b and Supplementary Fig. 3). To exclude the remote possibility that the corneal cells were reprogrammed to a limbal phenotype because they were transplanted in a limbal/conjunctival microenvironment, corneal transplants were implanted into the central cornea of recipient athymic mice; transplants ($n = 9$) were again able to restore the integrity of a wounded cornea (Fig. 2c). Transplanted cells often covered the entire distance (5.2 mm) to the opposite corneal side before limbal cells even had time to move out (Fig. 2b and

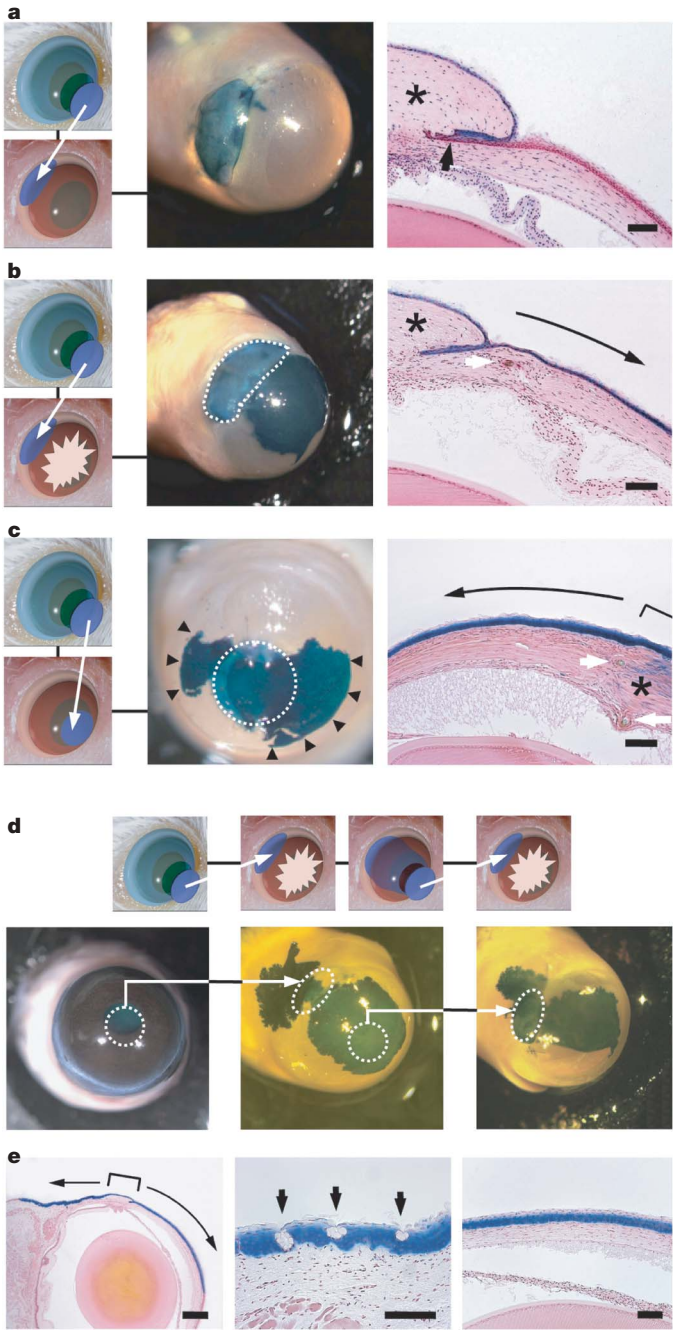


Figure 2 | The cornea of the mouse has extensive self-renewal capacity and contains oligopotent stem cells. Full-thickness biopsies were taken from the central cornea of 6-month-old ROSA26 (**a**–**c**) and ROSA26-GFPU (**d**, **e**) mice and transplanted at the limbus of a recipient athymic mouse as described in Fig. 1. **a**, The corneal transplant (asterisk, right panel) remained healthy for 4 months (β-gal staining); black arrow denotes that the transplanted epithelium is in continuity with the adjacent corneal epithelium. See also Supplementary Fig. 1b. **b**, The entire corneal epithelium of the recipient mouse was removed to challenge the corneal transplant 4 months after transplantation. After 7 days, cells that had migrated out of the transplant had almost entirely restored the corneal epithelium, whereas the limbal cells of the recipient contributed little. The β-gal-positive cells crossed the corneal midline almost reaching the opposite side (compare with Fig. 1d). See also Supplementary Fig. 3. **c**, The corneal transplant was challenged by a corneal wound. Same results as in **b**. Arrowheads (middle panel) denote the sharp boundary with the limbus of the recipient mouse. The white arrows point to a ligature knot (**b**, **c**, right panels), the black arrow indicates the direction of migration and the asterisk indicates the transplant. These experiments indicate that reprogramming of corneal cells by the limbal environment is unlikely and that the cornea of the mouse contains epithelial cells with extensive growth capacity. **d**, **e**, The corneal epithelium can be serially transplanted. The corneal transplant was challenged by wounding of the cornea and of the conjunctiva. Two months later, a biopsy was obtained from the healed cornea and transplanted onto a second athymic mouse. Corneal and conjunctival wounds challenged the transplant for a second time (**d**). The eye was collected 42 days later (β-gal staining). Labelled cells repaired the wounded cornea and conjunctiva (black arrows in the middle panel of **e** indicate goblet cells) (**e**). These experiments indicate that the central cornea of the mouse contains bona fide oligopotent epithelial stem cells. Scale bars, all 100 μm, except left panel **e**, 200 μm.

Supplementary Fig. 3), a further indication of the robust migratory capability of corneal cells[10]. Markedly, the boundary between the labelled cornea and the adjacent unlabelled limbus was sharply demarcated for the duration of the experiments (up to 9 months; Fig. 2c and Supplementary Fig. 3), another sign that limbal stem cells contribute little to corneal renewal.

We then performed serial transplantations of β-gal and GFP (green fluorescent protein) double-labelled central corneas (Fig. 2d) that were monitored by ultraviolet illumination on living mice and by β-gal staining on enucleated eyes. Transplanted corneal cells were again able to restore the corneal integrity for a minimum of 4 months (sum of serial transfers). Furthermore, when the transplants were challenged by either a conjunctival or a fornical wound, labelled corneal cells efficiently migrated towards the fornix generating β-gal-positive goblets cells, a landmark of the conjunctiva (Fig. 2d). Hence, corneal epithelial cells could adopt either a conjunctival or a

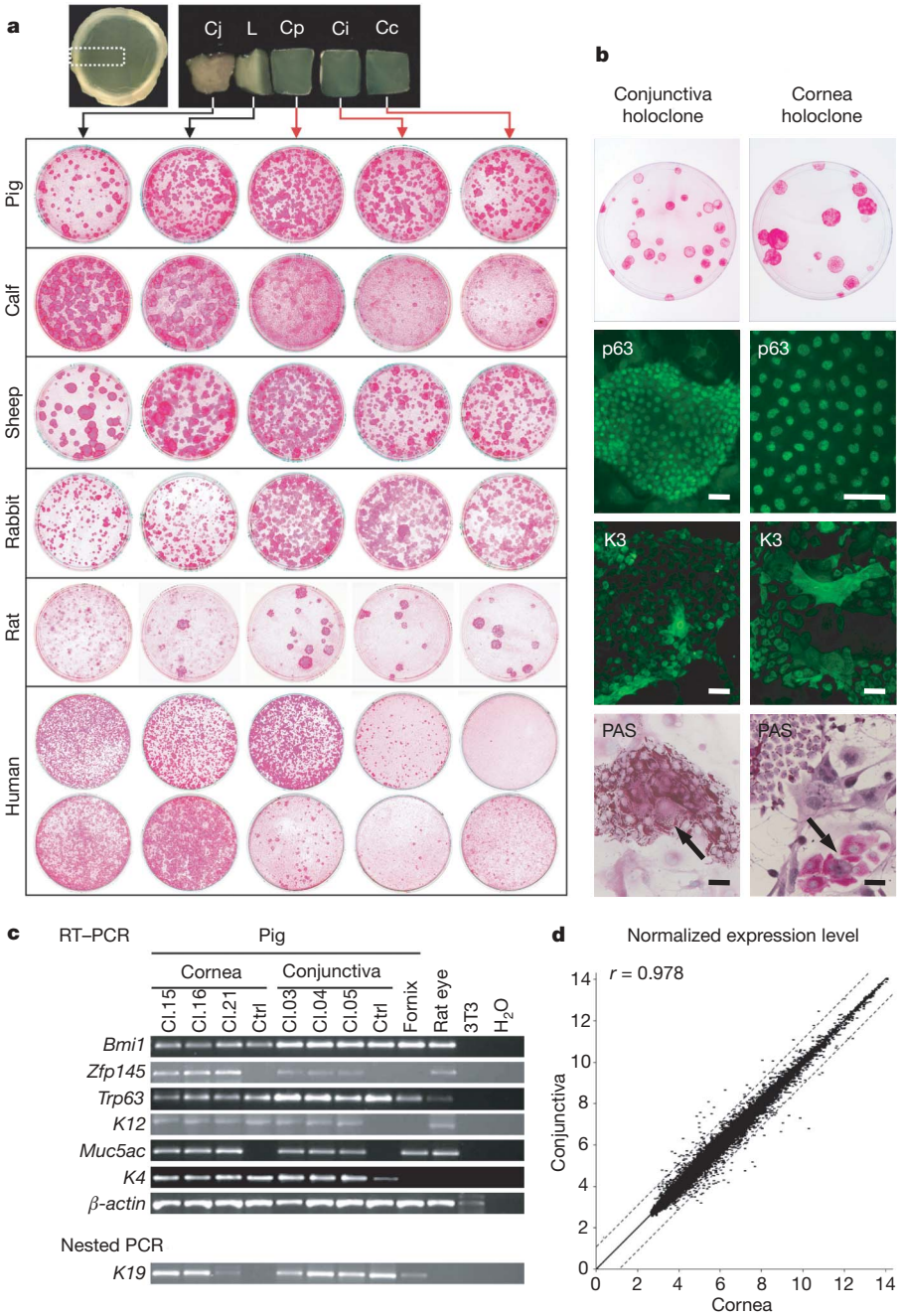


Figure 3 | The cornea of mammals contains clonogenic keratinocytes. **a**, The ocular surface was dissected and cut into several fragments: the conjunctiva (Cj), the limbus (L), the peripheral cornea (Cp), the intermediate cornea (Ci) and the central cornea (Cc). Cornea shown is from pig and each fragment is on average 4.5 mm^2. Each fragment was cultivated as described[8,28]; see also Supplementary Information. Several clonogenic keratinocytes were distributed through the entire ocular surface including the central cornea, although with species differences (for example, between pig and human). **b**, The cornea and conjunctiva of the pig contain holoclones[7] that formed colonies containing $p63^+$ and $K3^+$ corneal cells and large PAS^+ cells (arrows, bottom panels), suggestive of goblet cells. Scale bars, 50 μm **c**, The corneal and conjunctival holoclones (cl.) expressed *p63*, a putative marker of corneal stem cells[29], *Bmi1* and *Zfp145* (also known as *Zbtb16*), implicated in stem cell renewal[12], *K3* and *K12* (also known as *Krt3* and *Krt12*), markers of corneal differentiation[30], and *K19*, *K4* (also known as *Krt19*, *Krt4*) and *Muc5ac*, all markers of conjunctival differentiation[21]. Ctrl, control. **d**, DNA microarray analysis comparing three corneal clones (AP6cl15, AP6cl16 and AP1cl21) to three conjunctival clones (AP6cl3, AP1cl4 and AP1cl5). Only two genes out of 20,201 were differentially expressed. These experiments demonstrate that stem cells of the conjunctiva and cornea of the pig are oligopotent and share the same ocular signature.

corneal phenotype depending on which environment (conjunctival or a corneal) they were exposed to (Fig. 2e).

Together, our results challenge the prevailing opinion that the limbus is the sole niche for corneal stem cells, because long-term self renewal and the capacity to sustain serial transplantation are properties of stem cells and not of transit amplifying cells—the growth capability of which is limited to a few divisions by definition[11]. Furthermore, the capacity of corneal stem cells to generate goblet cells, the hallmark of conjunctiva, strongly suggests oligopotence. Our results further highlight the importance of function to demonstrate 'stemness' in agreement with recent observations in the epidermis and hair follicles[6,12], the bone marrow[13] and the gut[14]. Nonetheless, the limbal region undoubtedly contains stem cells that are efficient in restoring the corneal surface in the extreme situation of a massive corneal injury or for cell therapy as demonstrated by the impressive clinical results obtained with the transplantation of limbal holoclones[2,15].

The inefficacy of the cultivation of mouse keratinocytes and their inclination to undergo spontaneous immortalization seriously limited thorough single cell analysis. We therefore isolated keratinocytes from the ocular surface of several other mammals and cultivated them in conditions strictly identical to human cell therapy[15,16]. Despite the species differences (for example, between pigs and humans) the entire ocular surface of these mammals contains many keratinocyte colony-forming cells that can be serially passaged (Fig. 3a). The pig was selected for clonal analyses because of the robust growth of its ocular surface cells. Eyes, the integrity of which was controlled under a dissecting microscope or with a slit lamp, were obtained from two pigs (AP1 and AP6) soon after being killed. Several colonies (487 per mm^2 in the conjunctiva to 1,025 per mm^2 in the cornea) were obtained from 3–6 mm^2 fragments of bulbar conjunctiva and of the central cornea, located at least 1.5 and 6 mm from the limbus, respectively. A total of 277 single cells were then isolated from the central cornea[7,17] and 53 clones were obtained, of which 26 were passaged; 13 clones were holoclones, 8 were meroclones and 5 were paraclones. One holoclone and three meroclones were then passaged once a week for more than 16 weeks (over 150 doublings). In a separate experiment, 125 single conjunctival cells were isolated and 29 clones were obtained, of which 9 were holoclones. Both corneal and conjunctival holoclones maintained a diploid karyotype ($2n = 38$ chromosomes) in early passages, whereas various degrees of aneuploidy was observed in late passages. Notably, corneal and conjunctival holoclones had a similar, if not identical, phenotype in culture (Fig. 3b). Both initiated colonies containing $p63^+$ cells, keratin 3^+ cells and PAS^+ (periodic acid schiff) cells, reminiscent of goblet cells (Fig. 3b), and both had a similar ocular signature by polymerase chain reaction analysis with reverse transcription (RT–PCR; Fig. 3c) and DNA microarray analysis (Fig. 3d). The corneal epithelium of the pig thus contained bona fide oligopotent stem cells that expressed lineage markers as other tissue stem cells, for example, neural stem cells expressing markers of astrocyte differentiation [18]. Our finding that corneal and conjunctival stem cells have equal potency explains the successful use of both to treat corneal defects[19–21] and further supports the notion that stromal signals are determinant for epithelial stem cell fate[22]. Stromal dysfunction may then lead to aberrant epithelial differentiation, as shown by the formation of dermoid cysts on the ocular surface of epidermis in the wounded cornea of conditional Notch1 null mice[23], and of a corneal epithelium in the conjunctiva of humans[24].

To unify our observations with previous results[4], we propose a model that integrates cell dynamics and various constraints to which the corneal epithelium is submitted. In this model (Fig. 4), the limbus is a zone of equilibrium in which the expanding conjunctival and corneal epithelia are confronted in a mechanism reminiscent of tectonic plates. Rupture of the equilibrium—for example, because of an increased rigidity of the corneal stromal or because of an extensive corneal wound—results in migration of limbal stem cells onto the cornea[4,25,26].

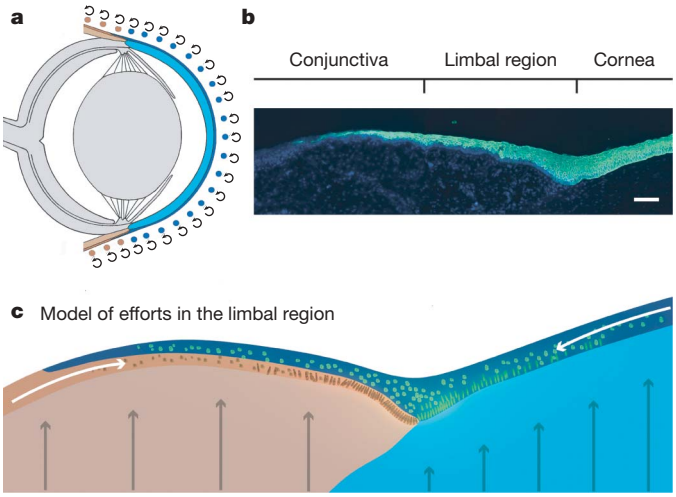


Figure 4 | A unifying model of ocular surface renewal. a, Epithelial stem cells of equal potency are distributed throughout the entire ocular surface. **b**, Suprabasal cells at the limbal region of the pig express $K3^+$ corneal-type differentiation whereas basal cells are negative. Differentiated corneal cells appear to slide over limbal basal cells, as suggested by their orientation. Scale bar, 100 μm. **c**, Schematic representation of the same region in **b**. The combination of blinking, variable stroma elasticity and pressure in the eyeball and within the corneal dome-shape (dark arrows) results in opposite forces (white arrows) in the conjunctival and corneal epithelia. Hence, the corneal and conjunctival epithelia, which are the site of intense cell multiplication, are continuously expanding but in opposite directions and confront at the limbus. Stem cells, sorted by an elutriation-like phenomenon, accumulate both on the conjunctival and corneal side of the limbus.

Most importantly, our model explains why resting stem cells accumulate at the limbus and why the limbus is a principal source of stem cells for cell therapy in humans[2,27]. It also explains why the limbus is a privileged site of cancer, as are other transitional zones—for example, the junction of the vaginal epithelium with the endocervical epithelium, the epidermis with the oral epithelium at the lips, and the oesophagus with the gastric epithelium.

METHODS SUMMARY

Detailed methods and surgical procedures are described in the Methods. Transplantations were authorized by the veterinarian authorities of the Canton de Vaud (authorization numbers 1524 and 1686) and performed by an ophthalmic surgeon (F.M.). ROSA26 or GFPU (modified GFP) donor mice were at least 4 months old, as were recipient mice (athymic or SCID). Keratinocytes were cultivated onto a feeder layer of lethally irradiated 3T3-J2 cells[28] and clonal analyses were performed as described[7,17]. Histological and immunological analyses were performed using standard procedures. Antibodies were mouse monoclonal anti-rabbit cytokeratin 3 (clone AE5, Chemicon), mouse monoclonal anti-human p63 (clone 4A4, Dako), mouse monoclonal anti-human cytokeratin 19 (clone RCK 108, Dako) and anti mouse monoclonal anti-mucin 5AC (clone 45M1, Neomarkers). Pig primers (listed in Supplementary Information) were selected using the primer 3-input software. PCRs were performed according to standard procedures and PCR products were sequenced to confirm the specificity of the reaction. DNA microarrays analyses were performed by the Lausanne DNA Array Facility (DAFL) using standard protocols and the Affymetrix GeneChip Porcine Genome Array. Raw and normalized microarray data are accessible through the NCBI Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo) public database (series record GSE12604).

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

Received 9 May; accepted 12 September 2008.
Published online 1 October 2008.

1. Whitcher, J. P., Srinivasan, M. & Upadhyay, M. P. Corneal blindness: a global perspective. *Bull. World Health Organ.* **79**, 214–221 (2001).
2. Pellegrini, G. *et al.* Long-term restoration of damaged corneal surfaces with autologous cultivated corneal epithelium. *Lancet* **349**, 990–993 (1997).

3. Haddad, A. Renewal of the rabbit corneal epithelium as investigated by autoradiography after intravitreal injection of 3H-thymidine. *Cornea* **19,** 378–383 (2000).
4. Cotsarelis, G., Cheng, S. Z., Dong, G., Sun, T. T. & Lavker, R. M. Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: implications on epithelial stem cells. *Cell* **57,** 201–209 (1989).
5. Jones, P. H., Simons, B. D. & Watt, F. M. Sic transit gloria: farewell to the epidermal transit amplifying cell? *Cell Stem Cell* **1,** 371–381 (2007).
6. Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446,** 185–189 (2007).
7. Barrandon, Y. & Green, H. Three clonal types of keratinocyte with different capacities for multiplication. *Proc. Natl Acad. Sci. USA* **84,** 2302–2306 (1987).
8. Pellegrini, G. *et al.* Location and clonal analysis of stem cells and their differentiated progeny in the human ocular surface. *J. Cell Biol.* **145,** 769–782 (1999).
9. Puangsricharern, V. & Tseng, S. C. Cytologic evidence of corneal diseases with limbal stem cell deficiency. *Ophthalmology* **102,** 1476–1485 (1995).
10. Buck, R. C. Cell migration in repair of mouse corneal epithelium. *Invest. Ophthalmol. Vis. Sci.* **18,** 767–784 (1979).
11. Potten, C. S., Schofield, R. & Lajtha, L. G. A comparison of cell replacement in bone marrow, testis and three regions of surface epithelium. *Biochim. Biophys. Acta* **560,** 281–299 (1979).
12. Claudinot, S., Nicolas, M., Oshima, H., Rochat, A. & Barrandon, Y. Long-term renewal of hair follicles from clonogenic multipotent stem cells. *Proc. Natl Acad. Sci. USA* **102,** 14677–14682 (2005).
13. Kiel, M. J. *et al.* Haematopoietic stem cells do not asymmetrically segregate chromosomes or retain BrdU. *Nature* **449,** 238–242 (2007).
14. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449,** 1003–1007 (2007).
15. Rama, P. *et al.* Autologous fibrin-cultured limbal stem cells permanently restore the corneal surface of patients with total limbal stem cell deficiency. *Transplantation* **72,** 1478–1485 (2001).
16. Ronfard, V., Rives, J. M., Neveux, Y., Carsin, H. & Barrandon, Y. Long-term regeneration of human epidermis on third degree burns transplanted with autologous cultured epithelium grown on a fibrin matrix. *Transplantation* **70,** 1588–1598 (2000).
17. Rochat, A., Kobayashi, K. & Barrandon, Y. Location of stem cells of human hair follicles by clonal analysis. *Cell* **76,** 1063–1073 (1994).
18. Merkle, F. T., Tramontin, A. D., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Radial glia give rise to adult neural stem cells in the subventricular zone. *Proc. Natl Acad. Sci. USA* **101,** 17528–17532 (2004).
19. Shapiro, M. S., Friend, J. & Thoft, R. A. Corneal re-epithelialization from the conjunctiva. *Invest. Ophthalmol. Vis. Sci.* **21,** 135–142 (1981).
20. Thoft, R. A. Keratoepithelioplasty. *Am. J. Ophthalmol.* **97,** 1–6 (1984).
21. Tanioka, H. *et al.* Establishment of a cultivated human conjunctival epithelium as an alternative tissue source for autologous corneal epithelial transplantation. *Invest. Ophthalmol. Vis. Sci.* **47,** 3820–3827 (2006).
22. Ferraris, C., Chevalier, G., Favier, B., Jahoda, C. A. & Dhouailly, D. Adult corneal epithelium basal cells possess the capacity to activate epidermal, pilosebaceous and sweat gland genetic programs in response to embryonic dermal stimuli. *Development* **127,** 5487–5495 (2000).
23. Vauclair, S. *et al.* Corneal epithelial cell fate is maintained during repair by Notch1 signaling via the regulation of vitamin A metabolism. *Dev. Cell* **13,** 242–253 (2007).
24. Kawasaki, S. *et al.* Clusters of corneal epithelial cells reside ectopically in human conjunctival epithelium. *Invest. Ophthalmol. Vis. Sci.* **47,** 1359–1367 (2006).
25. Di Iorio, E. *et al.* Isoforms of ΔNp63 and the migration of ocular limbal cells in human corneal regeneration. *Proc. Natl Acad. Sci. USA* **102,** 9523–9528 (2005).
26. Barbaro, V. *et al.* C/EBPΔ regulates cell cycle and self-renewal of human limbal stem cells. *J. Cell Biol.* **177,** 1037–1049 (2007).
27. Nakamura, T., Inatomi, T., Sotozono, C., Koizumi, N. & Kinoshita, S. Successful primary culture and autologous transplantation of corneal limbal epithelial cells from minimal biopsy for unilateral severe ocular surface disease. *Acta Ophthalmol. Scand.* **82,** 468–471 (2004).
28. Oshima, H., Rochat, A., Kedzia, C., Kobayashi, K. & Barrandon, Y. Morphogenesis and renewal of hair follicles from adult multipotent stem cells. *Cell* **104,** 233–245 (2001).
29. Pellegrini, G. *et al.* p63 identifies keratinocyte stem cells. *Proc. Natl Acad. Sci. USA* **98,** 3156–3161 (2001).
30. Schermer, A., Galvin, S. & Sun, T. T. Differentiation-related expression of a major 64K corneal keratin *in vivo* and in culture suggests limbal location of corneal epithelial stem cells. *J. Cell Biol.* **103,** 49–62 (1986).

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

Acknowledgements We are grateful to F. Radtke and S. Vauclair for Notch1$^{lox/lox}$ K14Cre ROSA26R mice, to A. Smith for critical reading of the manuscript, to T. Hoang-Xuan, L. Zografos, F. Munier and S. Kinoshita for continuous support and to L. Schnell, J. Vannod and S. Vermot for excellent technical help. The work was supported by grants to Y.B. from the Swiss National Science Foundation (grant 3100A0-104160), the EPFL, the CHUV and EuroStemCell. The early part of the work was supported by grants to Y.B. from the Institut National de la Santé et de la Recherche Médicale (INSERM), the Association pour la Recherche contre le Cancer and the Association Française contre les Myopathies. F.M. was supported by fellowships from the Fédération des Aveugles et Handicapés Visuels de France, the INSERM and then the CHUV.

Author Contributions F.M., A.R. and M.N. performed and assisted in the design of the experiments and the interpretation of results, A.R. and G.A.J. made the figures and G.A.J. and Y.B. contributed the new concept. Y.B. designed the experiments and wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to Y.B. (yann.barrandon@epfl.ch).

METHODS

Animals. Experiments were performed according to the Swiss guidelines and authorized by the veterinarian authorities of the Canton de Vaud (authorization numbers 1524 and 1686). Wild-type, OF1 (inbred), athymic (Swiss nu/nu, inbred) and SCID mice were from Iffa Credo-Charles River. ROSA26 mice[31] and GFPU mice[32] were from the Jackson laboratory. The donor male or female mice (ROSA26 or GFPU) were at least 4 months old, as were the recipient mice (athymic or SCID) when operated. Mice were anesthetized with a mixture of ketamine (0.104 mg per g of body weight) and xylazine (0.033 mg per g of body weight) in PBS. Operated mice systematically received 0.07 mg per g body weight of buprenorphine as post-operative analgesia.

Surgical procedures. All surgical procedures were performed under a sterile environment using a MZ 690 surgical microscope (Leica) or a dissecting binocular microscope (Zeiss) equipped with fluorescence and enhanced GFP filters. Instruments used were an ophthalmic cautery, Vannas scissors, a Troutman-O'Brien needle holder, Bonn and Barraquer forceps, a hockey spatula and a 1.5 mm diameter corneal trephine (all from Moria Surgical). Sutures were Biosord 10.0 (Alcon Surgical) and 11/0 nylon (Mani Inc) for transplants and Vicryl 8.0 (Ethicon) for lids. Special care was taken to keep the operated eyes moisturized during surgery. After they were collected, transplants were kept in DMEM supplemented with 10% calf serum (Hyclone) before they were stitched in place. Clinical monitoring of the transplants (transparency, shape of the eye, presence of ulcers) was performed under a dissecting microscope on anesthetized animals. Adverse reactions included loss of transplants, infection, formation of an opaque pannus with corneal vascularization and phtisis bulbi. Most of these reactions were the consequence of the surgical procedure.

Wounding of the ocular surface. Corneas were wounded according to standard protocols[4]. In brief, the entire surface of the cornea to the exclusion of the limbal region was exposed for 30 s to 20% ethanol applied with a cotton-tip applicator before it was washed with sterile PBS. Complete removal of corneal epithelial tissue was achieved using a Demarres spatula (Moria Surgical) with a 45° cutting angle. Efficacy of wounding was monitored using 0.05% ophthalmic fluorescein solution. Fornix and conjunctiva were wounded using 20% ethanol and an ophthalmic knife as described above.

Cell culture. Biopsies of the ocular surface were dissociated using 0.05% trypsin in 0.01% EDTA. Cells were cultivated onto a feeder layer of lethally irradiated 3T3-J2 cells in 3:1 DMEM/F12 medium, supplemented as described[17]. Cultures were fed every 3 to 4 days in medium supplemented with 10 ng ml^{-1} of human recombinant EGF (Upstate Biotechnology) starting at the first feeding, and were usually passaged once a week. Analyses of karyotypes were by Chrombios (Raubling).

Distribution of clonogenic epithelial cells in mammalian ocular surface. The ocular surface was dissected from eyes of different mammals soon after culling and cut into several fragments, the size of which was 2–6 mm^2. Each fragment was enzymatically dissociated and the cells were cultivated onto a feeder layer of irradiated 3T3-J2 cells[8,28]. Seeding density per dish was half of the cell suspension for the rat, 1/40 for the rabbit and 1/20 for other species. Cultures were fixed 9 to 12 days later. The shape, growth and differentiation of the colonies (as shown by the differing intensity of the rhodamine B staining) varied from species to species. Pig cells had the most robust growth whereas calf corneal cells stratified had the least. Human corneas were dissected from eyes obtained from children (1 year old and 4 years old) undergoing removal of an eye for retinoblastoma. A sample including the cornea, the limbus region and its adjacent conjunctiva was immediately dissected, placed in cold culture medium and transported to the laboratory where it was quickly processed as described. The corneal sample was then cut into five fragments (conjunctiva, limbal, peripheral, intermediate and central cornea); fragment sizes ranged from 4.5 to 6 mm^2. Each fragment was enzymatically dissociated and cells were then plated onto an irradiated feeder layer of 3T3-J2 cells and cultivated as described previously. Cultures were fixed at 12 days and stained with rhodamine B. Each Petri dish contained 30% of the cells present in each fragment.

Single-cell isolation. Single cells were isolated with an elongated Pasteur pipette under an inverted microscope and individually cultivated on a feeder layer of irradiated 3T3 cells as described[33]. Cultures were fed every 3 to 4 days as described previously. Clonal analyses were performed as described[7,17].

Histology. Tissue samples were fixed in Bouin's liquid, embedded in resin (Historesin, Leica) and 4-mm sections were stained with haematoxylin. β-gal staining was performed on 4% paraformaldehyde-fixed samples according to standard protocols. Samples were then post-fixed in Bouin's liquid and embedded in resin (Historesin, Leica). Four-millimetre sections were counterstained in nuclear fast red (CI 60760) and mounted in Eukitt. Sections were examined under an Axioskop miscrocope (Zeiss) equipped with a digital camera. Captured images were processed using Zeiss Axiovision 3.1 software. PAS staining was performed on cultured cells according to standard procedures.

Immunostaining. Tissue samples were snap-frozen in liquid nitrogen and 8-mm frozen sections were obtained using a cryostat (Leica) with the cutting chamber temperature set at −30° C. Sections or cultured cells were fixed for 10 min in 50:50 acetone/methanol (v/v) at −20° C. Antibodies were highly cross-adsorbed mouse monoclonal anti-rabbit cytokeratin 3 (clone AE5, Chemicon), mouse monoclonal anti-human p63 (clone 4A4, Dako), and mouse monoclonal anti-human cytokeratin 19 (clone RCK 108, Dako), used at dilutions of 1:500, 1:500 and 1:200, respectively. Samples were fixed for 10 min in 4% paraformaldehyde at 4 °C for staining with anti-mouse monoclonal anti-mucin 5AC (clone 45M1, Neomarkers). Secondary antibodies were goat anti-mouse Alexa Fluor 488 or Alexa Fluor 568 (Molecular Probes Inc), diluted to 1:200.

RT–PCR. Primers were selected to anneal with messenger RNA of selected rat or human genes using the primer 3-input software. Total RNAs were extracted using Trizol reagent (Invitrogen). Fifty to one-hundred nanograms of total RNA were amplified using the OneStep RT–PCR kit (Qiagen) according to the manufacturer's instructions. For nested PCR of keratin 19, 1 μl of the first amplification was added to a new reaction mix containing the inner primer for keratin 19 and amplified as described before for a total of 39 cycles. PCR products were resolved by agarose gel electrophoresis, stained with ethidium bromide and visualized under ultraviolet light. β-actin was used as a control for equal loading throughout the experiments. PCR products were sequenced to confirm the specificity of the reaction (Fasteris SA). Sequences were compared to human sequences from the GenBank database and were highly similar (90%).

List of primers. The following primers were used: *β-actin* forward, 5′-TCATGTTTGAGACCTTCAACACCC-3′; *β-actin* reverse, 5′-GTACTTGCGCTCAGGAGGAG-3′; *Bmi1* forward, 5′-TGTGCGTTACTTGGAGACCA-3′; *Bmi1* reverse, 5′-TCATTCACCTCCTCCTTTGG-3′; *K4* forward, 5′-CTCCAGCAAAAACCTTGAGC-3′; *K4* reverse, 5′-ACCTCGGCAATAATGCTGTC-3′; *K12* forward, 5′-GGTCCAGGTGAGGTCAATGT-3′; *K12* reverse, 5′-GACAGTTGGCAGCAGTACCC-3′; *K19* outer, forward 5′-CGCGACTACAGCCACTACT-3′; *K19* outer, reverse, 5′-GCTCACTATCAGCTCGCACA-3′; *K19* inner, forward, 5′-TTTGAGACGGAACAGGCTCT-3′; *K19* inner, reverse, 5′-TCAGTAACCTCGGACCTGCT-3′; *Muc5ac* forward, 5′-CTGCCAGGACTGCATCTGTA-3′; *Muc5ac* reverse, 5′-CCTCACACAGGCATCTCTCA-3′; *Trp63* forward, 5′-ACTGCCAGATTGCAAAGACC-3′; *Trp63* reverse, 5′-GCTGCTGTTGCACATGAAAT-3′; *Zfp145* forward, 5′-AAAGCAGAGGACCTGGATGA-3′; *Zfp145* reverse, 5′-TCATGGCTGAGAGACCAAAA-3′.

DNA microarray analysis. Quantitative analysis of RNA expression was performed using Affymetrix gene chip cDNA microarrays. Three central corneal clones and three conjunctival clones were used to prepare total RNA using the RNeasy Mini Kit (QIAGEN Inc). RNA quality was measured using the Agilent Bioanalysation system to ensure the integrity of the RNA. Complementary DNA synthesis, hybridization to the Affymetrix GeneChip Porcine Genome Array (23,937 probe sets representing 20,201 genes) and analysis was performed by the Lausanne DNA Array Facility using standard protocols. To identify differentially expressed genes from each group, P values were calculated using Bioconductor limma package[34] and probe sets with a false discovery rate[35] <0.05 were considered significant.

31. Friedrich, G. & Soriano, P. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. *Genes Dev.* **5**, 1513–1523 (1991).
32. Hadjantonakis, A. K., Gertsenstein, M., Ikawa, M., Okabe, M. & Nagy, A. Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech. Dev.* **76**, 79–90 (1998).
33. Barrandon, Y. & Green, H. Cell size as a determinant of the clone-forming ability of human keratinocytes. *Proc. Natl Acad. Sci. USA* **82**, 5390–5394 (1985).
34. Smyth, G. K. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.* **3**, article–3 (2004).
35. Hochberg, Y. & Benjamini, Y. More powerful procedures for multiple significance testing. *Stat. Med.* **9**, 811–818 (1990).

LETTERS

Stereocilin-deficient mice reveal the origin of cochlear waveform distortions

Elisabeth Verpy[1,2,3,4], Dominique Weil[1,2,3,4]*, Michel Leibovici[1,2,3,4]*, Richard J. Goodyear[5], Ghislaine Hamard[6], Carine Houdon[1,2,3,4], Gaelle M. Lefèvre[1,2,3,4], Jean-Pierre Hardelin[1,2,3,4], Guy P. Richardson[5], Paul Avan[7]* & Christine Petit[1,2,3,4]*

Although the cochlea is an amplifier and a remarkably sensitive and finely tuned detector of sounds, it also produces conspicuous mechanical and electrical waveform distortions[1]. These distortions reflect nonlinear mechanical interactions within the cochlea. By allowing one tone to suppress another (masking effect), they contribute to speech intelligibility[2]. Tones can also combine to produce sounds with frequencies not present in the acoustic stimulus[3]. These sounds compose the otoacoustic emissions that are extensively used to screen hearing in newborns. Because both cochlear amplification and distortion originate from the outer hair cells—one of the two types of sensory receptor cells—it has been speculated that they stem from a common mechanism. Here we show that the nonlinearity underlying cochlear waveform distortions relies on the presence of stereocilin, a protein defective in a recessive form of human deafness[4]. Stereocilin was detected in association with horizontal top connectors[5–7], lateral links that join adjacent stereocilia within the outer hair cell's hair bundle. These links were absent in stereocilin-null mutant mice, which became progressively deaf. At the onset of hearing, however, their cochlear sensitivity and frequency tuning were almost normal, although masking was much reduced and both acoustic and electrical waveform distortions were completely lacking. From this unique functional situation, we conclude that the main source of cochlear waveform distortions is a deflection-dependent hair bundle stiffness resulting from constraints imposed by the horizontal top connectors, and not from the intrinsic nonlinear behaviour of the mechanoelectrical transducer channel.

The cochlea is a highly sensitive and sharply tuned sound detector. It contains two types of sensory cells: outer hair cells (OHCs) that locally amplify and sharpen sound-induced mechanical stimulation of the cochlear partition, and inner hair cells (IHCs) that transmit sensory information to the brain. OHCs have been proposed to supply forces for amplification by changing the length and stiffness of their lateral wall in response to changes in membrane potential (a process known as electromotility)[8–11], or by active movement of their apically located mechanosensory hair bundle[12,13]. The OHC hair bundle is composed of actin-filled stereocilia arrayed in three rows of increasing height. The tallest stereocilia are embedded in the tectorial membrane, an acellular gel overlying the sensory epithelium. The stereocilia are coupled together by the tip link, which extends from the tip of a stereocilium to the side of the adjacent taller one and may gate the mechanoelectrical transducer (MET) channels[12], and by zipper-like horizontal top connectors that join the upper parts of adjacent stereocilia within and between rows[5–7] (Supplementary Fig. 1a).

Mutations in the gene encoding stereocilin (Supplementary Fig. 1b), a protein of the hair bundle, cause prelingual hearing impairment in humans[4]. We engineered mutant mice with a frame-shifting deletion in the stereocilin gene (Supplementary Fig. 2). In postnatal day 60 (P60) stereocilin-null ($Strc^{-/-}$) mice, auditory brainstem responses to tones in the 5–40 kHz frequency range showed increased thresholds, up to 60 dB (data not shown). As these threshold shifts suggested a possible failure of the cochlear amplifier[14,15], we studied cochlear sensitivity, frequency selectivity, as well as waveform distortions and masking interactions among spectral components of a complex sound[15], all ascribed to the functioning of the OHCs.

The signal recorded by a round-window electrode contains the compound action potential (CAP) from the cochlear nerve, representing synchronous neural activity in response to short tone bursts. CAP thresholds, by assessing cochlear sensitivity, allow estimations of amplification due to OHCs, normally ~60 dB at frequencies >10 kHz[16]. P14–P15 mice still have immature CAP thresholds (Fig. 1a versus 1b), but their mechanical and electrical cochlear responses are adult-like (see Figs 1d, 1e and 2a) indicating that the immaturity of their CAP thresholds is probably neural in origin. At P14, the difference between CAP thresholds in $Strc^{-/-}$ mice and those of $Strc^{+/+}$ mice was not statistically significant, less than 10 dB on average in the 5–40 kHz range (Fig. 1a and Supplementary Fig. 3). From P15 on, progressive hearing loss occurred in $Strc^{-/-}$ mice, increasing from a flat 25-dB to a ~60-dB ceiling above 10 kHz at P60 (Fig. 1b and Supplementary Fig. 3).

The round-window electrode also collects the cochlear microphonic potential, a phasic response reflecting MET currents from basal-coil OHCs. As low-frequency tones are not amplified by basal-coil OHCs, the amplitude of the cochlear microphonic response to these tones is proportional to the number of functional MET channels[17]. This number gradually decreased from P15 to P60 in $Strc^{-/-}$ mice (Fig. 1c). Models of active OHC operation (Supplementary Fig. 4) predict how much gain the residual active MET channels would generate[17]. A statistically significant difference was not found between the predicted and the observed auditory thresholds (Fig. 1c), which suggests that the decreased number of functional MET channels in OHCs accounts for the increased CAP thresholds in $Strc^{-/-}$ mice.

How well OHCs determine frequency selectivity, a task dependent on amplification, can be addressed by measuring how a continuous masker tone at varying frequency and level affects the CAP response to a test-tone burst presented simultaneously and just above threshold[18]. Masking occurs when the masker enters the frequency band of

[1]Institut Pasteur, Unité de Génétique et Physiologie de l'Audition, F75015 Paris, France. [2]Inserm UMRS 587, F75015 Paris, France. [3]Collège de France, F75015 Paris, France. [4]Université Pierre et Marie Curie, F75015 Paris, France. [5]University of Sussex, School of Life Sciences, Falmer, Brighton BN1 9QG, UK. [6]Institut Cochin, Plate-Forme de Recombinaison Homologue, F75014 Paris, France. [7]Université d'Auvergne, Laboratoire de Biophysique Sensorielle, F63001 Clermont-Ferrand, France.
*These authors contributed equally to the work.

the cochlear filter centred on the test tone. We studied masking at 10 kHz (Fig. 1d, e) and at 20 kHz (data not shown). In P14 *Strc*$^{-/-}$ mice, CAP tuning curves revealed the persistence of fine frequency selectivity, as the quality factor of tuning, $Q_{10\,dB}$, remained similar to that of controls. CAP tuning curves also provide information on the strength of masking: the more upshifted their tip relative to the test-tone level, the weaker the masking. Masking was weakened in P14 *Strc*$^{-/-}$ mice (Fig. 1d). Despite similar probe-tone levels, high-frequency maskers had to be >20 dB louder than in controls to provide the same CAP decrease. Simultaneous masking is due to two mechanisms[2,19]. One is neuronal and due to the masker producing action potentials that place neurons in the refractory period, and therefore unable to respond to the test tone (line-busy masking). Because stereocilin was not detected in the cochlear ganglion neurons (data not shown), this process should be preserved in *Strc*$^{-/-}$ mice. The other mechanism is due to nonlinear interactions between the masker and the test tones decreasing the mechanical response to the test tone (suppressive masking). This suppressive masking shows up in the basilar membrane mechanical response[18]. Of note, the extent of the decrease of the masking strength observed in *Strc*$^{-/-}$ mice is reminiscent of the suppression of the chinchilla basilar membrane motion induced by a high- or low-frequency tone at levels similar to those used here[20]. This suggests that *Strc*$^{-/-}$ mice totally lack suppressive interactions and that the persisting masking is therefore due to the line-busy mechanism only.

The mechanical nonlinearity that contributes to masking may also distort acoustic waveforms, so that in response to two-tone stimuli (frequencies f_1 and f_2), distortion-product otoacoustic emissions (DPOAEs) at intermodulation frequencies ($2f_1 - f_2$, $f_2 - f_1$, and so

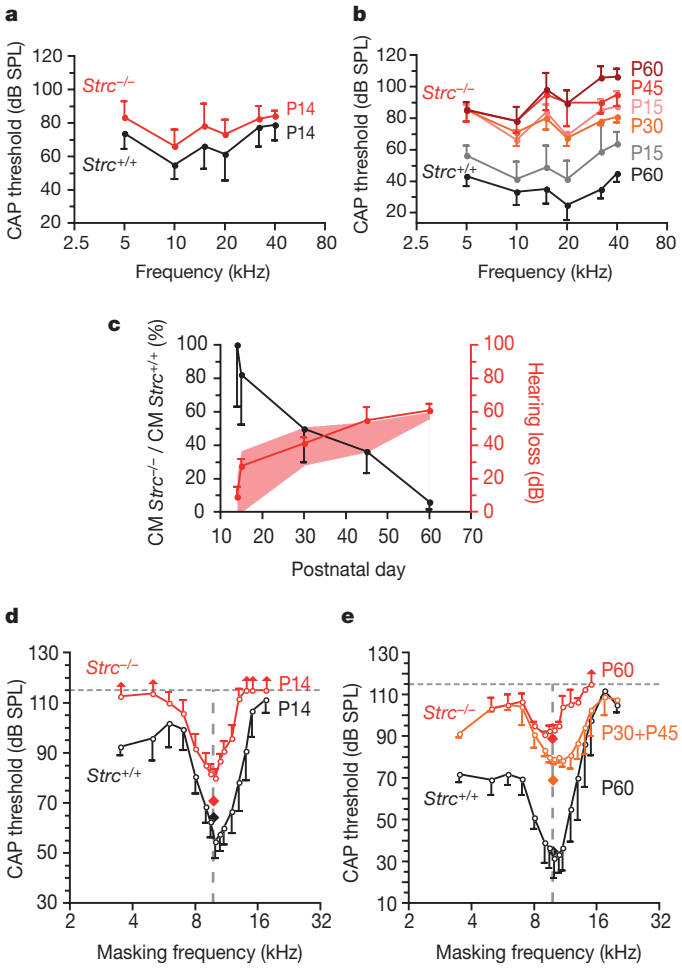


Figure 1 | Hearing sensitivity and frequency tuning in stereocilin-null versus wild-type mice. **a**, Mean CAP thresholds (±s.d., $n = 7$) in P14 mice. The difference between *Strc*$^{+/+}$ and *Strc*$^{-/-}$ mice is non-significant ($P > 0.05$). **b**, Mean CAP thresholds (±s.d., $n = 5$) in P15 to P60 mice. **c**, Ratio of cochlear microphonic (CM) amplitudes in *Strc*$^{-/-}$ versus *Strc*$^{+/+}$ mice at low frequencies (black curve), with predicted (pink shading) and actual (red curve) hearing loss. Differences, predicted versus actual, are non-significant ($P > 0.05$). **d**, CAP tuning curves in P14 mice (±s.d., $n = 7$). $Q_{10\,dB}$ (the ratio of test frequency to CAP tuning curve width at 10 dB above its tip) is similar in *Strc*$^{-/-}$ ($Q_{10\,dB} = 3.1$) and *Strc*$^{+/+}$ ($Q_{10\,dB} = 3.2$) mice. **e**, CAP tuning curves (±s.d.) in P60 *Strc*$^{+/+}$ mice ($Q_{10\,dB} = 3.4$, $n = 5$) and *Strc*$^{-/-}$ mice (P30 and P45, $Q_{10\,dB} = 2.2$, $n = 4$; P60, $Q_{10\,dB} = 3.1$, $n = 4$). The dashed vertical lines in **d** and **e** indicate the frequency of the test tones (10 kHz) and diamonds show the test-tone levels; vertical arrows indicate a lack of masking in some animals for the masking-tone maximal level (115 dB SPL, dashed horizontal lines).

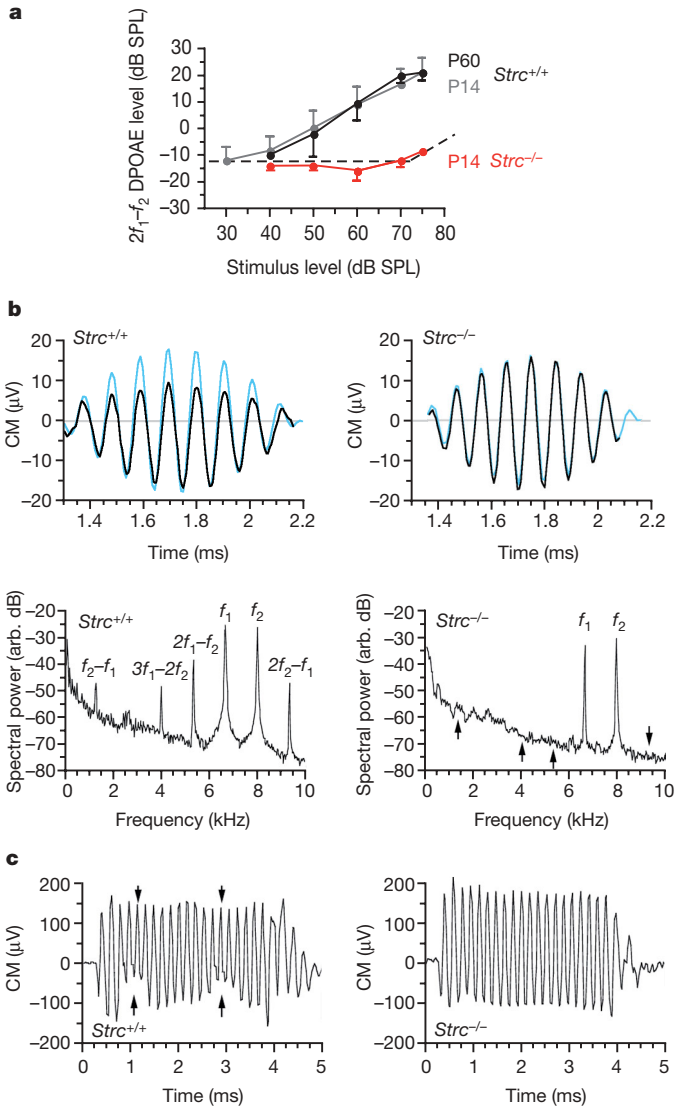


Figure 2 | Waveform distortions in wild-type and stereocilin-null mice. **a**, DPOAE amplitude (mean ± s.d.) at frequency $2f_1 - f_2$ as a function of sound pressure level common to stimuli f_1 and f_2 ($f_2/f_1 = 1.20$) in *Strc*$^{+/+}$ (P14 and P60) and *Strc*$^{-/-}$ (P14) mice ($n = 8$). Dashed line represents noise floor. **b**, Cochlear microphonic (CM) temporal waveforms and frequency spectra (black lines) in P14 mice, in response to one cycle of the beat of a pair of pure-tone stimuli (6.7 and 8 kHz, 65 dB SPL each, blue traces in top panels, scaled to the cochlear microphonic response of the *Strc*$^{+/+}$ ear). Arrows in the *Strc*$^{-/-}$ cochlear microphonic frequency spectrum (bottom right) mark positions of intermodulation frequencies; arb. dB denotes arbitrary dB (dB relative to an arbitrary reference). **c**, Amplitude modulation (arrows) of the cochlear microphonic response to a 5-kHz, 70-dB tone-burst by two periods of a 95-dB SPL infrasound (550 Hz) in P14 *Strc*$^{+/+}$ (left) and *Strc*$^{-/-}$ (right) mice.

on) are propagated backwards and detected as combination tones in the ear canal[21,22]. From P14 onwards, control mice presented DPOAE components well above the noise floor. In contrast, *Strc*$^{-/-}$ mice had no detectable DPOAEs up to 80 dB sound pressure level (SPL; Fig. 2a).

The complete absence of suppressive masking and DPOAEs in *Strc*$^{-/-}$ mice as early as P14, despite normal hearing thresholds and tuning curves, prompted a search for electrical distortion products. These show up as a levelling out of the cochlear microphonic waveforms that originate from MET currents through OHCs (Fig. 2b). In response to tonal stimuli at frequency f in P14 control mice, cochlear microphonic waveforms showed harmonics $2f$ and $3f$ at about −30 dB relative to the levels of the fundamental frequency (data not shown). For two-tone stimuli with neighbouring frequencies f_1 and f_2, P14 control mice had cochlear microphonic waveforms with several intermodulation components reaching from −12 to −15 dB relative to stimulus levels (Fig. 2b). Furthermore, when an infrasound bias (<0.6 kHz) was mixed with an audible frequency (5–10 kHz), the cochlear microphonic waveforms in P14 control mice showed periodic suppressive amplitude modulation, even for biasing levels as low as 85 dB SPL (Fig. 2c). In contrast, P14 *Strc*$^{-/-}$ mice showed none of these cochlear microphonic waveform distortions whatever the intensity of the acoustic stimulation up to 100 dB SPL (Fig. 2b, c).

Histological examination of the cochlea in P14 *Strc*$^{-/-}$ mice did not show any gross structural anomalies. Scanning electron microscopy (SEM) showed that the tectorial membrane extended across the sensory epithelium, as in wild-type mice. Hair-bundle imprints corresponding to the anchoring points of the tallest OHC stereocilia[23] were, however, not observed (data not shown). Nonetheless, cochlear microphonic waveforms had the same phase relative to sound in *Strc*$^{-/-}$ and control mice (Supplementary Fig. 5) indicating that, in both cases, OHCs were driven by the displacement of the cochlear partition and not its velocity as they are in *Tecta*$^{\Delta ENT/\Delta ENT}$ mice, in which the tectorial membrane is detached from the sensory epithelium[24]. Notably, in all OHCs, but not in IHCs, the tops of the stereocilia rows were less clearly aligned in *Strc*$^{-/-}$ compared to *Strc*$^{+/+}$ mice (Fig. 3a, b). In keeping with the apparently loose connection between adjacent stereocilia, the horizontal top connectors that normally develop from P9 onwards and are fully mature at P14 (ref. 7) could not be detected in P14 *Strc*$^{-/-}$ OHCs by scanning (Fig. 3c, d) and transmission (data not shown) electron microscopy. In contrast, the tip links were still present (Fig. 3d).

The distribution of stereocilin was analysed in P14 wild-type mice and compared to that in the *Tecta*$^{\Delta ENT/\Delta ENT}$ mice[24]. Stereocilin was detected in the hair bundles of OHCs only. It was present in the distal regions of all stereocilia rows, in both wild-type and *Tecta*$^{\Delta ENT/\Delta ENT}$ mice (Fig. 4a and data not shown). Scanning immunoelectron microscopy showed that stereocilin was distributed in a ring around the tip of each stereocilium from the tallest row in the OHCs of wild-type mice but not of *Tecta*$^{\Delta ENT/\Delta ENT}$ mice (Fig. 4b, c), suggesting that stereocilin is involved in the contact these stereocilia establish with the tectorial membrane. In addition, stereocilin was detected between all OHC stereocilia, both in wild-type and *Tecta*$^{\Delta ENT/\Delta ENT}$ mice (Fig. 4b, c). Using transmission immunoelectron microscopy, the protein was found to be associated with the horizontal top connectors (Fig. 4d). In OHCs, these links are characterized by a central density[5–7] (Supplementary Fig. 1a) that may be contributed by stereocilin, predicted to be an extracellular protein according to its amino acid sequence.

The presence of DPOAEs in *Tecta*$^{\Delta ENT/\Delta ENT}$ mice[25] indicates that the tectorial membrane is not essential for DPOAE production. The absence of all intermodulation and harmonic components in the responses of P14 *Strc*$^{-/-}$ mice, despite preserved amplification and tuning, thus suggests that their OHC hair bundles lack the very element at the origin of cochlear waveform distortions. This element is commonly thought to be the MET channel[16,26] for two reasons. First, the MET current relates to hair bundle deflection by the Boltzmann's distribution of the MET channels' open probability[27]. This sigmoidal transfer function converts hair bundle movement in response to sinusoidal sound pressure into distorted MET currents. Second, the thermodynamics of shuffling of the MET channels from open to closed states introduces a nonlinear, deflection-dependent stiffness[16,28]. In

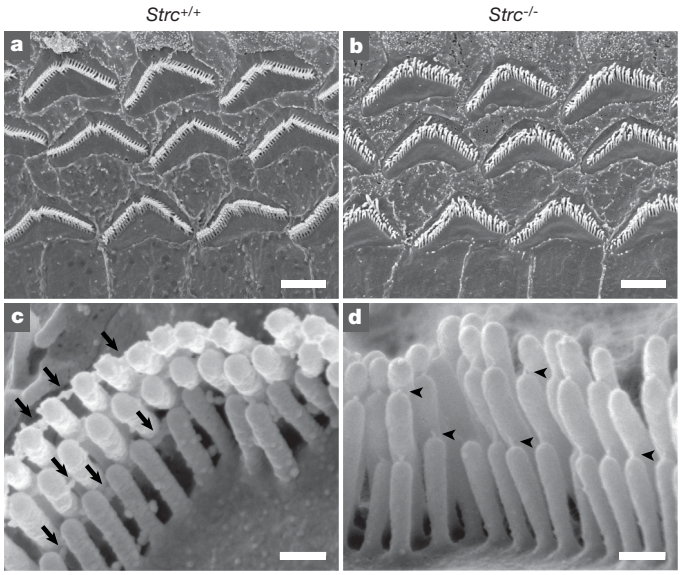


Figure 3 | OHC hair-bundle morphology in P14 wild-type and stereocilin-null mice. **a**, **b**, Low-magnification scanning electron micrographs showing the three rows of OHCs. The tops of the stereocilia are less clearly aligned in the *Strc*$^{-/-}$ mouse (**b**) than in the *Strc*$^{+/+}$ mouse (**a**). **c**, **d**, Detailed views of single OHC hair bundles. Top connectors (arrows in **c**) are not detected in the *Strc*$^{-/-}$ mouse, whereas tip links (arrowheads in **d**) are present. Scale bars, 2.5 μm (**a**, **b**) and 0.25 μm (**c**, **d**).

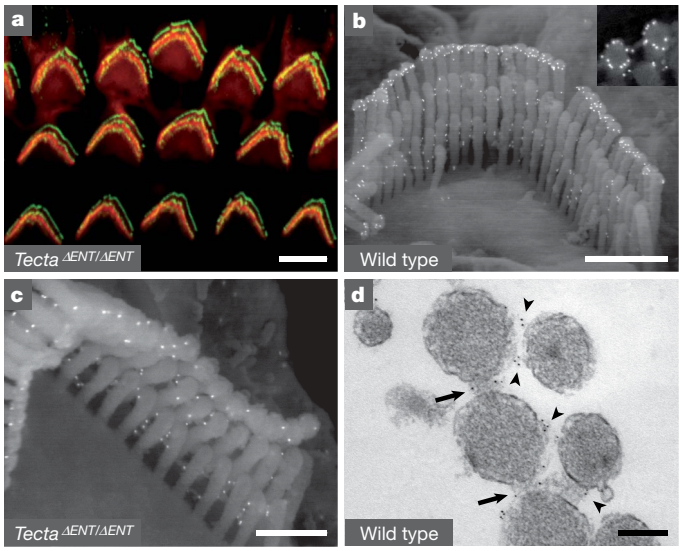


Figure 4 | Immunodetection of stereocilin in wild-type and *Tecta*$^{\Delta ENT/\Delta ENT}$ OHCs. **a**, Confocal image of *Tecta*$^{\Delta ENT/\Delta ENT}$ OHC hair bundles (P14) stained for stereocilin (green) and actin (red). **b**, **c**, Immunogold scanning electron micrographs of wild-type and *Tecta*$^{\Delta ENT/\Delta ENT}$ OHC hair bundles (P14) labelled for stereocilin. A ring-shaped labelling is seen around the tip of the tallest stereocilia in wild-type (**b**, inset) but not in *Tecta*$^{\Delta ENT/\Delta ENT}$ mice (**c**). In contrast, the labelling between stereocilia is seen both in the wild-type and the *Tecta*$^{\Delta ENT/\Delta ENT}$ mice. **d**, Transmission electron micrograph of a stereocilin-labelled OHC hair bundle (P22, transverse section). Top connectors within (arrows) and between (arrowheads) rows are labelled. Scale bars are 5 μm (**a**), 1 μm (**b**), 0.5 μm (**c**) and 0.2 μm (**d**).

Strc$^{-/-}$ mice, one might assume that the MET transfer function has become linear over the waveform distortion-free 30–100 dB SPL range. The corollary of this, however, is a reduction of the energy difference between the open and the closed states relative to the thermal energy (kT)[26] that would then prevent MET channels from operating normally, which they do in P14 *Strc*$^{-/-}$ mice. The nonlinear source of all cochlear waveform distortions that also permits suppressive masking should thus come not from the intrinsic working properties of the MET channels, but from a non-conductive hair-bundle stiffness impinging the response of the MET channels. Being the missing structure in waveform-distortion-free *Strc*$^{-/-}$ mice, horizontal top connectors qualify as an important source of nonlinear stiffness, either directly or indirectly by providing the necessary condition for another mechanical element to operate nonlinearly. The bending of neighbouring stereocilia towards each other that is observed when stereocilin-associated lateral links extend further towards the stereocilia base in mouse models of Usher syndrome type 1 (data not shown and Supplementary Fig. 6) supports the importance of these links in hair bundle non-conductive stiffness. The lateral links and pivots of OHC stereocilia are well-acknowledged contributors to the component of the hair bundle stiffness derived from the restoring force. Being reportedly large[13,27,29,30], this stiffness component would indeed provide a more suitable basis for the large waveform distortions and the strong suppressive masking observed in normal cochleae, than the conductive stiffness related to the MET channels.

METHODS SUMMARY

Production of *Strc* knockout mice. Stereocilin knockout mice (*Strc*$^{tm1Ugds/tm1Ugds}$, referred to as *Strc*$^{-/-}$) were produced using the Cre–*lox* system. We engineered mice in which exons 2 and 3 of *Strc* were floxed (*loxP*–*Strc*). *Strc*$^{-/-}$ mice were obtained by crossing these mice with a transgenic mouse expressing the Cre recombinase early and ubiquitously.

Recordings of cochlear microphonics, compound action potentials and otoacoustic emissions. An electrode was placed in the round-window niche of anaesthetized mice. The receptor currents of basal sensory cells and the synchronous activity of the cochlear nerve were recorded in response to tone-bursts and mixtures of tone-bursts and interfering tones. The sensitivity, frequency selectivity, susceptibility to masking and ability of waveforms to exhibit distortion were compared in *Strc*$^{-/-}$ versus *Strc*$^{+/+}$ mice. A miniature microphone probe in the external ear canal allowed combination tones in response to two-tone stimuli to be recorded as otoacoustic emissions.

Histological analysis and immunolocalization. The morphology of the cochlear sensory epithelia was studied in *Strc*$^{+/+}$ and *Strc*$^{-/-}$ mice using SEM. For immunolocalization of stereocilin, we used two affinity-purified polyclonal rabbit antibodies, anti-B and anti-D, raised against the synthetic peptides CFLSPEELQSLVPLSD (amino acids 970–985) and EQLAYLSPEQRRAVA (amino acids 1753–1767) derived from the mouse stereocilin sequence (AF375593), respectively (Supplementary Fig. 1b). The distribution of stereocilin was analysed using confocal immunofluorescence microscopy, scanning and transmission immunoelectron microscopy.

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

Received 27 April; accepted 28 August 2008.
Published online 8 October 2008.

1. Goldstein, J. L. Auditory nonlinearity. *J. Acoust. Soc. Am.* **41,** 676–689 (1967).
2. Delgutte, B. in *Auditory Computation* (eds Hawkins, H. H., McMullen, T. A., Popper, A. N. & Fay, R. R.) 157–220 (Springer, 1996).
3. Kemp, D. T. Stimulated acoustic emissions from within the human auditory system. *J. Acoust. Soc. Am.* **64,** 1386–1391 (1978).
4. Verpy, E. *et al.* Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus. *Nature Genet.* **29,** 345–349 (2001).
5. Furness, D. N. & Hackney, C. M. Cross-links between stereocilia in the guinea pig cochlea. *Hear. Res.* **18,** 177–188 (1985).
6. Tsuprun, V. & Santi, P. Structure of outer hair cell stereocilia side and attachment links in the chinchilla cochlea. *J. Histochem. Cytochem.* **50,** 493–502 (2002).
7. Goodyear, R. J., Marcotti, W., Kros, C. J. & Richardson, G. P. Development and properties of stereociliary link types in hair cells of the mouse cochlea. *J. Comp. Neurol.* **485,** 75–85 (2005).
8. Brownell, W. E., Bader, C. R., Bertrand, D. & de Ribaupierre, Y. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* **227,** 194–196 (1985).
9. Ashmore, J. F. A fast motile response in guinea-pig outer hair cells: the cellular basis of the cochlear amplifier. *J. Physiol. (Lond.)* **388,** 323–347 (1987).
10. He, D. Z. & Dallos, P. Somatic stiffness of cochlear outer hair cells is voltage-dependent. *Proc. Natl Acad. Sci. USA* **96,** 8223–8228 (1999).
11. Liberman, M. C. *et al.* Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature* **419,** 300–304 (2002).
12. Hudspeth, A. J. How the ear's works work. *Nature* **341,** 397–404 (1989).
13. Kennedy, H. J., Crawford, A. C. & Fettiplace, R. Force generation by mammalian hair bundles supports a role in cochlear amplification. *Nature* **433,** 880–883 (2005).
14. Dallos, P. & Harris, D. Properties of auditory nerve responses in absence of outer hair cells. *J. Neurophysiol.* **41,** 365–383 (1978).
15. Liberman, M. C. & Dodds, L. W. Single-neuron labeling and chronic cochlear pathology. III. Stereocilia damage and alterations of threshold tuning curves. *Hear. Res.* **16,** 55–74 (1984).
16. Robles, L. & Ruggero, M. A. Mechanics of the mammalian cochlea. *Physiol. Rev.* **81,** 1305–1352 (2001).
17. Patuzzi, R. B., Yates, G. K. & Johnstone, B. M. Outer hair cell receptor current and sensorineural hearing loss. *Hear. Res.* **42,** 47–72 (1989).
18. Dallos, P. & Cheatham, M. A. Compound action potential (AP) tuning curves. *J. Acoust. Soc. Am.* **59,** 591–597 (1976).
19. Moore, B. C. J. *Cochlear Hearing Loss* (Whurr Publishers Ltd, 1998).
20. Ruggero, M. A., Robles, L. & Rich, N. C. Two-tone suppression in the basilar membrane of the cochlea: mechanical basis of auditory-nerve rate suppression. *J. Neurophysiol.* **68,** 1087–1099 (1992).
21. Kim, D. O., Molnar, C. E. & Matthews, J. W. Cochlear mechanics: nonlinear behavior in two-tone responses as reflected in cochlear-nerve-fiber responses and in ear-canal sound pressure. *J. Acoust. Soc. Am.* **67,** 1704–1721 (1980).
22. Robles, L., Ruggero, M. A. & Rich, N. C. Two-tone distortion on the basilar membrane of the chinchilla cochlea. *J. Neurophysiol.* **77,** 2385–2399 (1997).
23. Kimura, R. S. Hairs of the cochlear sensory cells and their attachment to the tectorial membrane. *Acta Otolaryngol. (Stockh.)* **61,** 55–72 (1966).
24. Legan, P. K. *et al.* A targeted deletion in α-tectorin reveals that the tectorial membrane is required for the gain and timing of cochlear feedback. *Neuron* **28,** 273–285 (2000).
25. Lukashkin, A. N., Lukashkina, V. A., Legan, P. K., Richardson, G. P. & Russell, I. J. Role of the tectorial membrane revealed by otoacoustic emissions recorded from wild-type and transgenic *Tecta*$^{\Delta ENT/\Delta ENT}$ mice. *J. Neurophysiol.* **91,** 163–171 (2004).
26. Patuzzi, R. in *The Cochlea* (eds Dallos, P., Popper, A. N. & Fay, R. R.) 186–257 (Springer, 1996).
27. van Netten, S. M. & Kros, C. J. Gating energies and forces of the mammalian hair cell transducer channel and related hair bundle mechanics. *Proc. R. Soc. Lond. B.* **267,** 1915–1923 (2000).
28. Howard, J. & Hudspeth, A. J. Compliance of the hair bundle associated with gating of mechanoelectrical transduction channels in the bullfrog's saccular hair cell. *Neuron* **1,** 189–199 (1988).
29. Flock, A. & Strelioff, D. Graded and nonlinear mechanical properties of sensory hairs in the mammalian hearing organ. *Nature* **310,** 597–599 (1984).
30. Fettiplace, R. Active hair bundle movements in auditory hair cells. *J. Physiol. (Lond.)* **576,** 29–36 (2006).

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

Acknowledgements We thank M. Lenoir, S. Guadagnini and M.-C. Prévost for advice on SEM; E. Perret and P. Roux for advice on confocal microscopy, S. Chardenoux, S. Nouaille and A. Mallet for technical help, Y. Lallemand for providing us with PGK-Cre^{m} mice, N. Michalski for help with figure drawing and J. Boutet de Monvel for critical reading of the manuscript. This work was supported by the European Commission FP6 Integrated Project EuroHear, Fondation Raymonde et Guy Strittmatter, Région Ile-de-France (Programme Sésame), and the Wellcome Trust.

Author Contributions C.P. and P.A. contributed equally as co-senior authors. D.W. and M.L. are co-second authors. D.W. and G.H. produced the *Strc* knockout mice. P.A. conducted the auditory tests. M.L. performed the SEM experiments on wild type, *Strc*$^{-/-}$ and *Tecta*$^{\Delta ENT/\Delta ENT}$ mice, with help from C.H. G.M.L. carried out the SEM experiments in the mouse models of Usher syndrome. E.V. performed the immunofluorescence studies with help from C.H., and the RT–PCR analysis. G.P.R. and R.J.G. carried out the transmission electron microscopy analysis. E.V. supervised the stereocilin expression studies and the characterization of the morphological anomalies in the *Strc* knockout mice. C.P. supervised the whole project with help from G.P.R. and in collaboration with P.A. for the physiological studies. E.V., J.-P.H., G.P.R., P.A. and C.P. prepared the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.P. (cpetit@pasteur.fr) or E.V. (everpy@pasteur.fr).

METHODS

Generation of *Strc*$^{tm1Ugds/tm1Ugds}$ knockout mice. We designed a targeting vector in which exons 2 and 3 of *Strc* and the hygromycin selection cassette were flanked with *loxP* sites. Electroporation of the linearized vector into CK35 embryonic stem cells[31] (derived from a 129/Sv mouse embryo) plated on hygromycin selective medium resulted in approximately 300 recombinant embryonic stem cell clones, of which five were correctly targeted. Two independent recombinant embryonic stem cell lines carrying the floxed allele were selected and injected into C57BL/6J host blastocysts. Mating of male chimaeras with C57BL/6J females produced heterozygous animals. Integration of the recombinant DNA construct was confirmed by Southern blot analysis and PCR amplification of genomic DNA extracted from mice's tails. For Southern blot analysis, genomic DNA was digested with BglII and transferred on Immobilon-NY^{+} membrane (Millipore). The membrane was probed with a PCR-amplified fragment obtained by using primers 5′-GAGCTTCTGTCCAGTGATAGTTCAG-3′ and 5′-TGCTTAGGAAGCTTTCTGCAGCATGGG-3′. The PCR primers used to genotype the floxed *Strc* allele were HR (5′-TGGACGTAAACTCCTCTTCAGACC-3′) located in the hygromycin resistance gene, and *Strc*-R1 (5′-AGGCTGAGCCCACAGCACAAAG-3′) located in *Strc* intron 4 outside the targeting vector. Mice heterozygous for the floxed *Strc* allele were mated with the PGK-Cre^{m} transgenic mouse strain carrying the Cre recombinase gene driven by the early acting phosphoglycerate kinase-1 gene promoter[32]. Targeted deletion of exons 2 and 3 was confirmed by PCR analysis. Routine genotyping of animals for the *Strc*^{tm1Ugds} allele was carried out by two PCR amplifications, using forward primer *Strc*-F1 (5′-GGGCTCTGAGGAGGCTCTTTGGG-3′) located in exon 2, or *Strc*-F2 (5′-TGGGATTTGAACTCAGGTTGCTAGG-3′) located in intron 1, and reverse primer *Strc*-R2 (5′-CAGAGGCACACCTCTGCTCAGG-3′) located in exon 4. All studies were performed on mixed C57BL/6J × 129/Sv genetic backgrounds.

For RT–PCR analysis, total RNA was isolated from single inner ears of P15 *Strc*$^{+/+}$, *Strc*$^{+/-}$ and *Strc*$^{-/-}$ mice, and the forward primer 5′-53-TCTAGGCCCAGTGTGCACCT-3′ located in exon 1, and the reverse primer 5′-1192-GGCAGAGCAAGTAGATGGAGAAGTTGG-3′ located in exon 4, were used for amplification. The amplicons were gel purified and sequenced.

Recordings of cochlear microphonics, compound action potentials and otoacoustic emissions. For cochlear microphonics, CAPs and DPOAE measurements, mice were anaesthetized with ketamine and levomepromazin. CAP and cochlear microphonic responses were collected between a silver electrode inserted in the round window niche and a vertex needle electrode. CAP and cochlear microphonic potential were evoked by tone bursts (2-period rise-fall, 20-period plateau, repetition rate 20 s^{-1}, sound level varying from 10 to 115 dB SPL in 5-dB steps). Masking and biasing were produced by adding to the test-tone burst a continuous pure tone with variable frequency and level (2.5 dB steps). The response from the electrodes was amplified (gain 5,000), filtered (0.02–30 kHz), digitally converted, and averaged using a computerised data-acquisition system. Visual inspection was used to determine the CAP thresholds, or whether the CAP amplitude was halved in the presence of an increasingly loud masker. The DPOAE at frequency $2f_1 - f_2$ was recorded in response to two equal level primary tones, f_1 and f_2, with $f_2/f_1 = 1.20$. Ear-canal sound pressure was amplified (×100) and averaged (0.5 s). The amplitudes of the DPOAE at $2f_1 - f_2$ and of the background noise were extracted by fast Fourier transform. For cochlear microphonic waveform distortion measurements, the response from the electrode to one or two primary tones (levels ranging from 50 to 100 dB SPL) was Fourier-transformed and analysed at harmonic and intermodulation frequencies. Statistical differences were evaluated using the Student's *t*-test.

Antibodies and immunolabelling studies. Two rabbit immune sera, anti-B and anti-D, were produced against synthetic peptides B (NH_2-970-CFLSPEELQSLVPLSD-COOH) and D (NH_2-1753-EQLAYLSPEQRRAVA-COOH) derived from the mouse stereocilin amino-acid sequence, respectively (see Supplementary Fig. 1b). The antibodies were affinity purified on the corresponding peptides, and their specificity was verified by immunoblotting and immunostaining of transfected cells (Hela, Cos7, MDCK) producing full-length stereocilin, and by the loss of the inner ear immunolabelling in *Strc*$^{-/-}$ mice. Both antibodies produced similar staining patterns, but immunolabelling with anti-D required pretreatment with methanol (Triton X-100 was ineffective), which precluded subsequent phalloidin staining. The results presented are those obtained using the anti-B antibody. For immunofluorescence detection, we used Alexa-Fluor-488-conjugated goat anti-rabbit F(ab′)2 IgG fragment (1:800, Molecular Probes). Actin was labelled with TRITC-conjugated phalloidin (1 μg ml^{-1}, Sigma-Aldrich). Samples were processed as described[33], with the following modifications. Inner ears were fixed immediately on removal by immersion in 4% paraformaldehyde in PBS pH 7.4 (PFA/PBS) for 30 min at room temperature. After three rinses in PBS, cochlear sensory areas were microdissected and re-fixed in PFA/PBS for 30 min. For subsequent immunolabelling with anti-D, tissues were immersed in methanol (−20 °C) for 3 to 5 min. The tissues were preincubated in PBS containing 20% goat serum for 1 h at room temperature, before overnight incubation with the anti-stereocilin antibodies (~1 μg ml^{-1}) in PBS containing 1% bovine serum albumin (PBS/BSA). Whole-mount preparations were analysed using a laser scanning confocal microscope (LSM-510, Zeiss). Immunolabelling for SEM was performed as for immunofluorescence detection. Primary antibodies were detected with protein-A-conjugated 15 nm colloidal gold particles (EM Laboratory, Utrecht University, The Netherlands; diluted 1:60 in PBS/BSA). Finally, samples were post-fixed for 1 h in 2.5% glutaraldehyde in PBS before proceeding to SEM.

Scanning electron microscopy. Samples were processed as described[34]. When immunolabelling was coupled to SEM, dried specimens were mounted on colloidal silver adhesive (quick drying silver paint, Agar Scientific) to enhance their conductivity. Specimens were then coated with 30 nm of carbon and imaged using an Yttrium Aluminium Garnet detector.

Transmission electron microscopy. For immunogold electron microscopy, pre-embedding labelling using the anti-B antibody was carried out as described[33]. Procedures for transmission electron microscopy were as described[7].

31. Kress, C., Vandormael-Pournin, S., Baldacci, P., Cohen-Tannoudji, M. & Babinet, C. Nonpermissiveness for mouse embryonic stem (ES) cell derivation circumvented by a single backcross to 129/Sv strain: establishment of ES cell lines bearing the Omd conditional lethal mutation. *Mamm. Genome* **9**, 998–1001 (1998).
32. Lallemand, Y., Luria, V., Haffner-Krausz, R. & Lonai, P. Maternally expressed PGK-Cre transgene as a tool for early and uniform activation of the Cre site-specific recombinase. *Transgenic Res.* **7**, 105–112 (1998).
33. Michel, V. *et al.* Cadherin 23 is a component of the transient lateral links in the developing hair bundles of cochlear sensory cells. *Dev. Biol.* **280**, 281–294 (2005).
34. Delmaghani, S. *et al.* Mutations in the gene encoding pejvakin, a newly identified protein of the afferent auditory pathway, cause DFNB59 auditory neuropathy. *Nature Genet.* **38**, 770–778 (2006).

LETTERS

A key role for autophagy and the autophagy gene *Atg16l1* in mouse and human intestinal Paneth cells

Ken Cadwell[1], John Y. Liu[1], Sarah L. Brown[1], Hiroyuki Miyoshi[1], Joy Loh[1], Jochen K. Lennerz[1], Chieko Kishi[5], Wumesh Kc[1], Javier A. Carrero[1], Steven Hunt[2], Christian D. Stone[3], Elizabeth M. Brunt[1], Ramnik J. Xavier[6], Barry P. Sleckman[1], Ellen Li[3], Noboru Mizushima[5], Thaddeus S. Stappenbeck[1]* & Herbert W. Virgin IV[1,4]*

Susceptibility to Crohn's disease, a complex inflammatory disease involving the small intestine, is controlled by over 30 loci[1]. One Crohn's disease risk allele is in *ATG16L1*, a gene homologous to the essential yeast autophagy gene *ATG16* (ref. 2). It is not known how ATG16L1 or autophagy contributes to intestinal biology or Crohn's disease pathogenesis. To address these questions, we generated and characterized mice that are hypomorphic for ATG16L1 protein expression, and validated conclusions on the basis of studies in these mice by analysing intestinal tissues that we collected from Crohn's disease patients carrying the Crohn's disease risk allele of *ATG16L1*. Here we show that ATG16L1 is a bona fide autophagy protein. Within the ileal epithelium, both ATG16L1 and a second essential autophagy protein ATG5 are selectively important for the biology of the Paneth cell, a specialized epithelial cell that functions in part by secretion of granule contents containing antimicrobial peptides and other proteins that alter the intestinal environment[3]. ATG16L1- and ATG5-deficient Paneth cells exhibited notable abnormalities in the granule exocytosis pathway. In addition, transcriptional analysis revealed an unexpected gain of function specific to ATG16L1-deficient Paneth cells including increased expression of genes involved in peroxisome proliferator-activated receptor (PPAR) signalling and lipid metabolism, of acute phase reactants and of two adipocytokines, leptin and adiponectin, known to directly influence intestinal injury responses. Importantly, Crohn's disease patients homozygous for the *ATG16L1* Crohn's disease risk allele displayed Paneth cell granule abnormalities similar to those observed in autophagy-protein-deficient mice and expressed increased levels of leptin protein. Thus, ATG16L1, and probably the process of autophagy, have a role within the intestinal epithelium of mice and Crohn's disease patients by selective effects on the cell biology and specialized regulatory properties of Paneth cells.

Crohn's disease typically involves the distal ileum and in its most severe form is characterized by pathological changes including transmural acute and chronic inflammation. Crohn's disease is associated with increased expression of immunoregulatory cytokines including leptin[4] and adiponectin[5]. Multiple genetic factors predispose to Crohn's disease, but the specific relationship between the function of these genes and the diverse pathologies observed in Crohn's disease is not well understood. One Crohn's disease susceptibility allele is in the predicted autophagy gene *ATG16L1* (refs 6–9). Autophagy is an evolutionarily conserved process that recycles cellular components by means of delivery of double-membrane-bound vesicles containing cytoplasm and cytoplasmic organelles to the lysosome[10]. Autophagy has an important role in cell and tissue homeostasis, and has been implicated in a range of human diseases[10]. The mammalian ATG16L1 protein contains an amino-terminal domain that is homologous to yeast Atg16 (ref. 2), which functions in autophagy as part of a complex with autophagy proteins Atg5 and Atg12 (refs 2, 11). Atg16 is responsible, in yeast and mammalian cells, for proper subcellular localization of the autophagy machinery[8,11,12].

To determine the role of ATG16L1 and autophagy in the intestine, we generated two mouse lines with gene-trap-mediated disruptions of *Atg16l1* and a third line lacking *Atg5* in intestinal epithelial cells. Gene-trap vectors introduce a false splice acceptor into an intron, and can inhibit expression of intact messenger RNA[13] (Fig. 1a). This can result in decreased expression of a protein, potentially generating viable mice when full disruption of a gene is lethal. This approach is attractive because ATG5 is part of the ATG16L1 complex and full disruption of *Atg5* is lethal[2,14]. Mouse lines homozygous for gene-trap mutations were generated from commercially available embryonic stem cells carrying gene-trap mutations in the intron 3′ to either exon 6 or exon 10 of *Atg16l1* (ref. 13). *Atg16l1* mutant mouse embryonic fibroblasts (MEFs) expressed low levels of ATG16L1 protein (Fig. 1b) indicating that both mouse lines, $ATG16L1^{HM1}$ and $ATG16L1^{HM2}$, are hypomorphic (HM) for expression of the ATG16L1 protein. $ATG16L1^{HM}$ mice were born at Mendelian ratios (Supplementary Fig. 1a) and survive to adulthood; the characteristics of the two $ATG16L1^{HM}$ lines *in vivo* were similar across all experiments.

To determine whether ATG16L1 is an autophagy protein, we studied low-passage and transformed MEFs from $ATG16L1^{HM}$ mice compared to MEFs lacking the essential autophagy protein ATG5 (ref. 14). Rapamycin-induced and autophagy-dependent[15] degradation of the adaptor protein sequestosome (SQSTM1, also known as P62) was diminished in $ATG16L1^{HM}$ and $Atg5^{-/-}$ cells (Fig. 1c–e). Decreased degradation of P62 in $ATG16L1^{HM}$ cells was restored by expressing ATG16L1 (Supplementary Fig. 1b, c). $ATG16L1^{HM}$ MEFs also showed diminished rapamycin-induced production of LC3-II, the phosphatidylethanolamine-conjugated form of microtubule-associated protein 1 light chain 3 (LC3-I) generated during autophagy[10] (Fig. 1e, f). The induction of autophagosomes, as measured by LC3-positive dot formation after rapamycin treatment or starvation, was also decreased in $ATG16L1^{HM}$ MEFs (Supplementary Fig. 2), although the defect of $ATG16L1^{HM2}$ cells was more subtle in starved cells; this was confirmed in cells transfected with green fluorescent protein (GFP)–LC3 (Supplementary Fig. 3). Mammalian ATG16L1 is therefore an autophagy protein.

We next determined whether markers of autophagy were abnormal in the distal small intestine (ileum, a common site of Crohn's disease) of $ATG16L1^{HM}$ mice by measuring the expression of ATG16L1, LC3 and P62 proteins in ileal lysates (Fig. 1g–i). *Atg16l1*

[1]Department of Pathology and Immunology, [2]Department of Surgery, [3]Department of Medicine, [4]Department of Molecular Microbiology, Washington University School of Medicine, St Louis, Missouri 63110, USA. [5]Department of Physiology and Cell Biology, Tokyo Medical and Dental University Graduate School and Faculty of Medicine, Tokyo 113-8519, Japan. [6]Center for Computational and Integrative Biology and Gastrointestinal Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts 02114, USA.
*These authors contributed equally to this work.

mRNA is expressed throughout the crypt–villus axis (Supplementary Fig. 4). ATG16L1^{HM} mice expressed 23–37% of the expected level of ATG16L1 protein (Fig. 1g, h). Consistent with a role for ATG16L1 in ileal autophagy, both the ratio of LC3-I to LC3-II and the total amount of LC3-I and P62 were increased in lysates from ATG16L1^{HM} mice (Fig. 1g, i). To validate these results, we studied ileal lysates from mice generated by breeding *Atg5*$^{flox/flox}$ mice[16] with mice expressing the Cre recombinase under the control of the intestinal-epithelium-specific villin promoter (*Atg5*$^{flox/flox}$villin-*Cre* mice)[17]. ATG5 expression was diminished in ileal lysates of *Atg5*$^{flox/flox}$villin-*Cre* mice, and these mice exhibited changes in expression of LC3 and P62 similar to those observed in ATG16L1^{HM} mice (Supplementary Fig. 5). These data are consistent with a role for ATG16L1 and ATG5 in ileal autophagy.

Deficiency in ATG16L1 had no effect on the overall morphology of the ileum or the colon, as measured by analysis of crypt height or villus length (data not shown). However, there were obvious abnormalities in Paneth cells, leading us to focus our studies on these critically important intestinal innate immune cells. Paneth cells are ileal epithelial cells thought to have a role in control of intestinal microbiota by means of secretion of granule contents including antimicrobial peptides and lysozyme[18]. Staining of whole mounts of ileum revealed the expected co-localization of lysozyme and mucus (Fig. 2a). However, there was a notable lack of lysozyme staining in the mucus of ATG16L1^{HM} mice, suggesting an abnormality of Paneth cell secretion (Fig. 2b). We examined periodic acid-Shiff (PAS)/alcian-blue-stained sections and found extraordinary abnormalities in Paneth cells including aberrant, disorganized granules as well as decreased granule numbers (Fig. 2c, d and data not shown). Blinded analysis of these sections from 7 control and 12 mutant mice revealed a 100% concordance between ATG16L1^{HM} genotype and abnormal Paneth cell morphology. We quantified these Paneth granule abnormalities by staining sections for lysozyme, which is normally packaged efficiently in the granules[18] (Fig. 2e–h). In these sections, we observed a notable population of ATG16L1^{HM} cells with diffuse lysozyme staining (Fig. 2f, h). We also observed the presence of intact granules in the crypt lumen of

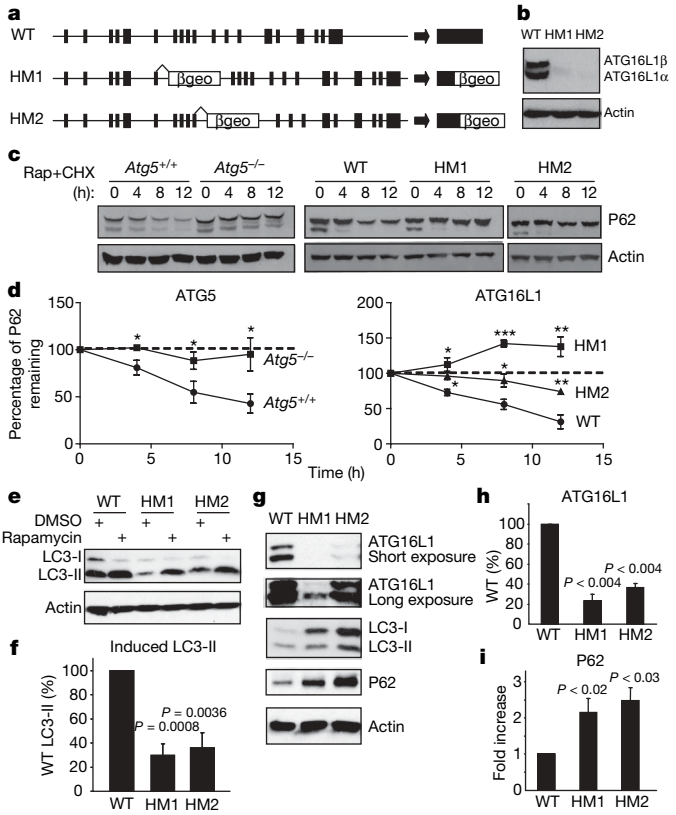


Figure 1 | Role of ATG16L1 in autophagy in cells and intestine from ATG16L1^{HM} mice. **a**, A gene-trap vector containing a splice acceptor followed by the βgeo (β-galactosidase/neomycin resistance fusion cassette) exon disrupts the *Atg16l1* gene locus by insertions within the intronic regions after exons 6 and 10 for the ATG16L1^{HM1} and ATG16L1^{HM2} lines, respectively. WT, wild type. **b**, MEFs were generated from day 12.5 embryos of *Atg16l1* mutant mice, and western blot analysis was used to detect the α and β isoforms of ATG16L1 from whole-cell lysates. **c**, **d**, *Atg5*$^{-/-}$ and ATG16L1^{HM} MEFs were grown in the presence of the chemical inducer of autophagy rapamycin (Rap, 50 μg ml^{-1}) and cyclohexamide (CHX, 5 μg ml^{-1}) for 0, 4, 8 and 12 h. Cell lysates were analysed by western blot for loss of P62 expression (**c**). P62 levels were quantified by densitometry normalized to actin ($n = 3$; **d**). $^{*}P < 0.04$; $^{**}P < 0.008$; $^{***}P = 0.0003$. **e**, **f**, Western blot analysis of LC3 expression in MEFs grown in the presence of rapamycin (50 μg ml^{-1}) or dimethylsulphoxide (DMSO) control for 4 h (**e**). LC3-II expression in the presence of rapamycin was quantified and normalized to actin ($n = 6$; **f**). **g**, Detection of ATG16L1, LC3 and P62 by western blot of ileal lysates from ATG16L1^{HM} mice. ATG16L1 can be detected in both mutant lines on longer exposures. **h**, **i**, Quantification of ATG16L1 (**h**) and P62 (**i**) normalized to actin detected in ileal lysates ($n = 3$). *P* values were calculated using two-tailed Student's *t*-test. Error bars, s.e.m.

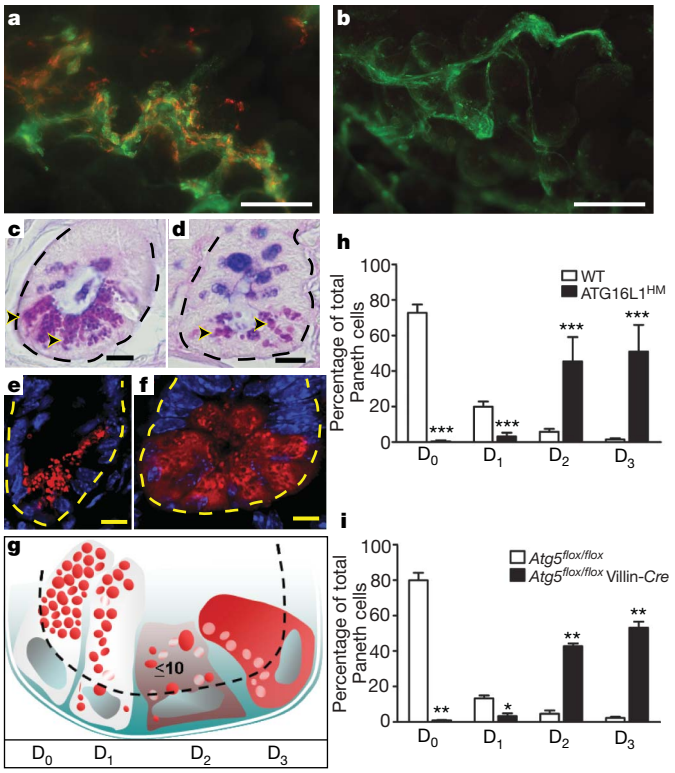
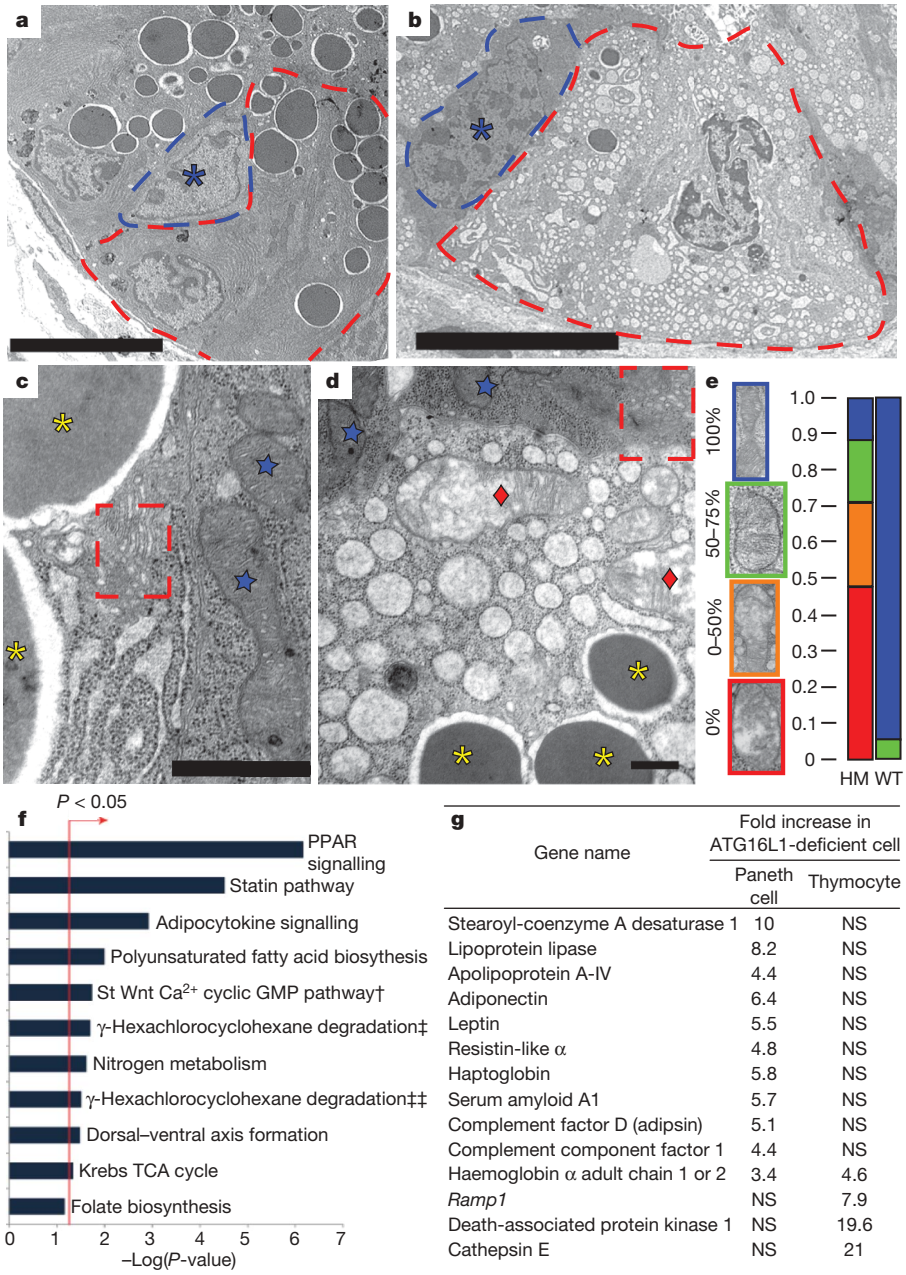


Figure 2 | Mutation of ATG16L1 or ATG5 leads to disruption of the Paneth cell granule exocytosis pathway. **a**, **b**, Whole-mount images taken immediately above the ileal mucosal surface from WT (**a**) and ATG16L1^{HM} (**b**) littermate mice stained with *Helix pomatia* lectin that labels mucus (green) and antisera directed against lysozyme (red). Images are representative of 3 WT and 3 ATG16L1^{HM} mice. **c**, **d**, Ileal sections from WT (**c**) and ATG16L1^{HM} (**d**) mice were stained with PAS/alcian blue to detect Paneth granules by light microscopy (dashed line denotes the crypt unit and arrowheads indicate Paneth cells). Images are representative of 7 WT, 7 ATG16L1^{HM1} and 5 ATG16L1^{HM2} mice; more than 100 crypts were analysed for each. **e**, **f**, Representative images of indirect immunofluorescence of sections stained for lysozyme (red) in WT (**e**) and ATG16L1^{HM} (**f**) mouse ileal crypts. **g**, Paneth cells display one of four patterns of lysozyme expression (represented in red; white represents areas that exclude lysozyme): normal (D_0), disordered (D_1), depleted (D_2) and diffuse (D_3). **h**, **i**, Number of Paneth cells from ATG16L1^{HM} (**h**) and *Atg5*$^{flox/flox}$villin-*Cre* (**i**) mice displaying each pattern of lysozyme expression ($n = 6{,}660$ cells from 5 ATG16L1^{HM} mice and 5,634 cells from 5 WT mice; $n = 1{,}756$ cells from 2 *Atg5*$^{flox/flox}$villin-*Cre* mice and 1,649 from 2 *Atg5*$^{flox/flox}$ mice). Scale bars: **a**, **b**, 200 μm; **c**–**f**, 10 μm. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. *P* values were calculated using two-tailed Student's *t*-test. Error bars, s.e.m.

whole mounts of ileum from ATG16L1^{HM} mice (Supplementary Fig. 6a–c), a finding confirmed by electron microscopy (Supplementary Fig. 6d, e), potentially explaining the absence of lysozyme staining in the mucus layer of the ileum (Fig. 3a, b). These observations indicate that ATG16L1 is required for maintaining the integrity of the Paneth cell granule exocytosis pathway. On the basis of these data, and the published role for NOD2 in resistance to infection[19], we orally challenged ATG16L1^{HM} mice with *Listeria monocytogenes*. We found no change in *L. monocytogenes* titres in spleen, liver and mesenteric lymph nodes (Supplementary Fig. 7), indicating that the marked changes in release of granules in ATG16L1^{HM} mice did not affect *L. monocytogenes* resistance. This argues that the phenotypes of mutations analysed so far in ATG16L1 and NOD2 are distinct.

Despite these profound alterations in the granule exocytosis pathway, we found no evidence for increased epithelial cell death or proliferation, as determined by quantification of apoptotic bodies and



g

Gene name	Fold increase in ATG16L1-deficient cell	
	Paneth cell	Thymocyte
Stearoyl-coenzyme A desaturase 1	10	NS
Lipoprotein lipase	8.2	NS
Apolipoprotein A-IV	4.4	NS
Adiponectin	6.4	NS
Leptin	5.5	NS
Resistin-like α	4.8	NS
Haptoglobin	5.8	NS
Serum amyloid A1	5.7	NS
Complement factor D (adipsin)	5.1	NS
Complement component factor 1	4.4	NS
Haemoglobin α adult chain 1 or 2	3.4	4.6
Ramp1	NS	7.9
Death-associated protein kinase 1	NS	19.6
Cathepsin E	NS	21

Figure 3 | Critical role of ATG16L1 in the structure and transcriptional profile of Paneth cells. a–d, Electron micrograph of littermate WT (**a**, **c**) and ATG16L1^{HM} (**b**, **d**) Paneth cells (for **a**, **b**, dashed lines denote individual cells, asterisk indicates normal progenitor cell; for **c**, **d**, blue stars indicate normal mitochondria in WT Paneth cells and ATG16L1^{HM} progenitor cells, red diamonds indicate degenerating mitochondria, the dashed box shows normal Golgi compartments, and yellow asterisks indicate granules; $n = 3$ mice per genotype). **e**, Mitochondrial integrity was quantified on the basis of the percentage of the visible mitochondrial section displaying intact cisternae ($n = 49$ WT and 71 ATG16L1^{HM} cells). **f**, Enrichment of pathways associated with human orthologues mapped from genes differentially expressed in ATG16L1-deficient Paneth cells relative to WT cells, using the canonical pathway collection from MSigDB[24]. The bar chart displays the negative log of the enrichment *P* values for each pathway using the hypergeometric distribution (see Methods). Pathways with only one gene assigned were excluded from the chart. † refers to the Wnt/Frizzled receptor-mediated cyclic GMP pathway obtained from the Signal Transduction Knowledge database (STKE, http://stke.sciencemag.org/cgi/cm/stkecm;CMP_12420). ‡ and ‡‡ denote similar pathways associated with γ-hexachlorocyclohexane degradation as curated by KEGG and GenMAPP, respectively. **g**, Chart of selected genes for which mRNA is enriched in either ATG16L1-deficient Paneth cells (baseline is WT Paneth cells) or ATG16L1-deficient thymocytes (baseline is WT thymocytes). Only haemoglobin is in common in these two sets of analyses. NS, not significant. Scale bars: **a**, **b**, 10 μm; **c**, **d**, 500 nm.

M-phase cells (not shown). Importantly, deletion of *Atg5* in the intestinal epithelium in *Atg5*$^{flox/flox}$villin-*Cre* mice led to Paneth cell and granule abnormalities similar to those observed in ATG16L1^{HM} mice (Fig. 2i and Supplementary Fig. 8), whereas other epithelial cells appeared normal. This indicates that, within the ideal epithelium, Paneth cells have a unique sensitivity to autophagy gene disruption.

To characterize better the effects of ATG16L1 deficiency in Paneth cells, we used transmission electron microscopy. We observed degenerating mitochondria, loss of granules and the frequent absence of apical microvilli in ATG16L1^{HM} Paneth cells (Fig. 3a–e and Supplementary Fig. 6d, e). Electron microscopy also showed ATG16L1^{HM} Paneth cells with marked increases in cytoplasmic vesicles (Fig. 3b, d); a similar abnormality has been reported in Paneth cells from a Crohn's disease patient[20]. Importantly, such notable findings were not present in epithelial progenitors or enterocytes (Fig. 3b, d and Supplementary Fig. 9), confirming that ATG16L1 deficiency selectively affects Paneth cells within the intestinal epithelium.

We next performed transcriptional profiling of Paneth cell RNA procured by laser capture microdissection (LCM)[21] (see methods and analysis in Supplementary Fig. 10). Consistent with a lack of cell death or degeneration, as detected by electron microscopy, microarray analysis revealed that less than 1.5% of probe sets detected changes in RNA levels using low-stringency criteria (≥1.3-fold difference; Supplementary Table 1 and Supplementary Fig. 11). Cluster analysis of significantly enriched transcripts in ATG16L1-deficient Paneth cells revealed a notable signature of genes involved in peroxisome proliferator-activated receptor (PPAR) pathways, adipocytokine signalling and aspects of lipid metabolism (Fig. 3f and Supplementary Fig. 12). Additionally, transcripts for several acute-phase reactants including serum amyloid A1, haptoglobin and complement factors D and I were increased (Fig. 3g). Of particular interest was the observation that the adipocytokines leptin and adiponectin, previously reported to be increased in Crohn's disease patients[4,5], were amongst the most highly enriched transcripts (Fig. 3g).

To determine whether these transcriptional changes are unique to ATG16L1-deficient Paneth cells, we examined a second primary cell type from ATG16L1^{HM} mice. Autophagy is important in the function of T cells[22]. Thymocytes from ATG16L1^{HM} mice were normal in number but expressed low levels of ATG16L1 and exhibited decreased conversion of LC3I to LC3II (Supplementary Fig. 13), demonstrating that ATG16L1 is important for autophagy in thymocytes. Microarray analysis of these cells revealed that, in marked contrast to Paneth cells, expression of only 27 genes was altered more than 1.3-fold (Supplementary Fig. 14). There was no notable change in expression of the most significant clusters of transcripts altered in ATG16L1^{HM} Paneth cells, and only one gene altered in thymocytes was also altered in Paneth cells (Fig. 3g). Therefore, the transcriptional signature of ATG16L1 deficiency is specific to Paneth cells, again emphasizing the unique effects of ATG16L1 deficiency on these cells. Taken together with the extensive morphologic abnormalities documented above, we conclude that ATG16L1 deficiency is associated with profound alterations in the specialized properties of Paneth cells including defective granule exocytosis and unexpected increases in expression of genes involved in regulating injury responses.

Given the importance of ATG16L1 and ATG5 in Paneth cells in mice, we examined the role of the human *ATG16L1* Crohn's disease risk allele by means of a retrospective analysis of ileocolic resection specimens from patients with Crohn's disease. We studied tissue sections from the uninvolved proximal margins, containing little or no inflammation, from 10 Crohn's disease patients homozygous for the *ATG16L1* risk allele compared to 7 Crohn's disease controls without the risk allele. All 17 patients lacked the three major *NOD2* risk alleles and the protective *IL23R* allele for Crohn's disease[7,23]. Independent blinded examination by T.S.S. and E.M.B. revealed that 100% of at risk patients and 0% of controls contained abnormal Paneth cells, similar to the results in ATG16L1^{HM} mice (Fig. 4a, b). We quantified lysozyme staining in Crohn's disease patient specimens, and, like ATG16L1^{HM} mice, found that patients carrying the *ATG16L1* risk allele contained an increased proportion of Paneth cells with disorganized or diminished granules or exhibiting diffuse cytoplasmic lysozyme staining (Fig. 4c–g). Moreover, consistent with transcriptional analysis in mice, an increased proportion of Paneth cells from at risk patients that exhibited diffuse lysozyme staining D_3 cells in Fig. 4g) also stained positive for leptin protein compared to similar cells in controls ($P < 0.05$; Fig. 4e–g). These findings demonstrate a concordance between the pathology and transcriptional profile of Paneth cells from ATG16L1^{HM} mice and Paneth cell abnormalities observed in Crohn's disease patients with the risk allele of *ATG16L1*.

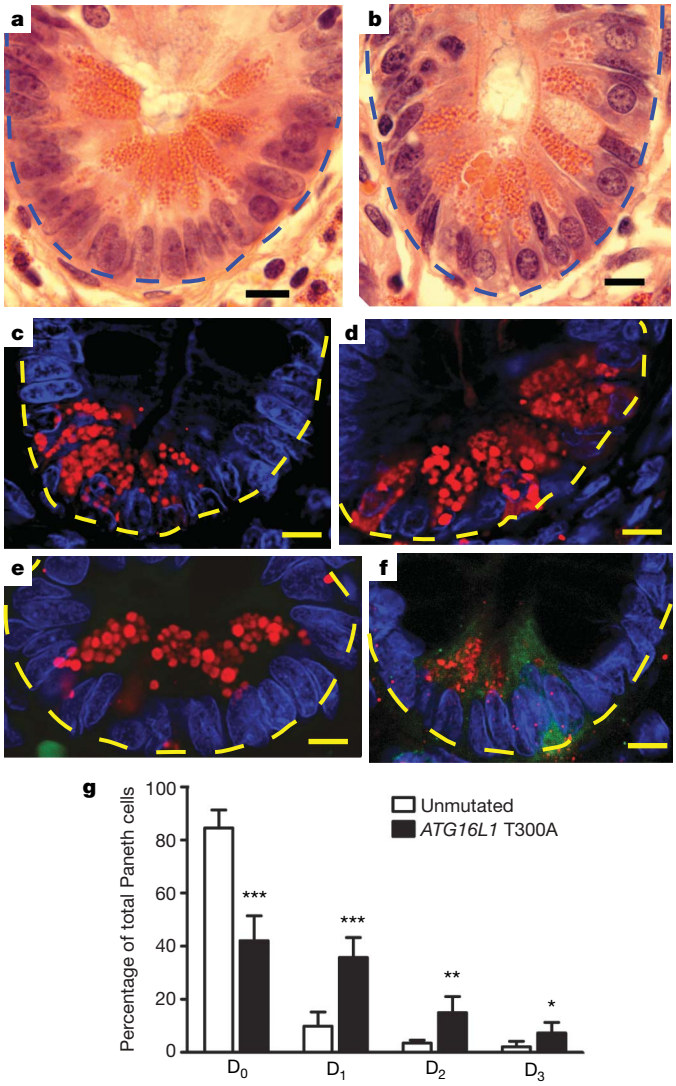


Figure 4 | Crohn's disease patients homozygous for the disease risk allele of ATG16L1 display Paneth cell abnormalities similar to ATG16L1^{HM} mice. **a**, **b**, Haematoxylin-and-eosin-stained sections of uninvolved areas from ileo–colic resections from patients with Crohn's disease homozygous for the safe (**a**) or risk (**b**) allele of *ATG16L1* (blue dashed line denotes crypt unit). **c**–**f**, Immunofluorescence images of Paneth cells from control patients (**c**, **e**) and patients with the risk allele (**d**, **f**) stained for lysozyme (red; **c**, **d**) and double-labelled additionally for leptin (green; **e**, **f**; yellow dashed line denotes the crypt unit). **g**, Aberrant lysozyme expression was quantified using the same criteria as that used for mouse sections in Fig. 3 ($n = 6{,}829$ cells for at risk patients and 8,182 for control, $n = 5$ patients for each genotype). Leptin-positive D_3 cells were quantified in patient samples ($n = 4$ patients per genotype) homozygous for the risk allele (76 out of 322 D_3 cells were positive from a total of 580 crypts examined) and homozygous for the safe allele (11 out of 93 D_3 cells were positive from a total 749 crypts examined). Scale bars: **a**–**f**, 10 μm. $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. *P* values were calculated using two-tailed Student's *t*-test. Error bars, s.e.m.

These data indicate that ATG16L1 has a specific role in humans and mice in regulating the specialized properties of Paneth cells, and provide a new and relevant mouse model that emulates one of the many diverse pathological hallmarks of human Crohn's disease. We show that ATG16L1 and a second autophagy protein ATG5 are critically important for the known role of Paneth cells in secretion of granule contents that may alter the intestinal microbiota. In addition, we demonstrate a previously unknown role of ATG16L1 in the regulation of Paneth cell expression of adipocytokines, previously associated with Crohn's disease. Within the intestinal epithelium, the marked effect of autophagy-protein deficiency on Paneth cells, but not on enterocytes, which share a common progenitor, indicates that autophagy can contribute to disease pathogenesis by means of a highly specific role within a single cell lineage. Indeed, the effects of hypomorphic expression of ATG16L1 were specific to Paneth cells and not seen in ATG16L1-deficient thymocytes. An important implication of this type of 'within-lineage' specificity is that future studies will need to focus on how ATG16L1 polymorphisms affect the function of differentiated Paneth cells. Because both environment and genotype have a role in human Crohn's disease pathogenesis[7], it will be interesting to determine whether environmental triggers including agents that damage the intestine or pathogens alter the pathological features associated with compromised ATG16L1 function.

METHODS SUMMARY

The embryonic stem cell lines BC0122 ($ATG16L1^{HM1}$) and XR0164 ($ATG16L1^{Hm2}$) containing insertions of the gene-trap vector were purchased from Bay Genomics. Founder chimaeric mice were mated with C57BL/6 mice, and heterozygous progenies were mated to each other to generate experimental mice. Littermate mice that are homozygous for the undisrupted allele were used as controls in all experiments.

For immunohistochemistry, paraffin-embedded tissue was washed for 5 min three times with xylene, twice with 100% EtOH, once with 95% EtOH, once with 70% EtOH, and once with PBS. Sections were boiled in Trilogy (Cell Marque) for antigen retrieval, rinsed in deionized water for 15 min, and washed with PBS. Sections were blocked in 1% BSA, 0.3% Triton X-100 PBS, and incubated with a 1:100 dilution of goat polyclonal anti-lysozyme C antibody (C-19), Santa Cruz Biotechnology. The specificity of the antibody was confirmed with a mouse monoclonal anti-lysozyme antibody (BGN/06/961), Abcam. Identical staining was observed when visualized with appropriate secondary antibodies conjugated to Alexa594 (1:500).

For whole-mount immunofluorescence, dissected ileal samples were carefully opened with fine scissors and the intestine was pinned on black wax in 10% buffered formalin for 1 h. The tissue was rinsed three times for 5 min in PBS and incubated with a 1:100 dilution of anti-lysozyme (both mouse and goat antibodies were equivalent) for one hour at 24 °C. After PBS washes, the tissue was incubated with a 1:500 dilution of Cy3-labelled anti-goat or anti-mouse antibody as well as a 1:100 dilution of FITC-labelled *H. pomatia* lectin for one hour at 24 °C. Tissue was mounted on glass slides after PBS washes and was viewed by an epifluorescence microscope.

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

Received 4 August; accepted 15 September 2008.
Published online 5 October 2008.

1. Barrett, J. C. *et al.* Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nature Genet.* **40**, 955–962 (2008).
2. Mizushima, N. *et al.* Mouse Apg16L, a novel WD-repeat protein, targets to the autophagic isolation membrane with the Apg12–Apg5 conjugate. *J. Cell Sci.* **116**, 1679–1688 (2003).
3. Ouellette, A. J. Paneth cell α-defensin synthesis and function. *Curr. Top. Microbiol. Immunol.* **306**, 1–25 (2006).
4. Barbier, M. *et al.* Overexpression of leptin mRNA in mesenteric adipose tissue in inflammatory bowel diseases. *Gastroenterol. Clin. Biol.* **27**, 987–991 (2003).
5. Yamamoto, K. *et al.* Production of adiponectin, an anti-inflammatory protein, in mesenteric adipose tissue in Crohn's disease. *Gut* **54**, 789–796 (2005).
6. The Wellcome Trust Case Control Consortium.. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
7. Xavier, R. J. & Podolsky, D. K. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* **448**, 427–434 (2007).
8. Rioux, J. D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genet.* **39**, 596–604 (2007).
9. Hampe, J. *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nature Genet.* **39**, 207–211 (2007).
10. Levine, B. & Kroemer, G. Autophagy in the Pathogenesis of Disease. *Cell* **132**, 27–42 (2008).
11. Fujita, N., Itoh, T., Fukuda, M., Noda, T. & Yoshimori, T. The Atg16L complex specifies the site of LC3 lipidation for membrane biogenesis in autophagy. *Mol. Biol. Cell* **19**, 2092–2100 (2008).
12. Kuma, A., Mizushima, N., Ishihara, N. & Ohsumi, Y. Formation of the approximately 350-kDa Apg12–Apg5.Apg16 multimeric complex, mediated by Apg16 oligomerization, is essential for autophagy in yeast. *J. Biol. Chem.* **277**, 18619–18625 (2002).
13. Stryke, D. *et al.* BayGenomics: a resource of insertional mutations in mouse embryonic stem cells. *Nucleic Acids Res.* **31**, 278–281 (2003).
14. Kuma, A. *et al.* The role of autophagy during the early neonatal starvation period. *Nature* **432**, 1032–1036 (2004).
15. Komatsu, M. *et al.* Homeostatic levels of p62 control cytoplasmic inclusion body formation in autophagy-deficient mice. *Cell* **131**, 1149–1163 (2007).
16. Hara, T. *et al.* Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* **441**, 885–889 (2006).
17. Madison, B. B. *et al.* Cis elements of the villin gene control expression in restricted domains of the vertical (crypt) and horizontal (duodenum, cecum) axes of the intestine. *J. Biol. Chem.* **277**, 33275–33283 (2002).
18. Porter, E. M., Bevins, C. L., Ghosh, D. & Ganz, T. The multifaceted Paneth cell. *Cell. Mol. Life Sci.* **59**, 156–170 (2002).
19. Kobayashi, K. S. *et al.* Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* **307**, 731–734 (2005).
20. Dvorak, A. M. & Dickersin, G. R. Crohn's disease: transmission electron microscopic studies I. Barrier function. Possible changes related to alterations of cell coat, mucous coat, epithelial cells, and Paneth cells. *Hum. Pathol.* **11**, 561–571 (1980).
21. Stappenbeck, T. S., Mills, J. C. & Gordon, J. I. Molecular features of adult mouse small intestinal epithelial progenitors. *Proc. Natl Acad. Sci. USA* **100**, 1004–1009 (2003).
22. Pua, H. H., Dzhagalov, I., Chuck, M., Mizushima, N. & He, Y. W. A critical role for the autophagy gene Atg5 in T cell survival and proliferation. *J. Exp. Med.* **204**, 25–31 (2007).
23. Duerr, R. H. *et al.* A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* **314**, 1461–1463 (2006).
24. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

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

Acknowledgements We thank N. Abe for technical assistance with fluorescence microscopy, V. Cavalli for microscope use, J. Eisenberg and A. Ng for help with bioinformatics analysis, A. J. Ouellette for technical advice, and I. Mysorekar and J. Mills for help with antibodies. This research was supported by grant U54 AI057160 Project 6 and the Broad Foundation (K.C., J. Loh and H.W.V.), training grant NIH T32 AR07279 (K.C.), the Lallage Feazel Wall Fellowship DRG-1972-08 (K.C.) from the Damon Runyon Cancer Research Foundation; the Pew Foundation (J. Liu, S.L.B., H.M., W.K. and T.S.S.); the Washington University Digestive Diseases Research Core Center P30 DK52574, Barnes Jewish Foundation, Johnson and Johnson Translational Seed Award, and the Crohn's Colitis Foundation of America (E.L., S.H. and C.S.); grants-in-aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan, and the Toray Science Foundation (C.K. and N.M.); NIH R01 AI062832 (J.A.C.); and NIH AI062773 and DK43351 (R.J.X.).

Author Contributions The original hypothesis of the article was formulated by K.C., H.W.V. and T.S.S. Autophagy experiments were performed by K.C. and C.K. Mouse lines were established by K.C., J. Loh and B.P.S. Mouse analysis experiments were performed by K.C., J. Liu, S.L.B., H.M., W.K., T.S.S. and J. Lennerz. Histological examination was performed by T.S.S, E.M.B. and J. Lennerz. *L. monocytogenes* experiments were performed by J.C. and K.C. Microarray analysis was performed by R.J.X. and T.S.S. Human tissue collection, genotyping and preparation were performed by E.L., S.H. and C.S. N.M. provided advice, the antibody to ATG16L1, and experiments by C.K. were performed in N.M.'s laboratory. The manuscript was written by K.C., H.W.V. and T.S.S., and all authors commented on the manuscript, data and conclusions before submission.

Author Information CEL files have been submitted to GEO for both LCM procured crypt base samples (accession numbers GSM318315 (ATG16 HM1-1), GSM318588 (ATG16 HM1-2), GSM318589 (WT1) and GSM318590 (WT2)) and the thymocyte samples (accession number GSE12707). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.W.V. (virgin@wustl.edu) and T.S.S. (stappenb@wustl.edu).

METHODS

Mice. Gene-trap mutant ATG16L1^{HM1} and ATG16L1^{Hm2} mouse tail DNA was genotyped by the presence of the gene-trap vector using primers with sequences TGGCTGGAGTGCGATCTTCC and CAGACGGCAAACGACTGTCCT. The presence of the undisrupted allele was determined using primers with sequences CAGGATCCTTCTGCACACATTT and CACCTGGTTACATTGGCAAACA for ATG16L1^{HM1} and GGTGCTGTATGTTACCATGTAC and GCACCCAAAGTC-TCCTAAGATT for ATG16L1^{Hm2}. *Atg5*$^{flox/flox}$ mice were previously described[16]. B6.SJL-Tg(Vil-cre)997Gum/J (villin-*cre*) mice were obtained from the Jackson Laboratory. Experimental mice were maintained by mating *Atg5*$^{flox/flox}$ mice with *Atg5*$^{flox/flox}$ mice that were heterozygous for the villin-*Cre* transgene. Genotyping of the *Atg5*$^{flox/flox}$ and the *Cre* gene has been described previously[25].

Antibodies and reagents. Rabbit polyclonal anti-ATG5 and anti-LC3 antibodies were from Novus Biologicals. Rabbit polyclonal anti-ATG16L1 antibody was described previously[2]. Mouse monoclonal anti-β-actin, rapamycin, cyclohexamide and FITC-conjugated lectin *H. pomatia* were from Sigma-Aldrich. Guinea pig polyclonal anti-P62 (GP-P62C) was from Progen. Goat polyclonal anti-lysozyme C antibody (C-19) was from Santa Cruz Biotechnology. Mouse monoclonal anti-lysozyme (BGN/06/961; IgG2a) antibody was from Abcam. Mouse anti-human leptin antibody (clone 44802; IgG1) was from Neuromics. Antibodies against insulin were used as a negative control for mouse IgG_1 (clone Z006, Zymed). Alexa594-conjugated secondary antibodies were from Molecular Probes. FITC-conjugated goat anti-rabbit secondary antibody was from Dako.

Cell culture. MEFs were grown in DMEM supplemented with 10% (vol/vol) FCS, MEM essential and nonessential amino acids, and penicillin and streptomycin. Low passage is defined as fewer than six passages. GFP–LC3 and cDNA encoding ATG16L1, cloned from template (IMAGE ID: 6813377, ATCC) by PCR reactions that generated SalI overhangs and cloned into the XhoI site of pMSCVpuro (Clontech), were introduced into MEFs by retroviral transduction as described previously[26,27]. Generation of SV40-T-antigen-transformed MEFs was described previously[14].

Immunoblot. Intestinal tissue from the distal ileum was resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.2% deoxycholic acid, 0.1% SDS, protease inhibitor cocktail) and homogenized by bead disruption in MagNA Lyser (Roche). A total of 2×10^5 MEFs per well of 12-well plates were grown for 12 h and lysed in 125 μl lysis buffer per well. 6× Laemmli buffer was added to lysates that were cleared of debris by centrifuging at 4 °C for 5 min and then boiled and subjected to western blotting using the indicated antibodies. Immunoblots were visualized by chemiluminescence. The optical density of the bands was quantified using Image J (National Institutes of Health).

Histology and microscopy. For tissue sections, the mouse ileum and colon were dissected from 4–8-week-old mice and fixed and prepared as described previously[21,28]. 2-cm strips of intestinal tissues were embedded in agar to enrich for well-oriented crypt-villus units. For dual detection of lysozyme and leptin, sections were incubated with the goat polyclonal anti-lysozyme (1:100) and anti-leptin (1:100), and visualized with appropriate secondary antibodies conjugated to Alexa594 (1:500). Sections were viewed with a Zeiss Axiovert 200 inverted fluorescence microscope and quantified on an Olympus AX70 epi-fluorescence microscope. For LC3 immunofluorescence, 5×10^4 MEFs were grown on 18×18 mm coverslips overnight and fixed with formaldehyde, permeabilized with 0.1% Triton X-100, blocked with 10% goat serum, and stained with anti-LC3 antibody and goat anti-rabbit secondary conjugated to FITC. Images were acquired on the Nikon Eclipse TE2002-E epifluorescence microscope using a ×20/0.75 Plan Apo objective and a CoolSNAP HQ2 CCD camera. For quantification of aggregates per cell, Matlab 2007a Gaussian filter was used to identify LC3$^+$ dots, and the area of the dots was quantified using Metamorph 6.2. GFP–LC3 dots were examined with Olympus IX81 equipped with a CCD camera (ORCA ER, Hamamatsu Photonics) and quantified as described previously[26]. Tissue fragments were prepared for electron microscopy as described previously[21].

LCM and quantitative PCR with reverse transcription analysis. Previously described techniques were used[21,28] with the following primers: *TNFa* (GCGGTGCCTATGTCTCAGC and CACTTGGTGGTTTGTGAGTGT), *ATG16L1* (GACCTGCTAAAAGTCATCGACC and AGTCAGAGCCGCATTTGAATC), selenoprotein P, plasma, 1 (TCACTTGCCCAGAGGAAGCT and CCTCGG-ACTCCTTAGACAACTTACAC), peripheral myelin protein 22 (TCCATG-TGTGAAGCCCTAAGC and GAGCCACCAGCTATTACTGTTTTTG), serine (or cysteine) peptidase inhibitor, clade B, member 1a (GTCCACAAGTCC-TTTGTGGAAGT and GCAACAACATACAGAATGTAGCAATG), serum/glucocorticoid regulated kinase 1 (CGTCCGAACGGGACAACAT and GTCCACCGTCCGGTCATAC), BTB and CNC homology 1 (CTGGCA-GGACCTAGGCAGAA and AGTCAAAACTCCTAGCAATACTGAGGATA), and protein tyrosine phosphatase, non-receptor type 9 (TGTGGATGA-GTCAGCCAACAG and TGCTTAACCAGTCAAGAAAAATGG).

Gene chip. RNA from LCM-procured cells from 2 WT and 2 ATG16L1^{HM} mice ($n = 2{,}500$ cells per mouse) was pooled, prepared and analysed as described previously[21,29]. Technical replicates were produced for each sample and Affymetrix 430.2 mouse microarrays were probed for each sample. In addition, to assess the statistical enrichment or over-representation of canonical pathways associated with human orthologues mapped from differentially expressed genes relative to their representation in the global set of unique genes on the microarray, *P* values were computed using the hypergeometric distribution, which was implemented in the R programming language. Pathway assignments were obtained from the canonical pathway collection of the Molecular Signatures Database (MSigDB)[24], which provides a catalogue of 639 gene sets classified into various pathway annotation groupings. Heat maps were generated by dCHIP analysis. CEL files have been submitted to GEO for both LCM procured crypt base samples (accession numbers GSM318315 (ATG16 HM1-1), GSM318588 (ATG16 HM1-2), GSM318589 (WT1) and GSM318590 (WT2)) and the thymocyte samples (accession number GSE12707).

Human tissue. Samples were from the Barnes Jewish Hospital collected between 2001 and 2007 from patients with ileal resections.

25. Miller, B. C. *et al.* The autophagy gene ATG5 plays an essential role in B lymphocyte development. *Autophagy.* **4,** 309–314 (2007).
26. Hara, T. *et al.* FIP200, a ULK-interacting protein, is required for autophagosome formation in mammalian cells. *J. Cell Biol.* **181,** 497–510 (2008).
27. Cadwell, K. & Coscoy, L. Ubiquitination on nonlysine residues by a viral E3 ubiquitin ligase. *Science* **309,** 127–130 (2005).
28. Brown, S. L. *et al.* Myd88-dependent positioning of Ptgs2-expressing stromal cells maintains colonic epithelial proliferation during injury. *J. Clin. Invest.* **117,** 258–269 (2007).
29. Pull, S. L., Doherty, J. M., Mills, J. C., Gordon, J. I. & Stappenbeck, T. S. Activated macrophages are an adaptive element of the colonic epithelial progenitor niche necessary for regenerative responses to injury. *Proc. Natl Acad. Sci. USA* **102,** 99–104 (2005).

LETTERS

Loss of the autophagy protein Atg16L1 enhances endotoxin-induced IL-1β production

Tatsuya Saitoh[1,3]*, Naonobu Fujita[4]*, Myoung Ho Jang[2], Satoshi Uematsu[1,3], Bo-Gie Yang[1,3], Takashi Satoh[1,3], Hiroko Omori[4], Takeshi Noda[4], Naoki Yamamoto[5], Masaaki Komatsu[6,7,8], Keiji Tanaka[6], Taro Kawai[1,3], Tohru Tsujimura[9], Osamu Takeuchi[1,3], Tamotsu Yoshimori[4,10] & Shizuo Akira[1,3]

Systems for protein degradation are essential for tight control of the inflammatory immune response[1,2]. Autophagy, a bulk degradation system that delivers cytoplasmic constituents into autolysosomes, controls degradation of long-lived proteins, insoluble protein aggregates and invading microbes, and is suggested to be involved in the regulation of inflammation[3–5]. However, the mechanism underlying the regulation of inflammatory response by autophagy is poorly understood. Here we show that Atg16L1 (autophagy-related 16-like 1), which is implicated in Crohn's disease[6,7], regulates endotoxin-induced inflammasome activation in mice. Atg16L1-deficiency disrupts the recruitment of the Atg12-Atg5 conjugate to the isolation membrane, resulting in a loss of microtubule-associated protein 1 light chain 3 (LC3) conjugation to phosphatidylethanolamine. Consequently, both autophagosome formation and degradation of long-lived proteins are severely impaired in Atg16L1-deficient cells. Following stimulation with lipopolysaccharide, a ligand for Toll-like receptor 4 (refs 8, 9), Atg16L1-deficient macrophages produce high amounts of the inflammatory cytokines IL-1β and IL-18. In lipopolysaccharide-stimulated macrophages, Atg16L1-deficiency causes Toll/IL-1 receptor domain-containing adaptor inducing IFN-β (TRIF)-dependent activation of caspase-1, leading to increased production of IL-1β. Mice lacking Atg16L1 in haematopoietic cells are highly susceptible to dextran sulphate sodium-induced acute colitis, which is alleviated by injection of anti-IL-1β and IL-18 antibodies, indicating the importance of Atg16L1 in the suppression of intestinal inflammation. These results demonstrate that Atg16L1 is an essential component of the autophagic machinery responsible for control of the endotoxin-induced inflammatory immune response.

Autophagy is a bulk degradation system, which controls the clearance and re-use of intracellular constituents, and is important for the maintenance of an amino acid pool essential for survival[3–5]. In addition, recent studies have disclosed multiple roles of autophagy in the regulation of cell death, differentiation and anti-microbial response in mammals[4,5]. Yeast genetic screening studies have identified a variety of essential components of autophagic machinery, called Atg proteins, which are phylogenetically highly conserved, and several mammalian counterparts, such as Atg5 and Atg7, have been reported[3–5]. Previously, we systematically characterized mammalian homologues of Atg proteins and identified Atg16L1 protein as an Atg5-binding protein[10]. Its coiled-coil domain, which mediates self-multimerization, is essentially required for starvation-induced autophagy in yeast, and this domain is conserved in mammalian Atg16L1 (refs 3, 10; Fig. 1a). We have proposed that the coiled-coil domain of Atg16L1 is required for the formation of an ~800 kDa high molecular weight protein complex with the Atg12-Atg5 conjugate and defines the site where LC3 (homologue of yeast Atg8) is conjugated to phosphatidylethanolamine (PE), an essential process for autophagy, by recruitment of an Atg3-LC3 intermediate to a source membrane of an autophagosome[10,11]. In addition, Atg16L1 has seven WD40 repeats at the carboxy terminus, which are absent in yeast Atg16 (ref. 10). Recent genome-wide association studies identified Atg16L1 as a candidate gene responsible for susceptibility to Crohn's disease[6,7]. However, the importance of Atg16L1 in autophagy and its role in inflammation have not been fully understood. Hence, we generated Atg16L1 mutant mice and examined the function of Atg16L1 in autophagosome formation as well as in the regulation of immune responses.

Atg16L1 mutant mice express deleted forms of Atg16L1 protein lacking the entire coiled-coil domain (Fig. 1a, b, and Supplementary Fig. 1a–c). However, such aberrant proteins do not act as dominant-negative molecules, because ectopic expression of truncated Atg16L1 protein lacking the coiled-coil domain (ΔCCD) in wild-type mouse embryonic fibroblasts (MEFs) did not interfere with autophagy (Supplementary Fig. 2a, b). Most Atg16L1-deficient mice died within 1 day of delivery, indicating that Atg16L1 is required for survival during neonatal starvation (Supplementary Fig. 1d, e). This phenotype is similar to that observed in Atg5- or Atg7-deficient mice[12,13]. Although Atg16L1 associates with Atg12-Atg5, Atg16L1 was dispensable for Atg12 conjugation to Atg5 (Fig. 1b). On the other hand, Atg16L1 was required for LC3 conjugation to PE (Fig. 1b). In Atg16L1-deficient MEFs, formation of the high molecular weight protein complex was disrupted and Atg12-Atg5 puncta were hardly observed (Fig. 1c, d, and Supplementary Fig. 3, 4a). On the other hand, GFP-Atg5 free from Atg12-conjugation formed puncta in Atg7-deficient MEFs or Atg5-deficient MEFs complemented with GFP-Atg5^{K130R}, although these puncta did not colocalize with LC3 (Fig. 1c, d, Supplementary Figs 4b, 5, data not shown). Formation of autophagosomes under the starved condition was not observed in Atg16L1-deficient MEFs, resulting in a decrease in the bulk degradation of long-lived proteins and the accumulation of p62/SQSTM1 (Fig. 1b–f). These results indicated that Atg16L1 is essentially required for autophagy by regulating the localization of the Atg12-Atg5 conjugate.

[1]Laboratory of Host Defense, [2]Laboratory of Gastrointestinal Immunology, WPI Immunology Frontier Research Center, Osaka University, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan. [3]Department of Host Defense, [4]Department of Cellular Regulation, Research Institute for Microbial Diseases, Osaka University, 3-1 Yamada-oka, Suita, Osaka 565-0871, Japan. [5]AIDS Research Center, National Institute of Infectious Diseases, Toyama 1-23-1, Shinjuku-ku, Tokyo 162-8640, Japan. [6]Laboratory of Frontier Science, Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo 113-8613, Japan. [7]Department of Biochemistry, Juntendo University School of Medicine, 2-1-1 Hongo Bunkyo-ku, Tokyo 113-8421, Japan. [8]PRESTO, Japan Science and Technology Corporation, Kawaguchi, Saitama 332-0012, Japan. [9]Department of Pathology, Hyogo College of Medicine, 1-1 Mukogawa-cho, Nishinomiya, Hyogo 663-8501, Japan. [10]CREST, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan.
*These authors contributed equally to this work.

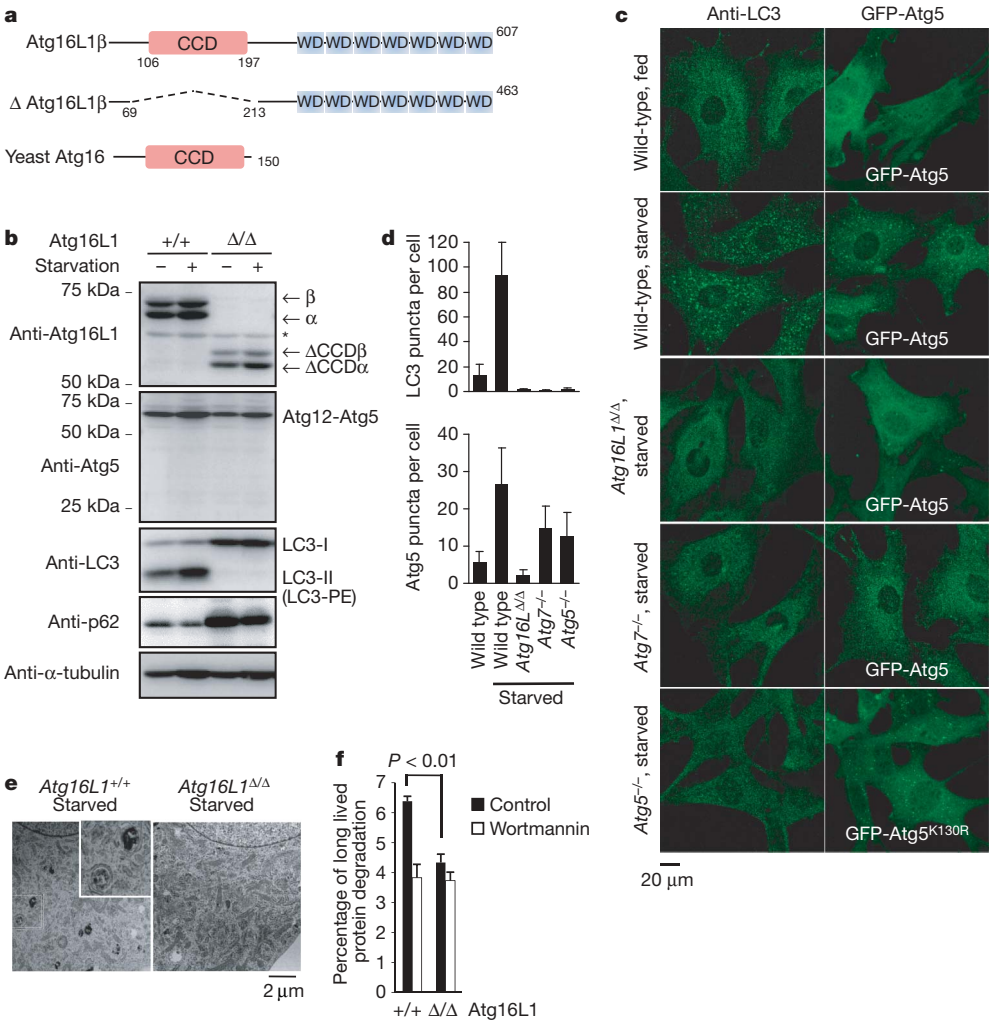


Figure 1 | Atg16L1 is essential for autophagy. a, Schematic structure of wild-type or truncated (Δ) Atg16L1 protein. CCD, Coiled-coil domain; WD, WD40 domain. Here and below, suffix α or β indicates isoform α or β. **b–d**, MEFs were cultured in nutrient-rich medium (fed) or Hanks' buffer (starved) for 1 h. Total cell lysates were prepared, and blotted with the indicated antibodies (**b**). α, β, indicate Atg16L1 isoform α, β; ΔCCD indicates Atg16L1ΔCCD; asterisk indicates a non-specific band. *Atg16L1*$^{\Delta/\Delta}$ indicates mice expressing ΔAtg16L1. The number of endogenous LC3 or GFP-Atg5 dots was counted (**c, d**). The results shown are mean + s.d. ($n > 20$). **e**, Electron micrograph of starved MEFs. **f**, Degradation of long-lived proteins in MEFs. Wortmannin is an autophagy inhibitor. The results shown are mean + s.d. Statistical significance (*P* value) was determined by the Student's *t*-test.

Although the involvement of autophagic machinery in the Toll-like receptor (TLR)-mediated antiviral response and phagocytosis have been reported[14,15], it is still unclear whether autophagy controls TLR-mediated inflammatory responses. We examined the role of Atg16L1 in the production of inflammatory cytokines, such as TNFα, IL-6 and IL-1β, in response to lipopolysaccharide (LPS), a major component of bacterial endotoxin[8]. Although both messenger RNA expression and production of TNFα, IL-6 and IFN-β were almost normal in Atg16L1-deficient fetal liver-derived macrophages, IL-1β production was highly elevated compared with that in wild-type macrophages (Fig. 2a, Supplementary Fig. 6). IL-1β mRNA synthesis was not impaired in Atg16L1-deficient macrophages, indicating that IL-1β production is enhanced at the post-transcriptional level in Atg16L1-deficient macrophages (Supplementary Fig. 6). Synthetic lipid A, an active component of LPS, also potently induced IL-1β production in Atg16L1-deficient cells (Fig. 2b). On the other hand, ectopic expression of the Atg16L1 protein lacking coiled-coil domain (ΔCCD) in RAW264.7 cells did not affect LPS-induced IL-1β production (Supplementary Fig. 2c, d). These results indicated that Atg16L1-deficiency is responsible for the elevated production of IL-1β.

We next generated chimaeric mice by transplantation of fetal liver cells into lethally irradiated CD45.1 mice to examine IL-1β production in other types of macrophage (Supplementary Fig. 7a–c). Following stimulation with LPS, peritoneal and bone-marrow macrophages deficient in Atg16L1 showed enhanced IL-1β production compared with wild-type macrophages (Supplementary Fig. 7d, e). Non-invasive Gram-negative bacteria, such as *Escherichia coli, Enterobacter aerogenes* and *Klebsiella pneumonia*, which are habitants in the commensal flora, also potently induced IL-1β production in Atg16L1-deficient cells (Supplementary Fig. 8a). On the other hand, the production of IL-1β and apoptosis induced by *Salmonella typhimurium*, an invasive Gram-negative bacterium, is almost normal in Atg16L1-deficient macrophages (Fig. 2c, Supplementary Fig. 9c). We also found that Atg16L1-deficient macrophages produced a high amount of IL-1β following stimulation by ATP or monosodium urate (MSU), an activator of the Nalp3 inflammasome[16,17] (Fig. 2c). Atg16L1-deficient macrophages normally produced inflammatory cytokine in response to muramyldipeptide, a ligand for NOD2 (ref. 16), indicating that Atg16L1 is not involved in signalling downstream of NOD2, whose de-regulation is also implicated in Crohn's disease[16] (Supplementary Fig. 10).

The expression levels of immature IL-1β protein following LPS stimulation in Atg16L1-deficient macrophages were almost comparable to those in wild-type cells, indicating an abnormality of post-translational regulation (Fig. 2d). Cleaved caspase-1, an activated

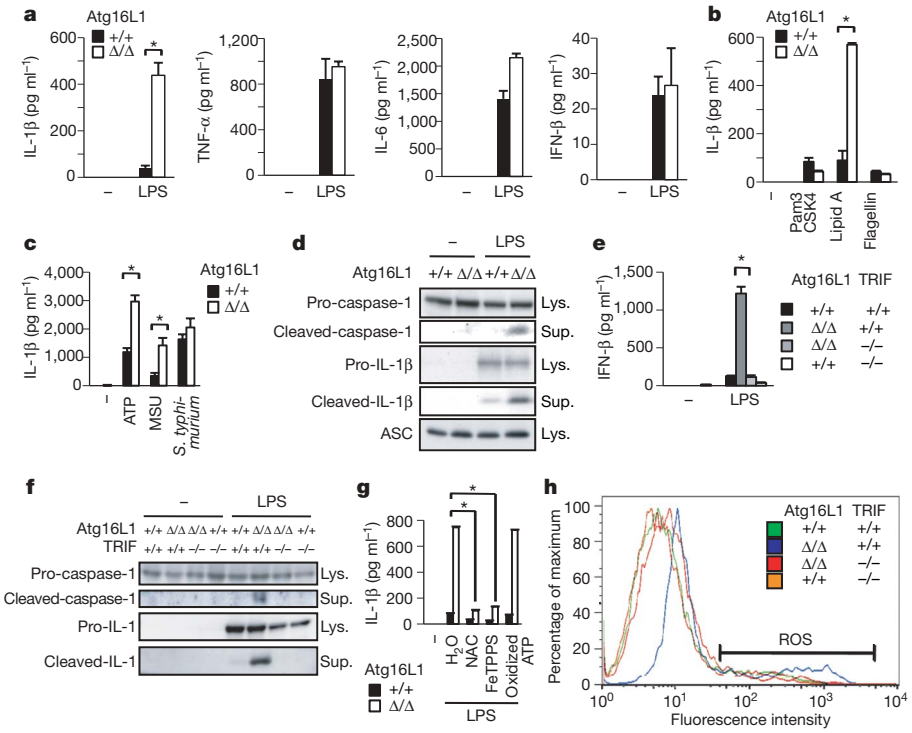


Figure 2 | Elevated endotoxin-induced IL-1β production from Atg16L1-deficient macrophages. **a**, Cytokine production from macrophages stimulated with LPS (100 ng ml^{-1}) for 24 h. Statistical significance was determined by the Student's *t*-test. $*P < 0.01$. **b**, IL-1β production from macrophages stimulated with indicated ligands. **c**, IL-1β production from LPS-primed macrophages infected with *S. typhimurium* (multiplicity of infection, m.o.i., 1), or stimulated with ATP or MSU for 1 h. **d**, Expression levels of caspase-1 and IL-1β in macrophages. Lys., cell lysates; Sup., culture supernatants. **e**, LPS-induced production of IL-1β from macrophages with the indicated phenotype. **f**, Expression levels of caspase-1 and IL-1β in macrophages treated as in **e**. **g**, Effect of the ROS scavenger FeTPPS (25 μM), N-acetyl-L-cysteine (NAC; 25 mM) or P2X7 receptor antagonist oxidized ATP (250 μM) on IL-1β production. **h**, ROS in LPS-stimulated macrophages were detected by CM-H_2DCFDA staining.

form that mediates processing of IL-1β and apoptosis[16,17], was detected in the culture supernatants of Atg16L1-deficient macrophages following LPS stimulation, and was responsible for the production of IL-1β and the induction of apoptosis (Fig. 2d, Supplementary Figs 9a, b, 11). IL-18 production, which is regulated by caspase-1-mediated cleavage[17], was also enhanced in response to LPS in Atg16L1-deficient macrophages (Supplementary Fig. 12). Recent studies have disclosed that NF-κB and p38 signalling pathways regulate the activation of caspase-1 and the induction of cell death in macrophages stimulated with LPS[18,19]. However, activation of NF-κB, p38 and IRF-3 signalling pathways by LPS was comparable between wild-type and Atg16L1-deficient macrophages (Supplementary Fig. 13). Among TLR family members, TLR2, TLR4 and TLR5 recognize bacterial components and play important roles in the anti-bacterial response[8]. Importantly, TLR4 ligand, but not ligands for TLR2 or TLR5, induced potent IL-1β production from Atg16L1-deficient macrophages (Fig. 2b, Supplementary Fig. 14). Enhancement of IL-1β production in Atg16L1-deficient macrophages was also induced by ligands for the viral nucleotide-sensing TLRs, TLR3, TLR7 and TLR9, although the production induced by these ligands was lower than that induced by LPS (Supplementary Fig. 14).

These findings prompted us to assess the involvement of the TRIF/IFN signalling, which is strongly triggered by the engagement of TLR4 in macrophages[8] and regulates apoptosis[18]. Consistent with this hypothesis, Atg16L1/TRIF double-deficient macrophages failed

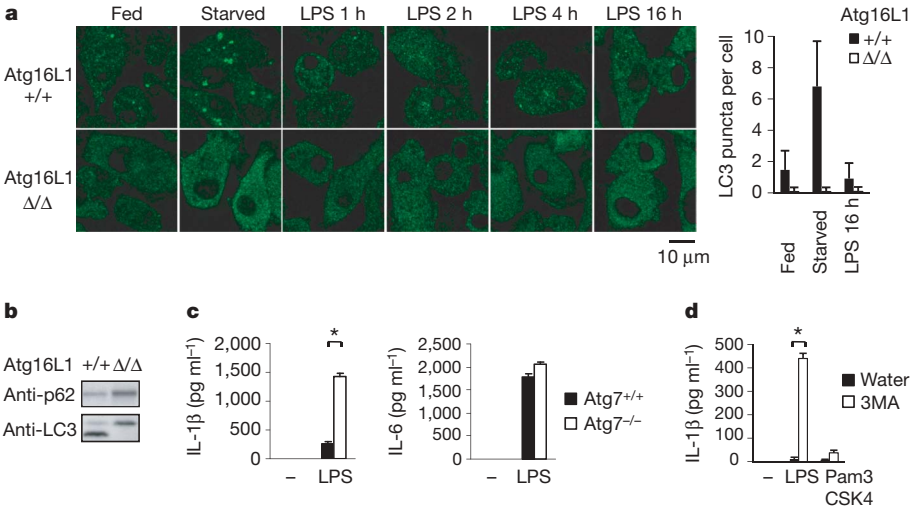


Figure 3 | Disruption of basal autophagy enhances LPS-induced IL-1β production. **a**, Macrophages were stimulated with LPS (100 ng ml^{-1}) for the indicated time period. The number of endogenous LC3 dots within each cell was counted. The results shown are mean + s.d. ($n > 100$). **b**, Expression levels of p62 and LC3 in macrophages. **c**, IL-1β and IL-6 production by wild-type or Atg7-deficient macrophages stimulated with LPS. Statistical significance was determined by the Student's *t*-test. $*P < 0.01$. **d**, Macrophages were pre-treated with or without 10 mM 3MA and then stimulated with the indicated ligands.

to produce IL-1β due to a lack of caspase-1 activation in response to LPS (Fig. 2e, f). The simultaneous stimulation of Atg16L1-deficient macrophages with IFN-β or IFN-γ enhanced IL-1β production and apoptosis induced by TLR2 ligand (Supplementary Figs 9d, 15a). Recent studies have disclosed that K^+-efflux and reactive oxygen species (ROS), especially peroxynitrate, play important roles in the production of IL-1β induced by ATP, MSU and asbestos[16,17,20,21]. Similarly, the enhanced IL-1β production from Atg16L1-deficient macrophages required K^+-efflux and ROS generation (Fig. 2g, Supplementary Figs 15b, 16). The level of ROS in Atg16L1-deficient macrophages was higher than that in Atg16L1/TRIF double-deficient or wild-type macrophages following LPS stimulation (Fig. 2h). Oxidized ATP, an antagonist for the P2X7 receptor, did not inhibit LPS-induced IL-1β production, indicating that extracellular ATP is not involved in its production (Fig. 2g). These results indicate that loss of Atg16L1 in macrophages causes aberrant LPS-induced IL-1β production in a TRIF-dependent manner. ROS might be accumulated in Atg16L1-deficient macrophages undergoing apoptosis and trigger caspase-1 activation following LPS stimulation.

The involvement of TLR signalling in the induction of autophagy has been recently reported[22,23]. Therefore we examined if stimulation of TLR4 or other TLRs induces puncta formation by endogenous LC3. In contrast to previous reports, LPS stimulation did not increase the number of LC3 puncta in primary macrophages, although nutrient deprivation induced the formation of autophagosomes (Fig. 3a, Supplementary Fig. 17a–c). Stimulation by other ligands for TLRs also failed to increase the number of puncta of endogenous LC3 in these macrophages (Supplementary Fig. 17b, d, e). Co-incubation with non-invasive bacteria did not increase the number of autophagosomes in macrophages (Supplementary Fig. 8b). On the other hand, infection with *S. typhimurium* resulted in Atg16L1-dependent formation of bacteria autophagosomes, even in the absence of both MyD88 and TRIF, essential adaptor molecules for TLR signalling pathways[8,9] (Supplementary Fig. 18a, b). These results indicated that TLR signalling is not associated with the formation of autophagosomes in primary macrophages.

Increasing evidence has revealed that basal autophagy plays critical roles under both physiological and pathological conditions, including neurodegeneration, hepatic dysfunction and the immune response[13,24–26]. In Atg16L1-deficient macrophages, autophagosomes were hardly detected and p62/SQSTM1 protein was accumulated under nutrient-rich conditions, indicating that basal autophagy is almost completely inhibited (Fig. 3a, b). Atg7-deficient macrophages also produced high levels of IL-1β in response to LPS, but produced normal levels of IL-6 (Fig. 3c). A chemical inhibitor of autophagy, 3-methyladenine (3MA), significantly enhanced production of IL-1β from wild-type peritoneal macrophages induced by stimulation with LPS, but not with ligand for TLR2 (Fig. 3d). Macrophages treated with 3MA underwent apoptosis following LPS stimulation (Supplementary Fig. 9e). Further, transient expression of inactive mutant of Atg4B, which inhibits the LC3 lipidation, enhanced LPS-induced IL-1β production in RAW264.7 cells (Supplementary Fig. 19a, b). These results indicate that inhibition of basal autophagy induces IL-1β overproduction.

Aberrant expression of inflammatory cytokines, including IL-1β and IL-18, has been shown to be involved in the development of colitis[27,28], and recent studies have reported that Atg16L1 is a candidate susceptibility gene for Crohn's disease[6,7]. Under specific pathogen-free conditions, Atg16L1-deficient chimaeric mice did not develop spontaneous colitis, and the colons of newborn Atg16L1-deficient mice were not inflamed (Supplementary Fig. 20a, b). The number of bacteria in the faeces of wild-type or Atg16L1-deficient chimaeric mice was almost same, and no bacteria were detected in spleen (Supplementary Fig. 20c, d). The number of $CD4^+Foxp3^+$ regulatory T cells, which suppress the inflammatory response and are required for immune homeostasis[29], was almost normal in the spleens and mesenteric lymph nodes of Atg16L1-deficient chimaeric mice (Supplementary Fig. 21a, b). We next assessed if Atg16L1-deficiency exacerbates inflammation in a dextran sulphate sodium (DSS)-induced experimental model of colitis. Strikingly, all chimaeric mice with Atg16L1-deficient haematopoietic cells died together with severe body weight loss following seven days of DSS exposure, whereas all chimaeric mice expressing wild-type Atg16L1 survived (Fig. 4a, b). Histological analyses revealed much severer inflammation in the distal colons of Atg16L1-deficient mice than in wild-type controls, with larger areas of ulceration and increased infiltration of lymphocytes

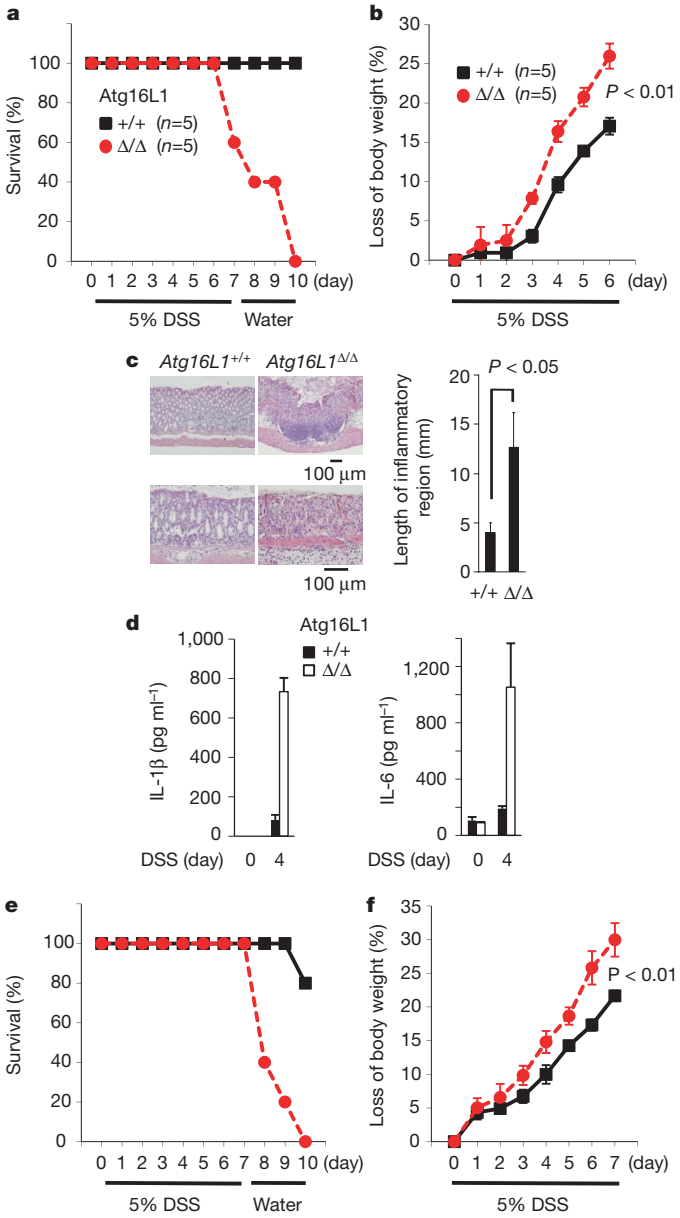


Figure 4 | Severe DSS-induced colitis in Atg16L1-deficient chimaeric mice. **a**, **b**, Fetal liver chimaeric mice were given 5% DSS in drinking water for 7 days. The survival (**a**) and weight loss (**b**) of each mouse genotype were plotted. The results shown are mean ± s.d. Statistical significance was determined by the Student's *t*-test. **c**, Typical distal colon appearance 6 days after the initiation of DSS administration. The results shown are mean ± s.d. ($n = 3$, each group). **d**, Expression levels of IL-1β and IL-18 in serum ($n = 5$, each group). **e**, **f**, Atg16L1-deficient chimaeric mice given 5% DSS in drinking water were intraperitoneally injected with both anti-IL-1β and anti-IL-18 neutralizing antibodies (squares; $n = 5$) or isotype control IgG (circles; $n = 5$) at days 1, 3, 5 and 7. The survival (**e**) and weight loss (**f**) of each mouse genotype were plotted.

(Fig. 4c). The levels of the proinflammatory cytokines IL-1β and IL-18 were significantly elevated in the sera of DSS-treated Atg16L1-deficient chimaeric mice relative to the levels in wild-type counterparts (Fig. 4d). Mortality and loss of body weight after DSS-exposure in Atg16L1-deficient chimaeric mice were improved by the injection of neutralizing antibodies for IL-1β and IL-18, showing the involvement of excessive production of these cytokines in the development of severe colitis (Fig. 4e, f). Administration of 3MA increased the level of IL-1β in serum and worsened the survival rate of mice treated with DSS, suggesting that autophagy protects mice from massive inflammation during colitis (Supplementary Fig. 22).

Our present study highlights a novel role for autophagy in the regulation of the inflammatory immune response. Autophagy controls inflammasome activation and limits production of the inflammatory cytokines IL-1β and IL-18. Given the importance of elevated expression of IL-1β and IL-18 caused by Atg16L1 deficiency in the pathology of chemical-induced colitis, it would be of interest to examine the involvement of autophagy in the pathogenesis of inflammatory bowel diseases such as Crohn's disease.

METHODS SUMMARY

Mice, reagents, cells and plasmids. Details are given in Methods.

Preparation of macrophages. E15.5 fetal liver stem cells from wild-type or Atg16L1-deficient littermates were cultured in the presence of GM-CSF ($10\,\mathrm{ng\,ml^{-1}}$) for 7 days to generate fetal liver macrophages. Unattached cells were removed on days 2, 4 and 6. Unless otherwise noted, fetal liver macrophages were used in the experiments. Bone-marrow-derived and peritoneal macrophages were prepared as described[9].

Histopathological analysis. The colon was removed and fixed with 4% PFA. The paraffin sections were stained with haematoxylin and eosin (H&E), and histologically analysed.

RT–PCR, immunoblotting, ELISA. Details of RT–PCR procedures are given in Methods. Immunoblotting was performed as described[11], and the experiments were repeated at least twice. The level of cytokine production was measured by ELISA according to the manufacturer's instructions. The results shown are means ± s.d. from three separate samples. The experiments were repeated at least three times.

Fluorescence microscopy analysis. Cells cultured on coverslips were fixed with 3% paraformaldehyde, and subjected to immunocytochemistry[11]. Samples were examined under a fluorescence laser scanning confocal FV1000 microscope (Olympus).

Detection of ROS. Macrophages were stimulated with LPS for 22 h, and then stained with CM-H_2DCFDA (10 μM; Molecular Probes), a fluorescent indicator for ROS, for 2 h. The level of fluorescence was determined by flow cytometry. The experiments were repeated at least three times.

Gel filtration, electron microscopy analysis, bulk protein degradation assay. Gel filtration analysis was performed as described[11]; electron microscopy analysis was performed as described[30]; details on the bulk protein degradation assay are given in Methods.

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

Received 7 August; accepted 26 August 2008.
Published online 5 October 2008.

1. Liu, Y. C., Penninger, J. & Karin, M. Immunity by ubiquitylation: A reversible process of modification. *Nature Rev. Immunol.* **5,** 941–952 (2005).
2. Wang, Y. *et al.* Lysosome-associated small Rab GTPase Rab7b negatively regulates TLR4 signaling in macrophages by promoting lysosomal degradation of TLR4. *Blood* **110,** 962–971 (2007).
3. Ohsumi, Y. Molecular dissection of autophagy: Two ubiquitin-like systems. *Nature Rev. Mol. Cell Biol.* **2,** 211–216 (2001).
4. Mizushima, N., Levine, B., Cuervo, A. M. & Klionsky, D. J. Autophagy fights disease through cellular self-digestion. *Nature* **451,** 1069–1075 (2008).
5. Levine, B. & Deretic, V. Unveiling the roles of autophagy in innate and adaptive immunity. *Nature Rev. Immunol.* **7,** 767–777 (2007).
6. Hampe, J. *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nature Genet.* **39,** 207–211 (2007).
7. Rioux, J. D. *et al.* Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nature Genet.* **39,** 596–604 (2007).
8. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124,** 783–801 (2006).
9. Yamamoto, M. *et al.* Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science* **301,** 640–643 (2003).
10. Mizushima, N. *et al.* Mouse Apg16L, a novel WD-repeat protein, targets to the autophagic isolation membrane with the Apg12-Apg5 conjugate. *J. Cell Sci.* **116,** 1679–1688 (2003).
11. Fujita, N., Itoh, T., Fukuda, M., Noda, T. & Yoshimori, T. The Atg16L complex specifies the site of LC3 lipidation for membrane biogenesis in autophagy. *Mol. Biol. Cell* **19,** 2092–2100 (2008).
12. Kuma, A. *et al.* The role of autophagy during the early neonatal starvation period. *Nature* **432,** 1032–1036 (2004).
13. Komatsu, M. *et al.* Impairment of starvation-induced and constitutive autophagy in Atg7-deficient mice. *J. Cell Biol.* **169,** 425–434 (2005).
14. Lee, H. K., Lund, J. M., Ramanathan, B., Mizushima, N. & Iwasaki, A. Autophagy-dependent viral recognition by plasmacytoid dendritic cells. *Science* **315,** 1398–1401 (2007).
15. Sanjuan, M. A. *et al.* Toll-like receptor signalling in macrophages links the autophagy pathway to phagocytosis. *Nature* **450,** 1253–1257 (2007).
16. Kanneganti, T. D., Lamkanfi, M. & Núñez, G. Intracellular NOD-like receptors in host defense and disease. *Immunity* **27,** 549–559 (2007).
17. Pétrilli, V., Dostert, C., Muruve, D. A. & Tschopp, J. The inflammasome: A danger sensing complex triggering innate immunity. *Curr. Opin. Immunol.* **19,** 615–622 (2007).
18. Hsu, L. C. *et al.* The protein kinase PKR is required for macrophage apoptosis after activation of Toll-like receptor 4. *Nature* **428,** 341–345 (2004).
19. Greten, F. R. *et al.* NF-κB is a negative regulator of IL-1β secretion as revealed by genetic and pharmacological inhibition of IKKβ. *Cell* **130,** 918–931 (2007).
20. Dostert, C. *et al.* Innate immune activation through Nalp3 inflammasome sensing of asbestos and silica. *Science* **320,** 674–677 (2008).
21. Hewinson, J., Moore, S. F., Glover, C., Watts, A. G. & MacKenzie, A. B. A key role for redox signaling in rapid P2X7 receptor-induced IL-1beta processing in human monocytes. *J. Immunol.* **180,** 8410–8420 (2008).
22. Xu, Y. *et al.* Toll-like receptor 4 is a sensor for autophagy associated with innate immunity. *Immunity* **27,** 135–144 (2007).
23. Delgado, M. A., Elmaoued, R. A., Davis, A. S., Kyei, G. & Deretic, V. Toll-like receptors control autophagy. *EMBO J.* **27,** 1110–1121 (2008).
24. Hara, T. *et al.* Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* **441,** 885–889 (2006).
25. Komatsu, M. *et al.* Homeostatic levels of p62 control cytoplasmic inclusion body formation in autophagy-deficient mice. *Cell* **131,** 1149–1163 (2007).
26. Paludan, C. *et al.* Endogenous MHC class II processing of a viral nuclear antigen after autophagy. *Science* **307,** 593–596 (2005).
27. Maeda, S. *et al.* Nod2 mutation in Crohn's disease potentiates NF-κB activity and IL-1β processing. *Science* **307,** 737–738 (2005).
28. Ishikura, T. *et al.* Interleukin-18 overproduction exacerbates the development of colitis with markedly infiltrated macrophages in interleukin-18 transgenic mice. *J. Gastroenterol. Hepatol.* **18,** 960–969 (2003).
29. Izcue, A., Coombes, J. L. & Powrie, F. Regulatory T cells suppress systemic and mucosal immune activation to control intestinal inflammation. *Immunol. Rev.* **212,** 256–271 (2006).
30. Nakagawa, I. *et al.* Autophagy defends cells against invading group A Streptococcus. *Science* **306,** 1037–1040 (2004).

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

Acknowledgements We are grateful to T. Kitamura, S. Yamaoka and N. Mizushima for providing materials. We thank K. J. Ishii, M. Yamamoto and members of the Laboratory of Host Defense for discussions; Y. Fujiwara, M. Shiokawa, R. Nakayama and N. Kitagaki for technical assistance; and M. Hashimoto and E. Kamada for secretarial assistance. This work was in part supported by grants from NIH (AI070167) and the Ministry of Health, Labour and Welfare of Japan, and by Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

Author Contributions T.S. generated the Atg16L1-deficient mice and performed the immunological experiments. N.F. performed the cell biology experiments. N.Y. generated the retroviral vector. M.K. and K.T. generated the Atg7-deficient mice. T.T. performed histological analysis of mice. M.H.J., S.U., B.-G.Y., T.S., H.O., T.N., T.K. and O.T. helped with experiments. T.Y. designed the cell biology research. S.A. supervised the overall research project.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.A. (akira@biken.osaka-u.ac.jp).

METHODS

Generation of Atg16L1-deficient mice. The fragment of the Atg16L1 gene was isolated from genomic DNA extracted from wild-type ES cells by PCR. A targeting vector was constructed by replacing exons 3, 4 and 5 of Atg16L1 with a neomycin-resistance gene cassette (neo), and a herpes simplex virus thymidine kinase driven by PGK promoter was inserted into the genomic fragment for negative selection. After the targeting vector was transfected into ES cells, G418 and gancyclovir doubly resistant colonies were selected and screened by PCR and Southern blotting. Homologous recombinants were micro-injected into C57BL/6 female mice, and heterozygous F1 progenies were intercrossed in order to obtain Atg16L1-deficient (Δ/Δ) mice. The Atg16L1-deficient mice used were on a 129Sv × C57BL/6 background.

For the generation of fetal liver chimaeric mice, fetal liver cells were harvested at E15.5 and injected into lethally irradiated CD45.1 or C57BL6J mice. Chimaeric mice were given antibiotics in drinking water for 2 months. The mice were analysed at least 3 months after reconstitution.

Mice were maintained in our animal facility and treated in accordance with the guidelines of Osaka University.

Reagents, mice and cells. Anti-IL-1β antibodies, recombinant mouse IFN-γ, ELISA kits for mouse IL-1β, human IL-1β, mouse IL-6 and mouse TNFα, were purchased from R&D Systems. Recombinant mouse IFN-β and the ELISA kit for mouse IFN-β were purchased from PBL InterferonSource. Anti-LC3 and anti-IL-18 antibodies and the ELISA kit for IL-18 were purchased from MBL International. GM-CSF and M-CSF were purchased from Peprotech. Anti-p62/SQSTM1 antibody was purchased from BIOMOL International. Anti-Atg12, anti-phospho-IRF3, anti-phospho-p38 and anti-phospho-ERK antibodies were purchased from Cell Signaling Technology. Anti-IκBα, anti-ERK and caspase-1 p10 were purchased from Santa Cruz Biotechnology. YVAD-fmk, wortmannin and FeTPPS were purchased from Calbiochem. ATP, oxidized ATP, LPS from *Salmonella minnesota* Re 595, 3-methyadenine, muramyl dipeptide and anti-α-tubulin antibody were purchased from Sigma. Lipid A was purchased from Peptide Institute. Flagellin was purchased from Invivogen. Uric acid crystals were purchased from Alexis. N-acetyl-L-cysteine was purchased from Nacalai Tesque. Poly(I)·poly(C) was purchased from GE Healthcare. R-848 was kindly provided by the Pharmaceuticals and Biotechnology Laboratory Japan Energy Corporation. CpG oligodeoxynucleotides (ODN1668) were synthesized at Hokkaido System Science Co. Anti-Atg16L1and anti-Atg5 antibodies were kindly donated by N. Mizushima (Tokyo Medical and Dental University).

K. pneumonia and *E. aerogenes* were kindly donated by the Research Institute of Microbial Disease, Osaka University. *E. coli* (DH5α) was purchased from TOYOBO. *Salmonella enteritica serovar typhimurium* (SR-11 × 3181) was provided by the Kitasato Institute for Life Science. Mice deficient in MyD88, TRIF or Atg7 were described previously[9,13]. Atg5-deficient MEFs and Plat-E cells were kindly donated by N. Mizushima and T. Kitamura (University of Tokyo), respectively. RAW264.7 cells were purchased from ATCC.

Plasmids. The retroviral expression constructs pMRX-ires-puro, pMRX-ires-bsr and pMRX-ires-EGFP (donated by S. Yamaoka) were derivatives of pMX (donated by T. Kitamura). Complementary DNA encoding human IL-1β was inserted into pMRX-ires-EGFP, generating pMRX-ires-IL-1β. Complementary DNA encoding Atg16L1 lacking in coiled-coil domain (Atg16L1 ΔCCDβ) was inserted into pMRX-ires-puro, generating pMRX-Atg16L1 ΔCCDβ-ires-puro. Complementary DNA encoding the GFP-Atg5 chimaeric protein was inserted into pMRX-ires-bsr, generating pMRX-GFP-Atg5-ires-bsr. Construction and inhibitory function details of pStrawberry-Atg4^{C74A} are described elsewhere[31].

RT–PCR. Total RNA was isolated using RNAeasy Mini kits (Qiagen) according to the manufacturer's instructions. Reverse transcription was performed using ReverTra Ace (TOYOBO) according to the manufacturer's instructions. For quantitative PCR, cDNA fragments were amplified by RealTime PCR Master Mix (TOYOBO) according to the manufacturer's instructions. Fluorescence from the Taqman probe for each cytokine was detected by a 7500 Real-Time PCR System (Applied Biosystems). To determine the relative induction of cytokine mRNA in response to LPS stimulation, the mRNA expression level of each cytokine was normalized to the expression level of 18S RNA. The experiments were repeated at least twice. The results were reproducible.

The primer pairs used in Supplementary Fig. 1 are as follows. Primer pair A (exon 1-2), 5′-GTTCGGCATGTCGTCGGGCCTG-3′ and 5′-ATTTCATGCC-TATTTGGCATGTCATGC-3′. Primer pair B (exon 6-7), 5′-GTCAAGCACG-GCTGCAGAAGGAGCTTG-3′ and 5′-GTAGCTGCTCTGCTGACAGCTCGG-3′. Primer pair C (exon 1-5), 5′-GTTCGGCATGTCGTCGGGCCTGN-3′ and 5′-GACCAGTTCCTGGTTCTCCTCAGTAG-3′. Primer pair D (exon 5-7), 5′-CGCCTCAATGCAGAGAATGAGAAGGAC-3′ and 5′-GTAGCTGCTCTGCT-GACAGCTCGG-3′. Primer pair E (exon 1-6), 5′-GTTCGGCATGTCGTCG-GGCCTG-3′ and 5′-CAAGCTCCTTCTGCAGCCGTGCTTGAC-3′.

Bulk protein degradation assay. Cells were exchanged into labelling medium containing ^{14}C-valine (1.5 μCi ml^{-1}) and incubated overnight. Cells were exchanged into chase medium (DMEM supplemented with 10% FBS and 10 mM unlabelled valine) and further incubated for 4 h to remove the contribution of short-lived proteins. After the chase period, cells were exchanged into HBSS containing 10 mM valine to induce autophagy. After a 2 h incubation, the media were collected and the trichloroacetic acid (TCA)-soluble fraction was analysed by scintillation counting. Cells were lysed in ice-cold RIPA buffer and the TCA-insoluble fraction was isolated and analysed by scintillation counting. To determine the rate of long-lived protein degradation, the count in the TCA-soluble fraction in the medium was divided by the equivalent TCA-insoluble count in the cell.

Detection of apoptosis. The appearance of mono/oligo-nucleosomes, markers for apoptosis, was detected by Cell Death Detection ELISA^{PLUS} (Roche). Chromatin condensation in apoptotic nuclei was detected using an ApoStrand ELISA Apoptosis Detection Kit (Biomol). The experiments were repeated twice, and the results were reproducible.

Determination of bacteria colony forming units (c.f.u.s). Levels of c.f.u.s in freshly isolated faeces or spleen were determined by homogenization of material in PBS/0.01% Triton X-100 followed by serial dilution plating on non-selective Luria-Bertani agar.

31. Fujita, N. *et al.* An Atg4B mutant hampers the lipidation of LC3 paralogues and causes defects in autophagosome closure. *Mol. Biol. Cell* advance online publication, doi:10.1091/mbc.E08-03-0312 (3 September 2008).

LETTERS

A fasting inducible switch modulates gluconeogenesis via activator/coactivator exchange

Yi Liu[1]*, Renaud Dentin[1]*, Danica Chen[2]*, Susan Hedrick[1], Kim Ravnskjaer[1], Simon Schenk[3], Jill Milne[4], David J. Meyers[5], Phil Cole[5], John Yates III[6], Jerrold Olefsky[3], Leonard Guarente[2] & Marc Montminy[1]

During early fasting, increases in skeletal muscle proteolysis liberate free amino acids for hepatic gluconeogenesis in response to pancreatic glucagon. Hepatic glucose output diminishes during the late protein-sparing phase of fasting, when ketone body production by the liver supplies compensatory fuel for glucose-dependent tissues[1–4]. Glucagon stimulates the gluconeogenic program by triggering the dephosphorylation and nuclear translocation of the CREB regulated transcription coactivator 2 (CRTC2; also known as TORC2), while parallel decreases in insulin signalling augment gluconeogenic gene expression through the dephosphorylation and nuclear shuttling of forkhead box O1 (FOXO1)[5–7]. Here we show that a fasting-inducible switch, consisting of the histone acetyltransferase p300 and the nutrient-sensing deacetylase sirtuin 1 (SIRT1), maintains energy balance in mice through the sequential induction of CRTC2 and FOXO1. After glucagon induction, CRTC2 stimulated gluconeogenic gene expression by an association with p300, which we show here is also activated by dephosphorylation at Ser 89 during fasting. In turn, p300 increased hepatic CRTC2 activity by acetylating it at Lys 628, a site that also targets CRTC2 for degradation after its ubiquitination by the E3 ligase constitutive photomorphogenic protein (COP1)[8]. Glucagon effects were attenuated during late fasting, when CRTC2 was downregulated owing to SIRT1-mediated deacetylation and when FOXO1 supported expression of the gluconeogenic program. Disrupting SIRT1 activity, by liver-specific knockout of the *Sirt1* gene or by administration of a SIRT1 antagonist, increased CRTC2 activity and glucose output, whereas exposure to SIRT1 agonists reduced them. In view of the reciprocal activation of FOXO1 and its coactivator peroxisome proliferator-activated receptor-γ coactivator-1α (PGC-1α, encoded by *Ppargc1a*) by SIRT1 activators[9–12], our results illustrate how the exchange of two gluconeogenic regulators during fasting maintains energy balance.

We compared the effects of short- and long-term fasting on hepatic CRTC2 activity in mice using an adenoviral CRE-luciferase (Ad-CRE-luc) reporter. Fasting for 6 h induced Ad-CRE-luc activity; these effects were augmented by intraperitoneal glucagon injection (Fig. 1a and Supplementary Fig. 1). Hepatic Ad-CRE-luc activity returned to near basal levels after 18–24 h of fasting, when circulating ketone bodies were highest and when hepatic gluconeogenesis was reduced (Fig. 1a, top and Supplementary Fig. 2)[13]. In keeping with the decrease in gluconeogenic gene expression, hepatic CRTC2 protein amounts were also downregulated in response to prolonged fasting (Fig. 1a, bottom, and Supplementary Figs 1 and 3).

The E3 ligase COP1 has been shown to silence the gluconeogenic program during refeeding by the ubiquitin-dependent degradation of CRTC2 (ref. 8). Although undetectable in 6-h-fasted mice,

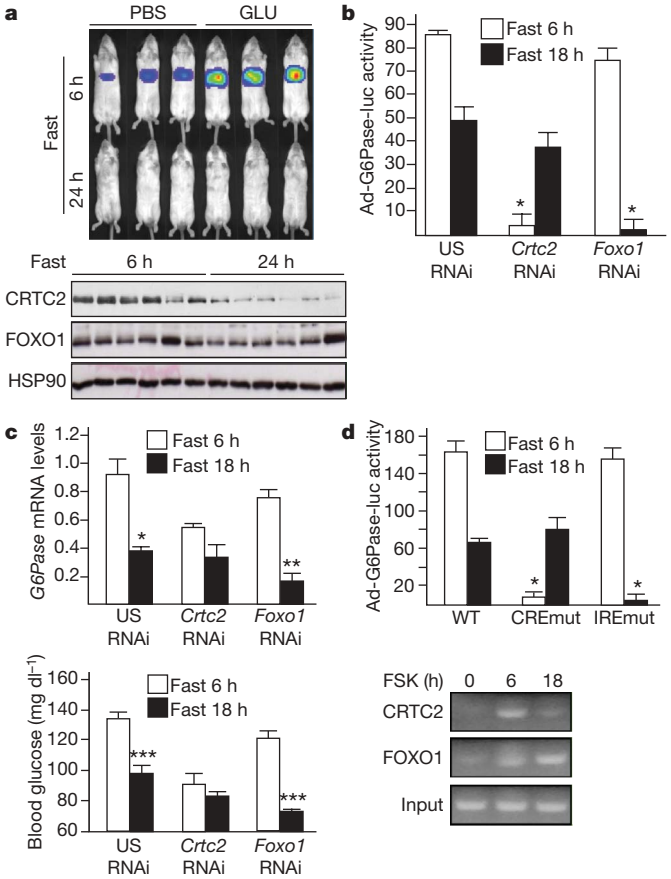


Figure 1 | Sequential activation of CRTC2 and FOXO1 during fasting. **a**, Ad-CRE-luc activity (top) and CRTC2 protein amounts (bottom) in mice fasted for 6 or 24 h. Intraperitoneal glucagon (GLU) injection is indicated, PBS was used as a control for glucagon injection. HSP90 was used as a protein loading control. **b**, **c**, The effect of 6 or 18 h fasting on Ad-G6Pase-luc activity (**b**), *G6Pase* messenger RNA amounts (**c**, top) and blood glucose concentrations (**c**, bottom) in mice infected with the adenoviruses *Ad-Crtc2* RNAi, *Foxo1* RNAi or the control virus US (unspecific) RNAi $n = 4$, $*P < 0.05$; $**P < 0.02$; $***P < 0.01$. **d**, Top, activities of wild-type or mutant Ad-G6Pase-luc reporters defective in CREB (CREmut) or FOXO1 (IREmut) binding. Mice were fasted for 6 or 18 h as indicated; $n = 4$, $*P < 0.05$. Bottom, chromatin immunoprecipitation assay showing binding of Myc-tagged FOXO1 or Flag-epitope-tagged CRTC2 to the G6Pase promoter in HepG2 hepatocytes exposed to FSK for 6 or 18 h. For panels **b**–**d**, data are mean ± s.e.m.

[1]The Salk Institute for Biological Studies, 10010 North Torrey Pines Rd, La Jolla, California 92037, USA. [2]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [3]Department of Medicine, University of California, San Diego, La Jolla, California 92093, USA. [4]Sirtris Pharmaceuticals Inc., 200 Technology Square, Cambridge, Massachusetts 02139, USA. [5]Department of Pharmacology and Molecular Sciences, Johns Hopkins University School of Medicine, 725 North Wolfe Street, 316 Hunterian Building, Baltimore, Maryland 21205, USA. [6]The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.
*These authors contributed equally to this work.

ubiquitinated CRTC2 protein amounts increased after 18 h of fasting (Supplementary Fig. 3). Indeed, prolonged exposure to glucagon triggered CRTC2 degradation in primary hepatocytes; these effects were blocked by treatment with the proteasome inhibitor MG132 and by RNA interference (RNAi)-mediated depletion of COP1 (Supplementary Figs 4 and 5).

In contrast to CRTC2, hepatic FOXO1 protein levels remained constant during fasting, suggesting that these transcriptional regulators are differentially regulated (Fig. 1a, bottom, and Supplementary Fig. 3). To test this idea, we used an adenoviral G6Pase (also known as G6pc)-luciferase (Ad-G6Pase-luc) reporter, which contains FOXO1 and CREB binding sites that mediate induction of the *G6Pase* gene during fasting[14–18]. Relative to feeding, hepatic Ad-G6Pase-luc activity increased markedly after fasting for 6 h (Supplementary Fig. 6). In contrast to the complete suppression of Ad-CRE-luc activity thereafter, however, Ad-G6Pase-luc activity decreased by only 50% after 18 h (Fig. 1b).

We performed knockdown studies to determine the regulatory contributions of CRTC2 and FOXO1 during fasting. RNAi-mediated depletion of hepatic CRTC2 reduced the gluconeogenic profile, which includes Ad-G6Pase-luc activity, gluconeogenic gene expression, and circulating blood glucose concentrations in short-term fasted mice, whereas depletion of FOXO1 had only modest effects at this time (Fig. 1b, c and Supplementary Fig. 7). In contrast, FOXO1 knockdown substantially reduced the gluconeogenic profile after 18 h of fasting, when CRTC2 was degraded.

We tested the relative importance of CREB and FOXO1 promoter binding sites for *G6Pase* gene expression during short- and long-term fasting. Mutation of the cyclic AMP response element (CRE) blocked Ad-G6Pase-luc induction during short-term fasting but had no effect during long-term fasting (Fig. 1d). Conversely, mutation of FOXO1 binding sites (insulin response elements, IREs) disrupted G6Pase reporter activity during long-term but not short-term fasting. We observed similar effects of CREB and FOXO1 binding sites on G6Pase promoter activity in cultured HepG2 cells exposed for 6 or 18 h to forskolin (FSK; Supplementary Fig. 8). Consistent with this activation profile, CRTC2 occupancy over the *G6Pase* promoter was maximal after short-term exposure to FSK, and returned to baseline levels after 18 h (Fig. 1d, bottom and Supplementary Fig. 8). In contrast, FOXO1 occupancy was low at 6 h and increased after 18 h, indicating that CRTC2 and FOXO1 probably regulate gluconeogenic gene expression sequentially in response to fasting.

We reasoned that CRTC2 activity during fasting may be regulated by lysine acetylation, because this modification has been shown to protect certain activators against ubiquitin-mediated degradation[19]. Supporting this idea, fasting led to CRTC2 acetylation after 8 h and to CRTC2 ubiquitination after 24 h (Fig. 2a, top). Exposure of primary hepatocytes to glucagon also triggered CRTC2 acetylation; these effects were reversed by subsequent exposure to insulin (Supplementary Fig. 9).

Using mass spectrometry to characterize residues in CRTC2 that undergo acetylation, we found a single site at Lys 628, also corresponding to the principal ubiquitination site in CRTC2 (Supplementary Fig. 10)[8]. We confirmed these findings using wild-type and Lys628Arg mutant CRTC2 constructs; exposure to FSK increased the acetylation of wild-type but not Lys628Arg mutant CRTC2 (Fig. 2a, bottom). Consistent with an important role for Lys 628 in modulating CRTC2 activity, Ad-CRE-luc activity, circulating glucose levels and CRTC2 protein amounts were increased in mice expressing mutant Lys628Arg CRTC2 compared to wild-type CRTC2 during prolonged fasting (Fig. 2b and Supplementary Fig. 11).

CRTC2 has been found to promote CREB target gene expression through an association with the histone acetyltransferase (HAT) paralogues CREB binding protein (CBP; also known as CREBBP) and p300 (ref. 20). Indeed, short-term fasting increased the CRTC2–p300 interaction in liver, whereas long-term fasting disrupted it (Fig. 2a). Exposure to glucagon or FSK also triggered this association in primary hepatocytes; these effects were blocked by subsequent exposure to insulin (Supplementary Figs 5 and 12).

During the course of studies to determine how insulin and glucagon regulate the CRTC2–p300 interaction, we noted that, similar to CRTC2, p300 and CBP contain a consensus recognition motif for the salk inducible kinase 2 (SIK2), which is at Ser 89 in p300 (ΨXBS/TXSXXXΨ, in which Ψ is a hydrophobic residue and B is a basic amino acid; p300, LLRSGSSPNL). Indeed, phosphorylation of p300 at Ser 89 has been reported to inhibit its transcriptional activity, although the underlying mechanism is unclear[21,22]. Under basal conditions, p300 was phosphorylated at Ser 89 in primary hepatocytes (Fig. 2c, top). Consistent with the upregulation of hepatic SIK2 activity during feeding and inhibition during fasting[8], the amounts of Ser 89-phosphorylated p300 increased when cells were exposed to insulin, and they decreased after treatment with glucagon. RNAi-mediated depletion of SIK2 reduced amounts of Ser 89-phosphorylated p300 in cells exposed to insulin, indicating that p300 is probably

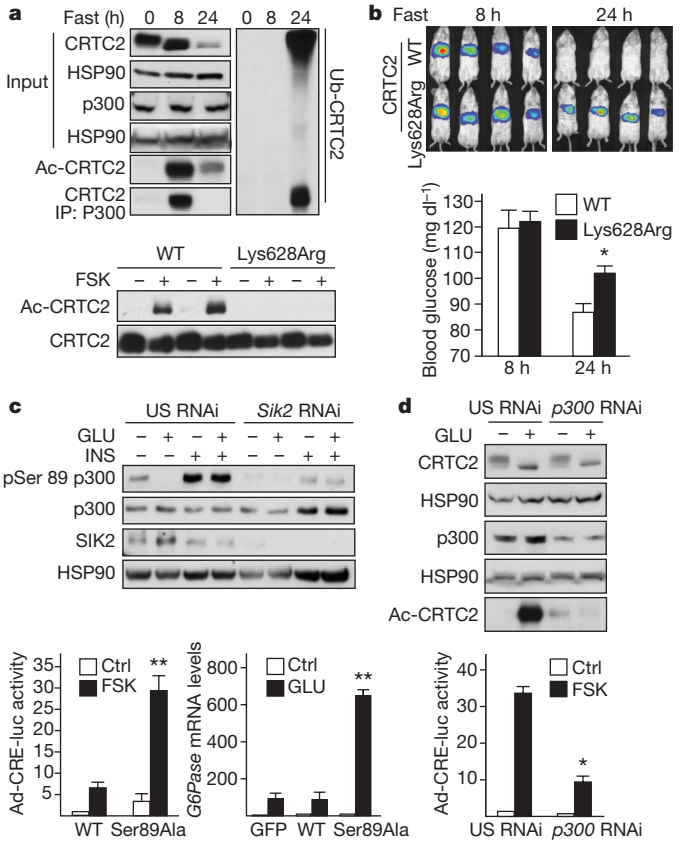


Figure 2 | CBP and p300 promote CRTC2 acetylation during fasting. **a**, Top, immunoblot of acetylated (Ac) or ubiquitinated (Ub) hepatic CRTC2 proteins in 8-h- and 24-h-fasted mice. Recovery of CRTC2 from p300 immunoprecipitates (IP) shown. Bottom, immunoblot of acetylated CRTC2 in HEK293T cells expressing wild-type (WT) or Lys628Arg-mutant CRTC2. Exposure to FSK is indicated. **b**, Hepatic Ad-CRE-luc activity (top) and circulating glucose levels (bottom) in 8-h- and 24-h-fasted mice expressing wild-type or Lys628Arg-mutant CRTC2. For blood glucose, $n = 3$, $*P < 0.05$. **c**, Top, immunoblot of phospho-Ser 89 (pSer 89) p300 protein amounts in primary hepatocytes exposed to glucagon (GLU; 2 h) followed by insulin (INS; 1 h). The effect of *Ad-Sik2* RNAi relative to control (US RNAi) is shown. Bottom, Ad-CRE-luc reporter activity (left) and *G6Pase* mRNA amounts (right) in primary hepatocytes expressing wild-type or Ser89Ala mutant p300. Exposure to FSK or glucagon (6 h) are indicated; Ctrl, control; GFP, green fluorescent protein; $n = 3$; $**P < 0.001$. **d**, Top, effect of *Ad-p300* RNAi on amounts of acetylated CRTC2 (top) in hepatocytes exposed to glucagon for 1 h. Bottom, effect of *Ad-p300* RNAi on Ad-CRE-luc reporter activity (bottom) in hepatocytes exposed to FSK for 6 h; $n = 3$, $*P < 0.001$. For **b**–**d**, data are mean ± s.e.m.

a direct substrate for this kinase. Moreover, SIK2 immunoprecipitates were able to phosphorylate wild-type but not mutant-Ser89Ala p300 *in vitro* (Supplementary Fig. 13).

To investigate the potential role of Ser 89 phosphorylation in regulating p300 activity, we prepared a phosphorylation-defective (Ser89Ala) p300 expression virus. Relative to wild-type p300, mutant-Ser89Ala p300 associated with CRTC2 more efficiently in cells that were exposed to glucagon and insulin (Supplementary Fig. 14). Furthermore, mutant-Ser89Ala p300 was more effective than wild-type p300 in potentiating Ad-CRE-luc reporter activity and gluconeogenic gene expression, confirming the importance of p300 dephosphorylation for CRTC2 induction (Fig. 2c, bottom).

Because they both have intrinsic HAT activity, CBP and p300 might be expected to modulate CRTC2 activity in part through acetylation. Indeed, overexpression of CBP increased the amounts of acetylated CRTC2 in cells exposed to FSK or to staurosporine, a SIK2 kinase inhibitor (Supplementary Fig. 15). The effects of p300 and CBP seem to be direct, because purified recombinant p300 protein was capable of acetylating a CRTC2 polypeptide containing the Lys 628 acetylation site *in vitro* (Supplementary Fig. 15). Conversely, RNAi-mediated depletion of p300 reduced CRTC2 acetylation and decreased Ad-CRE-luc activity in hepatocytes exposed to glucagon (Fig. 2d and Supplementary Fig. 16).

We evaluated the role of p300 and CBP in modulating gluconeogenesis through CRTC2. RNAi-mediated depletion of hepatic p300, and to a lesser extent CBP, reduced the gluconeogenic profile in 6-h-fasted mice (Fig. 3a). Hepatic CRTC2 protein amounts were also decreased in p300-depleted mice, suggesting that p300 is required to prevent hepatic CRTC2 degradation during fasting (Supplementary Fig. 16).

We tested the importance of p300 and CBP HAT activity for CRTC2-dependent gluconeogenesis. The addition of a cell-permeable p300 and CBP HAT inhibitor Lys-CoA-TAT[23,24] to cultured hepatocytes reduced amounts of acetylated CRTC2, and correspondingly enhanced CRTC2 ubiquitination and degradation (Fig. 3b, top). Ad-CRE-luc reporter activity and glucose output were consequently downregulated in hepatocytes exposed to Lys-CoA-TAT (Fig. 3b, bottom). These effects were blocked in cells expressing acetylation/ubiquitination-defective Lys628Arg CRTC2 but not wild-type CRTC2 (Fig. 3c and Supplementary Fig. 17). Lys-CoA-TAT administration also reduced the gluconeogenic profile in fasted mice (Fig. 3d and Supplementary Fig. 17), demonstrating the importance of p300 and CBP HAT activity for hepatic glucose production through CRTC2.

Having found that hepatic CRTC2 is deacetylated during prolonged fasting, we considered the involvement of a CRTC2 deacetylase in this process. Recently, the histone deacetylase SIRT1 has been shown to promote energy balance by modulating cellular gene expression in response to nutrient deprivation[25]. Indeed, SIRT1 activators have been found to improve glucose homeostasis in insulin-resistant animals by reducing hepatic gluconeogenesis, although, paradoxically, they increase the activity of FOXO1 and its coactivator PGC-1α[9,11,26,27]. Hepatic SIRT1 protein accumulated after 18 h of fasting, when CRTC2 acetylation and protein amounts were correspondingly reduced (Fig. 4a, top left)[9]. In line with these changes, we recovered CRTC2 from immunoprecipitates of SIRT1 prepared from livers of long-term- but not short-term-fasted mice. Moreover, Ad-SIRT1 overexpression in primary hepatocytes reduced the amounts of acetylated and total CRTC2 protein and decreased Ad-CRE-luc activity after exposure to glucagon (Fig. 4a, bottom, and Supplementary Figs 18 and 19). We observed similar inhibitory effects of Ad-SIRT1 on Ad-CRE-luc reporter activity and circulating blood glucose levels *in vivo* (Fig. 4a, right and Supplementary Fig. 19).

We proposed that SIRT1 activators may also attenuate the gluconeogenic program during fasting through the inhibition of CRTC2. Induction of SIRT1 activity with SRT1720 (ref. 26) or resveratrol reduced the amounts of acetylated CRTC2 and lowered Ad-CRE-luc activity in primary hepatocytes exposed to FSK or glucagon (Fig. 4b, top, and Supplementary Fig. 20). Consistent with its ability to lower circulating blood glucose concentrations in part through inhibition of hepatic gluconeogenesis[26], SRT1720 also decreased the amounts of acetylated hepatic CRTC2 in Zucker *fa/fa* rats (Fig. 4b, bottom). In contrast, SRT1720 did not alter Ad-CRE-luc activity or glucose output from hepatocytes expressing acetylation/ubiquitination-defective Lys628Arg CRTC2 (Supplementary Figs 17 and 21).

On the basis of these results, we tested, conversely, whether SIRT1 inhibitors increase CRTC2 activity. Exposure of primary hepatocytes to the SIRT1 antagonists sirtinol and nicotinamide enhanced CRTC2 acetylation and Ad-CRE-luc reporter activity in glucagon-stimulated cells (Fig. 4c, top and Supplementary Fig. 22). Intraperitoneal sirtinol administration also augmented hepatic Ad-CRE-luc reporter activity during prolonged fasting, when CRTC2 activity is normally downregulated (Fig. 4c, bottom).

We tested the role of SIRT1 further using mice with a liver-specific knockout of the *Sirt1* gene. Under basal conditions, CRTC2 protein

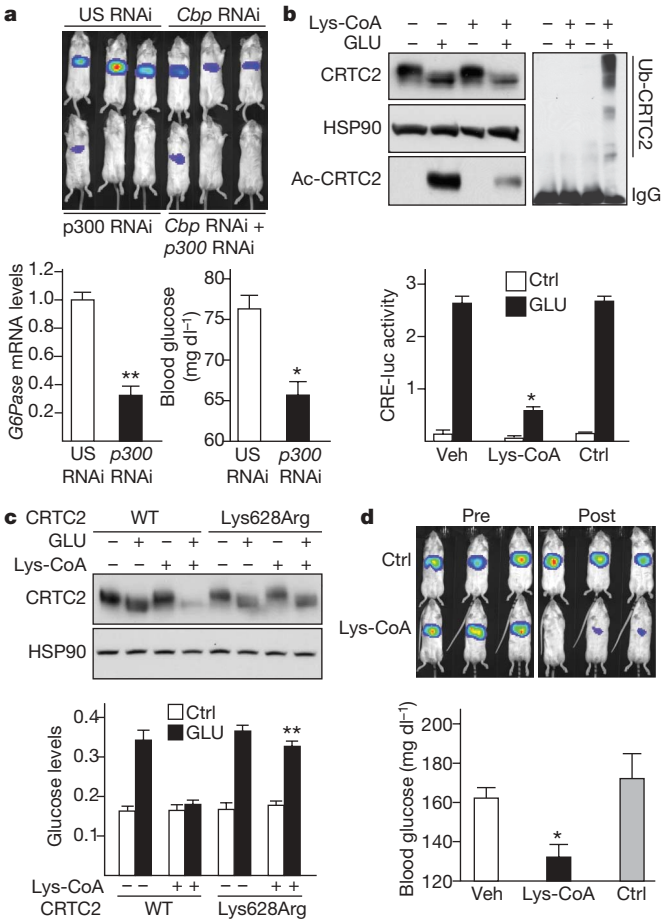


Figure 3 | p300 and CBP modulate hepatic CRTC2 activity. **a**, Top, Ad-CRE-luc activity in 6-h-fasted mice expressing *Ad-Cbp* RNAi, *Ad-p300* RNAi or Ad-US RNAi. Bottom, *G6Pase* mRNA amounts (left) and blood glucose levels (right) in mice expressing *Ad-p300* RNAi relative to control; $n = 3$; $*P < 0.001$; $**P < 0.01$. **b**, Effect of p300 and CBP HAT inhibitor Lys-CoA-TAT (Lys-CoA) or control (Ctrl) TAT peptide on CRTC2 acetylation (top) and Ad-CRE-luc activity (bottom) in primary hepatocytes exposed to glucagon. Ac, acetylated; Ub, ubiquitinated; $n = 3$; $*P < 0.001$. **c**, Effect of Lys-CoA-TAT on CRTC2 protein amounts (top) and on glucose output (bottom) from primary hepatocytes expressing wild-type (WT) or Lys628Arg-mutant CRTC2; $n = 3$; $**P < 0.001$. **d**, Top, Ad-CRE-luc activity in 6-h-fasted mice imaged before (pre) or after (post) intraperitoneal injection of Lys-CoA-TAT or TAT peptide. Bottom, blood glucose concentrations in 8-h-fasted mice injected with Lys-CoA-TAT or TAT control. Veh, vehicle control. For blood glucose levels, $n = 6$; $*P < 0.001$. All data are mean ± s.e.m.

amounts were comparable in primary cultures of *Sirt1*$^{-/-}$ and wild-type hepatocytes (Fig. 4d, top). In contrast to the degradation of CRTC2 after prolonged exposure of wild-type cells to glucagon, however, CRTC2 protein amounts remained elevated in *Sirt1*$^{-/-}$ cells. We observed similar differences in hepatic CRTC2 protein amounts between wild-type and *Sirt1*$^{-/-}$ mice during fasting (Supplementary Fig. 23). As a result, hepatic Ad-CRE-luc activity was elevated and unresponsive to sirtinol administration in fasted *Sirt1*$^{-/-}$ mice (Supplementary Fig. 23). Ad-CRE-luc activity and gluconeogenic gene expression were also increased in *Sirt1*$^{-/-}$ hepatocytes after glucagon exposure, demonstrating the importance of this deacetylase in modulating CRTC2 activity (Fig. 4d, bottom and Supplementary Fig. 24).

Taken together, these results indicate that fasting signals increase the gluconeogenic program transiently through the acetylation of CRTC2 by p300 and CBP (Supplementary Fig. 25). During prolonged fasting, SIRT1 deacetylates CRTC2 and promotes its ubiquitin-dependent degradation by COP1. The reciprocal upregulation of FOXO1 activity by SIRT1 during this period seems to be critical in maintaining energy balance through its effects on glucose metabolism[6]. Studies into the mechanism by which nutrient signals modulate p300 and SIRT1 activities should provide further insight into this process.

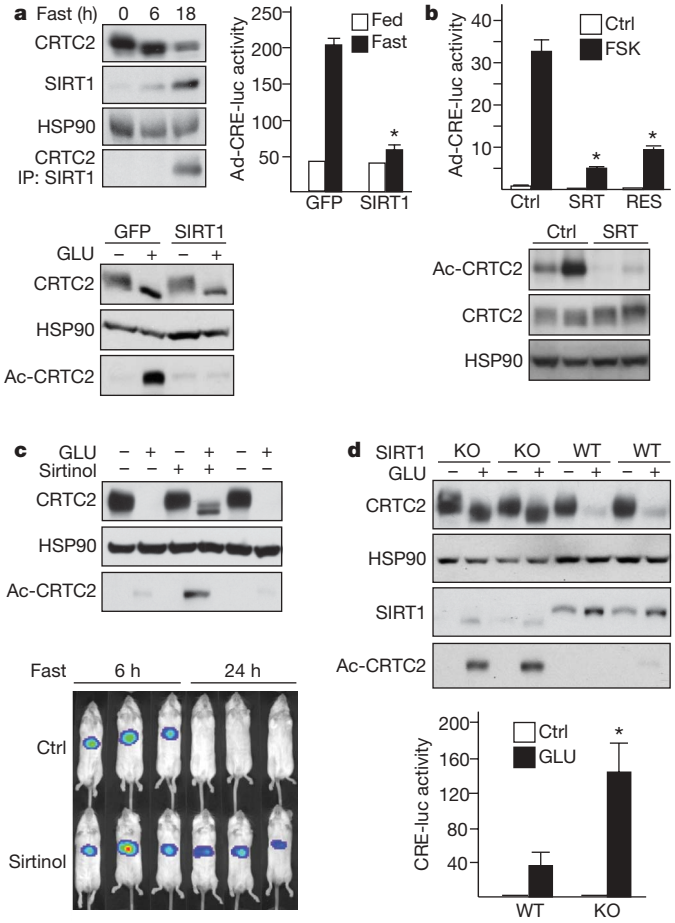


Figure 4 | SIRT1 attenuates CRTC2 activity during fasting. a, Top left, immunoblot of hepatic CRTC2 protein recovered from SIRT1 immunoprecipitates (IP) from *ad libitum* fed and 6-h- or 18-h-fasted mice. Right, effect of Ad-SIRT1 expression on Ad-CRE-luc activity in 8-h-fasted mice (top) and on CRTC2 acetylation in hepatocytes exposed to glucagon (GLU; bottom); $n = 3$; $*P < 0.001$. **b**, Top, Ad-CRE-luc activity in primary hepatocytes exposed to SIRT1 activators resveratrol (RES) or SRT1720 (SRT); $n = 3$; $*P < 0.001$. Bottom, immunoblot showing amounts of acetylated (Ac) CRTC2 in livers of Zucker *fa/fa* obese rats maintained on chow supplemented with SRT1720; $n = 4$; $*P < 0.05$. **c**, Top, immunoblot of acetylated and total CRTC2 protein in primary hepatocytes exposed to sirtinol and/or glucagon for 4 h. Bottom, Ad-CRE-luc activity in 6-h- or 24-h-fasted mice after sirtinol administration. **d**, Total and acetylated CRTC2 protein amounts (top) and Ad-CRE-luc activity (bottom) in primary hepatocytes from control (WT) or liver-specific *Sirt1*$^{-/-}$ (KO) mice. Exposure to glucagon (6 h) is indicated; $n = 4$; $*P < 0.05$. For **a**, **b** and **d**, data are mean ± s.e.m.

METHODS SUMMARY

In vivo imaging studies with Ad-CRE-luc and Ad-G6Pase-luc reporters were performed with an IVIS 100 Imaging System[8]. For nutritional studies, mice were evaluated under *ad libitum*, short-term-fasted (6–8 h), or long-term-fasted (18–24 h) conditions. Adenoviral expression vectors and RNAi constructs for *Sirt1*, *p300* and *Cbp* were generated as described[7]. Fasting gluconeogenesis was evaluated using ^{14}C-labelled Ala as reported[28]. Wild-type and *Sirt1*$^{-/-}$ primary hepatocytes were cultured, and glucose output was measured enzymatically after collection in medium containing lactate and pyruvate[8]. p300 and CBP HAT activity was inhibited by the administration of Lys-CoA-TAT *in vivo* and in cultured hepatocytes[24].

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

Received 7 July; accepted 20 August 2008.
Published online 5 October 2008.

1. Marliss, E. B., Aoki, T. T., Unger, R. H., Soeldner, J. S. & Cahill, G. F. Jr. Glucagon levels and metabolic effects in fasting man. *J. Clin. Invest.* **49**, 2256–2270 (1970).
2. Goodman, M. N. *et al.* Starvation in the rat. II. Effect of age and obesity on protein sparing and fuel metabolism. *Am. J. Physiol.* **239**, E277–E286 (1980).
3. Goodman, M. N., McElaney, M. A. & Ruderman, N. B. Adaptation to prolonged starvation in the rat: curtailment of skeletal muscle proteolysis. *Am. J. Physiol.* **241**, E321–E327 (1981).
4. Cahill, G. F. Jr. Fuel metabolism in starvation. *Annu. Rev. Nutr.* **26**, 1–22 (2006).
5. Puigserver, P. *et al.* Insulin-regulated hepatic gluconeogenesis through FOXO1–PGC-1α interaction. *Nature* **423**, 550–555 (2003).
6. Matsumoto, M., Pocai, A., Rossetti, L., Depinho, R. A. & Accili, D. Impaired regulation of hepatic glucose production in mice lacking the forkhead transcription factor Foxo1 in liver. *Cell Metab.* **6**, 208–216 (2007).
7. Koo, S. H. *et al.* The CREB coactivator TORC2 is a key regulator of fasting glucose metabolism. *Nature* **437**, 1109–1111 (2005).
8. Dentin, R. *et al.* Insulin modulates gluconeogenesis by inhibition of the coactivator TORC2. *Nature* **449**, 366–369 (2007).
9. Rodgers, J. T. *et al.* Nutrient control of glucose homeostasis through a complex of PGC-1α and SIRT1. *Nature* **434**, 113–118 (2005).
10. Baur, J. A. *et al.* Resveratrol improves health and survival of mice on a high-calorie diet. *Nature* **444**, 337–342 (2006).
11. Frescas, D., Valenti, L. & Accili, D. Nuclear trapping of the forkhead transcription factor FoxO1 via Sirt-dependent deacetylation promotes expression of glucogenetic genes. *J. Biol. Chem.* **280**, 20589–20595 (2005).
12. Gan, L. *et al.* FoxO-dependent and -independent mechanisms mediate SirT1 effects on IGFBP-1 gene expression. *Biochem. Biophys. Res. Commun.* **337**, 1092–1096 (2005).
13. Badman, M. K. *et al.* Hepatic fibroblast growth factor 21 is regulated by PPARα and is a key mediator of hepatic lipid metabolism in ketotic states. *Cell Metab.* **5**, 426–437 (2007).
14. Nakae, J., Kitamura, T., Silver, D. L. & Accili, D. The forkhead transcription factor Foxo1 (Fkhr) confers insulin sensitivity onto glucose-6-phosphatase expression. *J. Clin. Invest.* **108**, 1359–1367 (2001).
15. Barthel, A. *et al.* Differential regulation of endogenous glucose-6-phosphatase and phosphoenolpyruvate carboxykinase gene expression by the forkhead transcription factor FKHR in H4IIE-hepatoma cells. *Biochem. Biophys. Res. Commun.* **285**, 897–902 (2001).
16. Hornbuckle, L. A. *et al.* Selective stimulation of G-6-Pase catalytic subunit but not G-6-*P* transporter gene expression by glucagon *in vivo* and cAMP in situ. *Am. J. Physiol. Endocrinol. Metab.* **286**, E795–E808 (2004).
17. Ayala, J. E. *et al.* Conservation of an insulin response unit between mouse and human glucose-6-phosphatase catalytic subunit gene promoters: transcription factor FKHR binds the insulin response sequence. *Diabetes* **48**, 1885–1889 (1999).
18. Schmoll, D. *et al.* Identification of a cAMP response element within the glucose-6-phosphatase hydrolytic subunit gene promoter which is involved in the transcriptional regulation by cAMP and glucocorticoids in H4IIE hepatoma cells. *Biochem. J.* **338**, 457–463 (1999).
19. Freiman, R. N. & Tjian, R. Regulating the regulators: lysine modifications make their mark. *Cell* **112**, 11–17 (2003).
20. Ravnskjaer, K. *et al.* Cooperative interactions between CBP and TORC2 confer selectivity to CREB target gene expression. *EMBO J.* **26**, 2880–2889 (2007).

21. Yuan, L. W. & Gambee, J. E. Phosphorylation of p300 at serine 89 by protein kinase C. *J. Biol. Chem.* **275,** 40946–40951 (2000).
22. Yang, W. *et al.* Regulation of transcription by AMP-activated protein kinase: phosphorylation of p300 blocks its interaction with nuclear receptors. *J. Biol. Chem.* **276,** 38341–38344 (2001).
23. Guidez, F. *et al.* Histone acetyltransferase activity of p300 is required for transcriptional repression by the promyelocytic leukemia zinc finger protein. *Mol. Cell. Biol.* **25,** 5552–5566 (2005).
24. Zheng, Y. *et al.* Synthesis and evaluation of a potent and selective cell-permeable p300 histone acetyltransferase inhibitor. *J. Am. Chem. Soc.* **127,** 17182–17183 (2005).
25. Blander, G. & Guarente, L. The Sir2 family of protein deacetylases. *Annu. Rev. Biochem.* **73,** 417–435 (2004).
26. Milne, J. C. *et al.* Small molecule activators of SIRT1 as therapeutics for the treatment of type 2 diabetes. *Nature* **450,** 712–716 (2007).
27. Motta, M. C. *et al.* Mammalian SIRT1 represses forkhead transcription factors. *Cell* **116,** 551–563 (2004).
28. Ferre, P., Pegorier, J. P., Marliss, E. B. & Girard, J. R. Influence of exogenous fat and gluconeogenic substrates on glucose homeostasis in the newborn rat. *Am. J. Physiol.* **234,** E129–E136 (1978).

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

Acknowledgements We thank M. Kahn for the gift of phospho-specific p300 antiserum and L. Vera for technical assistance. We also thank R. Shaw and M. Mihaylova for sharing results on p300 phosphorylation. This work was supported by grants from the National Institutes of Health, the Clayton Medical Research Foundation, Inc., the Hillblom Foundation (to Y.L.), and the Kieckhefer Foundation.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details accompany the full-text HTML version of the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to M.M. (Montminy@salk.edu).

METHODS

Adenoviruses and animals. Wild-type CRTC2, mutant-Lys628Arg CRTC2, Ad-CRE-luc, Ad-RSV-β-gal (Rous sarcoma virus promoter) and Ad-SIRT1 adenoviruses have been described[8,9]. The Ad-G6Pase-luc reporter was constructed by the insertion of G6Pase-luc into pShuttle vector and by transferring this cassette to AdEasy by non-homologous recombination. *Ad-p300* RNAi was constructed using the sequence ACTTACCAGATGAATTAA. For live imaging experiments, 10^9 plaque forming units (p.f.u.) Ad-CRE-luc and 5×10^7 p.f.u. Ad-RSV-β-gal were delivered to 8–10-week-old male mice (Jackson Laboratories) by tail-vein injection. Mice were imaged on day 3–5 after adenovirus delivery. For other *in vivo* studies, 1×10^8 p.f.u. of overexpressing or RNAi adenovirus was used. SIRT1 liver-specific knockout mice were generated by crossing an SIRT1 allele containing a floxed exon 4 (ref. 29) with Cre-expressing mice driven by the liver-specific albumin promoter.

Measurement of gluconeogenesis rates *in vivo*. Fed, 6-h- and 24-h-fasted mice were injected intraperitoneally with labelled precursor (5mCi of [U-^{14}C]Ala) and blood samples were analysed after 30 min, when ^{14}C incorporation into glucose is still linear. Gluconeogenesis rates *in vivo* were determined as described previously[28]. Results are expressed as glucose synthesized in $\mu mol\, h^{-1} g^{-1}$ body weight. The separation of labelled glucose from labelled alanine and other charged compounds was performed by ion exchange chromatography. A neutralized aliquot of blood barium hydroxide-zinc sulphate was passed onto a layered bed resin column, consisting of Dowex-1-X8 (200–400 mesh) in the formate form and Dowex-50W-X8 (200–400 mesh) in the H^+ form. The sample was eluted with water, and labelled glucose contained in the eluate was assayed for radioactivity in a scintillation spectrometer. To determine whether glucose was indeed collected in the eluate, glucose was transformed into glucose 6-phosphate (G6P) by hexokinase in the presence of ATP. G6P was then absorbed into a column of Dowex-1-X8. By using this method, more than 95% of the counts found in this eluate were absorbed on the column, showing that 95% of labelled substrate on the glucose eluate was glucose, converted to G6P after incubation with hexokinase and ATP. The method thus provides a reliable estimate of the ^{14}C glucose formed from ^{14}C Ala.

***In vivo* imaging and analysis.** For imaging, mice were injected intraperitoneally with 50 mg kg^{-1} nembutal (Abbott Laboratories) and 100 mg kg^{-1} sterile firefly D-luciferin (Xenogen). Mice were imaged on the IVIS 100 Imaging System, and analysed with Living Image software (Xenogen). Luciferase activity from Ad-CRE-luc and Ad-G6Pase-luc reporters was normalized to β-glactosidase activity from co-injected RSV-β-gal plasmid as described previously[8]. For studies comparing mutant and wild-type G6Pase reporters (Fig. 1e), activities were also normalized to reporter DNA from hepatic lysates by quantitative-PCR analysis. For Lys-CoA-TAT studies, mice were injected intraperitoneally with Lys-CoA-TAT (10 nmol g^{-1} body weight) or control TAT at the beginning of a 6-h fast followed by intraperitoneal injection with 100 μg kg^{-1} glucagon (Sigma) 1 h before imaging. Lys-CoA-TAT and control TAT peptides were synthesized as described previously[24]. The absence of liver toxicity was confirmed by measuring plasma alanine aminotransferase activity using a kit from Teco diagnostics. Mouse tissues were prepared as previously described[8]. Blood glucose values were determined using a LifeScan automatic glucometer.

Cell culture. HEK293T cells were transfected as previously described[20]. For reporter studies, Ad-CRE-luc- and Ad-RSV-β-gal-infected hepatocytes (1 p.f.u. per virus per cell) were exposed to glucagon (100 nM) for ~6–12 h. For drug studies, hepatocytes were pre-treated with Lys-CoA-TAT (20 μM) for 18 h, sirtinol (100 μM, Calbiochem) or SRT1720 (1 μM, Sirtris) for 1 h and then incubated with glucagon (100 nM) for ~1–2 h. Glucose output from primary hepatocytes was determined enzymatically after 1 h collection in glucose-free M199 media supplemented with 10 mM lactate and 1 mM pyruvate. The corresponding hepatocytes were lysed in 100 μl buffer and protein concentration was determined using Bio-Rad protein assay reagent. Glucose production was expressed as nmol of glucose produced per h per total protein.

mRNA analysis. Total cellular RNAs from whole liver or from primary cultured hepatocytes were extracted using the RNeasy kit (Qiagen). mRNA levels were measured as described[8].

Immunoblot, immunoprecipitation and *in vitro* acetylation assay. Western blot and immunoprecipitation assays were performed as described[8]. CRTC2 antisera were described previously[7]. For *in vitro* acetylation assays, glutathione *S*-transferase (GST)–CRTC2 (amino acids 601–692) fusion protein (1 μg) was incubated in reaction buffer (50 μl final volume) containing 50 mM Tris-HCl, pH 8.0, 50 mM NaCl, 0.1 mM EDTA, 1 mM dithiothreitol and 2 μM acetyl-CoA, in the presence or absence of baculovirus-expressed and purified p300 protein (90 ng). The reactions were carried out at 30 °C for 30 min. GST–CRTC2 (amino acids 601–692) was pulled-down using glutathione sepharose 4B beads (Amersham). The acetylated proteins were detected by immunoblotting with the anti-acetyl-lysine antibody (Cell Signaling). Phospho-Ser 89 p300 antiserum has been described[30].

Statistical analyses. Results are reported as mean ± s.e.m. The comparison of different groups was carried out using a two-tailed unpaired Student's *t*-test. Differences were considered statistically significant at $P < 0.05$. All experiments were performed on at least two independent occasions.

29. Cheng, H. L. *et al.* Developmental defects and p53 hyperacetylation in Sir2 homolog (SIRT1)-deficient mice. *Proc. Natl Acad. Sci. USA* **100**, 10794–10799 (2003).
30. Miyabayashi, T. *et al.* Wnt/β-catenin/CBP signaling maintains long-term murine embryonic stem cell pluripotency. *Proc. Natl Acad. Sci. USA* **104**, 5668–5673 (2007).

CORRIGENDUM

doi:10.1038/nature07347

The delayed rise of present-day mammals

Olaf R. P. Bininda-Emonds, Marcel Cardillo, Kate E. Jones, Ross D. E. MacPhee, Robin M. D. Beck, Richard Grenyer, Samantha A. Price, Rutger A. Vos, John L. Gittleman & Andy Purvis

Nature 446, 507–512 (2007)

We have discovered a bug in the Perl script relDate v.2.2 that was used in part to date the nodes in the species-level mammalian supertree presented and analysed in our Article. The bug affected all but 80 of the 2,109 published dates, generally causing them to be slightly inflated, with the effect being stronger in more recent nodes. The absolute errors are mostly small (mean and median change of 1.32 and 0.70 million years, respectively), and a strong correlation between the two sets of dates exists ($r = 0.990$); however, 25 dates (all within Chiroptera) do change by more than 10 million years. Four of these dates are associated with the paraphyletic genus *Hipposideros*, whereas the remaining 21 cover most of Molossidae. The errors do not affect the results or overall conclusions of our paper qualitatively.

The Supplementary Information, including the tree files, has now been amended and can be accessed through the Supplementary Information link of the original Article. An additional file with a version of the amended Article can be accessed at http://www.uni-oldenburg.de/molekularesystematik/ under the 'Publikationen/Publications' link.

CORRIGENDUM

doi:10.1038/nature07432

STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling

Hiroki Ishikawa & Glen N. Barber

Nature 455, 674–678 (2008)

We inadvertently failed to notice that STING protein is encoded by the same gene as the previously described plasma membrane tetraspanner MPYS[1].

1. Jin, L. *et al.* MPYS, a novel membrane tetraspanner, is associated with major histocompatibility complex cass II and mediates transduction of apoptotic signals. *Mol. Cell. Biol.* 28, 5014–5026 (2008).

CORRIGENDUM

doi:10.1038/nature07514

A role for clonal inactivation in T cell tolerance to Mls-1[a]

Marcia A. Blackman, Hans-Gerhard Burgert, David L. Woodland, Ed Palmer, John W. Kappler & Philippa Marrack

Nature 345, 540–542 (1990)

In this Article, the name of Hans-Gerhard Burgert was incorrectly listed as Hans Gerhard-Burgert.

ADDENDUM

doi:10.1038/nature07566

Genes mirror geography within Europe

John Novembre, Toby Johnson, Katarzyna Bryc, Zoltán Kutalik, Adam R. Boyko, Adam Auton, Amit Indap, Karen S. King, Sven Bergmann, Matthew R. Nelson, Matthew Stephens & Carlos D. Bustamante

Nature 456, 98–101 (2008)

A related manuscript arriving at broadly similar conclusions based on partially overlapping data has recently been published[1]. Specifically, 661 of the 3,192 samples from the POPRES collection[2] analysed in our paper were also analysed by Lao *et al.*[1].

1. Lao, O. *et al.* Correlation between genetic and geographic structure in Europe. *Curr. Biol.* 18, 1241–1248 (2008).
2. Nelson, M. R. *et al.* The population reference sample, POPRES: a resource for population, disease, and pharmacological genetics. *Am. J. Hum. Genet.* 83, 347–358 (2008).

CORRIGENDUM

doi:10.1038/nature07515

Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains

Yusuke Sato, Azusa Yoshikawa, Atsushi Yamagata, Hisatoshi Mimura, Masami Yamashita, Kayoko Ookata, Osamu Nureki, Kazuhiro Iwai, Masayuki Komada & Shuya Fukai

Nature 455, 358–362 (2008)

In this Fig. 3c of this Article, Asp 324 was incorrectly labelled as Glu 324.

naturejobs

THE CAREERS MAGAZINE FOR SCIENTISTS

In response to a severe financial downturn, US president Franklin Roosevelt invested in the country's roads and infrastructure. And in the face of the current international slump, President-elect Barack Obama has pledged to do the same. Such projects generate jobs and assist in the transport and delivery of goods and services, providing a much-needed boost for a sluggish economy. But some say that science investment can pay similar dividends.

Edwin Mansfield from the University of Pennsylvania in Philadelphia calculated that every dollar invested in academic research yields $1.28 in social benefits (E. Mansfield *Res. Pol.* **20,** 1–12; 1991). He showed, for example, that some drugs would not have been developed without university research. Addressing biomedical research in particular, economists Kevin Murphy and Robert Topel from the University of Chicago in Illinois have claimed that reductions in the number of people dying from heart disease translated into an extra $1.5 trillion a year for the economy from 1970 to 1990, as people had longer and more productive working lives. They also calculated that reducing deaths from cancer by just 1% would save the economy around $500 billion (K. M. Murphy and R. Topel *The Economic Value of Medical Research*; Univ. Chicago, 1999). By this rationale, science should be a high budgetary priority for countries, even when discretionary spending is scarce.

And so it is in some nations. China, for instance, remains steadfastly committed to major environmental science and biomedical research projects, despite having seen a significant drop in growth (see page 155). And in the United States, Obama has advocated that the government kick-start a 'green economy' to generate jobs and, presumably, a self-sustaining industry that lessens dependence on fossil fuels. Although it may seem counterintuitive to some policy-makers, the careful spending of research money can boost the economy, rather than simply draining the coffers. Bridges and roads to scientific progress, it seems, are also worthwhile investments.

Gene Russo is editor of *Naturejobs*.

CONTACTS
Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
***Nature*events:**
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com
UK Corporate:
Nils Moeller (4953)
Southwest UK/RoW:
Alexander Ranken (4944)
Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)
France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Reya Silao (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com
***Naturejobs* web development:** Tom Hancock
***Naturejobs* online production:** Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2–37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

NATURE|Vol 456|13 November 2008

BRAIN WORK

Industry's urge to build on breakthroughs in neurotechnology could be a boon for business-minded scientists, says **Virginia Gewin**.

The brain may soon be even bigger business. Better treatments have long been needed to address Alzheimer's disease, chronic pain and depression. Now, breakthroughs in brain imaging, neuroinformatics and neurogenesis are creating an industry that is eager to lure cutting-edge academics. Diagnostics, products, brain-stimulation treatments and, above all, drugs are being designed to meet the ever-growing demand for treatment of neurological disease. Scientists with a knack for neuroscience and an interest in business could benefit.

Last year, global neurotechnology industry revenues rose 8.3% to US$130.5 billion, says NeuroInsights, a market-analysis firm based in San Francisco, California. Its *Neurotechnology Industry 2008 Report*, which profiled 500 public and private companies, divided the industry into three sectors: neuropharmaceuticals, neurodevices and neurodiagnostics. In 2007, venture capitalists invested a record $1.77 billion in budding neurotech firms (see graph). Drug companies such as AstraZeneca and Pfizer are investing heavily in developing treatments for diseases of the central nervous system (CNS).

"The conditions are ripe for neurotech to emerge as biotech did in the 1970s — if the economic downturn doesn't quash the momentum," says Christopher deCharms, chief executive of Omneuron, a neurotech company based in Menlo Park, California.

Major hurdles do remain. The global credit crisis is likely to have at least a short-term impact on investment and venture capital. And clinical trials for potential CNS treatments are notoriously difficult, expensive and prone to failure. There is reason for positivity, though, as some in the industry have already started to hedge their bets by increasingly pursuing unconventional collaborations with academia and government. The aim is to take advantage of academia's early-stage research and development (R&D), which is costly but necessary for better treatment design. Such arrangements should afford neuroscientists new types of career opportunities.

Sten Linnarsson: Karolinska graduate and entrepreneur.

Demand for new therapies is the primary reason for optimism at Medtronic in Minneapolis, Minnesota, according to Lothar Krinke, its vice-president of business development. Medtronic is developing deep-brain stimulation devices, the latest ones for epilepsy and depression. The company is increasing its workforce by roughly 10%, Krinke says. In particular, it wants to strengthen preclinical expertise as it expands its R&D. St Jude Medical, a neurotechnology company established by the renowned hospital in St Paul, Minnesota, is recruiting scientists, clinicians and engineers with device experience as it expands its own electrical-neuromodulation business. It hopes to use deep-brain stimulation to treat depression. Omneuron is hiring scientists to help turn real-time brain imaging into therapies — such as learning how to use cognition to 'turn off' brain states that cause chronic pain.

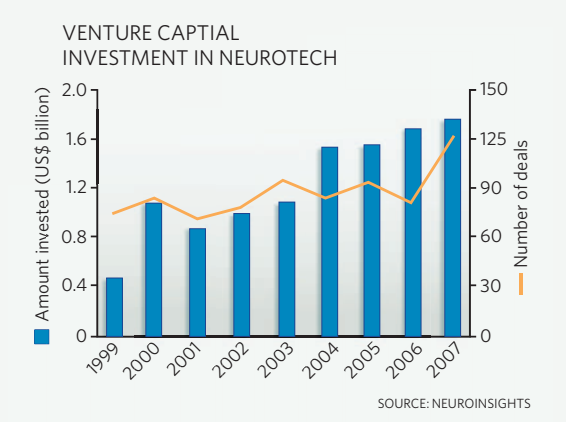
Search for talent

The budding area of neurogenesis research, meanwhile, has prompted drug companies to look to start-ups and academia for talent, says Dan O'Connell, venture capitalist at NeuroVentures in Charlottesville, Virginia. O'Connell calls the development of small molecules to encourage neurogenesis — in which endogenous stem cells mature into neurons — a "breakthrough area". Swiss-based drug giant AstraZeneca announced last month that it will collaborate with Columbia University's René Hen to explore novel neurogenesis-related depression and anxiety treatments.

Also in demand is expertise in the regulatory and clinical-trial arenas. "Ideally, we want to recruit those scientists in pharma who are frustrated and tired of the slow pace and want to come and work in a smaller company," says Ulf Ljungberg, chief executive of NeuroNova, a neurogenesis company in Stockholm, Sweden. But, he adds, most of the company's potential employees come from academia. St Jude Medical is

looking for scientists who have experience with the US Food and Drug Administration, an example of an ever-changing regulatory environment.

One company specializing in psychiatric patient evaluations and therapy responses, MedAvante, is recruiting people with expertise in clinical-trial design and statistics. MedAvante, based in Hamilton, New Jersey, builds assessment methodologies in an effort to improve the design of clinical trials for potential CNS treatments, which often show poor results. For example, a study of dozens of clinical trials for antidepressant treatments found that half the tested drugs failed (A. Khan *et al. Psychol. Med.* **35,** 743–749; 2005). A $20 million investment by Goldman Sachs in July will help MedAvante expand to meet demand from companies seeking to test drugs, by recruiting PhD clinical psychologists and senior scientists.



Head wins over heart

Another way companies hope to reduce the risk involved in developing neurotech products is through translational research, which is receiving a lot of industry interest — and funding. Although the drug industry overall is experiencing job cuts, neurotech-based research is a relatively bright spot. For example, Pfizer recently dropped its cardiovascular-disease efforts while increasing CNS research. "Big pharma is desperate for new products," says Scott Reines, recently retired senior vice-president, of Johnson & Johnson's CNS/Pain Center of Excellence.

Lennart Mucke, director of the non-profit Gladstone Institute of Neurological Disease in San Francisco, says companies that have been ineffective at producing CNS drugs realize they need to do more risky in-depth R&D on conditions and pathways, to come up with more rational drug designs. "Partnering with academia can mitigate that risk," he says. To facilitate this, Mucke says, Gladstone is recruiting scientists to work at the academic–industry interface in its Center for Translational Research. This was set up in 2006, in partnership with Merck, to find therapies for Alzheimer's, Parkinson's and other diseases. Its work builds on Gladstone's efforts to develop drugs targeting apolipoprotein E4, the main genetic risk factor for Alzheimer's.

Paul Matthews, head of GlaxoSmithKline's Clinical Imaging Centre at Imperial College London, says GlaxoSmithKline invested close to £50 million (US$80 million) in Imperial's academic health centre to foster a seamless partnership between industry and academia. He is seeking four experienced imaging scientists to help develop the use of imaging in evaluating molecules as potential neurological treatments. Few imaging specialists have experience in neuropharmacology, however — yet another reason to form partnerships with leading academic health centres.

X. HAN

Ed Boyden: business skills.

Entrepreneurial skills give an edge

Neurotech-industry managers are also having trouble finding people with the right business and regulatory knowledge. Last year, neurobiologist Ed Boyden created a course series called Neurotechnology Ventures, at the Massachusetts Institute of Technology in Cambridge. Students on the project-oriented course, which teaches the principles of building start-ups, create a business plan to solve a major problem through neurotechnology. Already in high demand, the series is simulcast to universities in Hong Kong and Taiwan.

In Sweden, the Stockholm School of Entrepreneurship (SSES) offers budding neuroscientists opportunities to explore business options. From one-day networking events to courses that teach business models and financial strategies, the SSES pulls expertise from its four academic institutions, which include the Karolinska Institute. Karolinska graduate Sten Linnarsson lauds SSES efforts, and plans to use his entrepreneurial experience to teach SSES courses. After graduating, Linnarsson started Global Genomics, which he recently sold to Genizon Life Sciences Company in Montreal, Canada. He has found it easy to migrate between industry and academia.

O'Connell advises academics interested in business to start with their local resources — including finding colleagues who have already established industry contacts. Academics with a venture idea should check out the university's tech-transfer office, he suggests. Then he recommends finding venture forums and other ways to network with investors — particularly at major meetings such as those of the Society for Neuroscience.

Linnarsson advises entrepreneurs to be independent for as long as possible. "As soon as a venture partner is brought in to build the business, a process begins that takes away some of the entrepreneur's control," he says.

Given the ongoing economic downturn, such efforts may be difficult. But the need for brain therapies should provide opportunities in the long term. James Cavuoto, editor of *Neurotech Reports*, a newsletter based in San Francisco, remains sanguine. "Even the most pessimistic among us think this is a sector that is somewhat immune to the overall economic climate," he says. ■

Virginia Gewin is a freelance science writer based in Portland, Oregon.

IMPERIAL COLLEGE LONDON

Imperial College London: hoping to foster a unique collaboration with industry.

ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG

FRIAS FREIBURG INSTITUTE FOR ADVANCED STUDIES

LIFE SCIENCES – LIFENET

The University of Freiburg, located in the center of the picturesque city at the slopes of the Black Forest, offers an exciting, research-driven academic environment, combined with a 550-year old tradition. This institution offers opportunities to work with leading international academics whose ideas will shape future scientific directions.

The flagship of the university is the Freiburg Institute for Advanced Studies (FRIAS) a newly established research center, funded by the German Federal and State Governments within the framework of the Excellence Initiative for the promotion of academic research (http://www.frias.uni-freiburg.de).

The FRIAS School of Life Sciences (LIFENET) invites applications for

External Senior Fellowships

in the field of Life Sciences/Medicine to study „complex systems“, the research topic of LIFENET. The school will bring together scientists involved in systems biology research to study how the complex properties of biological systems are defined by the interactions of their components.

We offer these positions to researchers of exceptional calibre at an advanced stage of their academic careers. They will get the opportunity to either pursue experimental and theoretical research leading to new horizons in their field, or to use the fellowship as an opportunity to develop integrative research activities together with local researchers.

We are interested in scientists with an interdisciplinary interest and focus on integrative systems biology, technology development in biology/medicine, emergent properties of biological systems, genetic networks, network-based data integration, design of biological databases, or statistical methods of software development. Interactions with the research focus of the Life Science departments of Freiburg University, in particular with the excellence cluster bioss (biological signalling studies - from analysis to synthesis), FRISYS (Freiburg Initiative of Systems Biology) and the collaborative research centers (e.g. SFB592, SFB620, SFB746, SFB780) are expected.

The duration of the stay at FRIAS can be negotiated, and can encompass short or long term visits (up to two years) to conduct research in Freiburg.

Applications should include a motivation letter, CV, list of publications, and a research proposal (about 3 pages, with a strong focus on the future research plan and already established or proposed interactions with local scientists). Please send applications in PDF format and requests for further information to the coordinator of LIFENET, Dr. Britta Küst (britta.kust@frias.uni-freiburg.de).

W172825R

STUDY*ing in Berlin:*
INFECTIOUS DISEASES AND IMMUNE REACTIONS

ZIBI Graduate School Berlin offers doctoral fellowships starting October 2009

Deadline for admission January 15, 2009

The mission of our International Graduate School is to develop students into creative, responsible and self-confident young researchers. We offer outstanding training and support in an excellent scientific network. Research projects aim to elucidate genetic, biochemical and organismal basis of infections and immune reactions.

The ZIBI Graduate School has an active and lively student community. The collegial atmosphere assures that everybody has a productive and memorable time in Berlin.

For further information and for application submission, please visit our website:

www.zibi-graduateschool-berlin.de

Contact: ZIBI Graduate School, c/o Dr. Susann Beetz,
Max Planck Institute for Infection Biology
Campus Charité Mitte, Charitéplatz 1, 10117 Berlin, Germany
Phone: +49 (0) 30 28460160, email: info@zibi-graduateschool-berlin.de

The ZIBI Graduate School is the roof of:

International Max Planck Research School for Infectious Diseases and Immunology (IMPRS-IDI) funded by the Max Planck Society (MPG)

Research Training Group “Genetic and Immunologic Determinants of Pathogen-Host-Interactions” (GRK1121) funded by the German Research Foundation (DFG)

W172937R

GOETHE UNIVERSITÄT FRANKFURT AM MAIN

At the Exzellenzcluster "Macromolecular Complexes" there is an opening for a

Project Implementation Manager

for National, European or International Structural/Molecular Biology Research projects. The Project Implementation Manager will act as liaison between the principal investigator and the partners of the networks, will supervise and prepare scientific reports, will be in charge of curating the projects' websites, and assist the principal investigator in the preparation of new grant proposals. The Project Implementation Manager will also be responsible for composing and disseminating Press Releases, high-quality figures and project material in the form of posters or PPT presentations.

The position is localized at the Cluster of Excellence Frankfurt in a highly collaborative, international culture. It fosters top quality, interdisciplinary research by promoting a vibrant environment consisting of independent research groups composed of outstanding graduate students and postdoctoral fellows. The scientific focus emphasizes on the experimental analysis at multiple levels of biological organization, from the molecule to the organism, as well as computational biology and systems biology.

The successful candidate should have a BSc/MSc or PhD in biological-related disciplines, excellent knowledge of English and German, as well as high organization and coordination skills. The successful candidate should have experience in manuscript editing, and previous experience in scientific journals will be advantageous.

Competitive remuneration (BAT IIa/Ib depending on the qualifications) and very good prospects for personal and professional growth are offered in an excellent working environment.

The position is initially limited up to three years, with the possibility of extension. The contract is based upon the regulations of the “Wissenschaftszeitvertragsgesetz" (law regulating time limited contracts for scientists) and the "Hessisches Hochschulgesetz" (Hessian Law for Higher Education).

The Goethe University Frankfurt is an inclusive, equal opportunity employer offering attractive conditions and benefits appropriate to an international research organization.
The University encourages particularly women to apply. Applications by handicapped people with the same qualifications will be treated preferentially.

Please send your application, a publication list, a detailed CV and the names of two referees willing to provide letters of recommendation to: **cryoem@biophysik.org**
For more informations you can contact Tanja Schuler at Exzellenzcluster "Macromolecular Complexes". Telephone: 0049 69 6303 4700 or email: office@cef-mc.de
Closing date: 15.12.2008

W173016R

www.uni-frankfurt.de

Max-Planck-Institut für demografische Forschung

Max Planck Institute for Demographic Research

www.demogr.mpg.de

The Max Planck Institute for Demographic Research seeks a path-breaking recent Ph.D. eager to develop a highly innovative

Independent Junior Research Group

at a frontier of mathematical, biological, historical, social, economic or qualitative demography.

The successful candidate will be given ample resources to establish and direct an independent research program. For this purpose, he or she will be guaranteed funding to recruit and lead a team of researchers, for operating expenses, scientific collaborators, and technical and secretarial support.

The successful candidate will start a five-year contract (with the possibility of a prolongation after positive evaluation) between July and October 2009. The payment corresponds to the W2 level on the German university scale, equivalent to an Assistant or Associate Professor.

Applicants should have completed a doctoral degree in the past decade. They should have an outstanding record - or show exceptional promise - as demographic scholars.

The Max Planck Society is committed to employing more handicapped individuals and to increasing the share of women in areas where they are underrepresented, and therefore expressly encourages applications from such qualified individuals.

The applicant should submit a three to five page description of a research program, along with a work plan, a complete CV, and three personal references.

Please send applications by January 31, 2009 to appl_irg@demogr.mpg.de.

MAX-PLANCK-GESELLSCHAFT

W173017R

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

FACULTY OF BIOLOGY

The **Faculty of Biology** of the Ludwig-Maximilian-University of Munich invites for applications for the position of a

Professorship (W2) for mycology

(6 years/tenure track).

The successful candidate is expected to establish a strong, competitively funded research program in the field of evolutionary biology and systematics of fungi. Teaching of mycology and systematic botany is required at the undergraduate and graduate levels of several curricula in biological sciences.

Prerequisites for the application are a doctoral degree, teaching skills and a record of research equivalent to the German "Habilitation". Scientific qualifications achieved in the private sector, outside Germany or as Junior professor will also be considered. The position will normally be granted for 6 years. After positive evaluation the appointment become permanent (but not earlier than after 3 years). In exceptional cases and outstanding qualification the appointment may be permanent from the beginning.

Applicants must be younger than 52 at the time of appointment. Exceptions to this rule may be considered if outstanding.

The Ludwig-Maximilian-University of Munich is interested in increasing the number of female faculty members and encourages women to apply. The University supports dual career couples.

The university is an equal opportunity employer and handicapped candidates with equal qualifications will be given preference. Further information concerning the equal opportunity policy of the LMU can be obtained from Prof. K. Jung (kirsten.jung@lrz.uni-muenchen.de), concerning teaching from Prof. J. Nickelsen (joerg.nickelsen@lrz.uni-muenchen.de), concerning the structure of the Faculty from Prof. J. Soll (dekanat19@lmu.de) and the scientific scope of the position from Prof. S. Renner (renner@lrz.uni-muenchen.de).

Applications (including CV, list of publications, a brief summary of present research and future plans, proof of teaching experience, five selected reprints of relevant publications and copies of other important documents) as well as an abbreviated version as pdf-file (including CV, list of publications, research plans and teaching experience) should be sent before **14.12.2008** to **Prof. Dr. Jürgen Soll, Dean of the Faculty of Biology, Ludwig-Maximilian-University of Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany**.

W172970R

MHH Hannover Medical School

The Department of Pediatric Hematology/Oncology, Hannover, Germany (Director Prof. Dr. Christoph Klein), in conjunction with the Excellence Initiative Regenerative Medicine and the DFG-sponsored Clinical Research Group Stem Cell Transplantation, will expand its scientific activities. We are looking for

2 PHYSICAN SCIENTISTS or POSTDOCTORAL FELLOWS

for projects aiming at preclinical investigations of hematopoietic stem cell gene therapy and molecular genetics of primary immunodeficiency syndromes. Experience in gene transfer technologies, cell biology (confocal laser microscopy), molecular or cellular immunology are a prerequisite.

Hannover Medical School, center of excellence for congenital bone marrow disorders, offers a stimulating research environment and excellent career perspectives.

Please send your CV, along with 2 letters of reference, to Christoph Klein MD PhD, Director - Department of Pediatric Hematology/Oncology, Hannover Medical School, Carl Neuberg Strasse 1, D-30625 Hannover (klein.christoph@mh-hannover.de; www.klein-lab.com)

W172709R

Philipps Universität Marburg

The **Faculty of Biology** invites applications for the following positions:

Professorship (W3) in Animal Physiology

to be filled by October 1st, 2009. Applicants should have a strong research background in the field of metabolic physiology (vertebrates or invertebrates). Priority will be given to candidates working on the hormonal or neuronal regulation of metabolic processes. Molecular research on signalling cascades/receptors should be complemented by systemic aspects of ecophysiology, evolutionary biology or behavioural physiology. We expect an integration into the emerging network "Sensors in organogenesis and organ function" at the Faculty of Biology. The position requires teaching within the bachelor, master and teacher (L3) degree programs. Excellent records of teaching international scientific reputation and grant funding are required.

The conditions of employment are set down in §§ 70, 71 of the Hessen State Law on Higher Education (HHG).

Junior Professorship (W1) in Developmental Biology

to be filled by April 1st, 2009. The successful candidate will initially be appointed for three years. Dependent on performance after three years, the position can be extended by up to three years.

The successful applicant is expected to strengthen the existing focus on organogenesis, and should preferably work on muscle development/myoblastfusion. Research should be carried out in an excellent genetic model organism (invertebrate). Applicants are expected to have expertise in genetics, molecular genetics and cell biology (CLSM, TEM) documented by publications. Successful applicants are expected to have a record of independent research funding. The position requires teaching in developmental biology (four semester hours) within the bachelor and master degree programs.

The conditions of employment are set down in § 74 of the Hessen State Law on Higher Education (HHG).

Philipps-Universität Marburg places great importance on the intensive mentoring of students and PhD candidates and therefore regular presence of the teaching staff is required.

Philipps-Universität is an equal opportunities employer, and applications from women are encouraged. Applications from candidates with children are welcome - the Philipps-Universität is dedicated to being a family-friendly university. Qualified disabled people are also encouraged to apply.

Applications (copies) should be submitted until December 12th, 2008 to the President of the Philipps-Universität Marburg, 35032 Marburg, Germany.

W172555R

UNIVERSITÄTS KLINIKUM FREIBURG

One postdoctoral position (Biology, Biochemistry or Medicine)

is available in the division of Infectious Diseases/Immunology at the Centre for Paediatrics and Adolescent Medicine, University Medical Centre Freiburg. We are looking for a highly motivated individual with interest in cell based innate immune mechanisms. The project is designed to generate innovative models of signal regulation in response to microbial components that integrate inflammatory signal transduction events (TLRs, inflammasome), antibacterial pathways and metabolic signaling. Experience in immunological techniques (e.g. immunoprecipitation, FACS) and eukaryotic genetics is a prerequisite, knowledge in imaging would be advantageous. An ex-perienced technician will be assigned to the successful candidate. The interdisciplinary group on "Innate Immunity" (www.uniklinik-freiburg.de/kinderklinik/live/forschung/angebimmun.html) is part of the newly founded Center for Chronic Immunodeficiency (CCI), which comprises groups from the Faculty of Biology and the Max-Planck-Institute of Immunobiology (www.cci.uniklinik-freiburg.de). The CCI offers access to state of the art equipment.

The initial employment will be for two years. Applicants should send a CV to the group leader Dr. Philipp Henneke, Centre for Paediatrics and Adolescent Medicine, Mathildenstr. 1, 79106 Freiburg, Germany, email philipp.henneke@uniklinik-freiburg.de, phone: +49 (0) 761 270 4300.

The University of Freiburg is a leading German University in the disciplines of Biology and Medicine. Freiburg is a historic city in the Black Forest in immediate vicinity to France and Switzerland. Freiburg is famous for both its cultural life and the superb opportunities for outdoor activities.

Full time positions will be dividable as long as it does not conflict with official or legal provisions. Severely disabled applicants will be favoured in case of equal qualification. The publishing institute will be liable for the content of this announcement. Hiring will be carried out by the human resources department of the University Medical Center.

W173052R

ZMBP ZENTRUM FÜR MOLEKULARBIOLOGIE DER PFLANZEN
Universität Tübingen

There is an opening for an

Independent Research Group Leader (Microbial effector biology)

at the Center for Plant Molecular Biology (ZMBP), University of Tübingen. The Center comprises four departments with a total of approximately 200 staff members. The successful candidate is expected to establish a competitive research program on the molecular, physiological and structural characterization of microbial (bacterial, fungal, oomycete) effector proteins that govern host plant susceptibility and immunity. Interested candidates are expected to interact with other research groups at the ZMBP and at other research institutes at the University of Tübingen, such as the X-ray crystallography unit at the Interfaculty Institute of Biochemistry (IFIB). This institution offers excellent conditions for atomic-scale resolution of protein structures (headed by Thilo Stehle).

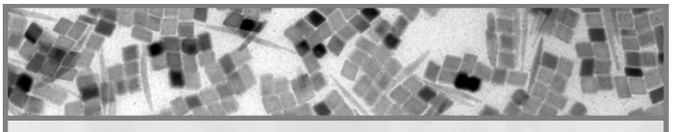
The position will be vacant from January 1, 2009 and is limited initially to 4,5 years. Salary and social benefits are according to the German E13 TV-L. The Center provides lab space with full equipment, funding for staff (1 PhD student, 1 technician), start-up money, an independent annual budget and shared equipment funds. The Research Group will benefit from the ZMBP central service units (administration, plant transformation, cytometry/FACS, greenhouse, analytics, microscopy, computer support) and the IFIB facilities.

For more information see http://www.uni-tuebingen.de/ZMBP.

Applications including CV, publication list and research program should be sent before December 15, 2008 to:

Prof. Thorsten Nürnberger · Universität Tübingen · Center for Plant Molecular Biology (ZMBP) · Auf der Morgenstelle 5 · 72076 Tübingen Germany · nuernberger@uni-tuebingen.de

W173068R



PhD Studies at the

International Helmholtz Research School (IHRS) of Biophysics and Soft Matter

At the interface between biology, chemistry, and physics, many new research fields are emerging, such as soft matter science, nanotechnology, biological physics, and quantitative biology. Forschungszentrum Jülich, Heinrich-Heine-Universität Düsseldorf, Universität zu Köln, and Forschungszentrum caesar in Bonn offer positions for three-year PhD projects together with an interdisciplinary graduate education in the IHRS BioSoft. We provide an intense training in experimental and theoretical techniques, and in transferable skills. Research topics include: Complex Fluids, Colloid and Polymer Physics, Dynamics of Macromolecules, Flow Dynamics and Microfluidics, Bioelectronics and Biomechanics, Cell Biophysics, Plant Biology, Structural Biology, Cellular Signalling Pathways, and many more.

Applications will be accepted at all times.

For further information and an application form: **www.ihrs-biosoft.de**

Contact: Forschungszentrum Jülich, IHRS BioSoft, 52425 Jülich, Germany
Email: admission_biosoft@fz-juelich.de, Tel.: +492461 61-1735

W173060R

GKSS FORSCHUNGSZENTRUM IN DER HELMHOLTZ-GEMEINSCHAFT

Unternehmen Großforschung

Grundlagen für morgen

The GKSS Research Centre is located in Geesthacht near Hamburg, Germany, with a further centre in Teltow near Berlin, and is a member of the Helmholtz Association of German Research Centres (HGF). With its approximately 800 employees it undertakes, in collaboration with universities and industry, research and development in the areas of coastal research, materials research, regenerative medicine, and structure research with neutrons and synchrotron radiation.

The Centre for Biomaterial Development in Teltow of the GKSS Research Centre and Berlin-Brandenburg-Centre for Regenerative Therapies (BCRT) invite applications for a

Senior Postdoc Position
Code-No. 2008/PB 4

in Biomaterial Characterization (Chemist, Material Scientist).

The position is initially limited to 2 years and can be extended depending on performance.
Important working fields of the group are characterization and synthesis of polymer-based biomaterials as well as the field of damage analysis for such materials. Biomaterial characterization will focus on in vitro and in vivo studies of the degradation behavior of different polymers especially the generated degradation products.
In cooperation with other research groups of GKSS, Max-Delbrück-Center, BAM and BCRT translational research project should be performed to move results from fundamental research into clinical applications.
As successful applicant, you should have proven expertise in one or more of the following research areas: synthesis and characterization of polymer- based biomaterials, degradation and damage analysis of such biomaterials or related implants, studies on the interaction of biomaterials with cells and tissues either by own research or in cooperation documented by publication record and a Ph.D. in a field relevant to these research areas.

You have a highly degree of self-motivation, enthusiasm, good soft skills and communication skills as well as a target-oriented working profile. You have at least two years experience as a Postdoc. You are interested in building new competencies at the centre and to participate in the development of internal and external co-operations. You have the ambition to work interdisciplinary with physicists, physicians and biologists in an international work environment offered at the centre. You use your experience in project management to attract third party funding. Publishing of your results in scientific journals and at international conferences is an important part of your work.

Further information can be obtained from Prof. Andreas Lendlein, institute director (andreas.lendlein@gkss.de).

We offer a salary related to the German collective labor agreement TV-AVH as well as the applicable public sector social benefits. In order to increase the number of female employees we ask interested and qualified women to apply for this job. Handicapped persons with equivalent qualifications will be preferred.

Please send your application indicating job offer **code No. 2008/PB 4** to **GKSS-Research Centre, Institute of Polymer Research, Kantstraße 55, 14513 Teltow, Germany** or to the e-mail address: personal.teltow@gkss.de. Please include your CV, names and e-mail addresses of three academic referees, statements of previous research and future research plans including career goals. For details of the institute, please visit the website at http://polymerresearch.gkss.de.

The closing date for applications is November, 28th, 2008.

GKSS • FORSCHUNGSZENTRUM GEESTHACHT
MAX-PLANCK-STRASSE 1 • 21502 GEESTHACHT

W173112R

KIT Karlsruhe Institute of Technology

KIT – Die Kooperation von Forschungszentrum Karlsruhe GmbH und Universität Karlsruhe (TH)

Forschungszentrum Karlsruhe in der Helmholtz-Gemeinschaft

Universität Karlsruhe (TH) Forschungsuniversität · gegründet 1825

Forschungszentrum Karlsruhe GmbH (FZK) is member of the Hermann von Helmholtz Association of National Research Centers and one of the largest public research institutions in Germany. FZK is a multi-disciplinary science and engineering research center that is active in four research areas defined by the Helmholtz Association: Energy, Structure of Matter, Key Technologies, and Earth and Environment. At present, the annual budget of the Forschungszentrum is about EUR 400 million, the total number of employees is approximately 3800.

Forschungszentrum Karlsruhe is presently merging with the Universität Karlsruhe to the Karlsruhe Institute of Technology (KIT). The KIT will be a scientific institution focusing on research, education, and innovation and it will represent a totally novel type of institution in the German science system.

Applications are invited for the position of

Head of the Executive Division

covering activities in one or several of the research areas of Key Technologies, Structure of Matter, and Materials Sciences.

The tasks to be assumed will be in line with the future conception of the KIT. The Head of the Executive Division is responsible for the research programs, institutes, infrastructure facilities, and KIT centers and focuses allocated to this division as well as for cooperation with the corresponding faculties and other scientific institutions. The complete scope of research will be covered, from purely fundamental research to application-oriented and close-to-industry development.

The Head of the Executive Division will be subordinate to and directly report to that member of the Executive Board that is responsible for research. The candidate must have proved his/her scientific competence, extensive expertise in science management, and vast experience in international cooperation. In addition, the candidate shall be capable of managing highly qualified scientists in a cooperative manner and of actively designing the overall development of KIT.

Forschungszentrum Karlsruhe and the Universität Karlsruhe (TH) are actively supporting the career of female scientists as well as the implementation of measures enhancing the compatibility of family and job. Applications of qualified female scientists are therefore particularly encouraged.

The salary will correspond to that usually offered for management functions in science.

Kindly address your application to

Vorsitzender des Vorstandes
der Forschungszentrum Karlsruhe GmbH
Prof. Dr. Eberhard Umbach
- persönlich -
Postfach 36 40, 76021 Karlsruhe, Germany

and send it by **December 05, 2008** at the latest.

W173110R

HOCHSCHULE BIBERACH

BIBERACH UNIVERSITY OF APPLIED SCIENCES

The University of Applied Sciences Biberach situated in the South of Germany in Baden-Württemberg is a small, well renowned and highly ranked University with 1400 students. The Faculty of Pharmaceutical Biotechnology was founded in 2006 as a public-private partnership with local companies such as Boehringer Ingelheim and Rentschler Biotechnologie. The curriculum is aimed at providing practical experience in the area of biopharmaceutical production processes, quality control and development.

The Faculty of Pharmaceutical Biotechnology at the University of Applied Sciences Biberach invites applications for a professorship for March, 2009

W 2 - Professorship

(Reference number PBT 08)

for **cell culture technology**.

The outstanding candidate should have a Ph.D. and extensive experience (an extensive track record) in current bioprocess technologies involving eukaryotic cell cultures. A good knowledge of the German language is essential.

The successful applicant will be expected to give lectures and lead practicals in cell culture technology, development of biotechnological processes and related disciplines.

This professorship is part of the new course on Pharmaceutical Biotechnology in the faculty with the same name. For further information about the curriculum please contact Prof. Dr. Hannemann, Tel. 00 49-(0) 73 51/5 82-4 50.

The requirements of employment arise from the »Landehochschulgesetz Baden-Württemberg«. More information and the complete job advertisement in German are available at www.hochschule-biberach.de/sections/service/stellenanzeigen

The successful candidates are expected to support the further development of the institute and to participate in administrative committees.

Applications with the usual documents (curriculum vitae in tabular form, photo, publications, reports, references etc.) including the reference number are requested **before December 31th, 2008** to the following address:

HBC Hochschule I Personalabteilung
Karlstrasse 11, 88400 Biberach, Germany
Tel. Nr. 00 49-(0) 73 51/5 82-1 20
www.hochschule-biberach.de

W173113R

CHARITÉ
UNIVERSITÄTSMEDIZIN BERLIN

A **postdoctoral position** is available in the Kramer laboratory (www.achim-kramer-lab.de) at Charité – Universitätsmedizin Berlin, Germany. The laboratory focuses on investigating the molecular mechanisms of circadian clocks and their impact on cellular and systemic functions in mammals.

The position is sponsored by Deutsche Forschungsgemeinschaft, the candidate is expected to focus on the biochemical investigation of key circadian proteins with a special emphasis on posttranslational modifications and protein interactions. Ideal candidate should have training and past research experiences in biochemistry and cell biology or related areas.

Requirements: A Ph.D. degree or equivalent in biological sciences obtained within 3 years and a good publication record in peer-reviewed journals. Past research experiences related to the areas of cell biology, biochemistry, molecular biology or chronobiology would be a plus. Candidates with strong motivation and good communication skills will be preferably considered.

Initial appointment is 1 year and can be renewed annually. Expected starting date is January 1, 2009 or earlier. Please submit application including CV and 3 names of references by email to **Prof. Dr. Achim Kramer, Laboratory of Chronobiology, Charité – Universitätsmedizin Berlin, Germany, email: achim.kramer@charite.de.**

W173116R

GEORG SPEYER HAUS

Institute for Biomedical Research, Frankfurt Main, Germany

The Georg-Speyer-Haus in Frankfurt am Main, Germany is an academic research institute of great tradition in the biomedical sciences. The institute maintains the status of an independent foundation, and is supported by the German Federal Ministry of Health and the Ministry for Sciences and Arts of the State of Hessen. The Georg-Speyer-Haus is dedicated to basic and translational research in the fields of cancer and infectious diseases, and cooperates closely with the University of Frankfurt. We are looking for

Group Leaders (m/f)

to establish their own laboratories. We are particularly interested in the study of the autonomous functions of tumor cells regulated by oncogenes and tumor suppressor genes, the signaling molecules which mediate the intercellular communication between tumor cells and their surrounding normal cells, the role of mesenchymal and lymphoid cells in the process of metastasis formation and immune evasion and the translational implications which these studies offer. We also would like to pursue the concept of tumor stem cells and its consequences for the understanding of therapy resistance and tumor recurrence.

Successful candidates should have several years of postdoctoral experience and an excellent publication record. Physicians and basic researchers are encouraged to apply. A core group will be funded by the Institute and additional funding should be acquired through competitive grants. We are encouraging young scientists, prepared to establish their own laboratory, assuming full responsibility for their projects and coworkers, to apply.

Please direct your application including CV, list of publications, summary of scientific achievements, a brief outline of planned research and the addresses for three letters of reference to the Director of the Georg-Speyer-Haus:

Prof. Dr. Bernd Groner • Georg-Speyer-Haus
Institute for Biomedical Research
Paul-Ehrlich-Straße 42-44 • D-60596 Frankfurt am Main, Germany
Telephone: +49 69 63395 180• Fax: +49 69 63395 185
E-mail: groner@em.uni-frankfurt.de

W173117R

RUPRECHT-KARLS-
UNIVERSITÄT
HEIDELBERG

The Faculty of Bioscicences of the University of Heidelberg, Germany, invites applications for a

Professorship (W 3) for Developmental Biology

in the context of the program „Hochschule 2012“. The professorship will represent the field of Organismic Developmental Biology (in animals or plants) in research and teaching. Cooperation with research groups within the Heidelberg life science campus, in particular with the collaborative research network SFB 488 and other SFB initiatives are highly appreciated. We expect adequeate participation in the teaching of Bachelor and Masters students. Additional information on the University and the Faculty can be found at http://www.uni-heidelberg.de.

The applicant holds a Ph.D. and has an excellent international reputation well documented by publications in leading journals. Habilitation or an equivalent scientific qualification, experience in teaching developmental biology, in the successfull acquisition of competitive third-party funding and in the leadership of a research group are required.

The University of Heidelberg is an equal opportunity employer. Applications of women are strongly encouraged. Handicapped candidates with equivalent qualifications will be given preference.

Complete applications along with CV, list of publications, statement of teaching experience and a listing of third party funding as well as a brief research outlook should be submitted within six weeks after publication of this add to the **Dekanat der Fakultät für Biowissenschaften, Im Neuenheimer Feld 234, D-69120 Heidelberg. +49 6221 54 4393.**

W173106R

Positions in joint experimental and theoretical systems biology projects

The Division of Theoretical Bioinformatics at the German Cancer Research Center (DKFZ) is inviting applications from highly qualified candidates to work in interdisciplinary systems biology projects with a strong focus on medical/ therapeutic applicability. The DKFZ is a world-leading institution in the field of applied and basic cancer research. Within DKFZ, systems biology research plays a leading role and is carried out in close cooperation with BioQuant the Center for Systems Biology at the University of Heidelberg, where part of the positions will be localized.

dkfz.
GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION

1 Junior Group Leader, Ref-No. 198/2008
1 Postdoc, Ref-No. 199/2008
1 PhD-Position, Ref-No. 200/2008

(All positions are initially limited to 3 years, with possibility to prolongation)

SBCancer.

The goal is to develop and improve predictive multi-scale (spatio-)temporal models to incorporate biological mechanisms spanning different time- and length scales from protein signaling to multi-cell communication and tissue/tumor development. The focus will be on elucidating cellular decisions and cell-cell communication. Applications will range from understanding life-death decisions in apoptosis/autophagy pathways, tumor-stroma interactions in cancer, development of tumor progression markers to the understanding of wound healing processes in skin. Candidates should hold a degree/PhD in physics, applied mathematics, engineering or biological sciences, having demonstrated expertise in interdisciplinary research, systems biology, biophysics and/or nonlinear dynamics. Experienced applicants will have the opportunity to lead an established subgroup on systems biology including the supervision of Postdocs and PhD-students.

HELMHOLTZ ASSOCIATION
Alliance on Systems Biology

Initial Deadline for applications is 15.12.2008 but further applications will be assessed until positions are filled. For further information please contact Prof. Dr. Roland Eils (00496221-423600, sekreils@dkfz.de).

The German Cancer Research Center is committed to increase the representation of women in science and encourages applications from qualified female scientists. Persons with disabilities will be given preference among equally qualified candidates.

Please send your application (mentioning the reference number) including area of interest, CV and copies of diploma certificates, two references with e-mail addresses, details of previous research experience and research interests to:

Deutsches Krebsforschungszentrum, Personal- und Sozialwesen
Im Neuenheimer Feld 280, D-69120 Heidelberg
or apply online: **www.dkfz.de**

BioQuant
MODEL base of LIFE

W173057R

UNIVERSITÄTS KLINIKUM FREIBURG

The Department of Neuropathology, University Medical Center Freiburg invites applications for the position of

Research Group Leader in Neuroimmunology / Neurodegeneration

Start: At earliest convenience

We are looking for young candidates with an internationally recognized track record in one of the following fields: neuroimmunology, neurodegeneration or immunology.The laboratory uses a variety of state-of-the-art techniques in genetics and immunology to uncover the function of the innate immune system within the brain during autoimmune and bacterial inflammation, demyelination and degeneration. A special focus is on brain macrophages (microglia) during disease. For references see Mildner, Nature Neuroscience, 2007, 10(12):1544-53, and Prinz, Immunity , 2008, 28(5):675-86.The position is initially guaranteed for five years, but not limited to this period. A tenure track for junior group leaders is in preparation.An attractive start-up package will include lab space, equipment and competitive funding for personnel and running costs. We expect that the group leader will obtain additional funding agencies.The University Medical Center offers in combination with the local Max-Planck-Institutes and the Biological Faculty highly attractive research facilities in both neuroscience and immunology.Applications containing a CV, list of publications, a short research plan proposal describing current project and future plans, as well as names and phone numbers of three referees including the current employer should be send by December 15th, 2008.

This position is limited to 2 years. The compensation is according to German Regulation.

To apply for his position please send your CV to: Dept. of Neuropathology, Neurozentrum , Breisacher Straße 64, D-79106 Freiburg, Germany. For forther information, please contact Prof. Dr. Marco Prinz unter Tel. +49 761 270 5105 or mail marco.prinz@uniklink-freiburg.de.

W172074R

ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG

The Faculty of Medicine of the Albert-Ludwigs-University Freiburg, Germany, and the Centre for Chronic Immunodeficiency of the University Hospital Freiburg (CCI) invite applications for a

Professorship (W 3) for Experimental Immunodeficiency

The professorship will represent the field of experimental immunodeficiency in research and teaching. Applications in all areas of immunology will be considered, but investigators specializing in the molecular basis of immune cell development and function with relevance for the understanding of human immunodeficiency diseases are especially encouraged to apply. The professorship is part of the executive board of the CCI, an integrated research and treatment center funded by the German Federal Government. In addition, the Collaborative Research Centre SFB620 (Immunodeficiency: Clinical Manifestations and Animal Models) provides excellent opportunities for interactive immunodeficiency research.

The applicant holds an M.D. or a PhD, has an excellent international reputation and broad experience in experimental immunology. Lecturing qualification (Habilitation) or equivalent scientific credits, experience in acquisition of competitive third-party funding and in the leadership of a research group are required.

The initial appointment will be for five years. After a successful evaluation, the position will be converted into a permanent (tenured) W 3-Professorhip.

The University of Freiburg is an equal opportunity employer. Applications of women are strongly encouraged. Handicapped candidates with equivalent qualifications will be given preference.

For application forms please contact the Dean of the Medical Faculty (E-mail: dekanat-professuren@uniklinik-freiburg.de). Completed applications along with all pertinent documents should be sent no later than December 15th, 2008 to the Dean of the Medical Faculty, Prof. Dr. Ch. Peters, University of Freiburg, D-79085 Freiburg im Breisgau, Germany (Phone: ++49-761 270-7235 or 7234; Fax: ++49-761 203 7236).

W173114R

MDC Berlin-Buch

HELMHOLTZ GEMEINSCHAFT

The MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE (MDC) BERLIN-BUCH is inviting applications for the following positions:

Postdoctoral and PhD Positions in experimental or computational Biology and Bioinformatics

The MDC Berlin-Buch is a member of the Helmholtz Association of National Research Centers. It is dedicated to interdisciplinary research in the areas of Cardiovascular and Metabolic Diseases, Cancer and the Function and Dysfunction of the Nervous System. The research group of Nikolaus Rajewsky and the **Berlin Institute for Medical Systems Biology** (BIMSB) at the MDC are expanding their research in Systems Biology based on the utilization of Cutting Edge High-throughput technologies and the application of computational biology.

Successful candidates will conduct experimental and/or computational research in small RNA biology (microRNAs, piRNAs) and post-transcriptional gene regulation in in vivo model systems (C. elegans, planaria), as well as selected vertebrate systems (microRNA function in T cell differentiation). Available technologies on site include next generation sequencing (Solexa, 454, SOLID) and quantitative proteomics (SILAC).

Please find further information about the MDC in Berlin-Buch at our web site (http://www.mdc-berlin.de) and visit the sites of the Rajewsky-Group and the BIMSB. Enquiries regarding these scientific positions and perspectives should be addressed to Prof. Dr. Nikolaus Rajewsky, via email to Nadine Ewald (nadine.ewald[at]mdc-berlin.de).

Positions are funded according to the German TVöD-System (EO13; 100% for Postdocs, 50% for PhD's).

***

Applications for Postdoctoral positions should be sent by December 15th, 2008, including CV, list of publications, outline of present and planned research and other relevant material in one pdf-file to Nadine Ewald (nadine.ewald[at]mdc-berlin.de).

***

Applications for PhD positions should be directed to the Helmholtz Graduate School "Molecular Cell Biology" 19. January 2008 through http://www.mdc-berlin.de/phd-apply, where you may choose to apply for Systems Biology research groups. Join the 200-member strong international PhD students community at the MDC and benefit from a structured interdisciplinary PhD training in a supportive environment!

W173142R

Postdoctoral Position / PhD Position

Institute of Pharmacology, University of Heidelberg & Max-Planck-Institute for Heart and Lung Research, Bad Nauheim

Two positions are available to study cellular signalling pathways in cardiovascular physiology and pathology. We seek enthusiastic, highly motivated postdocs and PhD students with a background in cardiovascular biology. Experience with genetically modified mice is appreciated. The working atmosphere is international.

For project-related literature: Wirth et al., *Nature Medicine* 14:64-68, 2008; Benyo et al. *J. Clin. Invest. 115:3634-3640, 2005* and Tunaru et al. *Nature Medicine 9:352-355, 2003*.

The positions will be available immediately. The group is currently located at the University of Heidelberg and will move in late summer 2009 to the Max-Planck-Institute for Heart and Lung Research in Bad Nauheim close to Frankfurt. For details of projects and for applications (incl. statement of interest, CV and two references), please contact Prof. Dr. Stefan Offermanns, Institute of Pharmacology, University of Heidelberg, INF 366, 69120 Heidelberg, Germany, (Stefan.Offermanns@pharma.uni-heidelberg.de).

W173139R

MDC Berlin-Buch

HELMHOLTZ GEMEINSCHAFT

The MAX DELBRÜCK CENTER FOR MOLECULAR MEDICINE (MDC) BERLIN-BUCH invites applications for the international PhD Program offering up to ten

Three to Four Year Fully Funded PhD Positions

Highly motivated candidates with outstanding research potential are invited to apply to join the new international PhD-Program of the **BERLIN INSTITUTE FOR MEDICAL SYSTEMS BIOLOGY** at the MDC in Berlin-Buch. The Program supports collaborative projects between the **Center for Genomics and Systems Biology at New York University** and places major focus on: RNA Biology, DNA variability, gene regulation, computational biology.

Successful candidates benefit from a unique international interdisciplinary education in molecular, cellular biology, computational biology and application of cutting edge technologies. The PhD program includes travel allowance for extensive research and training at the Center for Genomics and Systems Biology at NYU.

For more information please visit: http://www.mdc-berlin.de/en/bimsb and http://biology.as.nyu.edu/object/facilities.gsb.

Applications should be sent by December 31st 2008, including CV, publications, research interest and other relevant material in one pdf-file to Prof. Nikolaus Rajewsky via email to Nadine Ewald: nadine.ewald[at]mdc-berlin.de.

Join the 200-member strong international PhD students community at the Helmholtz Graduate School "Molecular Cell Biology" and benefit from a structured interdisciplinary PhD training in a supportive environment, summer schools, conference attendance, annual PhD retreats, etc.!

W173144R

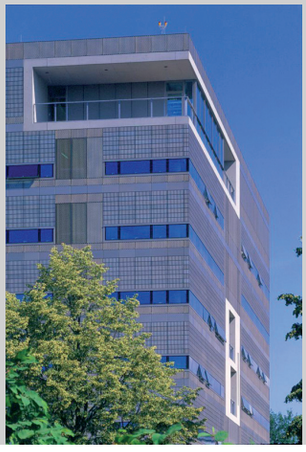
Postdoc Positions

Discovering Nature's Secrets and the Molecular Basis of Life

EXZELLENZCLUSTER
CellNetworks

The Cluster of Excellence CellNetworks, centered on world-renowned Heidelberg University, unites more than 100 leading research groups from the faculties of Medicine, Bioscience and Natural Science, DKFZ, EMBL, MPI for Medical Research and the Central Institute for Mental Health, in one common objective: to explore complex networks within and beyond the cell, to describe functional interaction in unprecedented detail and truly understand how it all fits into place.

We focus on life-science with expertise from biology, medicine, physical chemistry, biophysics, material science and advanced computation.



We offer several postdoc positions. To outstanding scientists we provide a distinguished two year position (German postdoc salary) within one of our research groups of your choice. Your application should include a CV, list of publications, a project proposal, two reference addresses, and a written agreement of the lab head you would like to work with.

The online application will be open from November 13, 2008. Please inquire at our website, find further information about the participating research groups of the cluster and contact your preferred group leader.

www.cellnetworks.uni-hd.de

W172342R

ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG

FRIAS FREIBURG INSTITUTE FOR ADVANCED STUDIES

SOFT MATTER RESEARCH

The University of Freiburg, located at the centre of the scenic city of Freiburg, is a traditional yet exciting research-led academic community. This institution offers opportunities to work with leading international academics whose ideas will shape future scientific directions.

The Freiburg Institute for Advanced Studies (FRIAS) is a university research centre funded through the German Excellence Initiative.

The FRIAS School of Soft Matter Research invites applications for

External Senior Fellowships

for the academic years 2010/11 and 2011/12

in the wider field of soft matter sciences. The School brings together scientists from experimental, theoretical or computational materials science, micro-engineering, physics, and chemistry to study different aspects of the exciting field of soft matter. The School offers the fellowships to researchers of exceptional calibre who are at an advanced stage of their academic careers. Fellows are given the opportunity to pursue experimental and/or theoretical research with new horizons in their field of competence.

The fellowships are typically granted for a period of one year. However, shorter phases of research in Freiburg distributed over a period of up to two years are also possible. The successful candidates are either provided with supplementary financial support to already existing sabbatical salaries, or full salaries at the German professorial level. Additional research funds, including postdoctoral support for the research program, may also be negotiated.

Interactions with researchers of the existing Soft Matter departments of Freiburg University, in particular with the department of microsystems engineering (IMTEK), the Freiburg materials research centre (FMF), and the departments of physics and chemistry, are expected. A documented interest in conducting research with a strong interdisciplinary focus is a further prerequisite.

Applications should include a motivation letter, CV, list of publications, and a research proposal (about 3 pages, with a strong focus on the future research plans and interactions). Applications in PDF format should be sent via email to Dr. Britta Küst (britta.kust@frias.uni-freiburg.de), who will also provide further information on the fellowships.

W172823R

University of Saarland School of Medicine

GK Ca^{2+} signaling and cellular nanodomains

PhD-Program Ca^{2+}-signaling and cellular nanodomains offers 12 PhD-Studentships

Central to the program are functional analyses of Ca^{2+}-transport proteins as well as Ca^{2+}-regulated processes focusing on one of the most attractive fields of research in life sciences. Cutting-edge research projects are combined with an intensive course program to train highly-motivated students in a stimulating environment together with the Sonderforschungsbereich 530 'Spatial and temporal interactions of cellular signaling molecules'.

The program starts on **May 1st 2009** with a research-oriented training period (lasting 10 weeks) that addresses current topics in cellular and molecular biology, biochemistry and neuroscience as well as state-of-the-art techniques, emphasizing 'hands-on' research experiences. Candidates with excellent results during the training qualify for admission to a 2.5-year PhD-project in one of the participating research groups. Participation in the training segment is certified. The program will support students during the training period with respect to costs for accommodation in a student dormitory, living and travel if the demand for financial aid is justified.

Applying

- Applicants must hold a Master's degree (or Diploma or second exam) in biosciences, pharmacy, physics, (bio-)chemistry or related fields
- Admission is only possible to the first segment, not directly to the PhD program
- Proficiency in the English language is required
- Application must be received by **January 07th 2009**
- For detailed information about the scientific program, the application and admission please visit our webpage at **http://www.uks.eu/gk-calcium**

W172919R

NATURE|Vol 456|13 November 2008

Quantum erat demonstrandum

Surfing the treacherous data wave.

C. N. Simms

The shrill ring of a phone pierces the slam and rattle of the Thursday afternoon table-football game. Last ball. I'm losing 5–4. The head of compliance is standing ready to pounce, her lines of men angled threateningly. I reach for the phone to delay the inevitable, and for a brief moment the office recaptures its mechanical silence.

"Simon Erat speaking," I answer.

"Mr Erat, this is Frank Hauptmann. It's good to speak to you again. Can you talk?"

"Of course. All our lines are secure."

"I need something brought from Tokyo to Zurich."

"You do remember that I'm based in London? It'd be cheaper for you to contact the Tokyo office direct."

"Commendably honest Mr Erat, but we want you to carry the data."

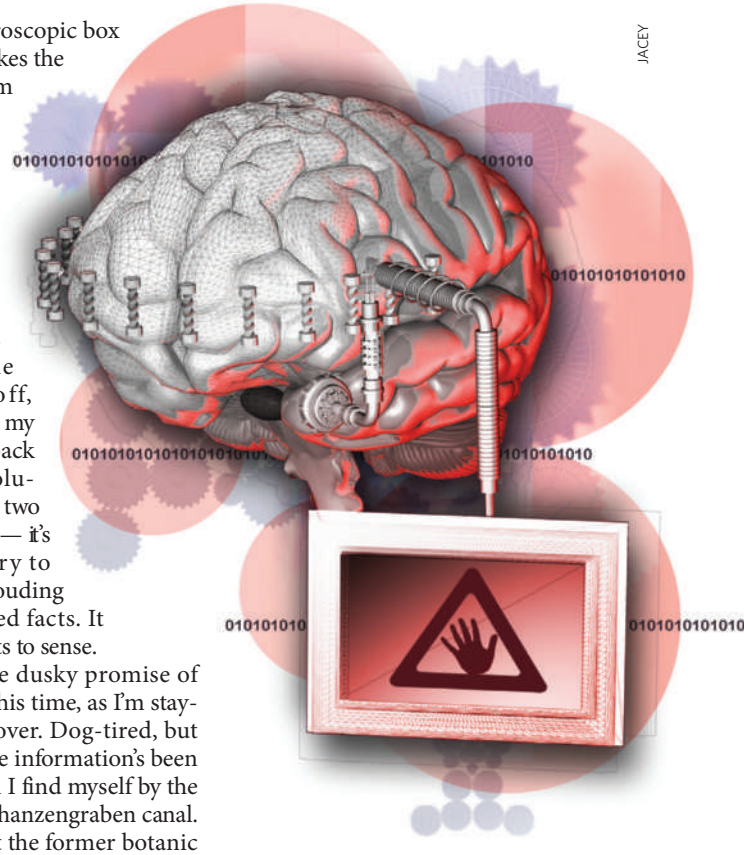
Four hours later I'm on the first BA flight to Narita. As there are no films I haven't seen, I while away the time before the first meal by playing chess against my entertainment screen. It doesn't stand a chance against a quantum mind.

Tokyo. A black Lexus whisks me from the airport to the skyscraper that is my collection point. On most trips, I'd get the Skyliner or the Narita Express, but I wasn't going to object to the red-carpet treatment. I'm only carrying a small bag, but Tokyo is not somewhere you want to be encumbered by luggage. After three intense hours in the basement plugged into a sleek, offline computer, half-sensing ethereal streams of glowing data points pouring in, the same car takes me back to the airport. Same car, but I never see the driver.

Dutifully packed onto the next plane, I loosen my tie, sink back into the seat and taste the Châteauneuf-du-Pape that the friendly German air stewardess hands me. Perhaps Swiss. I'm not good at accents, particularly with my head heavy with data. Air France this time. You can usually rely on them for a decent wine in business class. I've barely finished the drink before I pass out with exhaustion and I'm gently awoken to Zurich by the same stewardess.

New York wants to suck you in, London wants to spit you out and Paris wants to seduce you, but I like clean, efficient Zurich. The people are friendly, and the place doesn't seem to care whether I'm there, which suits me. I shake the sleep from my head, grab my small bag and I'm down the steps to the next car.

Retrieval is as painful as always. The microchip-sized gyroscopic box in my frontal lobe takes the energy it needs from my neural pathways as it does its trillions of calculations a second. Data exit is a cruel, wrenching twist of the beautiful stream that entered. Once it's done and the information is deleted, I flick the mental switch to off, unplug the jack from my temple and drop it back in its sterilizing solution. The past day or two have become a haze — it's almost painful to try to pierce the cloud shrouding the half-remembered facts. It hurts to think. It hurts to sense.



I step out into the dusky promise of an evening. No car this time, as I'm staying the night to recover. Dog-tired, but light-headed now the information's been lifted, I wander until I find myself by the dark waters of the Schanzengraben canal. Heading north, past the former botanic gardens, I glimpse an eerily lit bar above a pool. Crammed full, the bar's laughter travels to me on the breeze. Déjà vu. Did I come here the last time I was in Zurich? Rimini? Is that what it was called? Maybe. Threading my way between pot plants and people, I tread the faux-oriental mats and get a Bloody Mary and a raw kebab.

A girl bumps into me, laughing. She sees my predicament, grabs my arm and leads me towards one of the many grills. Was she here last time? For safety's sake, I should probably be in a hotel room by now — transfer leaves me like a lost puppy. And yet, I stay.

In no time, I'm absorbed into her group of friends and I attempt broken German and shattered French before they take pity and apologize for their expert English. The night drifts into an unlikely-to-be-remembered blur of noise, light and excitement. I can rarely lay down new memories immediately after data exit.

I awake in a hotel room. Mine? My bag is on a chair, unopened. I'm undressed and there's a note scrawled on the hotel paper on my pillow.

Ich hatte eine wundervolle Zeit.
Vielen Dank für alles. x

The phone is ringing.

"Yea …*Ja*?" I enquire, stealing a glance at the time. 7:21.

"Good morning Mr Erat," Hauptmann purrs. "I'd just like to thank you for a thoroughly professional and faultless service. *Quod erat demonstrandum* eh? Or rather *Quod erat faciendum*?" He laughs. "The balance will be transferred by midday."

"Always nice to hear," I smile back, the relief obvious to me, even if not to him.

Further words catch in my throat as I see the Air France hat on the bedside table.

"Until next time," I close off.

Anger and fear struggle for superiority. Thanks for everything? I flick to quantum. Of course there are no data — nothing to steal. Then they'd hired her to watch me. But why leave the hat?

"Thoroughly professional," I repeat under my breath with a hollow laugh. How had Hauptmann known where I'd go? Suddenly the idea of someone plugging a cord into my brain felt like violation.

Now Zurich had noticed me, it didn't feel quite the same. ■

C. N. Simms is a scientific subeditor with an unhealthy level of imagination.