nature

www.nature.com/nature

Vol 461 | Issue no. 7264 | 1 October 2009

Back on the map

Central and eastern European nations still lag behind Western countries in science. But they are slowly catching up.

Twenty years after the end of Soviet communism, many of the former satellite states in central and eastern Europe have joined the West as members of the European Union (EU). Yet by many measures, the science being done in those states still lags behind. Not only do their overall public and private scientific expenditures tend to be lower than those of their EU partners, as are their levels of participation in EU-funded research collaborations, but their very academic structures also remain a barrier to international competitiveness.

This is not to say that nothing has happened since 1989 — a year in which the fall of the Berlin Wall was paralleled by the near-collapse of academic and industrial science throughout the former Eastern bloc. The former German Democratic Republic, for example, has now caught up with the rest of western Europe owing to massive investment after the reunification of Germany in 1990. And the small Baltic country of Estonia, which spent that same period overhauling its Soviet-style academic structures and introducing a modern funding system based on grants, publications and peer review, has become a poster child for successful transition to Western-style science. Today it boasts considerable strengths in material, biomedical and environmental technologies.

But elsewhere, the academic hierarchy has often proved more resistant to change. Granted, countries including Hungary, Slovenia, the Czech Republic and Poland have seen the emergence of excellent research groups and institutes, many of them led by scientists who graduated around 1989 and were quick to grasp the opportunity to leave and gather experience abroad (see page 586). The International Institute of Molecular and Cell Biology in Warsaw, set up in 1999 with support from the United Nations Educational, Scientific and Cultural Organization (UNESCO), is a prominent example.

With the exception of Estonia, however, basic research in the region is still controlled by the national academies of science: government-funded organizations that independently operate numerous research institutes. Most of these academies have undergone evaluations of their respective strengths and weaknesses over the past decade, with the result that some of the least-productive institutes have been closed or restructured. And the knee-jerk opposition to the 'Westernization' of science is rapidly declining, as is political resentment and anti-Western ideology in general. Still, not all of these academy institutes are as well connected to the outside world of science as they should be. Young scientists there have good reason to complain about academic hierarchies whose existence is unjustified by scientific merit, and about funding channels that fail to reward the best research proposals.

"Science has benefited greatly from the flow of talent from eastern Europe and Russia."

Addressing their complaints will require the reinforcement of merit-based science in central and eastern Europe, which in turn will require that scarce resources are focused on existing and emerging strengths. The current economic crisis must not lead to a lasting decline in funding for science and higher education, either in Hungary, which has been hit particularly hard, or elsewhere in the region. The various national labs and institutions should reward mobility more than they have done in the past, as any upswing in their own scientific establishments will require that as many students and young scientists as possible gain experience in the best labs abroad. Finally, the region would gain substantial prestige and visibility if a large multinational research facility, such as the planned European Spallation Source, were to be built in one of the new EU member states.

Globally, science has benefited greatly from the flow of talent from eastern Europe and Russia over the past 20 years. That brain drain has not made the transition at home any easier. Although it has taken more time than anticipated to put central and eastern Europe back on the global map of science, the upcoming generation of young, energetic students and scientists should be able to complete the process. It would be to everybody's gain. The heart of Europe deserves good science, but the rest of the world needs good science from this culturally rich region just as much. ■

Containing risk

The US Department of Homeland Security should not be put in charge of biodefence research.

The past few months have brought a rich harvest of high-level reports on US biosecurity, with a particular focus on the nearly 1,400 American labs known to be doing research on human, plant and animal pathogens risky enough to require special containment. The reports were largely inspired by the announcement from the US Federal Bureau of Investigation in August 2008 that Bruce Ivins, a senior researcher at the US Army Medical Research Institute of Infectious Diseases in Fort Detrick, Maryland, was considered to be the "sole culprit" in the deadly 2001 anthrax attacks. Several of the reports — two of them issued last week — recommend that a single agency be put in overall charge of what is now a fragmented oversight system involving at least a dozen agencies. Their laudable goal is to tighten the safeguards against biological threats, including any that might emanate from within the labs, without stifling research on countermeasures against any future attacks.

As always, however, the devil lies in the details. Witness the Weapons of Mass Destruction Prevention and Preparedness Act of 2009 (see page 577), a bill introduced in the Senate last month by the two

top lawmakers on the Senate's Committee on Homeland Security, Joseph Lieberman (Independent, Connecticut) and Susan Collins (Republican, Maine). Perhaps not surprisingly given their committee's purview, Lieberman and Collins's bill would vest authority for regulation of the biodefence research complex in the Department of Homeland Security (DHS). But the DHS, a hotchpotch assembled from 22 existing agencies in 2003, combines a sprawling mandate with a paucity of biological expertise. It stands in stark contrast to the US Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, which has overseen labs working with a predetermined list of deadly human pathogens since 1997. (Animal and plant pathogens are handled by the US Department of Agriculture.) The Lieberman–Collins bill has accordingly alarmed scientists in this field, who say that their experience to date in executing contracts for the DHS has revealed a disturbing lack of expertise in its inspectors.

Opponents of the bill are also concerned by its designation of a new category of 'Tier I' agents for extra-tight regulation, including new levels of scrutiny to establish "personnel reliability". Lieberman said last week that only about eight microbes would fit this category. But the bill allows the secretary of homeland security to move any human pathogen into Tier I at any time, without having to provide a reason in public. This contrasts with the checks and balances — including consultation with scientists and public comment and discussion — now required to add a microorganism to the list of those overseen by the CDC or the agriculture department.

The fear is that the new Tier I requirements, and the costs they imply, would discourage labs from working on such pathogens. There is also the possibility that individual scientists will decide to choose microorganisms — or fields — more congenial to personal privacy. Already, researchers who work with any of the current 82 agents on the CDC's list must submit to a Department of Justice background check that involves fingerprinting and access to personal financial and medical records. At a recent American Society of Virology meeting in Vancouver, Canada, informal talk was about research being driven into Canada or the European Union, and the risk to careers associated with listed-agent work if labs begin to walk away from its costs.

Lieberman and Collins undoubtedly have good intentions with this bill. But they would be well advised to consult more carefully with front-line researchers before finalizing legislation in which so much is at stake. They should particularly consider the long experience and expertise resident in the CDC before making such a dramatic shift of lab oversight to an untried department. ■

Delimiting death

Procuring organs for transplant demands a realistic definition of life's end.

Prompted by the increasing practice of organ transplantation, and thus the need to procure donor organs that are as fresh as possible, many countries have modelled their legal definition of death on a US law passed in 1981 after extensive debate and thoughtful input from a specially appointed president's commission of experts.

The law seems admirably straightforward: "An individual who has sustained either (1) irreversible cessation of circulatory and respiratory functions, or (2) irreversible cessation of all functions of the entire brain, including the brain stem, is dead."

In practice, unfortunately, physicians know that when they declare that someone on life support is dead, they are usually obeying the spirit, but not the letter, of this law. And many are feeling increasingly uncomfortable about it.

In particular, they struggle with three of the law's phrases: 'irreversible', 'all functions' and 'entire brain', knowing that they cannot guarantee full compliance. They do know that when they declare a death — according to strict clinical criteria, the principles of which are outlined in the original report of the president's commission — that the person is to all intents and purposes dead. But what if, as is sometimes the case, blood chemistry suggests that the pituitary gland at the base of the brain is still functioning? That activity has nothing to do with a person being alive in any meaningful sense. But it undermines a claim that all functions of the entire brain have ceased. As do post-mortem observations that relatively large areas of tissue can be metabolically active in different brain areas at the time death is declared.

The criterion of irreversibility raises the question of how long one should wait to be sure that no function will re-emerge. Is the six hours recommended in the commission's report sufficient? Physicians who have been required by circumstance to wait much longer have occasionally observed a brainstem-mediated reflex — a cough, for example — up to 36 hours after they would have declared death.

The problem is that death is not a phase transition whereby a person stops being alive and becomes dead in an instant. It is a long process during which systems, networks and cells gradually disintegrate. At some point, the person is no longer there, and can never be made to return. But the kind of clear, unambiguous boundary assumed in the 1981 law simply does not exist.

Ideally, the law should be changed to describe more accurately and honestly the way that death is determined in clinical practice. Most doctors have hesitated to say so too loudly, lest they be caricatured in public as greedy harvesters eager to strip living patients of their organs. But their public silence was broken on 24 September at an international meeting that included physicians, transplant surgeons and bioethicists at the Italian Festival of Health in Viareggio. The meeting concluded that lawmakers in the United States and elsewhere should reconsider rigid definitions of death, and called for a wider public debate.

The time has come for a serious discussion on redrafting laws that push doctors towards a form of deceit. But care must be taken to ensure that it doesn't backfire. Learning that the law has not been strictly adhered to could easily discourage organ donation at a time when demand for organs already vastly exceeds supply. Physicians and others involved in the issue would be wise to investigate just how incendiary the theme might be, perhaps in contained focus groups, and design their strategy accordingly.

Few things are as sensitive as death. But concerns about the legal details of declaring death in someone who will never again be the person he or she was should be weighed against the value of giving a full and healthy life to someone who will die without a transplant. ■

NEWS BRIEFING

POLICY

Merkel wins: The **German general election** on 27 September returned Chancellor Angela Merkel to power. She aims to strike a new coalition with the Free Democratic Party, which has promised a liberal approach to contentious research issues such as human embryonic stem cells and genetically modified crops (see *Nature* **461,** 456–457; 2009). Merkel is also expected to review the country's plans to phase out its nuclear power stations over the next decade.

Nuclear curbs: The **United Nations Security Council** unanimously backed a non-binding resolution to bolster efforts aimed at slowing the spread of nuclear weapons. It endorsed strengthening the Nuclear Non-Proliferation Treaty, ratifying the Comprehensive Nuclear-Test-Ban Treaty (which would outlaw nuclear bomb tests) and improving nuclear security.

Emissions reporting: The US **Environmental Protection Agency** has announced its nationwide reporting system for greenhouse-gas emissions. Large facilities will have to disclose their emissions every year, as part of a programme that the agency said should cover 85% of US emissions. Data collection begins in January 2010, with disclosure for that year happening in 2011.

Stem-cell oversight: The **International Society for Stem Cell Research** has created a committee to weed out companies that offer unapproved stem-cell 'therapies'. The society's president, Irving Weissman of Stanford University in Palo Alto, California, launched the committee on 22 September at the World Stem Cell Summit in Baltimore, Maryland. The 18-member panel plans to create a **blacklist** of companies that don't provide documentation showing that their treatments have been reported in peer-reviewed literature, have been overseen by the institutional review board and have received regulatory approval. "That's the minimum beginning," says Weissman.

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

Pesticide review: A review committee assessing the safety of **methyl iodide** — a pesticide used to fumigate soils — began work on 24 September in Sacramento, California. The chemical was approved for agricultural use by the US Environmental Protection Agency in October 2007, prompting protests from activists and scientists. California, Washington and New York have yet to approve the fumigant, and a 2009 report from the California Department of Pesticide Regulation (DPR) warned of significant health risks. The DPR will wait to make a ruling until after the review committee reports its findings.

For a longer version of this story, see http://tinyurl.com/yeeuo3a

Agriculture appointment: Roger Beachy, the founding president of the Donald Danforth Plant Science Center in St Louis, Missouri, was last week appointed to direct the US Department of Agriculture's **National Institute of Food and Agriculture**, a new agency that will distribute the department's external basic research funding. For more, see page 580.

Unscientific review: Political influence led the US **Food and Drug Administration** (FDA) to approve a device to repair damaged knees against the recommendations of its own scientists, the agency has admitted. Scientific reviewers had twice turned down applications for **Menaflex**, made by ReGen Biologics of Hackensack, New Jersey, but FDA officials approved it last December after lobbying from four New Jersey congressmen. This was a "clear deviation from the principles of integrity", the FDA wrote in a report analysing its review process.

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

Carbon market confusion: The European Commission exceeded its authority in imposing tighter-than-requested caps for greenhouse-gas emissions on Poland and Estonia in the second period of the **Emission Trading System**, a court ruled last week. The price of carbon credits fell following the news, because traders thought the commission might have to grant additional allowances to Poland, Estonia and to six other countries that

BUDGET PROTEST

Thousands of students, professors and other staff at the ten University of California campuses staged mass protests on 24 September, objecting to deep budget cuts that have led to enforced periods of unpaid leave, higher fees and job losses. More than US$800 million has been sliced from $3.2 billion in state funds expected for the 2009–10 fiscal year, delaying the construction of new research facilities and halting recruitment efforts.

R. GALBRAITH/REUTERS

NUMBER CRUNCH

4°C

The 'increasingly plausible' rise in global warming by 2060 if greenhouse-gas emissions are not curbed quickly, according to Richard Betts, presenting a UK study from the Met Office's Hadley Centre.

are also appealing against their caps. The ruling, which is subject to appeal, is limited to 2008–12 emissions, however — and its long-term effect on European carbon emissions is unclear.

J. BRINON/AP

EVENTS

Uranium secret uncovered: **Iran** has admitted to having a second uranium-enrichment plant. On 25 September, the US, French and British premiers held a press conference to announce their intelligence on the site. Iran had informed the International Atomic Energy Agency in Vienna about the underground facility, near the city of Qom, four days earlier. The plant, not yet in operation, is thought to be big enough to hold around 3,000 gas centrifuges, which could enrich enough uranium for one nuclear weapon in a year.

Plagiarism allegation: Two **Iranian government ministers**, including Kamran Daneshjou, minister for science and education, have been accused of co-authoring plagiarized academic papers. Several of the papers in question have been retracted, and an Iranian parliamentary commission is considering an investigation. For more, see page 578.

RESEARCH

HIV vaccine: An experimental HIV vaccine has shown **moderate success** at preventing infection by the virus. A US$119-million study involving more than 16,000 HIV-negative men and women from Thailand found that a combination of two older drugs, which had failed to work individually, together reduced the risk of contracting HIV by nearly a third. "It's the largest step forward that's ever occurred in the HIV-vaccine field," says Dan Barouch of the Beth Israel Deaconess Medical Center in Boston, Massachusetts.
For a longer version of this story, see http://tinyurl.com/vachiv

Mars delay: The launch of **Russia's** Phobos-Grunt mission to study Mars and collect soil samples from one of its moons has been postponed to 2011, together with **China's** first Mars probe, the orbiter Yinghuo-1. Both crafts were supposed to take off in October this year, carried on a Russian rocket, but Russian space agency Roscosmos said last week that Phobos testing couldn't be completed in time to meet this year's launch window.

High-risk awards: The US **National Institutes of Health** announced 115 new awards for high-risk research on 24 September. The grants, which total US$348 million over five years, come in three varieties: Pioneer Awards for creative researchers at any career stage; New Innovator Awards for early-stage investigators; and a new category, Transformative R01 (T-R01) Awards to support unconventional projects. Improving on previous gender disparities, women researchers made up more than a third of Pioneer and New Innovator awardees — well above the 2008 R01 average of 26%. But only 15.5% of T-R01 awardees were women. The gender ratios of applicants were not available.

BUSINESS

Drug deal: The drug company **Abbott Laboratories** in Abbott Park, Illinois, is to buy **Solvay Group**'s pharmaceutical business for €4.5 billion (US$6.6 billion). The company will pay cash for the deal, which includes Solvay's vaccines business, based in Belgium. The acquisition will also see Abbott increase its annual US$2.7-billion pharmaceutical research and development investment by $500 million.

Climate rift: **Exelon** of Chicago, Illinois, the largest US nuclear-power provider, has become the latest company to quit the US **Chamber of Commerce** because of differences over climate change. In recent months the chamber, which is among the nation's most influential business groups, has opposed a government plan to regulate carbon dioxide emissions. Exelon's 28 September announcement that it would not renew its membership of the chamber came a few days after similar withdrawals by two other utilities: the Public Service Company of New Mexico and California's Pacific Gas and Electric.

NEWS MAKER

Irina Bokova

The Bulgarian diplomat is the first woman to be elected director-general of the United Nations Educational, Scientific and Cultural Organization (UNESCO).

THE WEEK AHEAD

5–7 OCTOBER
The 2009 Nobel prizes for physiology or medicine, physics and chemistry are announced.
http://nobelprize.org

5–7 OCTOBER
Singapore hosts the Stem Cells and Regenerative Medicine Congress Asia.
www.terrapinn.com/2009/stemcellsasia

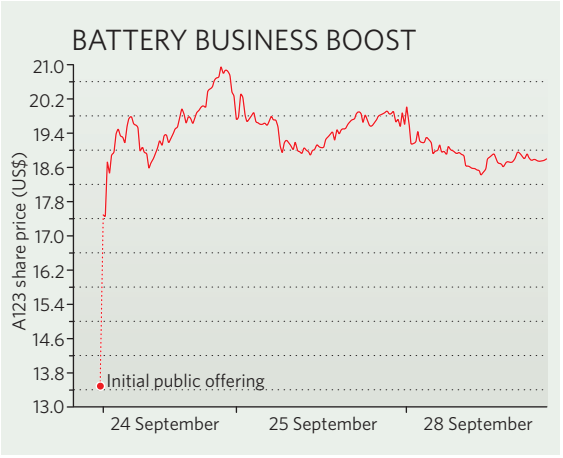
8–11 OCTOBER
The American Association for Cancer Research holds its 'Frontiers in basic cancer research' conference in Boston, Massachusetts.
http://tinyurl.com/aacrbost

BUSINESS WATCH

A university spin-off company has got battery investors buzzing. A123Systems, which makes rechargeable lithium-ion batteries, raised US$380 million at $13.50 a share at its initial public offering, before watching its share price rocket (see chart).

The cash boost, together with earlier federal grants and private investments, made A123 a billion-dollar company, even though it has yet to turn a profit. It has also applied for a $1.8-billion loan from the US government to build a mass-production facility in Michigan, and has signed a deal with US car manufacturer Chrysler.

The company, based in Waterford, Massachusetts, was founded in 2001 by materials scientist Yet-Ming Chiang and his colleagues from the Massachusetts Institute of Technology in Cambridge. It uses lithium iron phosphate in the cathodes, and is currently in a patent dispute with the University of Texas in Austin, where the material was developed. The compound is more stable than the lithium metal oxides used in today's laptops and mobile phones, but cannot store as much charge.
For a longer version of this story, see http://tinyurl.com/Libatteries



SOURCE: NASDAQ

NEWS

Booming biosafety labs probed

US lawmakers and scientists at odds over how to regulate high-containment labs.

Researchers are pushing back against moves in the US Congress to regulate more strictly the laboratories that handle the most dangerous pathogens. The scientists argue that adding rules to those already implemented since the anthrax attacks of 2001 could hobble work on countermeasures for killer pathogens, and drive researchers from the field.

"It took a number of years and substantial effort to arrive at the careful equilibrium that currently exists," Ronald Atlas, a past president of the American Society for Microbiology, told a panel at the House of Representatives on 22 September. "Precipitous, excessive policy changes could upset this delicate balance."

Days earlier, three top officials from the Galveston National Laboratory at the University of Texas Medical Branch wrote to express "grave concerns" about a bill introduced last month by senators Joseph Lieberman (Independent, Connecticut) and Susan Collins (Republican, Maine). The bill would transfer authority over work on human pathogens at high-biosecurity-level labs — Galveston is one of seven labs nationwide at the highest, biosafety level (BSL) 4 — from the Department of Health and Human Services to the Department of Homeland Security (DHS).

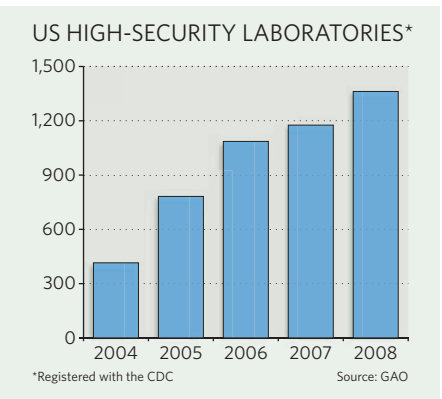
Members of Congress have raised concerns about oversight of the growing complex of biosafety labs (see 'Security measures'). Those concerns heightened after August 2008, when the Federal Bureau of Investigation identified Bruce Ivins, a scientist at the US Army Research Institute for Infectious Diseases at Fort Detrick, Maryland, as the "sole culprit" behind the anthrax attacks. Ivins had killed himself days earlier.

J. GATHANY/CDC

Legislators want to bring labs under the authority of the Department of Homeland Security.

In a report released last week, the Government Accountability Office (GAO) noted that the number of BSL-3 and BSL-4 laboratories registered with the Centers for Disease Control and Prevention (CDC) in Atlanta, Georgia, more than tripled between 2004 and 2008 (see chart). But this number doesn't include all such labs, because labs working with some dangerous pathogens, including those that can cause tuberculosis and severe acute respiratory syndrome, are not required to register.

"No federal agency knows how many such laboratories exist," Nancy Kingsbury, director for applied research and methods at the GAO, told a Senate hearing. Hence no one knows whether the current capacity is what is needed. Kingsbury recommended that the United States centralize oversight and regulation of its high-risk laboratories, as the United Kingdom did following a foot-and-mouth outbreak in 2007.



Any such centralization would be largely in the hands of Congress. Lieberman says he is "very sensitive" to scientists' concerns about restrictive regulations. "The whole thrust here is to not get in the way of research," he told *Nature* after a Senate homeland security committee hearing last week. "We want it to go forward. But we want it to go forward while we set up systems that will ferret out wrongdoers." He told the hearing that he and Collins will push forwards the bill this month.

But many take issue with the bill's proposal to consolidate lab oversight in the DHS. Jean Patterson, who heads a BSL-4 laboratory at the Southwest Foundation for Biomedical Research in San Antonio, Texas, recently decided to walk away from a US$1.2-million DHS grant after calculating that the costs of a DHS-imposed compliance review would outstrip the amount of the award. In addition, she says, during a site inspection visit, "the DHS reviewers were not as qualified as they needed to be. We would like only trained people to be inspecting us," she says, "and at this point, the CDC has that expertise."

Researchers and Congress are awaiting a new review on the topic, ordered by then-President George W. Bush and co-chaired by the secretaries of defence and health and human services. It is expected within weeks. ■

Meredith Wadman

For a Q&A with a researcher who stopped working with dangerous microbes because of regulations, see http://tinyurl.com/yhspcve

Security measures

The US Government Accountability Office (GAO) last week scolded the Centers for Disease Control and Prevention (CDC) for failing to insist on security improvements at two of the nation's biosafety level (BSL) 4 labs, which work with the most dangerous pathogens.

A year ago, the GAO found perimeter security seriously lacking at the two labs on the basis of 15 'key controls', such as the visible presence of armed guards and camera monitoring. Since then, Gregory Kutz, director of forensic audits and special investigations at the GAO, told a Senate hearing last week, the labs have made some improvements on their own. But both still had fewer than half of the 15 measures in place, according to a new GAO report.

"The CDC has taken limited action," Kutz said.

"There are significant differences in perimeter security" between BSL-4 labs, the CDC wrote in response to the first GAO report.

"This is because there are significant differences in the risks present at each [lab], which vary not only by physical plant and location, but by the agents possessed and how those agents are used and stored." **M.W.**

Iranian ministers in plagiarism row

Nature investigation reveals duplications in papers by government's science and transport chiefs.

EXCLUSIVE

Two Iranian government ministers have co-authored peer-reviewed papers that duplicate substantial amounts of text from previously published articles, according to an investigation by *Nature*.

Three journals have already confirmed that they will retract papers co-authored by Iran's science and education minister Kamran Daneshjou, a professor in the school of mechanical engineering at the Iran University of Science & Technology (IUST) in Tehran. Before being appointed science minister in early September, Daneshjou was also head of the interior ministry office overseeing the disputed presidential elections in June that kept Mahmoud Ahmadinejad in power. A further publication by Iran's transport minister and his deputy has also been called into question.

In an online story last week (see *Nature* doi:10.1038/news.2009.945; 2009), *Nature* revealed that substantial sections of text in a 2009 paper in the journal *Engineering with Computers*[1] by Daneshjou and IUST colleague Majid Shahravi were identical to a 2002 paper[2] by South Korean scientists in the *Journal of Physics D*. New York-based Springer, which publishes *Engineering with Computers*, has told *Nature* that it will retract the paper. The work also duplicates smaller amounts of material from papers given by other researchers at conferences[3–6], as well as a 2006 article[7] in the *International Journal of Impact Engineering*.

A. TAHERKENAREH/EPA/CORBIS

Kamran Daneshjou faces plagiarism questions.

Similar duplications also appear in other papers by the same co-authors in Springer's *Journal of Mechanical Science and Technology* (*JMST*)[8], the Taiwanese *Journal of Mechanics*[9] and the Iranian journal *Mechanical & Aerospace Engineering Journal*[10] — the text of which is almost identical to that in their paper in *Engineering with Computers*[1].

Mark de Jongh, publishing editor of the *JMST*, says that the journal intends to retract the second Springer article[8]. "I conclude that the paper in question contains about 50% identical content as the formerly published

Experts draw up ocean-drilling wish list

BREMEN

Earth scientists have laid the groundwork for the future of ocean drilling. More than 500 scientists — almost twice as many as organizers had initially expected — gathered last week in Bremen, Germany, to discuss priorities and research goals for the second phase of the Integrated Ocean Drilling Program (IODP), which is expected to begin in late 2013.

Since ocean drilling began in the 1960s, sediment and rock cores retrieved from the seabed have provided information about everything from plate tectonics to Earth's climate history. Much more remains to be discovered, scientists said at the meeting.

"We're not done," says Alan Mix, a marine geologist at Oregon State University in Corvallis. "Actually, we ain't seen nothing yet."

Researchers have generated a detailed wish list for new ocean-drilling projects, which will be boiled down into a science plan for the 2013–23 period by a group yet to be appointed. The finalized science plan will then be forwarded to funding agencies in Japan, Europe and North America, which currently support the IODP to the tune of around US$200 million per year.

> "We need to explain why what we are doing matters."

Mix says that targets might include the role of greenhouse gases in transitions between cold and warm climates, and the magnitude, speed and locations of resulting sea-level changes. A more ambitious project would be to relaunch the effort to drill through Earth's crust and into its mantle. A 1960s attempt to drill through the sea floor into this boundary, known as the Mohorovičić discontinuity, or 'Moho', failed.

The Japanese IODP vessel, the *Chikyu*, is already outfitted with technology to drill down some 7,000 metres into the crust, and there are plans to refit the vessel over the next three years with drilling and core-recovery technology to allow it to drill even deeper.

"Japan will lead the Moho project," says Asahiko Taira, the Yokohama-based executive director of the Japan Agency for Marine-Earth Science and Technology, which oversees Japan's ocean-drilling operations. "It's a classic geological quest, and definitely one of our prime targets."

"The journey down is equally important to the things we may find at the bottom," adds Benoît Ildefonse, a geologist at France's National Centre for Scientific Research (CNRS) based in Montpellier, France. "Recovering rocks from near Moho will, for the first time, allow us to test our ideas and models about how the crust forms."

The deepest sea-floor holes drilled so far include a 2,111-metre-deep hole drilled during the 1970s and 1980s off Nicaragua, and a 1,500-metre-deep hole drilled in 2005 in the Cocos plate in the eastern Pacific Ocean. The latter, performed by the US vessel *JOIDES Resolution*, was the first continuous retrieval of core from Earth's upper crust. Following a $130-million refurbishment, the 30-year-old ship is now capable of drilling 2,000 metres into the sea floor in waters as deep as 7,000 metres.

IODP leaders say they are increasingly aware of the need to explain the societal relevance of

MOUNTAINS SPEED EVOLUTION
Tectonics seem to fuel the appearance of new species.
www.nature.com/news

article" in the *Journal of Physics D*[2], he says.

The *Journal of Mechanics* intends to take similar action. "We have just finished the investigation of this serious case and the result showed that the paper by K. Daneshjou and M. Shahravi indeed plagiarized other works," says Yi-Chung Shu, executive editor of the journal and a researcher in mechanical engineering at the National Taiwan University in Taipei. "This paper will be definitely retracted."

Senior Iranian scientists have called for an inquiry into the affair, and an Iranian parliamentary commission is considering an investigation.

"This is a bitter blow to Iranian academic society, it's a scandal," says Ali Gorji, an Iranian neuroscientist based at the University of Münster in Germany.

"This is a bitter blow to Iranian academic society, it's a scandal."

Transport dispute

Nature's investigation has also revealed that a paper[11] co-authored by Hamid Behbahani, Iran's minister of roads and transportation, also contains large amounts of text from earlier articles by other researchers. The paper was co-authored by Hassan Ziari, Behbahani's deputy minister, and president of the national rail company, the Islamic Republic of Iran Railways. Both authors hold positions at the IUST, while the third author, Mohammed Khabiri, was a PhD student at the time. Behbahani was also the supervisor of Ahmadinejad's PhD in transportation engineering and planning.

Much of the text and the results of their 2006 article[11], in the journal *Transport*, is identical to sections from three earlier publications[12–14].

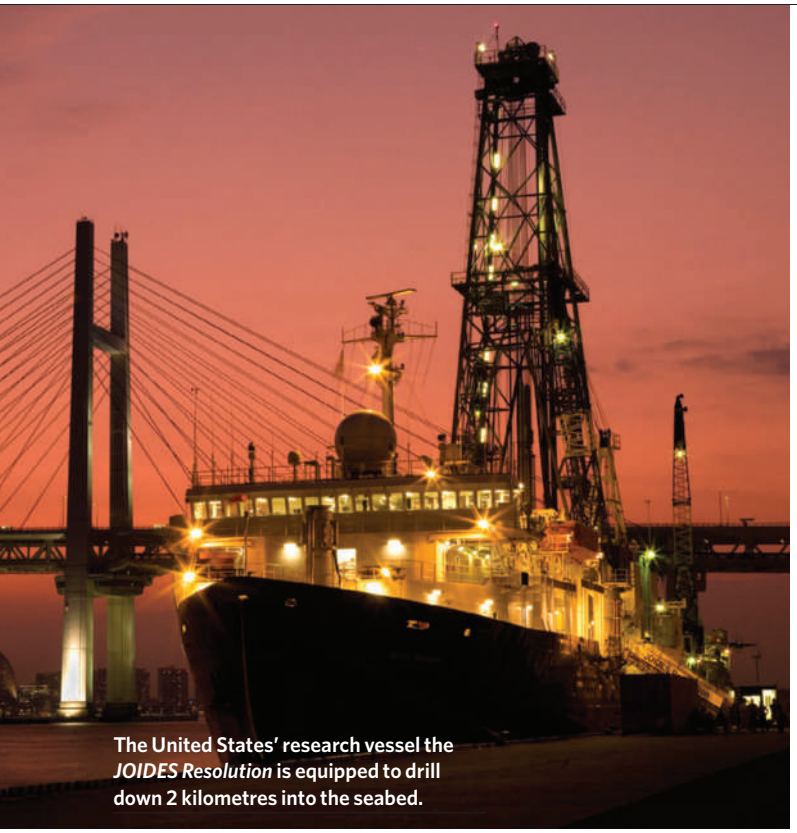
"Two of my papers were copied-and-pasted by the plagiarizing paper," says Bin Jiang, a researcher at the University of Gävle in Sweden. "This is outrageous."

"The plagiarism is obvious," concurs Jiang's co-author, Christophe Claramunt, a scientist at the Naval Academy Research Institute in Brest, France. "We look forward to appropriate action from the editor of *Transport*."

Nature has attempted, without success, to contact each of the Iranian authors of all the disputed papers. However, two Iranian news websites have published a response attributed to Majid Shahravi (see http://www.tabnak.ir/fa/pages/?cid=65586 and http://alef.ir/1388/content/view/54040). That statement refutes any plagiarism and defends the originality of the paper in *Engineering with Computers*[1], based on the fact that it had passed peer review and had cited the 2002 Korean paper[2]. ■

Declan Butler

1. Daneshjou, K. & Shahravi, M. *Eng. Comp.* **25,** 191–206 (2009).
2. Lee, W., Lee, H.-J. & Shin, H. *J. Phys. D* **35,** 2676–2686 (2002).
3. Quan, X. *et al.* in Fifth Asia–Pacific Conference on Shock and Impact Loads on Structures (eds Jones, N. & Brebbia, C. A) (WIT Press, 2003).
4. Yaziv, D., Mayseless, M. & Reifen, Y. in Proc. 19th Int. Symp. Ballistics (ed. Crewther, I. R.) (Technomic Publishing Company, 2001).
5. Sauer, S., Hiermaier, S. & Scheffer, U. Paper given at 10th International Symposium on Interaction of the Effects of Munitions with Structures, San Diego, California, 7–11 May 2001; available at http://hsrlab.gatech.edu/AUTODYN/papers/paper126.pdf.
6. Quan, X. & Birnbaum, N. in Proc. 18th Int. Symp. Ballistics (ed. Reinecke, W. G.) (Technomic Publishing Company, 1999).
7. Segletes, S. B. *Int. J. Impact Eng.* **32,** 1403–1439 (2006).
8. Daneshjou, K. & Shahravi, M. *J. Mech. Sci. Tech.* **22,** 2076–2089 (2008).
9. Daneshjou, K. & Shahravi, M. *J. Mech.* **25,** 117–128 (2009).
10. Daneshjou, K. & Shahravi, M. *Mech. Aerosp. Eng. J.* **3,** 69–86 (2008).
11. Ziari, H., Behbahani, H. & Khabiri, M. M. *Transport* **XXI,** 207–212 (2006).
12. Forbes, G. *Transportation Research Circular Issue Number: E C019* B-6/1 (Transportation Research Board ISSN, 2000).
13. Jiang, B. & Claramunt, C. Paper given at 5th AGILE Conference on Geographic Information Science, Palma, Spain, 25–27 April 2002; available at http://www2.hig.se/~bjg/s6_Jiang.pdf.
14. Jiang, B. & Claramunt, C. *GeoInformatica* **8,** 157–171 (2004).



J. BECK/IODP

The United States' research vessel the *JOIDES Resolution* is equipped to drill down 2 kilometres into the seabed.

their work. "We need to explain very well why what we are doing matters," says Catherine Mével, director of the European Consortium for Ocean Research Drilling, which coordinates the activities of 16 European IODP members and Canada. "We haven't always been able to make this clear enough in the past."

Meanwhile, China — an associate IODP member alongside South Korea, Taiwan, India, Australia and New Zealand — has announced plans to build a deep-sea drilling vessel of its own.

"The small Chinese deep-sea research community is rapidly growing in numbers, and our formerly reluctant government is increasingly convinced about the significance of ocean drilling," says Wang Pinxian, a marine geologist at Tongji University in Shanghai and vice-chair of the Chinese IODP science committee.

As an associate member with limited ship time and managerial rights, China pays the reduced IODP membership of US$1 million. The Chinese government and the Chinese Academy of Sciences are currently discussing whether China should apply for full membership beyond 2013, for which it would need to pay around $6 million per year, says Wang.

IODP rules and overall programme architecture are unlikely to undergo any substantial changes after 2013, says Rodey Batiza, section head of marine geosciences at the US National Science Foundation in Arlington, Virgina. New members will be welcome to join, but membership will not be linked to specific national drilling preferences, he says.

"We'd like to seamlessly continue drilling after 2013 with a programme designed to deliver the best science at the lowest cost," he says. "We do already have some very exciting questions that we can ask, and perhaps answer, in the next ten years." ■

Quirin Schiermeier

AGSTOCK IMAGES/CORBIS

US agriculture research gets priority plan

Federal restructuring aims to lessen the influence of pork-barrel politics.

US agricultural research is getting a makeover. The National Institute of Food and Agriculture (NIFA) opens its doors on 1 October, with plant biotechnologist Roger Beachy at its helm.

Many scientists hope that the new institute will improve the focus, effectiveness and visibility of basic agricultural research, long criticized for being driven by regional politics instead of national priorities. But it's too early to tell how big an impact NIFA will have.

"Overall, there's been a negative perception of agriculture research, especially on [Capitol] Hill," says Beachy, founding president of the non-profit Donald Danforth Plant Science Center in St Louis, Missouri. "This is an opportunity to re-evaluate the role of research and research organization."

The institute is part of a major reshuffle at the US Department of Agriculture (USDA), which in April enlisted health and agriculture specialist Rajiv Shah as the department's first chief scientist. Both Shah and his boss, agriculture secretary Tom Vilsack, have stressed the importance of basic agriculture research in addressing national and global needs, including climate change, biofuels, nutrition, and food safety and security.

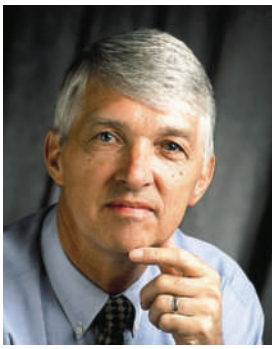
Yet Beachy's arrival also underscores the often-close ties between US agribusiness and federal research. At Washington University in St Louis in the 1980s, Beachy worked on virus resistance in plants, and later collaborated closely with Monsanto, the leading producer of genetically engineered seed, on transgenic crops. The Danforth Center was founded with grants from Monsanto's philanthropic arm among others, and the president and chief executive of Monsanto is on the centre's board of trustees. It was also Bill Danforth, the centre's founder, who originally proposed NIFA in 2004 when he was chair of a USDA task force; Congress approved it in 2008.

Ignacio Chapela, an environmental scientist at the University of California, Berkeley, worries that such close industry links might result in narrow-minded policies. "Imprinting NIFA with such a clearly biased character is likely to exacerbate the discredit of US policy-making bodies in a national and international environment," he argues.

But unless the country sets unified national goals in awarding research grants, Danforth says, agriculture will continue to lose out, as it did in the economic stimulus package passed by Congress in February. "The National Institutes of Health got $10 billion, the National Science Foundation got $3 billion and agriculture got zip," he says. "The reason in my view is sophisticated people thought the money would be wasted, and deservedly so, because too much went to pork."

Beachy says he hopes that NIFA and its competitive grants programme, the Agriculture and Food Research Initiative (AFRI), will cut back on congressional demands to earmark funds for specific projects. One big question will be their budget: in the current fiscal year, AFRI received just over $200 million, although that figure could reach $295 million for next year if numbers being worked through in congressional committees hold up. Beachy eventually hopes for $700 million annually.

Despite its ambitious goals, NIFA's ultimate impact is uncertain.



Roger Beachy, NIFA chief.

COURTESY OF THE DANFORTH CENTER

"It's hard to know what's going to happen," says animal scientist Jim Ireland at Michigan State University in East Lansing, who has called for more funding for agriculture research, particularly for farm animals. "It's good to see that they have new people, but the budget's a problem they're going to be facing regardless of how smart they are."

Animal scientist Michael Roberts, who was chief scientist of a previous USDA grant-giving body from 1998 to 2000, is also cautious. "It seems like there hasn't been a great deal of change, except for the name," he says. "This is a new organization that seems to be operating on a similar budget as the previous organization." But Beachy might have more political influence because of the structure of the new position, Roberts says: "He's a good scientist, very well respected and pretty savvy politically." ■

Lizzie Buchen

Instant climate model gears up

A climate simulator that started life in a doctoral dissertation is being adopted by negotiators to assess their national greenhouse-gas commitments ahead of December's climate summit in Copenhagen.

Dubbed C-ROADS — for Climate Rapid Overview and Decision-support Simulator — the tool translates complex climate modelling into readily digestible predictions. Using data on greenhouse-gas emissions input by country or region over a given period, the simulator projects temperatures, sea level and atmospheric carbon dioxide concentrations to 2100.

The tool hit the headlines last week when Robert Corell, chairman of the Washington-based Climate Action Initiative, made a dire assessment: even with all the international pledges to reduce greenhouse-gas emissions, including a yet-to-be-enacted commitment by the United States, by the end of this century global average temperature will still outpace the 2 °C increase targeted by the G8 countries and others. "We're headed to a 4-degree world," Corell told reporters at a press conference in Washington DC. "We don't want to go there."

C-ROADS has its origins in 1997 doctoral work by Thomas Fiddaman, now a modeller with Ventana Systems in Harvard, Massachusetts. The current version represents a collaboration between Ventana, researchers at the Massachusetts Institute of Technology (MIT) in Cambridge and the Sustainability Institute in Hartland, Vermont.

The tool made its debut in the policy world last year during a global-warming war game in Washington DC (see http://tinyurl.com/climgame), and has since been picked up by climate negotiators in the United States and Europe. Earlier this month, modellers on the C-ROADS team travelled to Beijing to train Chinese negotiators on the software.

The goal is to distribute C-ROADS to all parties so that everybody is working off of the same baseline in evaluating proposals, says Andrew Jones, who heads up the initiative at the Sustainability Institute. "We want to get a lot of the bickering over the different numbers out of the way," he says.

The team has calibrated C-ROADS against global climate models used by the Intergovernmental Panel on Climate Change (IPCC). Unlike those, it can be run on a laptop and produces instant results. Users can adjust dates and emissions from all of the major emitters and groups of developing countries, providing immediate feedback on the likely effects of any given policy commitments.

An independent team led by climatologist Robert Watson of the University of East Anglia in Norwich, UK, reviewed the model and recommended in March that the United Nations consider adopting it as a formal tool to support the climate negotiations. That hasn't happened yet, but modellers with the C-ROADS team are attending UN climate meetings to help negotiators assess policy proposals.

C-ROADS tracks historical data well and generally performs in line with average IPCC modelling results, says John Sterman, an MIT management professor who works on the project.

Sterman says a key problem with global-warming policy is the time lag between today's emissions and the problems they cause, which can be decades down the road. The model attempts to get around that by allowing policy-makers — and the public, through a simplified version on the Climate Interactive website (http://climateinteractive.org) — to see the likely consequences of their decisions immediately. "It's not that the other models are flawed," Sterman says. "They are opaque to the policy-makers."

Jones says that the model produces warming of 4.5 °C by 2100 for business as usual and 3.8 °C based on the targets announced in March. Taking into account all the latest pledges by countries, including a commitment from Russia this summer, the model's current reading is 3.5 °C.

"The doom and gloom story is getting 90% of the attention," he says, "but now we're at 3.5 °C. The global climate deal is getting better over time."

Still, global leaders made little headway on climate last week, at both a UN summit in New York and a meeting of the G20 in Pittsburgh, Pennsylvania. Leaders of the G20 did, however, agree to phase out fossil-fuel subsidies while providing "targeted support for the poorest". ■

Jeff Tollefson

For more, see nature.com/roadtocopenhagen.

Cellulosic ethanol hits roadblocks

The third of four weekly articles looks at how the financial crisis is slowing efforts to commercialize next-generation ethanol.

THE BUSINESS OF BIOFUELS

This year was supposed to be a big year for BlueFire Ethanol. The company, based in Irvine, California, had planned to start operating its first commercial cellulosic biofuel plant by the end of the year, converting waste from a neighbouring landfill into ethanol. The refinery, located near Lancaster, California, would chew up everything from cardboard to hedge clippings and produce more than 11 million litres of biofuel a year.

But after a delay due to permit issues, BlueFire found itself facing a formidable roadblock: the global economic crunch. "It put us right in the middle of the meltdown," says chief executive Arnold Klann. With more financing still needed for construction, he says, the company has now pushed the plant's start-up date to late 2010 or early 2011.

A couple of years ago, Klann would have had reason to be optimistic. Cellulosic ethanol — a biofuel that can be produced from agricultural residue, grasses and municipal waste — was touted as superior to maize (corn) ethanol because it produces fewer greenhouse-gas emissions and does not compete with food-supply needs. In February 2007, the US Department of Energy selected BlueFire and five other companies to negotiate for up to US$385 million in funding for commercial-scale plants. And later that year, Congress issued a federal mandate to produce 61 billion litres of cellulosic biofuels annually for transportation by 2022.

Today, a combination of the recession, heightened investor wariness brought on by financial losses in maize ethanol and other hurdles have slowed the commercialization rush. According to ThinkEquity, an investment bank based in San Francisco, California, the United States will have the capacity to produce less than 13 million litres of cellulosic ethanol this year, and it will almost certainly fail to meet the US Environmental Protection Agency's (EPA) projection of 381 million litres of cellulosic biofuels in 2010. Two of the six companies selected by the Department of Energy to negotiate for commercial plant funding have dropped out of the programme, and several plants belonging to other companies have been delayed. The industry has "slowed down tremendously", says Sean O'Hanlon, executive director of the American Biofuels Council in Miami, Florida.

Part of the trouble began when last year's high maize prices and falling ethanol prices led to several bankruptcies in the industry. Investors are now nervous about investing in cellulosic ethanol, says Laurence Alexander, an analyst at the investment bank Jefferies in New York. And the industry is rapidly approaching the regulatory 'blend wall', which says that petrol for conventional automobiles can contain no more than 10% ethanol by volume. With US maize-ethanol production expected to surpass 39 billion litres in 2009, and petrol consumption at 523 billion litres per year, cellulosic ethanol doesn't have much room to enter the market, says energy economist Wallace Tyner at Purdue University in West Lafayette, Indiana.

M. STRAVATO/POLARIS/EYEVINE

A joint venture between the fuel firms Verenium and BP aims to build a commercial biofuels plant.

Other countries have continued to lag behind the United States, partly because they have fewer government incentives and support programmes. The European Union (EU), which requires that 10% of transport energy come from renewable sources by 2020, has not set a specific goal for cellulosic fuel. China has restricted maize-ethanol production and aims to produce 2.5 billion additional litres of ethanol annually from non-food-grain sources by 2010. But the government has not offered specific subsidies to cellulosic-ethanol producers, says John Romankiewicz, an analyst in the Beijing office of London-based New Energy Finance. "They're just waiting for the signal from the government," he says. "And they may be waiting for a while."

Companies that were expected to lead commercialization in the United States have changed or scaled back their plans. Range Fuels, based in Broomfield, Colorado, originally planned to complete the final phase of construction on a Georgia commercial plant in 2011 but has delayed that until late 2012, says chief executive David Aldous. Iogen, based in Ottawa, Canada, suspended operations on an Idaho plant to focus its resources on a possible plant in Saskatchewan, says company spokesperson Mandy Chepeka. Iogen has applied for Canadian government funding of up to Can$200 million (US$182 million) and is still evaluating the Saskatchewan site, she says.

Industry link-ups

Some companies are hopeful that government agencies and the oil industry will see them through the downturn. Verenium, based in Cambridge, Massachusetts, secured US$90 million from oil company BP in August 2008, and the two companies announced a joint venture this year. The joint venture has applied for a loan guarantee from the US energy department to build a commercial-scale plant in Florida, which is scheduled to begin construction in 2010 and will produce 136 million litres of ethanol a year, says Verenium chief executive Carlos Riva.

But whether cellulosic-ethanol producers

MELTING MEMORY CHIPS
Phase-change memory enters mass production.
www.nature.com/news

SAMSUNG

can be profitable remains an open question. In February, researchers at Sandia National Laboratories in Livermore, California, and Albuquerque, New Mexico, and General Motors estimated that cellulosic ethanol could compete with petrol in 2030 only if oil was $90 a barrel or higher. The price of oil is around $65–70 a barrel.

One company looking to bring down costs is Novozymes, based in Bagsværd, Denmark. In February, Novozymes launched its first commercially available enzymes for cellulosic-ethanol production, claiming that it had reduced the cost of the enzymes by roughly half since 2008, and aims to halve the cost again by 2010. The reductions were achieved by breeding fungal strains that produce more efficient enzymes, raising protein production levels and optimizing process design, says global biomass business development manager Cynthia Bryant. The company has partnered with the agricultural products supplier COFCO and oil refiner Sinopec in Beijing to develop a process that uses maize stalks and leaves, and it opened a research unit in Brazil this year to study the conversion of sugarcane residues to ethanol.

Others are investigating ethanol production techniques that do not require buying enzymes from other companies. Coskata, a company based in Warrenville, Illinois, in which General Motors has an undisclosed stake, relies on gasification to turn biomass into hydrogen and carbon monoxide. Proprietary bacteria then ingest the gases and produce ethanol, says chief marketing officer Wesley Bolsen. Mascoma, a start-up based in Lebanon, New Hampshire, that announced a partnership with Chevron Technology Ventures on 14 September, claims to have engineered microorganisms whose enzymes can break cellulose down into simple sugars and immediately convert those into ethanol.

Upcoming policy decisions could give cellulosic ethanol a boost. The EPA must decide by 1 December whether to increase the blend wall to 15%, in response to a request submitted by advocacy group Growth Energy based in Washington DC. The industry is also waiting to see whether the EPA will put annual targets for cellulosic-biofuel production into effect on 1 January 2010 as originally proposed. In the EU, member states must submit action plans to meet renewable-energy goals by the end of June 2010, which could include more support for cellulosic ethanol.

Attention has already begun to turn from ethanol to other cellulosic biofuels. A Perspective published in *Science* in August noted that biomass-derived hydrocarbon fuels, such as petrol and diesel replacements, could be more attractive because they offer higher mileage and do not require new distribution infrastructure (J. R. Regalbuto *Science* **325**, 822–824; 2009). Once the first few cellulosic-ethanol plants have been built, people might move on to other fuels such as biobutanol or biomass-derived petrol, says Alexander. And the plant delays are giving alternative technologies a chance to catch up.

But the slowdown could also benefit those who produce cellulosic ethanol, says David Berry, a partner at venture-capital firm Flagship Ventures in Cambridge, Massachusetts. "The economic downturn has made it more difficult but also helped bring a sense of reality to the field," he says. "Companies will pull through stronger." ■

Roberta Kwok

NEXT WEEK: BIOMASS TO LIQUIDS

Innovation strategy

The Obama administration has good proposals but needs a more systematic approach, **David Goldston** argues.

PARTY OF ONE

US President Barack Obama travelled to a struggling city in upper New York state last week to unveil a 22-page "Strategy for American Innovation". The Obama administration is genuinely interested in developing policies to foster innovation, but a close reading of the strategy indicates that it is just beginning to figure out how to do that.

Innovation is a broad and amorphous, yet highly contested, area of policy. Federal efforts to spur "the development of new products, services, and processes" — the strategy document's definition of "innovation" — always provoke debates about the appropriate role of government, as well as charges from conservatives that programmes are being designed to "pick winners and losers".

The topic has been such a perennial political flashpoint that even choosing a name for such efforts is no simple matter. The Obama administration no doubt chose "innovation policy", in part, because it sounds hard to dispute — who could be against progress? — and it breaks from past nomenclature, such as the "industrial policy" of the 1980s or President Bill Clinton's "technology policy", which did not succeed in keeping the conservatives at bay. The Obama phrase is also a quiet nod to the sometimes controversial field of "innovation economics", which argues that governments do need policies explicitly directed at promoting innovation; merely creating a healthy economic environment is not enough.

But "innovation policy" is also a useful phrase politically because it's broad enough to encompass almost anything the government is doing and to frame any activity as a forward-looking way to create jobs. The document highlights not only most of Obama's science and technology initiatives, such as increasing research spending, advancing energy technology and expanding broadband access, but also more tangentially related proposals, such as reforming health insurance.

Indeed, the main problem with the strategy is that it reads like a compendium that was put together by asking every federal agency to send the White House a list of what they were doing that could be construed as innovation, rather than an exercise in mapping barriers to innovation and then designing policies to overcome them. The last section of the document, "Grand Challenges", offers an interesting, but random, assortment of possible technical advances, and has the air of notes taken at a meeting at which smart people sat in a room tossing around ideas. And the whole paper feels a bit rushed.

All this is a little disappointing, but much less damning than it may sound. First, the repackaged flavour is typical of initial strategy documents in every administration — the documents are largely political — and it is early to expect Obama to have a fully articulated innovation plan. Indeed, what's striking is that the White House chose to issue the strategy now and that the president built a speech around it; innovation does not seem to be a back-burner concern for Obama. And the list of initiatives is quite impressive, if previously announced; through the American Recovery and Reinvestment Act and other measures, the administration has been pouring money into research and development (R&D) and education, for instance.

The question the document begs is what the rationale will be for a systematic innovation plan going forwards. For example, the strategy notes that the president has proposed a goal that public and private R&D account for at least 3% of gross domestic product.

This figure has floated through US innovation debates for years largely because it was the investment level achieved by some competing nations in the past. But it's not clear why 3% is the right target. Why does the percentage matter more than the actual dollar level? And how much does the impact depend on the projects that are being funded? The entire relationship between research and innovation, innovation and economic growth, and growth and jobs is unsettled intellectual terrain.

The strategy also generally avoids discussing any policy area that Obama has not yet addressed. For example, the only mention of immigration talks about improving "the processing of high-tech visas". And the document seems to omit any proposal that might needlessly stir up old antagonisms. It leaves out, for example, the signature technology programmes of the Clinton administration, such as the Technology Innovation Program (formerly the Advanced Technology Program), which provides grants to firms for high-tech projects, even though Obama has proposed increasing their budgets.

The plan, however, does not shy away from controversy on items that are already in play. Notably, it cites the proposed cap on carbon dioxide emissions as part of the energy innovation strategy. Here the breadth of the strategy is a plus; in the past, innovation discussions too often focused solely on increasing the supply of new technologies, ignoring the need to create demand for them.

So at its best, the strategy can be seen as a broad inventory of existing proposals and programmes that might promote innovation, at least as a secondary effect. A useful next step would be a probing analysis of what impedes innovation in the United States and a list of proposals that are targeted directly at removing those hurdles.

For instance, a recurrent and central complaint in innovation discussions is that financial institutions in the United States are interested only in projects with quick pay-offs. But no strategy ever seems to include a remedy for this. That is partly because strategies often draw on ideas from lobbying groups, whose innovation agendas usually consist of proposals they were already pushing to further their members' interests.

Lobbyists and think tanks also often push for a reorganization of the federal bureaucracy; two ideas circulating now are creating a manufacturing "czar" or a White House office to review policy for its impact on innovation. But government reorganization can be a drain on time and energy, and the goal should be to infuse concern about innovation throughout the agencies, not to carve out backwater niches that can focus on it. Moreover, the heart of the problem doesn't seem to be a lack of government concern, but a dearth of clear, disinterested analysis and fresh solutions.

Some additional innovation in policy will be needed to help promote greater innovation in the US economy. ■

David Goldston (partyofonecolumn@gmail.com) is the director of government affairs at the Natural Resources Defense Council in Washington DC. Views expressed are his own.

R. DEPARDON/MAGNUM PHOTOS

SCALING THE WALL

The collapse of communism opened up the world to scientists from eastern Europe. **Quirin Schiermier** talks to researchers about what changed.

On the 9 November 1989, the lights were out early for Pavel Jungwirth. Jungwirth had recently graduated from Charles University in Prague, Czechoslovakia, and he was hoping eventually to pursue a PhD in physics. But his plans had been interrupted when he was drafted into the army, and for now he was quartered in barracks outside the city. With a 10 p.m. curfew, Jungwirth had little opportunity to watch the historic spectacle that was captivating audiences around the world, as an ecstatic crowd breached the Berlin wall.

That year, Soviet-controlled communist governments throughout central and eastern Europe had been stumbling in a historic chain reaction. In May, Hungary had begun to dismantle the Iron Curtain. In June, the Polish anti-communist Solidarity party led by shipyard worker Lech Wałęsa had won by a landslide in free elections. But for many, the most powerful symbol of communism's collapse was on the November day when East Germany's rulers surrendered to weeks of peaceful rebellion by its citizens and announced that people from East Berlin could pass through the formidable barrier that had divided the city since 1961.

One of those watching the events on her family television was Alicja Józkowicz, a 22-year old then studying zoology at the Jagiellonian University in Kraków, Poland. "It was a very emotional moment," she says. Like Jungwirth, Józkowicz was set on becoming a scientist. The drama unfolding on the screen made her realize, she says, "it would change my professional opportunities".

Neither Jungwirth nor Józkowicz could anticipate how dramatically their professions would change. The region's political transformation meant that universities and research institutes that had been isolated from international science were now expected to take part in it. Some researchers found this a formidable challenge, grappling for the first time with having to publish their work and compete for funding. But many others, such as Jungwirth, Józkowicz and a generation of eastern European students who in 1989 were embarking on a career in science, saw it as an opportunity. These researchers soon became a common sight at labs in western Europe and North America, and many of them have gone on to establish competitive laboratories.

"I consider it a small miracle that within the past 20 years sprouts of excellence have grown in my country," says Jungwirth, who now runs a chemistry laboratory in Prague. "Compared with 1989, our academic community is profoundly different — inspiring, self-confident and international." Józkowicz, now a group leader in the Department of Medical Biotechnology at Jagiellonian University, agrees: "Personally I have a feeling I am a rightful member of an international society of scientists."

"Compared with 1989, our academic community is profoundly different — inspiring, self-confident and international."
— Alicja Józkowicz

Lost tradition

Yet something could have been lost in the internationalization of eastern Europe, says Wolf Lepenies, a sociologist and former director of the Institute of Advanced Study in Berlin. When he and others used to visit the East, "institutions were poor, instruments were lacking, but intellectual traditions had prevailed that we had barely taken notice of", he says. "Debates were conducted with an earnestness and a responsibility that were largely unfamiliar to us." Lepenies says that the 'we will help you' approach that western Europe has often used in its interactions with the East carries a whiff of condescension. "Eastern Europe is a good place to shake the complacency of the West," he says. "'We need each other!' would have been a more appropriate guideline."

Jungwirth's family expected him to study science. His father, a physicist himself, went frequently to Novosibirsk, a strong centre of Soviet

physics and chemistry, and he would years later become a vice-president of the Academy of Sciences of the Czech Republic in Prague. But in 1970, Jungwirth's father was kicked out of the Communist Party because of his open disagreement with the 1968 Soviet occupation. This stance affected his son's plans. Jungwirth had hoped to study medicine, but according to an unwritten rule, children of politically unreliable citizens should not enrol in subjects through which they might exert an undue influence on people. Medicine was one such discipline.

Jungwirth instead opted for physics, which was considered socially neutral. As a student, he followed the 1989 political events with eager anticipation. When writing his diploma thesis he had access to a computer and a printer — a rarity in a country where the few functioning photocopiers were rigidly controlled for fear of dissident activity. Some like-minded friends asked him to print copies of a letter they had written in support of Václav Havel, the playwright and essayist who had become the political figurehead of the anti-communist movement and had been taken into custody. Jungwirth helped them out, but a copy of the letter fell into the wrong hands and he was summoned to the dean. "I thought 'that's it, I'm out'," he says. He faced the dean over a massive rosewood desk surrounded by insignia of communism and the portrait of the Czech party leader Gustáv Husák. Then something unexpected happened. After lecturing him for a minute or two, the dean crumpled up the ill-fated letter and flung it into a corner of his study. Perestroika, Jungwirth thought, had definitely arrived in Prague.

By November, huge momentum was building for Havel's 'Velvet Revolution' and a student rebellion had expanded into a nationwide general strike. On 25 November, Jungwirth slipped out of the barracks to join the millions shouting for freedom on the streets of Prague. Husák resigned on 10 December, paving the way for democracy, and on 29 December Havel was installed as the new President of Czechoslovakia.

Jungwirth switched from military to civil duty as soon as this option was made possible in 1990 and resumed his studies a year later. He was interested in atmospheric chemistry and, at the J. Heyrovský Institute of Physical Chemistry in Prague, started putting together molecular simulations of hydrophobic molecules such as methane. Soon he was embarking on his long-awaited PhD.

Shades of grey

During Józkowicz's school years in Kraków in the 1980s, the Western world and Western science seemed unimaginably far away. "I never even dreamed of going there," she says. "I remember my childhood and youth as a time when everything was grey, and each year things just seemed to get worse," she says. "The stores were empty all the time."

Józkowicz was curious about the natural world and she remembers her science teachers — and textbooks — as being excellent. "Much better actually than the science education my son is getting now," she says. The high standard of science education was true of many eastern European countries.

Józkowicz was aware early on of the Solidarity movement that began in the early 1980s and that was to become the nucleus for political changes in Poland. Solidarity, led by Wałęsa and the first independent trade union in the Soviet-controlled Eastern bloc, quickly turned into an openly anti-communist political and social movement supported by millions of Poles. For Józkowicz, the most significant time was when Solidarity was elected and Tadeusz Mazowiecki became prime minister. "We really felt that a new era was about to begin," she says.

Lech Wałęsa founded the Polish Solidarity party that won free elections in 1989.

A. KEPLICZ/AP PHOTO

But when capitalism eventually arrived in Poland it also brought problems, as it did in other countries. The political and intellectual freedom was invigorating but, at least in the early days, disillusionment was never far behind. "The stores filled up," she says. "But salaries dropped just as quickly."

Scientists found the transition particularly harsh. Until that time their research money had been guaranteed by the government. With the collapse of communism, many countries were close to bankrupt and science funding dropped abruptly to almost zero. It was impossible in the early years for most scientists to make a living without getting second jobs, and many gave up altogether.

Those who managed to hold onto their labs could travel freely to other countries for the first time, meet their counterparts there and begin to integrate with the international scientific community. But this came with tough international competition and new codes of conduct. Conferences were rare, peer review was unknown, and few researchers spoke English. Many developed something of an inferiority complex. The national science academies, which operate hundreds of basic-research institutes throughout central and eastern Europe, today have a disproportionate number of members close to or above the age of retirement who have continued to pursue their research and have domestic influence but have never really entered the international science arena.

Jungwirth and other young scientists, though, had little to lose. Pitifully low salaries, poorly equipped labs and a lack of grant money quickly led to thousands of mainly

The anti-communist movement led to the 1989 election of Václav Havel as Czechoslovakia's President.

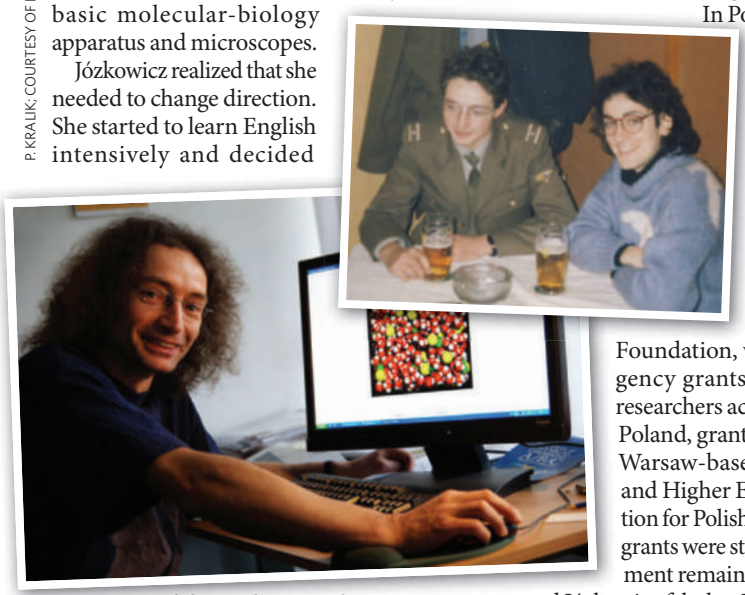
AP PHOTO

young scientists leaving for the West. Jungwirth was lucky: his PhD supervisor, chemist Rudolf Zahradník, had good contacts and used them to send Jungwirth to Switzerland in 1992. He spent a year there working with Thomas Bally at the University of Fribourg, where his 3,000 Swiss-franc (US$2,290) monthly salary made him feel "rich beyond measure". (During his PhD Jungwirth also became casual acquaintances with another former PhD student of Zahradník's, Angela Merkel).

Later, Jungwirth worked as a postdoc for a few months at the Hebrew University of Jerusalem in Israel and at the University of California, Irvine, with the same supervisor, Benny Gerber. He became interested in computational modelling of atmospheric chemistry. "In Switzerland, and later in the United States, I saw that scientists acquired positions because they knew something, not because they knew somebody," says Jungwirth. He was impressed by how his supervisor "developed new methods in amazing outbursts of creativity. I felt on top of the world."

Józkowicz was also learning about science abroad. Her first international meeting — the 1991 world immunology congress in Budapest — was a shock. Not only did she realize that she could hardly understand the English-language talks, she also found that her group's science was hopelessly behind. She had recently started a PhD comparing the amphibian immune system to those of mammalian species in the department of evolutionary immunology at Jagiellonian University. "Our research was completely outside the mainstream," she says. "Worse, the methods we used were just totally outdated." Her lab had no cell sorters and only the most basic molecular-biology apparatus and microscopes.

Józkowicz realized that she needed to change direction. She started to learn English intensively and decided that as soon as she had completed her PhD she would switch from zoology to medicine. She did this in 1996, moving to the Department of Clinical Biochemistry in the medical school at Jagiellonian University. She studied the control of blood-vessel growth — angiogenesis — by the gas nitric oxide. But she was still not happy with the pace at which things were improving.



P. KRALIK; COURTESY OF P. JUNGWIRTH

From army to lab: Pavel Jungwirth now runs a chemistry group in Prague.

In Poland, as in other countries, grant systems for science were being established and foreign aid and philanthropic aid were providing some relief. In 1992, for example, Hungarian-born billionaire George Soros founded the International Science Foundation, which provided emergency grants to more than 20,000 researchers across eastern Europe. In Poland, grants were provided by the Warsaw-based Ministry of Science and Higher Education and Foundation for Polish Science. But individual grants were still small, modern equipment remained mostly unaffordable and Józkowicz felt that Polish science generally lacked international flair and recognition.

An opportunity presented itself in 1997 when a poster of hers on *in vitro* gene therapy was shown at a congress in Germany. Diabetes researcher Lawrence Chan of the Baylor College of Medicine in Houston, Texas, noticed it and invited her to join his group for a year. Four weeks later she landed in Houston, leaving behind her husband and six-year-old son. She threw herself into her work. "I felt painfully lonely," she says.

Culture shock

Texas was quite an experience for the then 30-year-old, now officially a postdoc (a job term that at the time hadn't been much used in eastern Europe). Józkowicz found the style with which Chan ran his 40-strong research team inspiring. "Each postdoc got a gene, human, mouse or monkey, to work on, and within one week of their arrival everybody was busy doing experiments and producing results," she says. Józkowicz got mice, and her experiments with them identified a way to perform gene therapy that protected against atherosclerosis for the animal's lifetime. It was to become her most cited paper[1]. "The year with Chan taught me what real teamwork can achieve — it was really the first crucial step in my development." She later signed up for three years of postdoctoral research at the Medical University of Vienna.

Heading home

Many scientists from eastern Europe choose to stay in their native countries — or return to them — even though salaries there are often much lower and the labs are less-generously funded than those in western Europe, the United States or other areas (see page 682).

Deep personal attachment to home is one of the reasons. Another is tradition. Under communist rule, many scientists spent their entire career in the same lab. The mobility of researchers has greatly increased since, but students and scientists from the region are still more reluctant to change labs, or spend time abroad, than those from China, western Europe and the United States.

In the United Kingdom and Denmark, for example, nearly 10% of the science and technology workforce changed jobs in 2005; the figure for most eastern European countries is less than half that.

On the flip side, only a few universities and institutes in eastern Europe are capable of attracting foreign scientists. In Poland, the International Institute of Molecular and Cell Biology in Warsaw is one such example. The best institutes run by the national science academies in Poland, Hungary and the Czech Republic also employ foreign talent, particularly from Germany, but senior-level North American or Asian researchers are still a rare sight in labs in eastern Europe.

Some researchers think that eastern European scientists' preference to stay at home is sometimes too pronounced — and it may be one reason that academic institutions in these countries struggle to compete on the international stage.

Only five universities in the region — Charles University in Prague, the University of Szeged in Budapest, the University of Ljublijana in Slovenia, Poland's University of Warsaw and Jagiellonian University in Kraków — are ranked among the world's top 500 by an often-cited academic ranking drawn up by China's Jiao Tong University in Shanghai. The other reasons are many: across the region, public expenditure on science is notably below the European Union average, and industrial research capacities are generally underdeveloped (see graphic, page 590). Moreover, the region is under-resourced when it comes to large research infrastructures such as synchrotron machines, super-computers and particle accelerators. **Q.S.**

M. JÓZKOWICZ J. DULAK

1980s to 2000s: Alicja Józkowicz took postdocs abroad to pursue biomedical research.

Just as Józkowicz was leaving home, Jungwirth was returning. Like many young scientists who had taken the opportunity to leave eastern Europe, he decided to go back there (see 'Heading home'). In 1995 he took up a group-leader position in his native country — by this time the Czech Republic after its split with Slovakia — at the Institute of Organic Chemistry and Biochemistry of the Academy of Sciences of the Czech Republic in Prague.

With the help of a 200,000-Deutschmark (US$143,000) grant from the German Volkswagen Foundation, Jungwirth was able to quickly establish a reasonably competitive group. He also secured grants from the US National Science Foundation in Arlington, Virginia, and the North Atlantic Treaty Organization (NATO) Science Programme in Brussels. "Buying serious equipment was still out of the question," Jungwirth says. "But the students were excellent and the computers were just fine." As a theorist, computers were the main thing he needed and Jungwirth started to establish a reputation with his molecular models[2].

Over time Jungwirth started to feel that things were getting better. The Czech government invested relatively generously in science, and the Academy of Sciences was less resistant to reform than its counterparts in some other countries. Jungwirth expanded his research interests. By the late 1990s he was exploring atmospheric chemistry, simulating ions such as chloride at aqueous surfaces to try to work out how they could react with ozone and other pollutants[3]. He also ventured into biology, studying how salt ions influence the properties of proteins[4]. "I had a lot of luck," he says. "The timing was just right. I see many people 5–10 years older than me who are now in a much worse situation."

> **"I saw that scientists abroad acquired positions because they knew something, not because they knew somebody."**
> — Pavel Jungwirth

In 2003, Józkowicz returned home too — to Kraków, where her son had just celebrated his twelfth birthday. She had maintained a formal affiliation with Jagiellonian University, and was eager to teach a group of her own the new methods and styles she had learned abroad.

But although the world had been changing, Polish academic circles had not. For 18 months all her applications for independent funding came to nothing. The academic establishment was still very hierarchical, with advancement based on favouritism rather than merit. Changing labs was — and sometimes still is — considered an act of disloyalty. Józkowicz found that the years spent abroad pretty much debarred her from the national funding system.

Her break came in 2004 when she won a prestigious Central European Senior Research Fellowship worth some £350,000 (US$570,000) awarded by the British Wellcome Trust, which is open to researchers in Estonia, Poland, the Czech Republic and Hungary. The money supported projects aimed, among other things, at studying blood vessels in tumours and during cardiovascular disease.

Home for good

The grant seemed to make it easier for her to convince Polish grant-givers of the quality of her research and since 2007 she has won funding totalling €330,000 (US$488,000) from the Polish Science Ministry, the Wellcome Trust and from Adamed, a Warsaw-based pharmaceutical company. Her small group — now with access to modern fluorescent microscopes and other equipment — has become the nucleus for Jagiellonian's Department of Medical Biotechnology, created in 2005, where she hopes to receive tenure soon. "My plans?" she says. "Make the most of the new money, produce papers in leading journals — and live in Kraków with my family for the rest of my life."

Some of the region's catch-up has been aided by the expansion of the European Union (EU), which was joined in May 2004 by ten countries, including Poland and the Czech Republic, followed by Bulgaria and Romania in 2007. Józkowicz's department this year received around €6,000,000 from EU structural funding, which is aimed at generating equal living conditions across the EU and can be used to fund research infrastructure and equipment. The university received an extra 45 million zloty (US$16 million) to establish a new Jagiellonian Center for Experimental Therapeutics, due to open later this year. The European Research Council has also begun to make generous grants available to young investigators across Europe.

Both Jungwirth and Józkowicz believe that the East–West gap will narrow further as old cultural habits die away and scientific borders dissolve. "Don't forget we started from empty walls," says Józkowicz. "But we now have a young and energetic environment for science, with excellent students who are eager to work. I do believe that at least some labs in Poland and elsewhere will very soon become attractive to foreign scientists."

Some of that change is palpable in the attitudes of the scientists that Jungwirth now meets. "Until the turn of the decade or so I could still play the game of coming from a poor country," says Jungwirth. "People had been eager to see us, and keen to help. The feeling of being welcomed with open arms made things a lot easier for us when times were not so good." Not so now, he says, as the history starts to fade from memory: scientists who are starting their PhDs today were born after the Berlin Wall fell. "We cannot say now that we do not publish well," says Józkowicz, "or we do not make valuable research because of 'them' or 'circumstances beyond our control'. I remember that, when I was a teenager, the often repeated sentence was 'in the West it would be obvious or easy but here — forget about it. Never-ever would it be possible in Poland'. And it is possible."

If any further evidence was needed of what is possible for scientists in the Eastern bloc, it was apparent at a party last year in Prague. Jungwirth was attending the 80th birthday of his former supervisor, Zahradník, who served as president of the Czech Academy of Sciences between 1993 and 2001.

Late in the evening a black limo with a German licence plate pulled up. Merkel, now the German Chancellor, had spontaneously decided to come down from Berlin to congratulate her aged PhD supervisor. Free from rules of protocol, she chatted the hours away, in fluent Russian, with her science colleagues of old. The word is that it was a good party. ■

Qurin Schiermier is *Nature*'s Germany correspondent.

1. Kim, I. H., Józkowicz, A., Piedra, P. A., Oka, K. & Chan L. *Proc. Natl Acad. Sci. USA* **98,** 13282–13287 (2001).
2. Jungwirth, P. & Tobias, D. J. *J. Phys. Chem. B* **106,** 6361–6373 (2002).
3. Knipping, E. M. *et al. Science* **288,** 301–306 (2000).
4. Vrbka, L., Vondrášek, J., Jagoda-Cwiklik, B., Vácha, R. & Jungwirth, P. *Proc. Natl Acad. Sci. USA* **103,** 15440–15444 (2006).

For more on eastern Europe, see http://tinyurl.com/eeurope.

BEYOND THE BLOC

From the 1950s, science has been a priority in the Soviet-controlled Eastern bloc. Space science and nuclear physics, in particular, received generous support, and achievements such as the Soviet space missions served as proof of the alleged superiority of the communist system. Then in the late 1980s, collapse of communist regimes and their replacement by democracy and market economies led to a dramatic drop in science expenditure across the region.

Reliable data for the early 1990s — the period of worst hardship — are unavailable because national statistical services were not yet up and running or had no reported science figures. Between 2003 and 2004 the European Commission included the ten countries in central and eastern Europe that have since joined the European Union in its regular reports on science and technology indicators. Figures from the 2008–09 report show that research intensity — the percentage of gross domestic product spent on research and development (R&D) — is below the European average of around 1.8% in all ten countries.

But there are strong differences within the region. Although the Czech government has in recent years invested substantially in science, the research bases of countries such as Bulgaria and Romania are still underdeveloped. Notable improvements, mainly driven by the business sector, have been made in Estonia and Hungary — although the recent economic crisis has hit Hungary harder than it has hit others, threatening the upward trend.

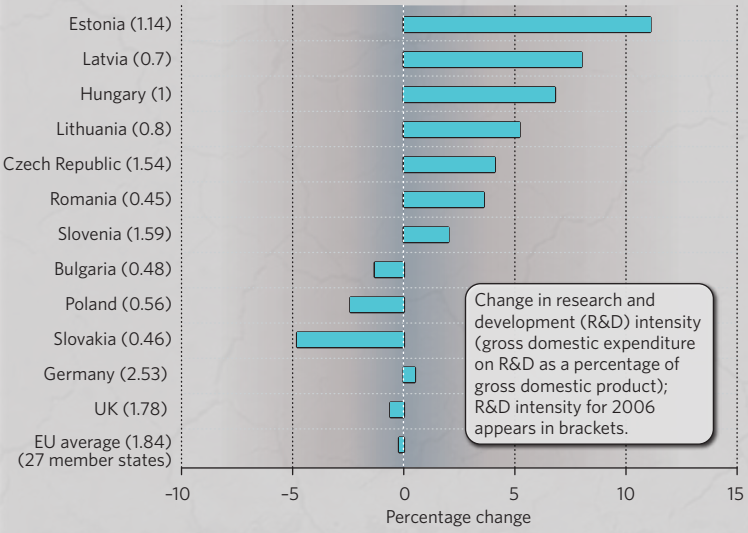
Poland, the largest country in the region, has in recent years fallen further behind in reaching EU and national goals for R&D intensity, as have Slovakia and Bulgaria.

There is a general lack of highly qualified scientists and technical workers, whose share of the overall workforce ranges between 9.8% in Romania and 16.8% in Estonia.

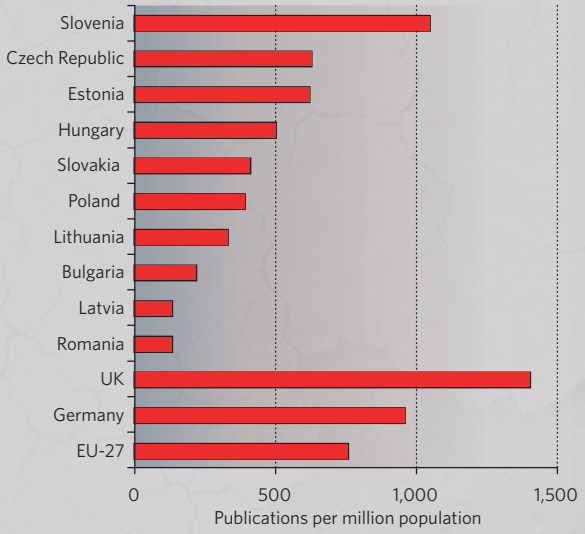
The region also still fails to attract much foreign scientific talent. Only 1% or so of participants in the EU's Marie Curie Actions programme — which promotes and facilitates the mobility of young scientists in Europe — choose a university, research institute or industry lab in the new member states as their host institution, possibly due to the resources and reputations of other member states. ■

Quirin Schiermeier

GROWTH IN R&D INTENSITY 2000–06



SCIENTIFIC PUBLICATIONS



SOURCE: EUROPEAN COMMISSION DIRECTORATE-GENERAL FOR RESEARCH

HISTORY IN THE MAKING

Key points in eastern European independence and research and development.

B. BISSON/SYGMA/CORBIS

1945 **August:** Signing of the Potsdam Agreement dividing Germany and Berlin between Allied forces

1957 **October:** Launch of Sputnik

1961 **August:** Completion of the Berlin Wall

1980 **September:** Formation of the Polish Solidarity movement

P. TURNLEY/CORBIS

1989 **June:** Poland hosts first free elections in the Eastern bloc
23 October: Hungary's **Mátyás Szűrös** declares a republic
9 November: Fall of the Berlin Wall
10 November: Bulgarian president Todor Zivkov (BKP — Bulgarian communist party) is deposed
17 November: Student demonstration in Prague brutally suppressed, sparking Velvet Revolution
29 December: Václav Havel elected President of Czechoslovakia
25 December: Romanian dictator **Nicolae Ceaușescu** is executed

1990 **March:** Lithuania gains independence from the Soviet Union
June: First free elections in Bulgaria
July: Estonian Science Foundation created — first Western-style funding agency in the region

G. FOUET/AFP/GETTY

LIE OF THE LAND

Ten former communist states that have joined the European Union since 2004. United Kingdom and Germany shown for comparison.

SOURCE: INTERNATIONAL MONETARY FUND 2007

ESTONIA
GDP: **$20.9 billion**
Population: **1.3 million**
Tallin

LATVIA
GDP: **$28.8 billion**
Population: **2.3 million**
Riga

LITHUANIA
GDP: **$38.9 billion**
Population: **3.4 million**
Vilnius

UNITED KINGDOM
GDP: **$2.80 trillion**
Population: **60.9 million**

GERMANY
GDP: **$3.32 trillion**
Population: **82.2 million**

POLAND
GDP: **$425 billion**
Population: **38.1 million**
Warsaw

SLOVAKIA
GDP: **$75.2 billion**
Population: **5.4 million**
Bratislava

CZECH REPUBLIC
GDP: **$174 billion**
Population: **10.3 million**
Prague

ROMANIA
GDP: **$169 billion**
Population: **21.6 million**
Bucharest

SLOVENIA
GDP: **$47.2 billion**
Population: **2.0 million**
Ljubljana

HUNGARY
GDP: **$138 billion**
Population: **10.1 million**
Budapest

BULGARIA
GDP: **$39.6 billion**
Population: **7.6 million**
Sofia

AP PHOTO

1991 **February:** Foundation for Polish Science established (a grant-giving agency along Western lines)
June: Slovenia gains independence from Yugoslavia
August: Estonia and **Latvia** gain independence

1992 **January:** Collegium Budapest, Institute for Advanced Study, takes up operation

1993 **January:** Czech Republic and Slovakia become independent states

B. DAVID/CORBIS SYGMA

1999 **January:** International Institute of Molecular and Cell Biology in Warsaw opens with support of UNESCO and the Polish Academy of Sciences

2000 **March:** European Council endorses the creation of the European Research Area, an initiative to bolster European research

2004 **May:** Estonia, Latvia, Lithuania, Poland, Czech Republic, Slovakia, Hungary and Slovenia gain European Union membership

2007 **January:** Romania and Bulgaria join the European Union

CORRESPONDENCE

Call from China for joint nanotech toxicity-testing effort

SIR — In response to your News story 'Nanoparticle safety in doubt' about lung damage in Chinese factory workers (*Nature* **460,** 937; 2009), we would like to stress that China has been paying close attention to research into and documentation of the risks of working with nanomaterials.

As in most Western countries, industrial use of nanoscale products has been proliferating in China over the past decade. These are widely used in weaving, dyeing, cosmetics and medicine, for example, as well as in waste-water treatments. The need to develop international standard-analysis protocols to assess nanoparticle toxicities is therefore becoming increasingly urgent.

China is taking measures to address the issues (for a review, see G. Q. Zhou *et al. Prog. Biochem. Biophys.* **35,** 998–1006; 2008). These have included a series of national conferences and symposia on the safety of nanomaterials: for example, the 243rd Xiangshan conference in 2004. Research projects such as the '973 programme' of basic research and development have been initiated by the Chinese ministry of science and technology. And several important projects have been funded by the National Natural Science Foundation.

A recently established Chinese journal, *Asian Journal of Ecotoxicology*, publishes primary research papers and reviews on nanoparticle safety and toxicity in almost every issue. There are regular updates on developments in the field and a scientific platform for data sharing and policy discussion (for example, see N. Wang *et al. Asian J. Ecotoxicol.* **2,** 252–264; 2007).

We urge the relevant international scientific organizations to join forces and work out a scheme for establishing high-speed tests that will sort out safety and toxicity issues for different industrial nanomaterials (R. F. Service *Science* **321,** 1036–1037; 2008). Such a collaborative venture would accelerate progress in this rapidly expanding field.

Shuping Bi, Jing Zhang, Jiongjia Cheng
Chemistry Department, Nanjing University, Nanjing 210093, China
e-mail: bisp@nju.edu.cn

Consent: criteria should be drawn up for tissue donors

SIR — The drive to develop new human pluripotent stem-cell lines has attracted a new, exuberant cohort of researchers who may not be familiar with the regulations and standards governing donation of human tissue. Scientists should ask donors to agree to some basic rules for research involving induced pluripotent stem (iPS) cell lines derived from their tissue. This would help to prevent road blocks of the kind highlighted in your Editorial (*Nature* **460,** 933; 2009) about Sabine Conrad and colleagues' Corrigendum to their paper 'Generation of pluripotent stem cells from adult human testis' (*Nature* **460,** 1044; 2009).

The rules would cover sharing cell lines with other investigators, as in this case; carrying out large-scale genome sequencing; injecting iPS cells or their derivatives into animals; and patenting discoveries or commercial uses arising from iPS cells or derivatives, with no sharing of royalties with donors. Additional specific consent would be needed for allogeneic human transplantation or reproductive research using gametes derived from iPS cells. Permission might be needed to recontact donors about new research proposals.

The advent of human adult germline stem cells and iPS cells avoids ethical issues over the use of early human embryos. So, unlike human embryonic stem cells, there is not likely to be a shortage of donors for germline stem cells or iPS cells (except in the case of rare diseases). Given that iPS cell lines can be propagated indefinitely, they are likely to become widely used. But applications still require proper consent from the tissue donor.

The proposed minimum-consent criteria from tissue donors should help to maximize the scientific value of cell lines in realizing the promise of this technology.

Bernard Lo Program in Medical Ethics and Department of Medicine, University of California, San Francisco, California 94143, USA
e-mail: bernard.lo@ucsf.edu
Bruce R. Conklin Gladstone Institute of Cardiovascular Disease and Departments of Medicine and Molecular Pharmacology, University of California, San Francisco

Consent: a need for guidelines to reflect local considerations

SIR — As you point out in your Editorial (*Nature* **460,** 933; 2009) on the distribution of human cell lines, withholding scientific material from the broader research community contravenes the basic norms of science. We do not believe, however, that standard international consent guidelines for tissue donors are the solution to this problem, and suggest that these should instead be devised on a local scale, in collaboration with ethics committees.

Far from research being "hindered by restrictions from donors" as you suggest, people are generally willing to donate tissue for research, and even to give open-ended consent to unspecified future applications. This willingness is underpinned by donors' faith in medical research and in their right to protection and confidentiality; the assumption is that their tissue will be used only for 'ethical' investigations.

But problems can arise, for example, over whether consent covers the proposed usage (at present there are many different models of consent, ranging from specific to general), and when and how tissue should be discarded (K. Aalto-Setälä *et al. PLoS Biol.* **7,** e42; 2009).

The answers may not always be obvious, and ethics committees (in collaboration with donors or their representatives) need to take into account the type of tissue involved, as well as the demographics and potential vulnerability of the donor or donor community, in judging the acceptability of a research proposal.

None of this precludes distribution of tissue in the name of scientific progress, nor should it if the wishes of donors are to be respected. However, it does challenge any unqualified presumption among researchers about access to human material; it also calls into question the ethical acceptability of using internationally standardized consent forms, as recommended in your Editorial. Rather, we would argue for international standards to ensure that tissue distribution is not thwarted by ethics committees, accompanied by a plurality of local approaches to obtaining consent.

This strategy would address the problems you outline, while demonstrating respect for moral decisions made by individuals and groups and preserving donors' trust in biological medicine.

Wendy Lipworth, Rob Irvine, Bronwen Morrell Centre for Values, Ethics and the Law in Medicine, Medical Foundation Building, University of Sydney, NSW 2006, Australia
e-mail: w.lipworth@usyd.edu.au

Contributions may be sent to correspondence@nature.com. Please see the Guide to Authors at http://tinyurl.com/373jsv.

OPINION

Science and the Stasi

The acquisition of scientific and technological secrets was at the heart of East Germany's foreign espionage operations before the fall of the Berlin Wall, reveals **Kristie Macrakis**.

In 1992, three years after the fall of the Berlin Wall, a spy walked into the US embassy in Warsaw and offered to sell the CIA the real and code names of all intelligence agents from the HVA (*Hauptverwaltung Aufklärung*) — the foreign department of the Stasi, the East German Ministry for State Security. The CIA bought the highly sensitive information for a mere US$75,000.

The spoils — released to the Berlin Stasi archive and made available to me in 2005 — have the potential to alter popular perceptions of the activities of the East German intelligence agency and secret police. An analysis of the CIA material reveals that about 40% of all HVA sources planted in West German companies, research institutions and universities were stealing scientific and technical secrets.

It is not uncommon for nations seeking to catch up in technology to spy — they are often seduced by the possibility of turning secrets into products. History abounds with examples of countries that copied and then improved on foreign technology. In the eighteenth century, France spied on, and stole, Britain's textile technology. The Soviet Union stunned the world with its theft of American atomic-bomb secrets in the twentieth century. Since the end of the cold war there has been an increase in state-sponsored and orchestrated industrial, economic and scientific espionage. Recent cases and convictions of Chinese technological espionage against US and UK high-tech industries illustrate the persistence of this quest.

Boosting scientific, technical and economic prowess through espionage is a protracted process. Ultimately, success lies in a country integrating stolen ideas or products into its research and development system. But even with a highly perfected espionage operation such as that of East Germany, pilfering nations often forget that a scientific establishment based on pirated and cloned technology is rarely a leader, especially in fast-moving fields such as computing. In fact, every smuggled document weakens true scientific innovation by maintaining dependence on espionage and the state security regime.

The Sector for Science and Technology was the largest and most important unit in the Ministry for State Security for acquiring technical blueprints, plans and hardware. It was also considered the most successful because it saved East Germany millions of marks in research and development costs through its efficient acquisition pipeline of secrets from the West. Modelled on the KGB's science and technology directorate, the sector was founded in 1971 from three operational units for biology and physics, computers and military technology and the economy, plus an evaluation unit for grading stolen material. Major General Horst Vogel, an unassuming bear of a man and a working-class patriot from Saxony, headed the unit from 1975 to 1989.

J. BURTON

Technology on demand

The sector was almost like a secret mail-order operation. Scientists could place orders to acquire goods from the West. The goods were then evaluated and graded by Stasi-approved scientists and passed on to a research institute. Typical was the case of Wolfgang Biermann, general director of Carl Zeiss Jena, a manufacturer of optical systems, placing an order for documents on, and a sample of, a one-megabit chip.

During the cold war, the HVA managed to penetrate West Germany's most important institutions including prestigious scientific and technical establishments, as well as managing to recruit important American sources on German soil (the HVA had agents all over the world, but only the CIA material on German citizens was returned to Germany). By 1989, the sector had planted agents at internationally competitive companies such as IBM, Siemens, AEG/Telefunken, Standard Elektrik Lorenz, Texas Instruments and Digital Equipment Corporation. The aerospace defence contractor Messerschmitt-Bölkow-Blohm was particularly hard hit throughout the cold war (see 'All in the family'). Its sensitive military secrets were quickly passed on to the Soviet Union.

Contrary to expectations, the sector ran only a sprinkling of agents at centres of research such as the Max Planck Institutes, Fraunhofer and universities or technical colleges such as TH-Aachen and TH-Berlin. University agents tended to be students who were groomed as future sources in industry. As a result there were only a handful of agents who were university professors. The HVA targeted industry rather than basic research at universities, particularly after the spectacular 1979 defection to West Germany of Werner Stiller. An officer from the sector's basic-research unit, Stiller unmasked more than a dozen scientists working on basic and applied nuclear research. Several agents worked at smaller, often independently owned businesses and there was a large contingent of self-employed businessmen.

Thousands of Western scientists, engineers

and businessmen betrayed their country to work for Eastern bloc intelligence. Unexpectedly, the agents were not well-known leaders in their fields, rather they were anonymous male, salaried company employees, the majority working in electronics. Many were also secretaries, students, production workers and repairmen with access to secret material. Only ten agents working in West German institutions were women, five secretaries and five salaried employees, including one with a PhD.

HVA secretary spies seduced by male 'Romeo' agents made headline news during the cold war, but the science sector does not seem to have used this strategy much. Many West Germans spied for money; but an equal number spied for ideological reasons, and some for love or for adventure.

Open secret

Good espionage does not necessarily lead to good science or technology. Take the case of the computer industry. In the 1960s, and again in the 1980s, East German spies and smugglers successfully bypassed the Western embargo prohibiting Eastern bloc countries from importing sensitive dual-use technologies. The Eastern bloc countries stole plans, copied technologies and reverse-engineered machines through agents planted at key institutions.

This was no secret in the West. Not only did Western scientists write messages on microchips such as "When are you going to stop stealing?", but they sometimes altered embargoed goods to fail on arrival. When Werner Scheele, a West German businessman codenamed 'Rhein', smuggled a US high-current ion implanter into East Germany through Indonesia, Western intelligence agencies doctored the machine with self-destructive parts. Delivered to East Germany in December 1989, it exploded around the same time as the country imploded.

> "Every smuggled document weakens true scientific innovation."

Computer technology was alluring, but the state overestimated the power of secrets to propel its computer industry forwards. The effort to copy the IBM 360 was deemed a success during the 1960s because the Robotron Company created a clunky clone, called the RYAD series. But East Germany's quest to latch on to the 1980s computer revolution through espionage and illegal technology transfer failed miserably. They just couldn't catch up with powerhouse competitors from Japan and the United States in such a fast-moving industry.

Despite the vaunted reputation of East German intelligence, peeling away at its layers of mystery as new material has become available over the past ten years has revealed an incredibly bloated human intelligence-gathering operation. The Stasi maintained an enormous back-up staff, much larger than those of Western spy agencies, to recruit, run and handle agents. And, unlike other spy agencies that do much of their business through diplomatic channels and the embassies, East Germany did not have embassies in Western countries until 1972 when it was recognized as a separate German state. The Berlin Wall also hindered meetings with Western agents in East Germany. Much human effort was expended to run operations with a small return.

Ultimately, the stopgap measures offered by spies did not provide long-term solutions for East Germany's economic ills during the cold war. After the fall of the Berlin Wall only about a dozen science spies were convicted, and with very light sentences, despite thousands of investigations — West German prosecutors considered the damage to their industry minimal. As a former intelligence officer once told me, the sector was no "magic shop".

After 1989 there was a large-scale takeover of East German science and industry by the West. Despite pockets of innovation, the rotting East German infrastructure needed a massive infusion of modernity to be internationally competitive. Notable is the success of Silicon Saxony, one of Europe's leading microelectronics locations with hundreds of firms and tens of thousands of employees, built on the ashes of East Germany's computer-industry efforts.

This does not mean that science through espionage always fails. During the cold war, Japan proved itself far more adept at turning secrets into products in the electronics industry. But it had a much stronger economy.

Despite the HVA's lack of long-term success, one thing should be remembered: East German intelligence did not act alone. Science and technology had been at the top of the KGB's shopping list since the beginning of the cold war and it continues to be a priority for the KGB's successor foreign intelligence service, the SVR. Although the Eastern bloc countries did not always successfully integrate stolen secrets into their industrial bases, military secrets, such as plans from the Tornado fighter jet, could have been used against the West had the cold war turned hot. ■

Kristie Macrakis is professor of history, technology and society at the Georgia Institute of Technology in Atlanta. Her recent books include: *Seduced by Secrets: Inside the Stasi's Spy-Tech World* (Cambridge University Press, 2008) and she is co-editor of *East German Foreign Intelligence* (Routledge, 2009).
e-mail: macrakis@gatech.edu

For online special, see http://tinyurl.com/eeurope.

All in the family

PICTURE ALLIANCE/PHOTOSHOT

One of the Stasi's most prized agents was Dieter Feuerstein (pictured), a West German national born in 1955. During the 1970s he became a leftist, but his father cryptically told him there were "other ways to achieve political goals". When Feuerstein's father died shortly before he graduated from high school, he contacted his parent's 'friends' in Berlin who confessed their real jobs as Stasi officers.

He was recruited in 1974, given the code-name 'Petermann' and carried on the family tradition of spying. His wife also worked for the Stasi. Although Feuerstein had excelled in the humanities and social sciences, his case officers steered him towards natural sciences. They suggested he enrol in a machine-building course at the Technical University of Berlin and then switch to air and space technology. It took him eight years to complete his undergraduate degree because he found the subject so difficult.

In 1984 he got a job at aerospace company Messerschmitt, Bölkow-Blohm in Munich. Even though he worked for a weapons manufacturer, his case officer at the time frequently reminded him that he was working for peace. "When the enemy is not there anymore, they can't disturb the peace," the case officer told him. Feuerstein worked on the high-speed anti-radiation missile project and the Tornado fighter jet. Because East Germany couldn't use the military material he stole, it was passed on to the Soviet Union.

According to West German counter-intelligence, in 1990 the Soviet Union wanted to take over the HVA's best agents, including Feuerstein. In a scene reminiscent of a cold war spy novel, a KGB-driven black limousine pulled up beside him on a Berlin street but he never got in. Instead, his case officer defected to the West and turned him in. Feuerstein was convicted of espionage and spent four years of an eight-year sentence in prison. He now works as a patent researcher. **K.M.**

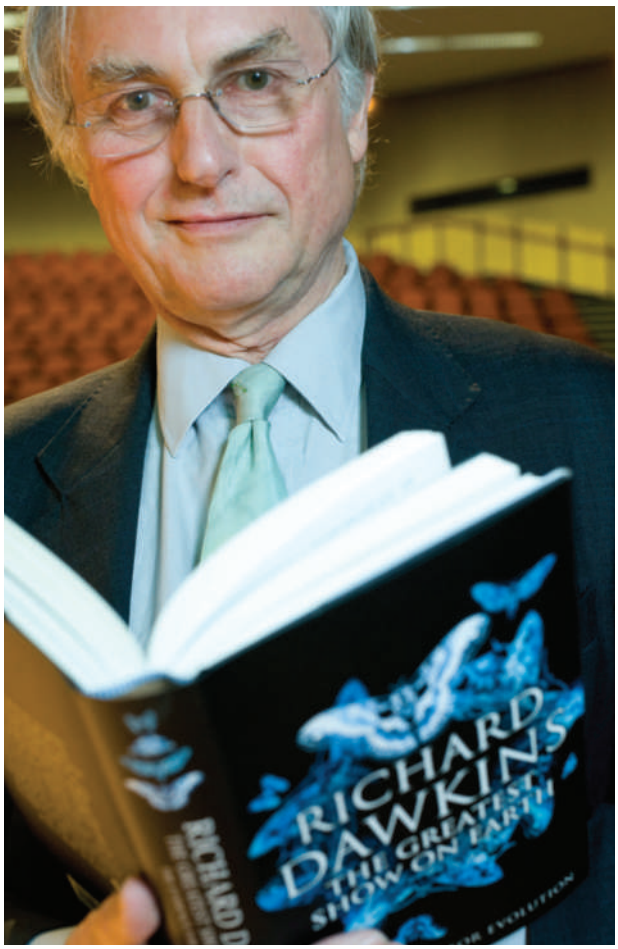
BOOKS & ARTS

Showcasing the evidence for evolution

Laurence D. Hurst compares two seasoned authors' strategies for explaining the difference between evolution fact and fantasy — Richard Dawkins's thunder and Carl Zimmer's poise.

I am usually reluctant to claim anything to be a 'fact'. Likewise, I shy away from the word 'proof' beyond the certainties of mathematics. But, as Richard Dawkins makes plain in *The Greatest Show on Earth*, cautious scientists like me should get off the fence. Our enterprise has established facts and we should have the confidence to say so. Evolution is one such fact, and the evidence for it is laid out in two new books by Dawkins and Carl Zimmer.

The Greatest Show on Earth is Dawkins on top form: unambiguous, beautifully argued, with prose flowing like quicksilver. He starts, as Charles Darwin did in *On the Origin of Species*, by coaxing the reader down a logical path, from artificial selection and domestication, through the ubiquity of the heritability of traits, to the inevitability of evolution by natural selection. Unlike many of his previous books, the logic of evolution plays second fiddle to the evidence for it and how that evidence is obtained. He explains what we have learnt from studying the relatedness between species (phylogenetics) and the way organisms develop (embryology), from the fossil record (palaeontology) and from the geographical distribution of species (biogeography). He is especially good on the unintelligent design of organisms.



Richard Dawkins's latest work is about more than just the theory.

N. CUNARD/EYEVINE

As Dawkins makes clear, evolution is testable and has survived every test. Sometimes we need, he says, to be a detective, uncovering the incriminating evidence of an unseen event. We can also, through genetics and through experimental evolution in the laboratory and field, witness evolution directly and in real-time. Darwin surely would have been thrilled to read of *in vitro* evolution of new complex traits, such as citrate metabolism in bacteria.

The Greatest Show on Earth is not just about the facts. Dawkins asserts that his book is not anti-religious, although he hardly hides his opinions. He never wastes an opportunity to knock what he sees in creationists as intellectually corrupt. Some will criticize him for being uncompromising, but I applaud him for taking such a categorical — and entertaining — stance. Dawkins emerges like a prize-fighter, knocking out of the ring all objections.

The Greatest Show on Earth: The Evidence for Evolution
by Richard Dawkins
Free Press/Bantam Press: 2009. 480 pp/406 pp. $30/£20

The Tangled Bank: An Introduction to Evolution
by Carl Zimmer
Roberts: 2009. 394 pp. $59.95/£29.99

Carl Zimmer's approach to explaining evolution in *The Tangled Bank* is rather different. In a non-confrontational way, he lays out the evidence for all to see. His prose, while authoritative and easy to read, is poised rather than animated. Dense with facts, the book is billed as the first textbook on evolution for the general reader, and in that framework it excels. Zimmer doesn't counterpoint the facts of evolution with creationist assertions but biblical literalism stalks the pages like the elephant in the room.

By necessity both authors cut corners. Both, for example, fail to explain why proteins accumulate changes at a steady rate regardless of an organism's lifespan, by using the same simplified explanation that predicts more rapid changes in short-lived species. While arguing that changes can accumulate in genomes by chance and irrespective of selection, Dawkins argues that 95% of our genome is effectively useless. This ignores unexpected recent findings that the majority of our genome is active making RNA transcripts, even though only a few of these go on to specify proteins.

A greater weakness is that neither book offers much perspective on unanswered questions. A reader could be forgiven for thinking that evolutionary biologists today are doing little more than dotting the proverbial 'i's. Although the fact of evolution is beyond doubt, the recent explosion of genetic sequence data means that the skills of evolutionists are needed now more than ever.

So which is the better strategy for explaining the difference between fact and fantasy: that of the quiet American or that of the British Rottweiler? Zimmer makes the facts palatable but Dawkins reminds us that some people persist in their beliefs even if they are profoundly contradicted by observation and logic. He relates his interview with a creationist who, when asked whether she had seen hominid skulls in the museum, asserts that such skulls simply do not exist. Such blithe disrespect for scientific evidence and for historical fact should concern us all. ■

Laurence D. Hurst is a professor of evolutionary genetics at the University of Bath, Bath BA2 7AY, UK.
e-mail: l.d.hurst@bath.ac.uk

Taxonomy comes of age

Naming Nature: The Clash Between Instinct and Science
by Carol Kaesuk Yoon
W. W. Norton: 2009. 352 pp. £19.99/$27.95

Faced with the bewildering diversity of nature, humans have long attempted to classify it. In her personal and readable perspective on taxonomy, Carol Yoon argues that the basic instinct to recognize natural groups has gradually been replaced by an increased rationality.

Yoon explores the formative environment, motives and personality of the field and scientists within it, from its origins to maturity. She argues that taxonomy generates new views of the world that are counterintuitive, for example that fungi are more like animals than plants; reductionist, depending on molecular data and statistical techniques; and bewildering, such as the prolific yet invisible domains of bacteria and archaea. She is concerned that the professionalization of biology has distanced humanity from nature. But her incursions into anthropology, neuroscience and psychology to explain our empathy with nature are disappointingly superficial.

The first major step in taxonomy was taken by Carl Linneaus some 250 years ago. In a magnificent work, *Systema Naturae*, he ordered the known natural world. Although the places of some organisms in his scheme have changed, his binomial system for naming organisms has stood the test of time.

A century later, in 1859, came Darwin's theory of evolution, with its seismic impact on science and society. Surprisingly, its effect on taxonomy was minimal. Taxonomists continued for another century to explore the world and discover new species as before, now drawing tree diagrams to represent the relationships between organisms and their evolution. Some of these evolutionary trees were insightful, others fanciful. Unfortunately none was testable, repeatable or objective. So taxonomy got a bad name.

Yoon rightly recognizes the role of the 1960s movement of quantitative experimentalists in being the most significant first step in transforming the field into a full science. These numerical taxonomists started a serious debate on how to articulate earlier opinions and experience, and distil them into testable conclusions. As a result, data and transparent methodology came to challenge 'expert' opinion. However, these researchers treated all data equally, without considering their relative evolutionary or biological significance.

A more rational and evolution-based approach was needed. For some this was found in DNA sequences. But this approach attracted a similar criticism, namely that a complex organism couldn't be reduced to a string of nucleotides.

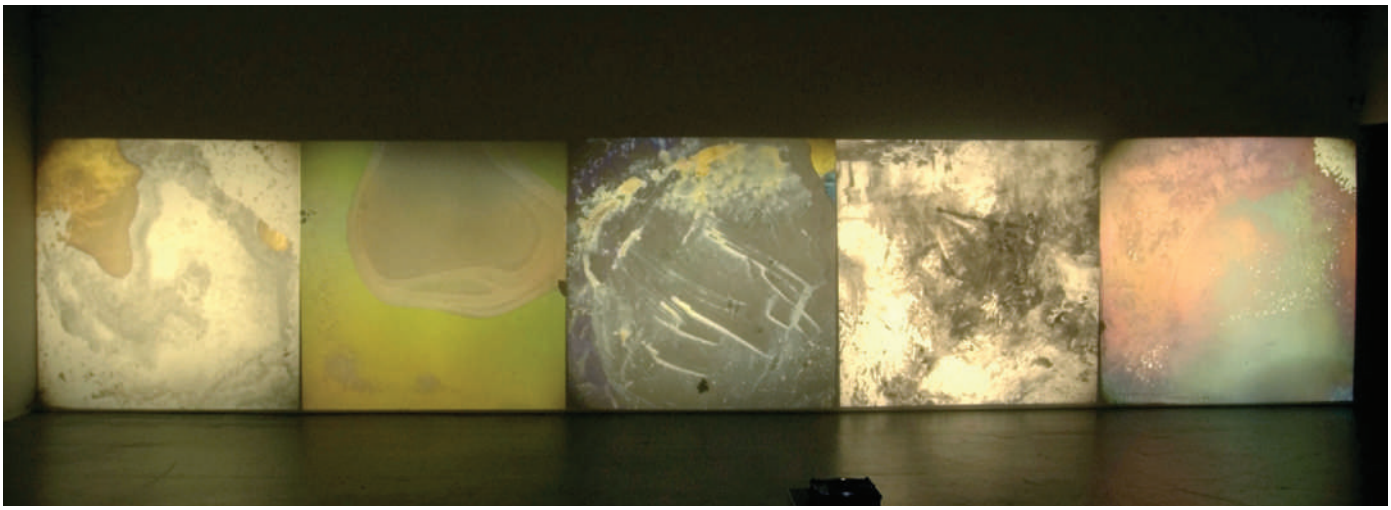
Over the following period of great change in biology and computational technology, taxonomists struggled to find their way. In the 1970s, the phylogenetics movement arose with a simple message: to seek close relatives with similarities that are shared uniquely by their descendants and no other groups. With this powerful new methodological tool came fervent disciples, called 'cladists', who wanted to purge the field of non-adherents. Taxonomy again turned in on itself, ignoring the huge strides being made in areas such as developmental genetics, population genetics and population biology that could have crucially informed their interpretations of nature. Describing these ideological waves, Yoon's book comes alive with her tales of meeting influential figures, including the God-like Ernst Mayr.

Yoon leaves us with taxonomy having passed though its rebellious adolescence into maturity. But what of its future? Taxonomy has never had so many tools at its disposal and so many connections to the rest of biology. The tragedy is that there are too few taxonomists who can utilize these advances and tackle the challenging questions of our time. ■

Richard Lane is director of science at the Natural History Museum, London SW7 5BD, UK, and author of the recent report *Taxonomy in Europe in the 21st Century*.
e-mail: r.lane@nhm.ac.uk

N. MICHALKE/PHOTOLIBRARY

Collections such as this at Berlin's Museum of Natural History may inspire taxonomists of the future.



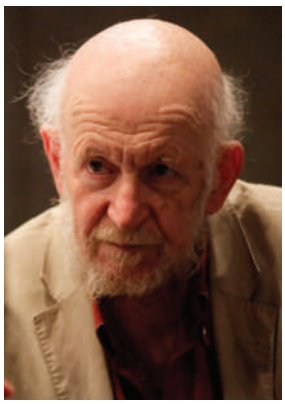
PHOTOGRAPHY FBM STUDIO ZURICH/2009 GUSTAV METZGER

Q&A: Gustav Metzger on destruction

Gustav Metzger's monumental and technical artworks comment on the capacity of human society to obliterate itself. From displays that eat themselves with acid to liquid-crystal patterns projected onto performing bands such as The Who in the 1960s, he questions environmental degradation, nuclear war and capitalism. As a major retrospective of his work opens, Metzger argues that scientists should be more active in counteracting society's tendency to seek oblivion.

Why is the destruction of nature an important theme for you?
My childhood experience of visiting the huge forest near Nuremberg, where I was born in 1926 before moving to England when I was 13, had a very strong impact on me. And a Jewish orthodox upbringing stresses value judgement and concern with life. Being a Jew, making judgements, arguing about what to do and how to live: all of this has formed me.

Do you think science is broadly constructive or destructive?
We human beings wouldn't have survived without science and technology, so it is constructive. But the physics that has led to nuclear weapons has occupied me in particular. Thinking about how humans ever came to conceive let alone construct nuclear weapons, I concluded that there must be destructive forces inside us.



It is a desperate situation — more desperate now than in the 1960s. Many small states are acquiring nuclear weaponry and the biggest units on the planet, China and India, are forging away in a direction that we recognize will destroy us. Science is in the middle of it all, but it is capitalism that has fuelled its destructive capacity. Technology is paid for by people who want to grow, to catch up, and in the process they are destroying life.

Scientists are trained to be impartial observers. Do you think more scientists should become politically active in protecting the planet?
We are each responsible and need to interact with society to try to stem its destructivity. I appeal to scientists to do just that. Some scientists have warned us, like those who got together after the Second World War to try to stop atomic war. But we need to look at the crisis in society from the standpoint of the existing system. Without bringing in a critique of the capitalism at the centre, nothing will happen. Scientists should take a stand against the dangers inherent in the capitalist system.

Why does the technical aspect of working with materials appeal to you?
I moved into technology as separate from the act of painting at the end of 1959. In order to develop auto-destructive art I had to involve myself with certain techniques. The liquid-crystal work started from seeing them on a cover of *Scientific American* in the spring of 1964. It led me to project the colour pictures onto screens. The years 1966–67 were among the best of my life, with the excitement of new techniques, of learning from science and applying the technology.

This year at the Manchester International Festival, you planted 21 willows upside down in concrete, their roots taking on the appearance of stunted branches. Why?
It is not just a question of showing destruction, how we misuse nature. Beyond that there is a potential to find beauty in this chaos, in this overturning of normality. In the period that the work was up, the roots were covered with a kind of brown skin, and it began to fall off and reveal a white structure underneath. As the years go by, it will change.

B. JOHNSON

Do you have hope for the future — will nature triumph?
I am afraid the answer is no. Let us take the word hope away. As I see the future, it is exceedingly grim. ■

Interview by **Emma Marris**, a writer based in Columbia, Missouri.

Gustav Metzger: Decades 1959–2009
Serpentine Gallery, London
Until 8 November

NEWS & VIEWS

DRUG DISCOVERY

Propping up a destructive regime

Randall T. Peterson

The Wnt signalling pathway balances the opposing activities of two proteins to transmit signals within cells. An inhibitor that stabilizes one of these proteins reveals a new target for anticancer drug development.

Cancer is often caused by unwanted activation of the Wnt signalling pathway, a group of proteins that work together to transmit signals from the outside of a cell to its nucleus. This is especially true in the colon, where mutations in a Wnt-pathway component called APC (adenomatous polyposis coli) activate signalling and are associated with the majority of colon cancers[1]. Compounds that prevent Wnt signals from reaching the cell nucleus are therefore attractive candidates for treating colon cancer and other Wnt-dependent cancers.

Few promising candidates have been discovered, however, largely because the known components of the Wnt pathway are difficult drug targets. On page 614 of this issue, Huang *et al.*[2] reveal a compound, known as XAV939, that potently inhibits Wnt signalling through a hitherto-unknown component of the Wnt pathway. The finding of a 'druggable' member of the pathway could change our fundamental understanding of Wnt signalling and provide an entrance point for finding drugs that target Wnt-dependent cancers.

The Wnt signalling pathway is fascinating, not only because of the biological functions it controls, but also because of its unusual mechanism for transmitting information to the cell nucleus. Most pathways do this through a chain of enzymes known as kinases, which work by transferring phosphate groups to molecules, and which are relatively easy to inhibit. By contrast, the Wnt pathway transmits signals by controlling the relative stabilities of two proteins[3]. The first of these is axin, a scaffold protein on which is built a 'destruction complex' that destabilizes the second protein, β-catenin. When a cell comes into contact with Wnt proteins, axin is destabilized, the destruction complex is inactivated, and so β-catenin is stabilized. The stabilized β-catenin can then migrate to the nucleus, where it turns on a specific set of Wnt target genes. Axin and β-catenin thus exist in opposition to each other, with stabilization of one typically linked to destabilization of the other[3].

Given the dearth of kinases and other obvious drug targets along the Wnt pathway, Wnt inhibitors have been hard to come by. Nevertheless, compounds that disrupt interactions between Wnt-pathway proteins have been identified, and some anti-inflammatory drugs, known as COX inhibitors, have been shown to diminish Wnt signalling, albeit by poorly understood mechanisms[4]. Compounds that affect the production of Wnt proteins and axin stability were also reported earlier this year[5]. Each of these classes of compound has its own strengths and shortcomings as potential therapeutic agents, but perhaps most impressively, COX inhibitors have been shown to reverse the growth of colorectal polyps (precursors of colorectal cancer). This seems to validate the Wnt pathway as an important anticancer target, and provides an impetus for discovering compounds that act elsewhere in the pathway.

To identify new inhibitors of Wnt signalling, Huang *et al.*[2] used a human cell line that was engineered to glow when Wnt signalling is activated. This enabled the authors to screen a large collection of drug-like molecules to identify those that could block Wnt signalling. One of the compounds to do so was XAV939, which seemed to be a potent stabilizer of axin. But how exactly does it work?

To determine the stabilization mechanism, Huang *et al.* performed an impressive proteomics analysis of 700 proteins that bound to XAV939-coated beads. They found that the binding of 18 of these proteins to the beads could be disrupted by excess XAV939 in solution, indicating specific binding of these proteins to the inhibitor. Of these eighteen, two — the related proteins tankyrase 1 and 2 — bound directly to axin. When Huang and colleagues knocked down tankyrase function in cells using small interfering RNAs, this stabilized axin. Taken together, these results suggest that the tankyrases might be the targets of XAV939, and could thus represent previously unknown components of the Wnt signalling machinery.

Tankyrases are enzymes that modify proteins by attaching several ADP nucleotides to them, a process called PARsylation. These enzymes had previously been implicated[6] in such cellular processes as protecting the ends of chromosomes (telomeres), insulin responsiveness and spindle assembly during mitosis, but no prior link to the Wnt pathway had been reported. In the case of telomere protection, the PARsylation of tankyrase substrates leads to their ubiquitination — tagging with the protein ubiquitin, which marks the substrate

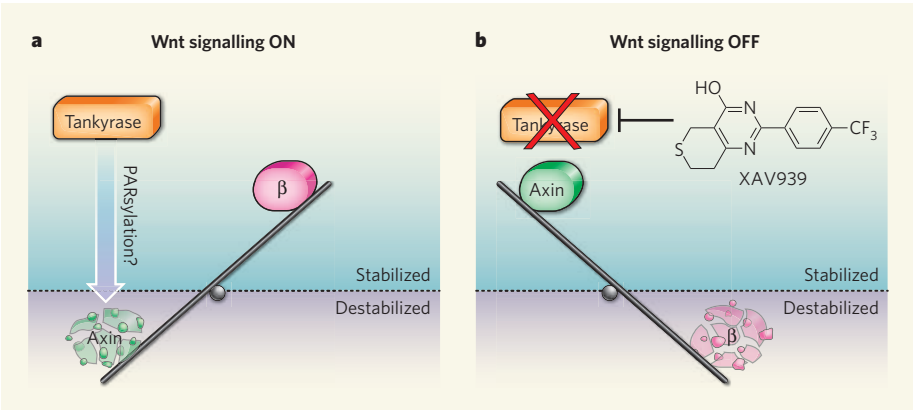


Figure 1 | The Wnt-signalling seesaw. The Wnt signalling pathway consists of a group of proteins that transfers signals from the outside of a cell to its nucleus. A delicate balance exists between two components of the pathway, axin and β-catenin (β). **a**, When axin is destabilized and degraded, β-catenin is stabilized and Wnt signalling is turned on. Huang *et al.*[2] report that tankyrase enzymes are components of the Wnt pathway. The tankyrases destabilize axin, perhaps through a biochemical modification called PARsylation. **b**, The tankyrase inhibitor XAV939 seems to block tankyrase-mediated destabilization of axin. This stabilizes axin, so that β-catenin is destroyed and Wnt signalling is turned off.

for subsequent degradation — hinting that tankyrases might also regulate axin stability by PARsylation and ubiquitination. Consistent with this idea, Huang *et al.*[2] demonstrated that tankyrase 2 directly PARsylates axin *in vitro*, and that axin PARsylation occurs *in vivo*. They also found that treatment of cells with XAV939 reduced the PARsylation, ubiquitination and degradation of axin. These findings led the authors to propose that tankyrases are essential components of the Wnt signalling pathway, and that they promote axin degradation by PARsylating it directly (Fig. 1).

Naturally, questions remain about the details of tankyrase involvement in Wnt signalling. For example, is tankyrase activity constitutive — does it simply remove axin from an active pool at a constant rate — or is it regulated? If regulated, what factors influence tankyrase activity, and how? It will also be crucial to determine how widespread the effects of tankyrase inhibition are in cells. The answer to this question will probably influence the usefulness of tankyrase inhibitors as drugs: if tankyrase activity is widespread and affects several biological processes, it may be difficult to achieve specific, therapeutic inhibition of Wnt signalling without causing undesirable side effects. In fact, a growing list of tankyrase substrates indicates the enzymes' involvement well beyond the Wnt pathway, as does the fact that the embryos of genetically engineered mice that lack both tankyrase 1 and tankyrase 2 die before birth[7].

Conversely, several of the reported functions of tankyrases — such as their involvement in the elongation of telomeres — favour cancer progression, so inhibition of non-Wnt tankyrase activities could theoretically complement Wnt inhibition in cancer chemotherapy[8]. The results of future animal tests of tankyrase inhibitors will therefore be of great interest. In the meantime, Huang and colleagues' demonstration[2] that XAV939 can inhibit proliferation of APC-deficient colorectal cancer cells serves as an encouraging hint of the therapeutic possibilities of tankyrase inhibitors. ■

Randall T. Peterson is at the Cardiovascular Research Center, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts 02129, USA, and at the Broad Institute, Cambridge, Massachusetts.
e-mail: peterson@cvrc.mgh.harvard.edu

1. Powell, S. M. *et al. Nature* **359,** 235–237 (1992).
2. Huang, S.-M. A. *et al. Nature* **462,** 614–620 (2009).
3. Tolwinski, N. S. & Wieschaus, E. *Trends Genet.* **20,** 177–181 (2004).
4. Barker, N. & Clevers, H. *Nature Rev. Drug Discov.* **5,** 997–1014 (2006).
5. Chen, B. *et al. Nature Chem. Biol.* **5,** 100–107 (2009).
6. Hsiao, S. J. & Smith, S. *Biochimie* **90,** 83–92 (2008).
7. Chiang, Y. J. *et al. PLoS ONE* **3,** e2639 (2008).
8. Seimiya, H. *Br. J. Cancer* **94,** 341–345 (2006).

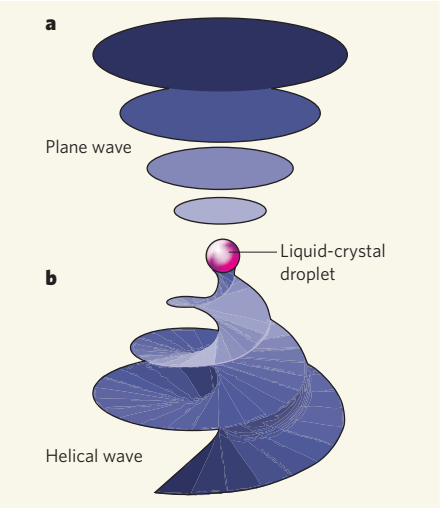


Figure 1 | Light with a twist. a, In plane waves of light, the wavefronts form planes that are separated from each other by a distance of one optical wavelength. **b,** Brasselet *et al.*[2] pass plane waves of circularly polarized light (in which the optical field of the light rotates around the axis of the beam) through specially prepared droplets of liquid crystals to make helical waves. Here, the wavefront forms a helix, which confers orbital angular momentum on the light. (Figure adapted from ref. 2.)

OPTICS

Droplets set light in a spin

Miles Padgett

Fusilli pasta is made by extruding dough through an appropriately shaped hole. A new method for making similar shapes in the optical field of light involves passing laser beams through droplets of liquid crystals.

The fact that certain forms of polarized light can carry a spin angular momentum has been known since the early twentieth century[1] — today we associate this quantity with individual photons. But light can also carry an orbital angular momentum arising not from its polarization, but from its phase profile. Such beams have many uses in optical manipulation, imaging and even information processing, but generating or measuring these beams requires specialist lenses or holograms. Brasselet *et al.*[2] now report in *Physical Review Letters* that such beams can be made simply by focusing a laser through a microscopic liquid-crystal droplet.

The spin angular momentum of light is associated with circular polarization, in which the optical field of a light beam rotates around the beam's axis. Circularly polarized light is described as being right-handed or left-handed, depending on the direction of rotation. At every point in the cross-section of a normal laser beam that has circular polarization, the waves of the optical field are in step and rotate together. This produces a 'plane wave', in which the wavefronts are parallel planes, each separated from the next by a distance of one optical wavelength (Fig. 1a). But in a beam carrying orbital angular momentum, the wavefronts instead form one or more continuous helices (Fig. 1b). If there is a single helix, the wavefront looks like a screw thread; if there are two helices, the wavefront looks like DNA; and for three helices, the wavefronts have the shape of fusilli pasta. In fact, beams can be made with any number of helices — the more helices there are, the larger the orbital angular momentum.

The first laser beams that carried orbital angular momentum were made in 1992, by passing a normal laser beam through a system of lenses[3]. This opened the door to further studies, and led to the discovery a few years later that light beams carrying orbital angular momentum can act as 'optical spanners' that rotate microscopic objects[4]. Orbital angular momentum is now known to underpin many phenomena, including the Doppler shifts of spinning bodies, certain forms of Heisenberg's uncertainty principle and manifestations of quantum entanglement[5].

Orbital angular momentum is still almost always introduced into normal laser beams using converter devices — usually, large optical components such as lens systems, holograms or precisely machined spiral wedges of plastic or glass. On a smaller scale, micromachining techniques have been used to fabricate miniature converters, primarily for use in optical tweezers[6] (laser beams that can trap and move microscopic objects). The beauty of Brasselet and colleagues' system[2] is that the high purity of the beam it produces is a natural consequence of the internal orientation of the molecules in the liquid-crystal droplet. This orientation arises from the conditions and reagents used to prepare the droplets, and so no complicated set-ups or machining techniques are required.

So how exactly does it work? Brasselet and colleagues' liquid-crystal material is birefringent, which means that horizontally polarized light travels through it at a different speed from vertically polarized light. This effect is used widely in optics to make waveplates that transform the polarization state of light. A half-waveplate reflects the electromagnetic field of a beam about the optic axis of the crystal, transforming right-handed circularly polarized light into left-handed, and vice versa. Brasselet and colleagues' droplets act as half-waveplates, but with an added twist. Not only do they interconvert left- and right-handed circularly polarized

light, but the transmitted light also undergoes a geometric Pancharatnam–Berry phase delay — a change in phase that depends on the orientation of the optic axis of the liquid crystal. Because of the way in which the optic axes are orientated within the droplets, laser beams emerge with a helical wavefront (Fig. 1b), and hence with an orbital angular momentum. Pancharatnam–Berry phase delays have previously been used in macroscopic light-mode converters based on liquid crystals[7], but never before has the effect been a natural consequence of microscopic droplet structure.

A surprising feature of Brasselet and colleagues' microscopic converter is that it works over a wide range of optical wavelengths — a feat previously made possible only using combinations of optical components[8]. In their present form, however, the inherent structure of the droplets[2] means that the resulting beam contains only two intertwined wavefronts, whereas traditional approaches can generate any number of them. The challenge now will be to extend the droplet approach to yield larger numbers of intertwined wavefronts, and to construct a robust, miniature converter that can be used in practical applications. Given the apparent purity of the beams produced using Brasselet and colleagues' strategy, this is a challenge well worth pursuing. ■

Miles Padgett is in the Department of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, UK.
e-mail: m.padgett@physics.gla.ac.uk

1. Poynting, J. H. *Proc. R. Soc. Lond. A* **82,** 560–567 (1909).
2. Brasselet, E., Murazawa, N., Misawa, H. & Juodkazis, S. *Phys. Rev. Lett.* **103,** 103903 (2009).
3. Allen, L. *et al. Phys. Rev. A* **45,** 8185–8189 (1992).
4. He, H., Friese, M. E. J., Heckenberg, N. R. & Rubinsztein-Dunlop, H. *Phys. Rev. Lett.* **75,** 826–829 (1995).
5. Franke-Arnold, S., Allen, L. & Padgett, M. *Laser Photon. Rev.* **2,** 299–313 (2008).
6. Knoner, G. *et al. Opt. Express* **15,** 5521–5530 (2007).
7. Marrucci, L., Manzo, C. & Paparo, D. *Phys. Rev. Lett.* **96,** 163905 (2006).
8. Leach, J. & Padgett, M. J. *New J. Phys.* **5,** 154 (2003).

PALAEONTOLOGY

Feathered dinosaurs in a tangle

Lawrence M. Witmer

A dramatic feathered dinosaur fossil from the Jurassic of China resolves a 'temporal paradox'. But it adds intriguing complications to the debates on the evolution of feathers and flight in birds.

Birds are dinosaurs. That's hardly the stuff of headlines any more, as data have streamed in revealing anatomical similarities between birds and the theropod dinosaurs from the tips of their noses to the tips of their feathered tails. More elusive have been the details of the transition to birds and the evolution of flight. On page 640 of this issue, Hu and colleagues[1] present a spectacular new specimen of the feathered theropod *Anchiornis huxleyi* that solves some problems. But it simultaneously creates new ones, revealing what a gloriously messy business it is to tease apart the evolutionary tangles that we retrospectively anoint as an 'origin'.

Anchiornis is a small, crow-sized theropod, assigned to a group known as the troodontids (Fig. 1), which in life was covered with long bird-like feathers. The new fossil, like other, more poorly preserved specimens, was collected from the Tiaojishan Formation of Liaoning, China. Liaoning Province has yielded many specimens of feathered theropods and true birds[2], and so it might seem that yet another feathered dinosaur shouldn't merit much attention. But what's important about the fossils of *Anchiornis* is their age — they are from the Jurassic period, and at about 155 million years old are much older (by about 25 million to 35 million years) than the other feathered Liaoning theropods, which come from Early Cretaceous rocks. Even more significantly, *Anchiornis* is older (by 5 million to 10 million years) than the iconic 'first bird' *Archaeopteryx*, which comes from younger Jurassic rocks in Germany.

One lingering problem with the hypothesis that birds descended from dinosaurs had been that the most bird-like theropods occurred later in time than did *Archaeopteryx*. It has been argued that this 'temporal paradox' (how can a 'descendant' arise before an 'ancestor'?) both invalidates the theropod ancestry of birds[3] and, reversing the ancestor–descendant relationship, suggests that some of the Cretaceous bird-like theropods actually descended from Jurassic *Archaeopteryx*-like birds[4,5]. In truth, the temporal paradox never seriously challenged the theropod hypothesis, because it essentially assumed that fossils like *Anchiornis* wouldn't be found — arguments based on negative evidence are always dicey. However, the notion of some Cretaceous theropods being secondarily flightless descendants of early birds remains a valid hypothesis given the common and repeated evolution of flightlessness in birds[6].

Anchiornis resets that whole debate. By predating *Archaeopteryx*, *Anchiornis* shows that bird-like feathered theropods were around 'early enough' to serve as ancestors, although no one is suggesting that *Anchiornis* itself was that ancestor. Thus, there is no temporal paradox. At the same time, the idea that some Cretaceous theropods might be flightless descendants of early birds can now be assessed on the basis of evidence rather than being mandated by temporal congruence. But *Anchiornis* does more than refute the temporal paradox, in that the distribution of feathers on its body suggests that we now need to revise our thoughts on the evolution of flight.

In 2003, the description of *Microraptor gui* as a 'four-winged dinosaur'[7] rocked the palaeontological world. This creature, a basal member (that is, on an early-branching twig) of a different theropod lineage, the dromaeosaurids, had elongate, bird-like feathers not just on its arms but also on its legs and feet. This find was entirely unexpected in that, although feathered forelimbs and tails had been reported in various non-avian theropods, there was little reason to suspect that elongate pennaceous feathers (that is, with shaft and vanes, as in the flight feathers of living birds) occurred on the legs, let alone the feet. Later, elongate feathers were found on the legs and feet in *Pedopenna*, a basal member of the avialans, the group that includes *Archaeopteryx* and other birds[8].

Anchiornis, a basal troodontid, also has long feathers on its legs and feet to match those on its arms and tail, so the family Troodontidae now joins the Dromaeosauridae and Avialae

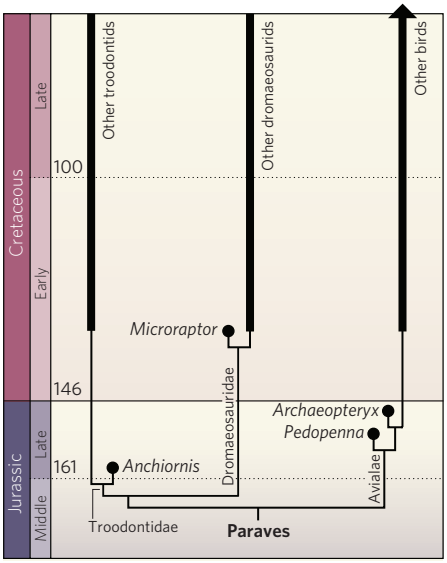


Figure 1 | *Anchiornis huxleyi* in context. The fossil described by Hu *et al.*[1] is assigned to the family Troodontidae, which together with the closely related Dromaeosauridae and Avialae comprise the Paraves (itself a subgroup of the theropod dinosaurs). One significant aspect of *Anchiornis* is that it predates *Archaeopteryx*, the iconic 'first bird', by some 5 million to 10 million years. Another is that it shows that basal members of all three of the Paraves groups — *Anchiornis*, *Microraptor* and *Pedopenna* — had long pennaceous feathers on their lower legs and feet, as well as on their hands and tail. The implication is that avian evolution conceivably went through a 'four-wing' stage. Numbers are approximate ages of the geological divisions in millions of years ago.

on the list of theropods with 'hind wings'. Among modern birds, only a few species have long feathers on their legs (the tibiotarsal region), and none has long, aerodynamically relevant feathers on their feet. So palaeontologists have been scrambling to make sense of what feathered legs and feet in basal birds and dromaeosaurids mean for the evolution of flight.

When we just had *Microraptor*, it was easier to dismiss the long foot feathers as potentially a mere early experiment in aerodynamics that was independent of the evolution of avian flight[9,10]. And indeed, it potentially biases the functional argument to refer to these elongate leg and foot feathers as 'flight feathers' that formed a lift-generating 'hind wing', in that other functions are conceivable[9] (such as display). Still, these elongate feathers would have had aerodynamic effects even if they had not evolved originally as flight adaptations, and credible aerodynamic models have been proposed[10]. But *Anchiornis* shows that *Microraptor* was no one-off, and that basal members of three separate theropod groups had long foot feathers.

This association of Troodontidae, Dromaeosauridae and Avialae is no chance occurrence. These groups together form a branch of the theropod evolutionary tree known as Paraves — they're each other's closest relatives (Fig. 1). The fact that basal members of all three groups had long pennaceous feathers on their lower legs and feet strongly suggests that the paravian common ancestor also had feathered feet. It's not at all certain yet whether these feathers comprised an aerodynamically competent flight surface that provided lift and/or thrust. But it's hard to imagine that they wouldn't have had some aerodynamic effects (drag, for instance).

More to the point, it now looks as if we'll have to accept that avian evolution indeed went through — at the risk of overstatement — a four-wing stage, only to eventually lose the long foot feathers. What this means for the evolution of the avian flight stroke[9] is now an open question. Likewise, we'll need to seriously consider how these otherwise seemingly very adept and agile runners (*Anchiornis* has extremely long and slender hindlimbs) could manage with long feathers on their feet.

Even apart from feathers and aerodynamics, Hu and colleagues' analysis[1] of *Anchiornis* reveals just how similar these early paravians were to each other: the basal members of each of the three groups seem to mix-and-match their anatomical similarities, so that their unique attributes are becoming more and more subtle. It's getting hard to tell members of one group from another. On the bright side, in this year of Darwin, that fact provides a comforting affirmation of the evolutionary prediction that species in different groups will become increasingly similar as we approach their common origin. ■

Lawrence M. Witmer is in the Department of Biomedical Sciences, Ohio University College of Osteopathic Medicine, Athens, Ohio 45701, USA.
e-mail: witmerL@ohio.edu

1. Hu, D., Hou, L., Zhang, L. & Xu, X. *Nature* **461,** 640–643 (2009).
2. Norell, M. A. & Xu, X. *Annu. Rev. Earth Planet. Sci.* **33,** 277–299 (2005).
3. Feduccia, A. *The Origin and Evolution of Birds* (Yale Univ. Press, 1996).
4. Paul, G. S. *Dinosaurs of the Air* (Johns Hopkins Univ. Press, 2002).
5. Feduccia, A. *et al. Auk* **124,** 373–380 (2007).
6. Witmer, L. M. in *Mesozoic Birds* (eds Chiappe, L. M. & Witmer, L. M.) 3–30 (Univ. California Press, 2002).
7. Xu, X. *et al. Nature* **421,** 335–340 (2003).
8. Xu, X. & Zhang, F. *Naturwissenschaften* **92,** 173–177 (2005).
9. Padian, K. *BioScience* **53,** 450–452 (2003).
10. Chatterjee, S. & Templin, R. J. *Proc. Natl Acad. Sci. USA* **104,** 1576–1580 (2007).

SUPRAMOLECULAR CHEMISTRY

Molecular crystal balls

Seth M. Cohen

Sorcerers have long gazed into crystal balls to conjure up information. Chemists are also getting in on the act, using porous crystals to trap unstable reaction intermediates and to reveal their structures.

The reactions of molecules with one another often proceed through intermediates before the final products are formed. Such intermediates are frequently short-lived and unstable, which restricts our ability to characterize them. Typically, the identification of such fleeting species is limited to fast, time-resolved spectroscopic measurements. But in this issue (page 633), Kawamichi *et al.*[1] report that they have trapped an unstable chemical intermediate in a porous crystalline material, and were thereby able to characterize the structure of the intermediate unambiguously by X-ray crystallography. The authors suggest that such porous crystals can act as protective matrices within which chemical reactions can be performed, allowing us to peer into the details of reaction mechanisms in an unprecedented way.

Kawamichi *et al.* examined a reaction familiar to every student of organic chemistry: the combination of an amine and an aldehyde to form a Schiff base (Fig. 1a). Although the mechanism of this fundamental reaction has been extensively studied, direct observations of the ephemeral intermediate — a hemiaminal — are rare. The crystal structure of a hemiaminal trapped in the active site of an enzyme has been reported[2], but structure determination in protein crystals is not a general approach for characterizing reaction intermediates.

The authors[1] use a 'coordination network' of organic ligand molecules and metal ions to trap the elusive hemiaminal. Coordination networks — also known as porous coordination polymers or metal–organic frameworks — are solid, crystalline materials that generally

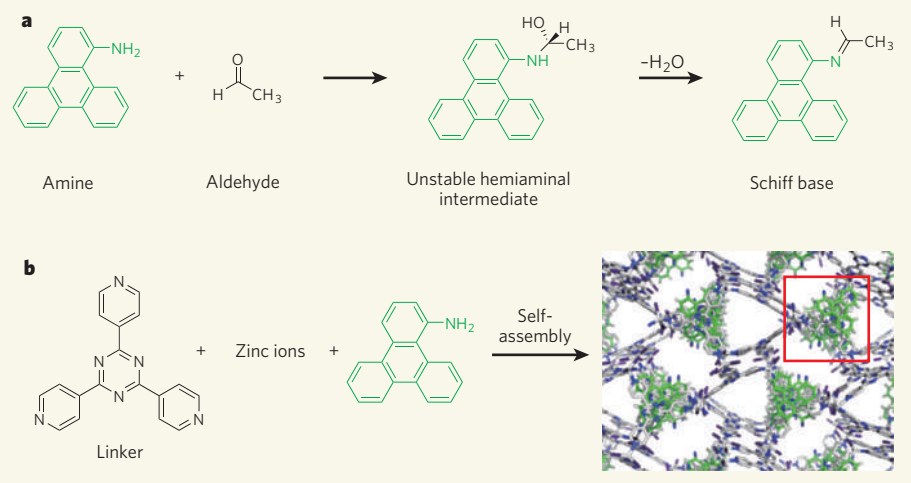


Figure 1 | Caught in a trap. a, Kawamichi *et al.*[1] have stabilized and observed the crystal structure of an elusive hemiaminal intermediate that is transiently formed during the reaction of an amine and an aldehyde to form a Schiff base. **b**, They did this by performing the reaction within the restrictive and orderly confines of a crystalline coordination network (a cage-like molecular structure) at low temperatures. The authors trapped amine reactants as guest molecules in the network as it self-assembled from its constituent parts — organic 'linker' molecules and zinc ions that act as nodes between the linkers. In the network structure, the amines are shown in green, and one is highlighted in the red box.

self-assemble from their molecular components (Fig. 1b). They have lattice structures that resemble climbing frames, in which the organic ligands (known as linkers) are the bars of the climbing frame, and the metal ions (nodes) are the rivets that connect the bars together. The result is a grand structure that consists largely of empty space. Just as children play within the lattice of a climbing frame, so molecules can diffuse and interact within the lattice of a coordination network.

Successful observation[1] of the hemiaminal relied on two key features of the network. First, it could trap the intermediate as part of a 'guest' molecule in its pores at low temperatures, stabilizing the species within a shielded, confined space. And second, the highly crystalline structure of the network restricts the trapped molecules to a specific periodic arrangement that allowed them to be characterized by X-ray crystallography. To generate the hemiaminal, Kawamichi *et al*. gently passed a solution of an aldehyde over an intact crystal of their coordination network in which an amine was tightly bound as a guest molecule. This was done with the crystal already mounted on a diffractometer at low temperature, allowing the hemiaminal to be trapped and its structure determined directly. Confined within the network, identification of the elusive species by X-ray crystallography was as easy as determining the structure of the network itself. The authors then raised the temperature of the crystal and re-determined its structure, whereupon they observed that the reaction had proceeded within the crystal, yielding the expected Schiff base product.

The process used by Kawamichi and coworkers — modifying pristine crystals of a coordination network with soluble chemical reagents — is an example of postsynthetic modification, a strategy that has garnered increased attention from chemists in recent years[3,4]. A growing number of research groups have been investigating the types of network and scope of chemical reactions that such networks can undergo without losing their highly ordered structure. But unlike most reports, which have used postsynthetic modification to generate networks possessing new features or functions, Kawamichi *et al*. have cleverly focused on the ability of coordination networks to provide a well-defined, ordered and confined space within which to perform chemical reactions. In previous studies[5], the same group highlighted this property of the networks by referring to them as "single-crystalline molecular flasks". The current use[1] of coordination networks logically and creatively extends their earlier work[6] on discrete molecular capsules, also comprised of organic ligands and metal ions, within which the authors exerted unusual control over chemical reactivity.

Kawamichi and colleagues' strategy will not be a panacea for characterizing all reaction intermediates. For example, only reactions that do not degrade the somewhat fragile coordination network (or those that occur in conditions that do not degrade the network) can be studied by this approach. Nevertheless, Kawamichi *et al*.[1] show that postsynthetic modification of coordination networks is a broadly applicable tool for isolating and characterizing transient intermediates that form during chemical reactions. It is likely that further studies will reveal the structure of other transient intermediates — organic and perhaps organometallic — that have been immobilized within porous networks. Performing chemical reactions within these crystals may also reveal new patterns of chemical reactivity, distinct from that observed in solution[7]. Furthermore, postsynthetic modification of coordination networks will probably have an increasing role in enhancing the stability, functionality and uses of these materials in real-world applications — for example, in the storage tanks of vehicles that use alternative transportation fuels such as hydrogen. ■

Seth M. Cohen is in the Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0358, USA.
e-mail: scohen@ucsd.edu

1. Kawamichi, T., Haneda, T., Kawano, M. & Fujita, M. *Nature* **461,** 633–635 (2009).
2. Heine, A. *et al. Science* **294,** 369–374 (2001).
3. Wang, Z. & Cohen, S. M. *Chem. Soc. Rev.* **38,** 1315–1329 (2009).
4. Song, Y.-F. & Cronin, L. *Angew. Chem. Int. Edn* **47,** 4635–4637 (2008).
5. Kawamichi, T., Kodama, T., Kawano, M. & Fujita, M. *Angew. Chem. Int. Edn* **47,** 8030–8032 (2008).
6. Yoshizawa, M., Klosterman, J. K. & Fujita, M. *Angew. Chem. Int. Edn* **48,** 3418–3438 (2009).
7. Jones, S. C. & Bauer, C. A. *J. Am. Chem. Soc.* **131,** 12516–12517 (2009).

BEHAVIOURAL ECOLOGY

Winged warnings

Graeme D. Ruxton

Alarm signals emitted by animals may not be all that they seem. But a good example has been identified in the whistling sound of a crested pigeon's wings when it takes flight in response to a predator.

In keeping with the principle that there's safety in numbers, animals gain various anti-predatory benefits from forming groups. One such benefit is that an individual in a group can be alerted to danger by a group-mate, rather than by detecting the predator itself. As they recount in *Proceedings of the Royal Society*, Hingee and Magrath[1] have compiled convincing evidence of an unusual example of this behaviour.

Intra-group alarm signalling obviously requires communication about predators to occur between group members. But the reliability of these alarm calls is often unclear[2], in that they could be given deceptively to reduce competition for food[3]. Although the evidence for deception is equivocal, false alarms may be common if there is little cost to emitting an alarm call when no predator is in fact present. In many species, group members that detect a predator do not give an obvious alarm vocalization, but simply flee.

In the limited previous experimentation[2] carried out on this subject, group-mates have apparently struggled to differentiate between individuals leaving the group because they've seen a predator and those leaving for other reasons. Hingee and Magrath[1], however, show that the rattle-like whistling sound generated by the flapping of a fleeing crested pigeon (*Ocyphaps lophotes*) can be reliably associated with flight in response to a predator, and that this information is used by group-mates to trigger their own anti-predatory behaviour.

The reliability of this signal stems from its nature: the whistling sound is generated by the movement of air across the wing (an audio file accompanies the article online). Because birds must take flight to produce the signal, the high energetic cost of flapping flight will discourage false alarms. Flock-mates can differentiate between the noises made by different types of flight, and in the authors' experiments they took flight in response to a playback of sound emitted by birds in alarmed flight but not to the sounds that birds make in other, non-predator-driven, departures from the flock. Test flocks took flight immediately in 11 of 15 trials with an alarm whistle and in none of 15 cases with a non-alarm whistle. Such non-alarm whistles tend to be of lower tempo and amplitude than alarm whistles.

Hingee and Magrath[1] also performed studies with alarm whistles whose amplitude had been experimentally reduced before being played to test flocks. In these experiments, flocks generally did not flee, but showed higher levels of vigilance immediately after the playback than did flocks subjected to control playbacks. However, even non-alarm whistles may be useful to flock-mates, because such flight sounds still induced an increase in vigilance. Indeed, there was a trend for vigilance to increase with increasing tempo of such flight sounds (that is, the more similar they became to alarm whistles). So the birds may not simply get information on whether or not an attack is imminent, but may also be able to extract useful information about the severity of risk.

Hingee and Magrath provide morphological evidence that the shape of the eighth primary

NATURE|Vol 461|1 October 2009

50 & 100 YEARS AGO

50 YEARS AGO

A critical review of the methods of determination of the temperatures of ancient seas by the measurement of the oxygen isotopes ratio in fossil calcareous organisms is given by Y. A. Birstein (*Priroda*, **5**, 21; 1959). It is based partly on the work published in the Soviet Union and it leads to certain new ideas regarding the origin of the deep sea fauna. Thus the author of this review is casting doubt upon the conclusions of C. Emiliani and C. Edwards ... regarding the sharp changes of sea temperatures during the late Tertiary era, and also about those of A. Fr. Bruun ... regarding the extinction of the deep sea fauna. According to the author all the deep oceanic regions must be considered to be regions of a relatively constant temperature, affording a place of refuge to many animal species which have eventually died out in the waters of a lesser depth.

From *Nature* 3 October 1959.

100 YEARS AGO

The discovery of Halley's comet at a time so far preceding the date of perihelion passage adds another proof of the great capacity of the photographic method. The interesting point to many observers is as to when the comet will become visible to them as a telescopic object. This must, of course, depend in a large measure upon the diameter of their glasses and on their powers of vision. After the present moon has left the sky, say during the second week in October, the comet ought to have increased in light sufficiently for it to be observed in a 12-inch telescope. The calculated magnitude of the comet will be 14½ on October 15, and its distance from the earth about 230 millions of miles ... The comet will be visible in an excellent position nearly all night during most of the winter, but will continue small and faint until it blazes out next April.

From *Nature* 30 September 1909.

feather in this species is unusual, and they link this feature to the characteristic noise of the bird in flight. Experimental manipulation to confirm this link would now be valuable, as it could have a bearing on the likely generality of this mode of anti-predatory communication and on its evolution within the crested pigeon.

If the unusual feather structure of this species is essential for signal production, then the whistle probably evolved because it modifies the behaviour of the signal receivers in such a way as to benefit the whistler. This might come about because the original detector of a predator benefits (through dilution of risk or sensory confusion of the predator) if other flock-mates flee along with it. There may also be particular benefits from warning a mate or kin. Or perhaps the intended receiver of the signal is the predator, if the signal reliably indicates that that whistler is already taking rapid avoidance flight and so the predator would be better focusing its attention on other flock-mates. Thus, other flock-mates might be eavesdroppers that gain useful information from a signal that evolved for other reasons. Finally, it could even be that the noise production originally evolved because it benefited the signaller in a context unrelated to predation; it might be relevant that the courtship behaviour of the crested pigeon involves opening and closing the wings.

As for the generality of this type of anti-predatory communication, it is certainly true that birds taking flight (Fig. 1) or mammals taking to their heels can both be noisy. The limited previous empirical study[2] does not provide strong evidence of a general ability of group-mates to distinguish between predator-induced and other departures from a group. However, Hingee and Magrath's research suggests that we should return to this question, and perhaps concentrate on the noise of fleeing as a signal that is difficult to fake, and on species that gather in structurally complex environments (such as scrubland) where noise might be easier to detect than visual evidence of a fleeing group-mate. No matter what, we have not heard the last of the whistling pigeon. ■

NATURE PRODUCTION/NATUREPL.COM

Figure 1 | Ambiguous signal. Are flock-mates taking flight in response to a predator?

Graeme D. Ruxton is in the Department of Ecology and Evolutionary Biology, University of Glasgow, Glasgow G12 8QQ, UK.
e-mail: g.ruxton@bio.gla.ac.uk

1. Hingee, M. & Magrath, R. D. *Proc. R. Soc. Lond. B* doi:10.1098/rspb.2009.1110 (2009).
2. Lima, S. L. *Anim. Behav.* **50,** 1097–1108 (1995).
3. Munn, C. A. *Nature* **319,** 143–145 (1986).

APPLIED PHYSICS

Lasers go nano

Francisco J. Garcia-Vidal and Esteban Moreno

Two experiments that produce laser light by exploiting the collective wave-like motion of free electrons on a metal surface bring the science and technology of lasers into the nanoland.

Let us imagine what the impact would be of having a fully operational laser of nanoscale dimensions — a nanolaser. On the applied front, myriad prospects would arise, among them the fabrication of ultra-fast photonic nanocircuits with unprecedented performance and the improvement of techniques such as nanolithography and single-molecule biochemical sensing. We could even envisage nanolasers as key elements in quantum-information technologies. On a more fundamental level, we would have a powerful tool with which to study strong interactions between light and matter. Two papers, one published in *Nature* recently[1] and the other in this issue (page 629)[2], report significant experimental advances towards the implementation of functional nanolasers.

In contrast to an ordinary torch, a laser emits an intense and strongly collimated beam of coherent, monochromatic light. These properties arise from the interaction between two of the laser's components, an optical cavity that sustains a photonic resonance and a gain medium that amplifies it. Energy is externally pumped into this amplifying medium and is later released as photons that, through a mechanism called stimulated emission and owing

to the laser cavity's optical-feedback property, are all emitted coherently — that is, they have the same frequency and are in step with one another.

The external input of energy typically involves an electric current (electrical pumping) or light illumination (optical pumping). In certain respects, lasing bears some resemblance to the tuning up of an orchestra before the start of a concert. Musicians 'pump' their instruments to emit a single frequency, 'stimulated' by the notes played by their fellows. As the tuning goes on, the concert hall is increasingly filled with sound (albeit not coherent) that may eventually spread out through the open doors of the hall.

The quest for ultra-compact lasers of nanometre scale has faced a fundamental obstacle: the diffraction limit. By conventional means, light cannot be focused down to a spot smaller than around half its wavelength. For visible light, whose wavelength ranges between 300 and 800 nanometres, the smallest dimension of both the spot and the laser cavity is limited to a few hundred nanometres. A feasible way to overcome this limitation is by using a metallic cavity, which can confine light within dimensions markedly smaller than the light's wavelength.

This ability stems from the existence of collective, wave-like motions of free electrons on a metal surface, termed surface plasmons. These oscillations generate electromagnetic waves that are strongly localized at the surface. In this context, and by analogy with the laser, the idea of a spaser (surface plasmon amplification by stimulated emission), in which plasmons play the part that photons do in conventional lasers, was proposed[3,4]. Importantly, surface-plasmon oscillations can be coherently coupled to light outside the metallic cavity, and hence can be exploited to realize ultra-small lasers.

The main obstacle to attempts to produce a plasmon-based laser has been the (resistive) energy losses in the metal from which the cavity is made. Several experimental studies[5,6] have, however, demonstrated that the incorporation of an optical-gain medium into the system can compensate for such losses. Along these lines, Hill and co-workers[7] reported lasing from an electrically driven, metal-coated cavity with dimensions of 200–300 nm.

The present investigations, carried out by Noginov *et al.*[1] and by Oulton *et al.*[2], provide a major step forward for plasmon-based nanolasers. In Noginov and co-workers' experiments[1], gold nanoparticles are encapsulated in silica spheres of 44 nm diameter (Fig. 1a). The optical gain is supplied by organic dye molecules embedded in the silica shells, which are optically pumped; the surface plasmons are provided by the gold cores. The energy pumped into the shells is transferred to the surface plasmons and stimulates the coherent emission and amplification of surface-plasmon waves. Finally, these waves are converted into visible laser light of wavelength 531 nm, which radiates outside the spheres. Although Noginov and colleagues' experiments were performed using an ensemble of silica spheres instead of a single one, the authors argue that the observed laser light was produced by the single spheres independently, rather than being a collective phenomenon.

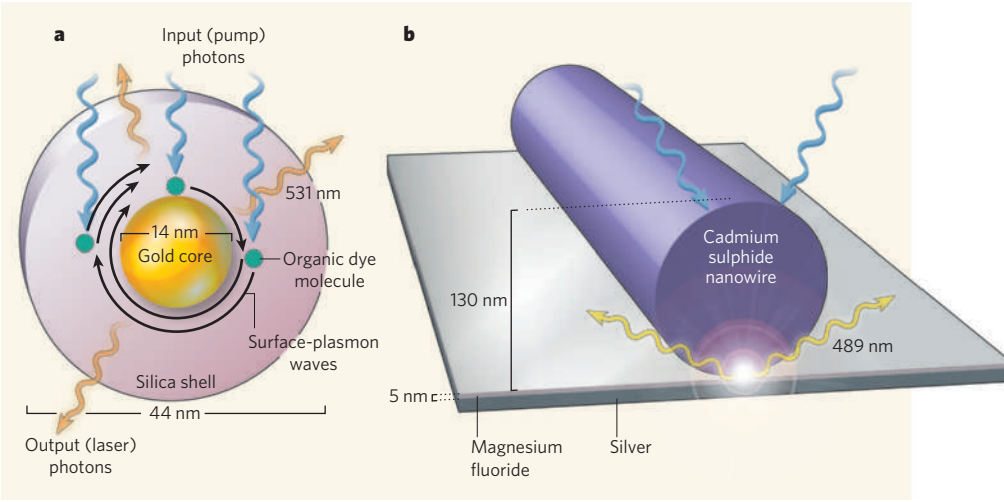


Figure 1 | Plasmon-based nanolasing. a, The laser structure developed by Noginov and colleagues[1] consists of a gold core surrounded by a silica shell in which organic dye molecules are embedded. The molecules provide the laser's optical gain. The energy (photons) pumped into the system is transferred to the collective motion of the electrons on the gold core's surface and stimulates the coherent emission and amplification of so-called surface-plasmon waves. These waves are ultimately converted into laser light of wavelength 531 nm. **b,** Oulton and co-workers' laser[2], whose working principle is also based on surface-plasmon waves, consists of a cadmium sulphide nanowire separated from a silver surface by a 5-nm insulating gap made of magnesium fluoride. The emergent, 489-nm-wavelength laser light is emitted from a strongly confined spot within the gap region.

The laser developed by Oulton and co-workers[2] is a single hybrid surface-plasmon cavity-waveguide structure that consists of a high-gain cadmium sulphide semiconductor nanowire separated from a flat silver surface by a 5-nm insulating gap (Fig. 1b). The authors' experiments and theoretical simulations suggest that the relevant dimension of the cavity, and hence the size of the laser spot, is of the order of 25 nm, around 20 times smaller than the laser's operating wavelength of 489 nm. The laser spot size is controlled by the width of the insulating gap and the diameter of the nanowire. In contrast to standard nanowire lasers[8], there is no cut-off laser size below which the strongly confined surface-plasmon waves cease to exist. Therefore, the expectation is that even smaller plasmon-based lasers can be realized if the width of the gap is reduced further.

Although the two studies[1,2] offer considerable advances in plasmonic-laser science, issues remain that must be tackled before a fully operative plasmonic nanolaser can be realized. One issue is the degree to which the emitted laser light is coherent, a key laser property that has not been explicitly measured by the authors[1,2]. Another troublesome point is related to the lack of collimation of the emergent laser light — in the experiments, converting surface-plasmon waves into conventional laser light resulted in light propagation in all directions.

As far as realistic applications go, electrical pumping would be more compatible with standard technologies than the optical pumping used in the two experiments[1,2]. Nevertheless, the presence of a semiconductor material in the hybrid scheme developed by Oulton and colleagues ensures that their plasmonic laser could also be driven by electric currents. In both experiments, surface plasmons were exploited merely to implement ultra-small, nanometre-scale lasers. An exciting perspective to be explored is using the devices as pure spasers — that is, as sources of coherent surface plasmons. This would pave the way to a new type of nanocircuitry based entirely on surface plasmons. ■

Francisco J. Garcia-Vidal and Esteban Moreno are in the Departamento de Física Teórica de la Materia Condensada, Universidad Autónoma de Madrid, Madrid 28049, Spain.
e-mail: fj.garcia@uam.es

1. Noginov, M. A. *et al. Nature* **460,** 1110–1112 (2009).
2. Oulton, R. F. *et al. Nature* **461,** 629–632 (2009).
3. Bergman, D. J. & Stockman, M. I. *Phys. Rev. Lett.* **90,** 027402 (2003).
4. Zheludev, N. I., Prosvirnin, S. L., Papasimakis, N. & Fedotov, V. A. *Nature Photon.* **2,** 351–354 (2008).
5. Seidel, J., Grafström, S. & Eng, L. *Phys. Rev. Lett.* **94,** 177401 (2005).
6. Noginov, M. A. *et al. Phys. Rev. Lett.* **101,** 226806 (2008).
7. Hill, M. T. *et al. Nature Photon.* **1,** 589–594 (2007).
8. Duan, X., Huang, Y., Agarwal, R. & Lieber, C. M. *Nature* **421,** 241–245 (2003).

Correction
In "Gamma-ray bursts: Maybe not so old after all" by Enrico Ramirez-Ruiz and William Lee (*Nature* **460,** 1091–1092; 2009), an editorial error introduced the statement that GRB 070714B is "slightly more than 8 billion parsecs away". The correct units are light years, not parsecs.

ARTICLES

Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling

Shih-Min A. Huang[1], Yuji M. Mishina[1], Shanming Liu[1], Atwood Cheung[1], Frank Stegmeier[1], Gregory A. Michaud[1], Olga Charlat[1], Elizabeth Wiellette[1], Yue Zhang[1], Stephanie Wiessner[1], Marc Hild[1], Xiaoying Shi[1], Christopher J. Wilson[1], Craig Mickanin[1], Vic Myer[1], Aleem Fazal[1], Ronald Tomlinson[1], Fabrizio Serluca[1], Wenlin Shao[1], Hong Cheng[1], Michael Shultz[1], Christina Rau[2], Markus Schirle[2]†, Judith Schlegl[2], Sonja Ghidelli[2], Stephen Fawell[1], Chris Lu[1], Daniel Curtis[1], Marc W. Kirschner[3], Christoph Lengauer[1]†, Peter M. Finan[1], John A. Tallarico[1], Tewis Bouwmeester[2]†, Jeffery A. Porter[1], Andreas Bauer[2]† & Feng Cong[1]

The stability of the Wnt pathway transcription factor β-catenin is tightly regulated by the multi-subunit destruction complex. Deregulated Wnt pathway activity has been implicated in many cancers, making this pathway an attractive target for anticancer therapies. However, the development of targeted Wnt pathway inhibitors has been hampered by the limited number of pathway components that are amenable to small molecule inhibition. Here, we used a chemical genetic screen to identify a small molecule, XAV939, which selectively inhibits β-catenin-mediated transcription. XAV939 stimulates β-catenin degradation by stabilizing axin, the concentration-limiting component of the destruction complex. Using a quantitative chemical proteomic approach, we discovered that XAV939 stabilizes axin by inhibiting the poly-ADP-ribosylating enzymes tankyrase 1 and tankyrase 2. Both tankyrase isoforms interact with a highly conserved domain of axin and stimulate its degradation through the ubiquitin-proteasome pathway. Thus, our study provides new mechanistic insights into the regulation of axin protein homeostasis and presents new avenues for targeted Wnt pathway therapies.

The evolutionarily conserved Wnt/β-catenin signal transduction pathway controls many biological processes[1]. A key feature of the Wnt pathway is the regulated proteolysis of the downstream effector β-catenin by the β-catenin destruction complex. The principal constituents of the β-catenin destruction complex are adenomatous polyposis coli (APC), axin and glycogen synthase kinase 3α/β (GSK3α/β). In the absence of Wnt pathway activation, cytosolic β-catenin is constitutively phosphorylated and targeted for degradation. On Wnt stimulation the β-catenin destruction complex dissociates, leading to the accumulation of nuclear β-catenin and transcription of Wnt pathway-responsive genes.

Inappropriate activation of the Wnt pathway has been observed in many cancers[2,3]. Notably, truncating mutations of the tumour suppressor APC are the most prevalent genetic alterations in colorectal carcinomas[4–6]. The efficient assembly of the multi-protein destruction complex is dependent on the steady-state levels of its principal constituents. Axin has been reported to be the concentration-limiting factor in regulating the efficiency of the β-catenin destruction complex[7,8] and overexpression of axin induces β-catenin degradation in cell lines expressing truncated APC[9–11]. Thus, it is likely that axin protein levels need to be tightly regulated to ensure proper Wnt pathway signalling. In fact, Wnt signalling itself regulates the level of axin at several steps, with *AXIN2* being a major transcriptional target of the β-catenin–T-cell factor (TCF) complex and Wnt signalling promoting the degradation of axin[12,13]. However, the molecular mechanisms that regulate protein homeostasis of destruction complex components and complex assembly remain elusive.

In this study we used chemical-genetic and -proteomic approaches to search for novel modulators of the Wnt signalling pathway. We identified a low molecular mass compound that can prolong the half-life of axin and promote β-catenin degradation through inhibiting tankyrase (TNKS). Our study uncovers a new mechanism that controls axin protein stability and Wnt pathway signalling, and its therapeutic exploitation holds promise for treating Wnt-pathway-dependent cancers.

XAV939 inhibits Wnt signalling by increasing axin levels

XAV939 was identified as a small molecule inhibitor of the Wnt/β-catenin pathway from a high-throughput screen using a Wnt-responsive Super-Topflash (STF) luciferase reporter assay in HEK293 cells (Fig. 1a). XAV939 strongly inhibited Wnt3a-stimulated STF activity in HEK293 cells, but did not affect CRE, NF-κB or TGF-β luciferase reporters (Fig. 1b). In contrast, LDW643, a close structural analogue of XAV939 (Fig. 1a), had no effect on the Wnt3a-induced STF reporter (Fig. 1b). XAV939 treatment blocked Wnt3a-induced accumulation of β-catenin in HEK293 cells (Fig. 1c), indicating that the compound modulates Wnt signalling upstream of β-catenin. Interestingly, XAV939 also inhibited STF activity in SW480 cells, a colorectal cancer cell line harbouring a truncated APC (Fig. 1d). XAV939 decreased β-catenin abundance, but significantly increased β-catenin phosphorylation (S33/S37/T41) in SW480 cells (Fig. 1e), indicating that XAV939 promotes the phosphorylation-dependent degradation of β-catenin by increasing the activity of the destruction complex.

To explore how XAV939 may increase the activity of the destruction complex, we investigated whether compound treatment alters the protein levels of known Wnt pathway components. Notably, the

[1]Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA. [2]Cellzome AG, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [3]Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA. †Present addresses: Novartis Institutes for Biomedical Research, CH-4002 Basel, Switzerland (T.B., A.B.); Novartis Institutes for Biomedical Research, Cambridge, Massachusetts 02139, USA (M. Sc.); Sanofi-Aventis, 94403 Vitry-sur-Seine, France (C.L.).

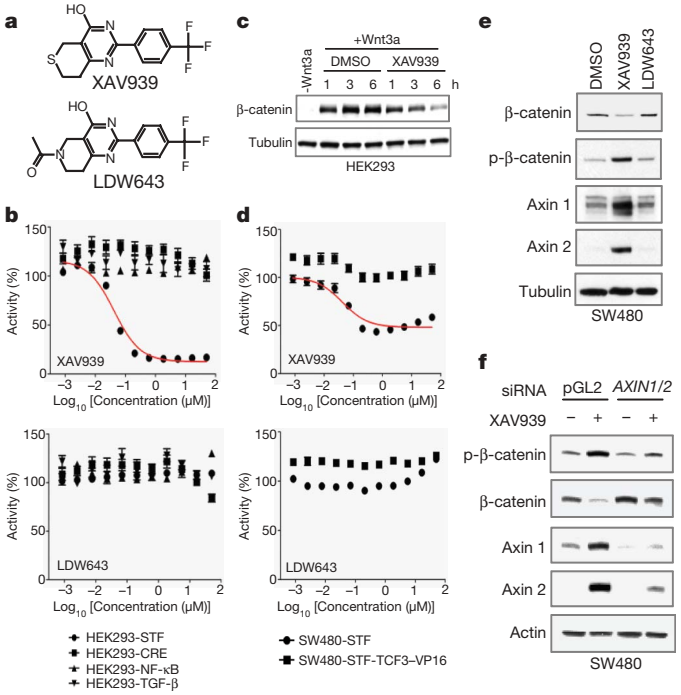


Figure 1 | XAV939 inhibits Wnt/β-catenin signalling by increasing axin protein levels. **a**, Compound structures of XAV939 and LDW643. **b**, XAV939 (24 h) specifically inhibits STF reporter, but not CRE, NF-κB or TGF-β reporter in HEK293 cells ($n = 12$). The corresponding reporter activity was normalized to DMSO. Error bars represent s.d. here and throughout the study. **c**, XAV939 (1 μM) reduces Wnt3a-induced stabilization of β-catenin in HEK293 cells. **d**, XAV939 (24 h) inhibits STF activity in APC-deficient SW480 cells ($n = 12$). SW480 cells expressing TCF3–VP16 were used as a control as overexpression of TCF3–VP16 fusion protein largely bypassed the requirement of β-catenin on STF activity. **e**, XAV939 (1 μM, 16 h) decreases the abundance of β-catenin and increases the abundance of axin and phospho-β-catenin (p-β-catenin) in SW480 cells. **f**, The effect of XAV939 on β-catenin in SW480 cells is eliminated by *AXIN1/2* siRNAs.

protein, but not messenger RNA, levels of axin 1 and axin 2 were strongly increased after XAV939 treatment (Fig. 1e and data not shown; Supplementary Fig. 1a). In addition, we noted a strong increase in axin–GSK3β complex formation, presumably because of increased axin protein levels (Supplementary Fig. 1b). Similar effects of XAV939 were also observed in DLD-1 cells, another colorectal cancer cell line with truncated APC (Supplementary Fig. 1c, d). Importantly, short interfering RNA (siRNA)-mediated depletion of axin 1/2 in SW480 cells reversed the effect of XAV939 on β-catenin degradation (Fig. 1f and Supplementary Fig. 1f) and diminished the inhibitory activity of XAV939 on the STF reporter (Supplementary Fig. 1e), indicating that XAV939 inhibits Wnt signalling by increasing axin 1/2 protein levels. Together, these findings support the hypothesis that XAV939 increases the concentration of axin–GSK3β complex, thereby promoting phosphorylation and degradation of β-catenin.

XAV939 regulates axin levels through tankyrase inhibition

To identify the cellular efficacy target(s) through which XAV939 upregulates axin protein levels, we used a three-channel iTRAQ quantitative chemical proteomics approach. This strategy is based on the immobilization of a bioactive analogue of XAV939 (Supplementary Fig. 2) to affinity-capture cellular proteins from HEK293 cell lysates spiked with an excess amount (20 μM) of XAV939, the inactive analogue LDW643 or DMSO. Specific binding to the immobilized compound should be competed with XAV939, but not with LDW643. Of 699 proteins quantified (Supplementary Table 1), 18 proteins were significantly and specifically competed-off (>65%, >2σ of the mean) with soluble XAV939 (Fig. 2a), including the poly(ADP-ribose) polymerases PARP1, PARP2, TNKS1, TNKS2 and several known PARP1 substrates, presumably co-purified with PARP1.

To establish the affinity of binding to the identified PARP proteins, we performed a compound competition experiment, and showed that XAV939, but not LDW643, blocks TNKS binding at 0.1 μM and blocks PARP1/2 binding at 1 μM (Fig. 2b). We further characterized the compound binding using Cy5-labelled XAV939 and recombinant PARP proteins. We found that XAV939 binds tightly to the catalytic (PARP) domains of TNKS1 and TNKS2 ($K_d = 0.099$ and 0.093 μM, respectively) (Fig. 2c). XAV939 also binds to recombinant PARP1, although with a significantly lower binding affinity ($K_d = 1.2$ μM).

To determine which PARP family member(s) are the actual efficacy targets of XAV939, we assessed their siRNA-mediated loss-of-function phenotypes. Co-depletion of TNKS1 and TNKS2 phenocopied the effect of XAV939 by increasing the protein levels of axin 1 and 2, whereas combinatorial *PARP1/2* knockdown did not (Fig. 3a and Supplementary Fig. 3a, b). In addition, ABT-888, a potent PARP1 and PARP2 inhibitor[14] that has minimal activity on TNKS1 and

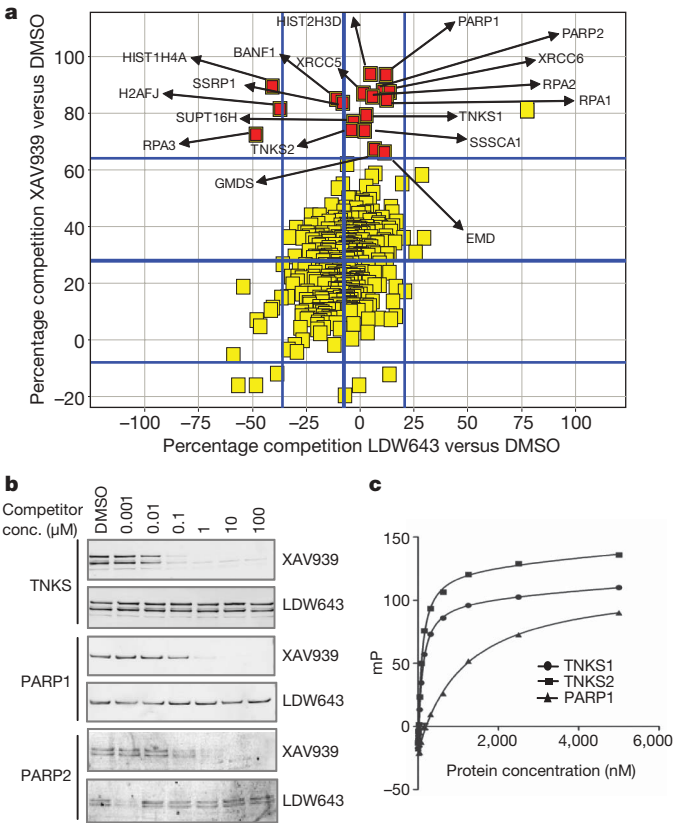


Figure 2 | Identification of the cellular efficacy targets of XAV939. **a**, Scatter plot depicting all proteins ($n = 398$) identified and quantified with >3 unique spectra in a three-channel iTRAQ quantitative chemical proteomics experiment. Proteins are plotted as a function of the percentage competition with the active compound XAV939 relative to the vehicle (DMSO), plotted on the *y*-axis, versus the percentage competition with the inactive compound LDW643 relative to DMSO, plotted on the *x*-axis. The blue border lines on each axis signify the 2 s.d. from the average competition value across all identified proteins with >3 spectra, which is indicated by the middle blue line. All proteins that are specifically competed with the active compound are highlighted in red. **b**, Immunoblot analysis for TNKS1/2, PARP1 and PARP2 on lysates from a compound competition experiment. **c**, Cy5-conjugated XAV939 directly binds the PARP domain of TNKS1 and TNKS2 with high affinity. Raw millipolarization (mP) data were analysed with a one-site total binding saturation algorithm.

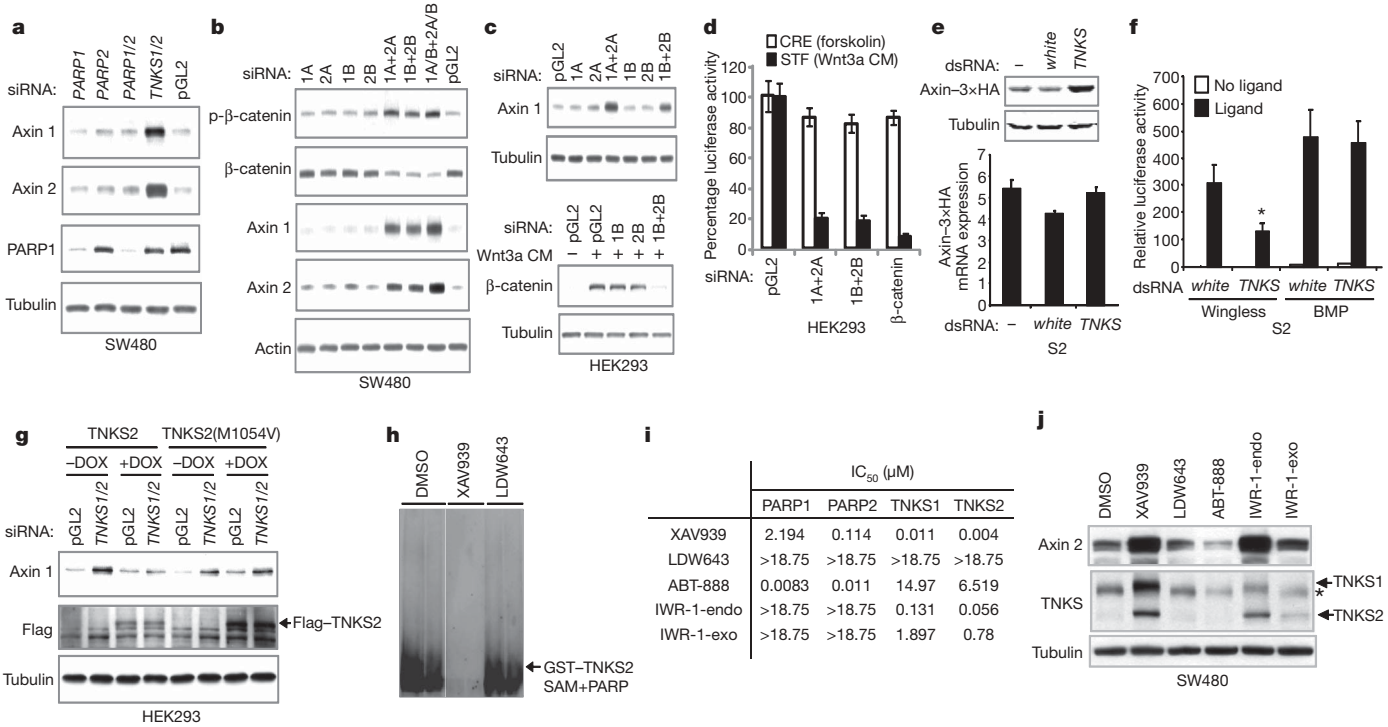


Figure 3 | Tankyrase modulates axin protein levels. **a**, **b**, Simultaneous depletion of TNKS1 and TNKS2 phenocopies XAV939 by increasing axin protein levels and decreasing β-catenin protein levels in SW480 cells. For both *TNKS1* and *TNKS2*, two independent siRNAs were generated from unique target sequences, labelled A and B (1A, 2A, 1B and 2B). **c**, Co-depletion of TNKS1 and TNKS2 increases the protein level of axin 1 (upper panel) and blocks Wnt3a-induced β-catenin accumulation (lower panel) in HEK293 cells. **d**, Co-depletion of TNKS1 and TNKS2 inhibits STF reporter, but not CRE reporter, in HEK293 cells ($n = 4$). **e**, Knockdown of TNKS increases the protein level (upper panel), but not the mRNA level (lower panel), of haemagglutinin-conjugated *Drosophila* axin (Axin–3×HA) in *Drosophila* S2 cells. dsRNA against *white* was used as control ($n = 4$). **f**, Depletion of TNKS specifically inhibits Wnt reporter (LEF–Luc), but not BMP reporter (BRE–Luc), in S2 cells ($^*P < 0.001$, $n = 4$). **g**, Doxycycline (DOX)-induced expression of wild-type TNKS2, but not catalytically inactive mutant TNKS2(M1054V), rescues *TNKS1/2* siRNA-induced accumulation of axin 1 in HEK293 cells. **h**, XAV939 inhibits auto-PARsylation of TNKS2. Recombinant TNKS2 protein used in this experiment contains the sterile alpha motif (SAM) domain and the PARP domain. **i**, Compound activity in PARP1/2 and TNKS1/2 biochemical assays. **j**, Effects of compounds (5 μM, 24 h) on the protein levels of axin 2 and TNKS1/2 in SW480 cells. A background band that migrated right below the TNKS1 band is indicated by an asterisk.

TNKS2 (Fig. 3i and Supplementary Fig. 3i), did not affect the protein levels of axin and TNKS (Fig. 3j). Collectively, these results indicate that TNKS1 and TNKS2 are the cellular efficacy targets of XAV939.

Using additional siRNAs, we further demonstrated that co-depletion of TNKS1 and TNKS2 increases β-catenin phosphorylation, decreases β-catenin abundance, and inhibits the transcription of β-catenin target genes in SW480 cells (Fig. 3b and Supplementary Fig. 3c, d). Notably, depletion of TNKS1 or TNKS2 alone did not increase axin 1/2 protein levels (Fig. 3b), indicating that TNKS1 and TNKS2 function redundantly in regulating axin protein levels. Co-depletion of TNKS1 and TNKS2 also phenocopied the pharmacological effect of XAV939 in HEK293 and DLD-1 cells (Fig. 3c, d and Supplementary Fig. 3e). In addition, combined XAV939 and *TNKS1/2* siRNA treatment increased axin protein levels even further (Supplementary Fig. 3f).

We examined whether regulation of axin and Wnt signalling by TNKS is evolutionarily conserved. We found that a double-stranded RNA (dsRNA) targeting *Drosophila TNKS* increased protein levels, but not mRNA levels, of exogenously expressed *Drosophila* axin in S2 cells (Fig. 3e and Supplementary Fig. 3g), and specifically inhibited a Wnt reporter ($P < 0.001$) (Fig. 3f). We also showed that treatment of zebrafish embryos with XAV939, but not its inactive analogue LDW643, significantly decreased the signal from a β-catenin/LEF1 reporter (TOP–GFP; green fluorescent protein) (Supplementary Fig. 4a) and inhibited the mRNA expression of the β-catenin target gene *axin2* (Supplementary Fig. 4c). Furthermore, concomitant knockdown of two zebrafish TNKS genes (*tnks1a* and *tnks1b*) with morpholinos specifically decreased TOP–GFP expression (Supplementary Fig. 4b). Together, these findings indicate that the regulatory function of TNKS in Wnt signalling is evolutionarily conserved.

Wnt signalling is required for regenerative processes in adult zebrafish, including fin regeneration[15,16]. *Inks* genes are expressed in caudal fin tissue, and no transcriptional changes are detectable after injury (Supplementary Fig. 4d, panel iii). XAV939, but not LDW643, inhibited fin regeneration (Supplementary Fig. 4d, panels i and ii) and Wnt-dependent activation of *axin2* transcription (Supplementary Fig. 4d, panel iii). This phenotype is remarkably similar to what was observed for IWR-1, an axin stabilizer recently identified[17], which we have found to also inhibit TNKS (as described in detail below). The fact that two structurally unrelated TNKS inhibitors block Wnt-dependent fin regeneration indicates that TNKS inhibition can antagonize Wnt signalling *in vivo*. However, we did not observe obvious early developmental defects associated with Wnt inhibition. Given that axin accumulation requires several hours, it is possible that the kinetics of axin stabilization are too slow to have an impact on early developmental phenotypes in this model system. Alternatively, TNKS function in Wnt pathway modulation may be cell-type- or organ-specific (for example, tail fin).

TNKS1 and TNKS2 modify their substrates through the addition of several ADP-ribose units, referred to as poly-ADP-ribosylation (PARsylation)[18]. To determine whether the PARsylation activity of TNKS is essential for regulating axin protein levels, we performed siRNA-rescue experiments. In transient transfection experiments, TNKS overexpression at high levels induces β-catenin stabilization in a PARP-domain-independent fashion. However, this is an overexpression artefact probably mediated by sequestration of endogenous axin by overexpressed TNKS (see Supplementary Fig. 6 for detailed discussion). We therefore used lentiviral transduction at a low multiplicity of infection for the rescue experiments, in which the

expression of TNKS is markedly lower than in the transient transfection experiments (data not shown). Doxycycline-induced expression of wild-type TNKS2, but not the catalytically inactive TNKS2(M1054V) mutant[19], rescued the effect of *TNKS1/2* siRNAs on axin 1 protein expression (Fig. 3g). Similar results were obtained with TNKS1 (Supplementary Fig. 3h). These data indicate that the catalytic activity of both tankyrases is required for the regulation of axin protein levels. Consistent with this finding, we showed that XAV939, but not LDW643, inhibits auto-PARsylation of TNKS2 *in vitro* (Fig. 3h). In biochemical activity assays, XAV939 strongly inhibited TNKS1 and TNKS2, with half-maximal inhibitory concentration values of 0.011 and 0.004 μM, respectively, but displayed much weaker effects on PARP1 and PARP2 (Fig. 3i). Auto-PARsylation of TNKS has been reported to promote its own degradation through the ubiquitin-proteasome pathway[20]. We found that XAV939 treatment led to a significant increase in TNKS protein levels (Fig. 3j), indicating that XAV939 also inhibits TNKS auto-PARsylation *in vivo*.

A recent study described two compound series (IWR-1, -2, and IWR-3, -4, -5) that increase axin protein levels and inhibit Wnt signalling[17]. It was shown that IWR-1-endo (also referred to as IWR-1) increased axin protein levels, whereas its close analogue IWR-1-exo had much weaker activity. Although the authors suggested that IWR-1 might stabilize axin by binding to a region in the carboxy-terminal half of axin, we examined whether IWR-1 compounds may in fact stabilize axin by inhibiting TNKS. Notably, IWR-1-endo strongly inhibited TNKS1 and TNKS2 in biochemical assays, with IWR-1-exo being approximately tenfold less active (Fig. 3i and Supplementary Fig. 3i), consistent with their potency in axin stabilization assays (Fig. 3j). In contrast, neither of these two compounds inhibited PARP1/2 (Fig. 3i). Consistent with the activity of IWR-1-endo in TNKS1/2 biochemical assays, IWR-1-endo, but not IWR-1-exo, significantly stabilized endogenous TNKS1, TNKS2 and axin 2 (Fig. 3j and Supplementary Fig. 3j), indicating that IWR-1-endo inhibited auto-PARsylation of TNKS *in vivo*. These results indicate that IWR-1-endo stabilizes axin at least in part through TNKS inhibition. Together, our data demonstrate that the PARsylation activity of TNKS1/2, but not PARP1/2, is required for the regulation of axin protein levels.

TNKS interacts with axin to regulate axin levels

We next explored how TNKS regulates axin protein levels. Using co-immunoprecipitation experiments, we found that endogenous TNKS1 and TNKS2 associate with axin 2 in SW480 cells (Fig. 4a). In addition, we detected a strong interaction between axin 1/2 and TNKS1/2 in the yeast two-hybrid assay (Fig. 4b and data not shown). The ankyrin repeat domain of TNKS1 was required and sufficient for the interaction with axin 1 (Supplementary Fig. 6a). Strikingly, a small amino-terminal region of axin 1 (amino acids 19–30), which encompasses the most conserved stretch of amino acids within axin (Fig. 4c), was both required and sufficient to interact with TNKS1 (Fig. 4b). The specific interaction of axin 1 with TNKS1 through this domain, which we named tankyrase-binding domain (TBD), was further substantiated by glutathione *S*-transferase (GST) pull-down and co-immunoprecipitation assays (Fig. 4d, e).

We next assessed the functional consequences of disrupting the physical interaction between axin and TNKS. If binding between axin and TNKS promoted axin degradation, deletion of the TBD should result in the stabilization of axin. Indeed, whereas cells expressing wild-type GFP–axin 1 demonstrated low basal protein levels that were strongly increased in response to XAV939 treatment, cells expressing GFP–axin 1(Δ19–30) already exhibited high basal protein levels that did not further respond to compound treatment (Fig. 4f and Supplementary Fig. 5). Importantly, restoring the TNKS1–axin 1 interaction by fusing the heterologous TNKS binding domain of either IRAP (insulin-regulated aminopeptidase, also known as LNPEP) or TRF1 (telomeric repeat binding factor, also known as TERF1) to GFP–axin 1(Δ19–30) fully restored its response to XAV939 (Fig. 4f). We further hypothesized that overexpression of the N-terminal domain of axin may compete with endogenous axin for the binding of TNKS and thus stabilize axin. Indeed, overexpression of GFP–axin 1N (amino acids 1–87), but not the mutant with the

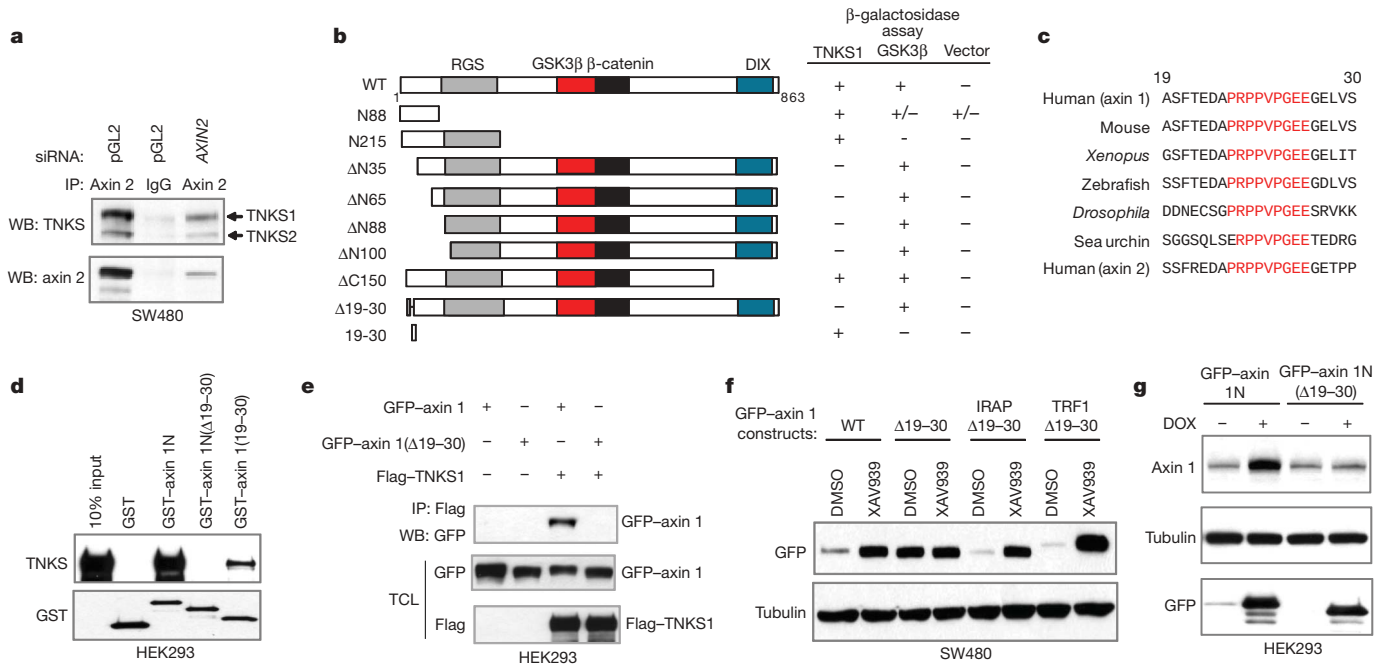


Figure 4 | Tankyrase physically and functionally interacts with axin. a, Co-immunoprecipitation of endogenous axin 2 and TNKS in SW480 cells. **b**, Binding between axin 1 fragments and TNKS1 in the yeast two-hybrid assay (+, strong binding; +/−, weak binding; −, no binding). The N88 fragment retains partial self-activation activity. **c**, The TBDs of axin proteins are evolutionarily conserved. Highly conserved amino acid residues are depicted in red. **d**, The N-terminal fragment of axin 1 (amino acids 1–87) binds to TNKS1 in a GST pull-down assay. **e**, The TBD is required for the interaction between axin 1 and TNKS in a co-immunoprecipitation assay. **f**, The TBD is required for XAV939-induced axin 1 protein accumulation in SW480 cells. The expression of GFP–axin 1 was under the control of the metallothionein promoter to ensure low transcription. **g**, DOX-induced overexpression of GFP–axin 1N (amino acids 1–87) leads to accumulation of endogenous axin 1 in HEK293 cells.

deleted TBD, substantially increased endogenous axin 1 protein levels while not affecting its mRNA expression (Fig. 4g and data not shown). Together, these findings demonstrate that the physical interaction between axin and TNKS, which is mediated by the evolutionarily conserved TBD, is critical for regulating axin protein levels.

Axin is degraded through PARsylation and ubiquitination

The increase in axin protein levels in response to XAV939 treatment could be due to modulation of translation or protein stability. Consistent with the latter possibility, XAV939 treatment prolonged the half-life of endogenous axin 2 in SW480 cells (Fig. 5a). The degradation of axin is probably mediated by the ubiquitin-proteasome pathway, because the polyubiquitination of axin 1 increased significantly after addition of the proteasome inhibitor MG132 (Fig. 5b). In contrast, co-treatment of XAV939 with MG132 significantly diminished axin 1 and axin 2 polyubiquitination (Fig. 5b and Supplementary Fig. 7a), indicating that XAV939 may stabilize axin by preventing its polyubiquitination.

Auto-PARsylation of TNKS or TNKS-mediated TRF1 PARsylation leads to ubiquitination and degradation of TNKS or TRF1, respectively[20,21]. Together with our findings, this raised the possibility that axin degradation may be facilitated through direct PARsylation by TNKS. Indeed, TNKS2 was able to PARsylate an axin 1 fragment (amino acids 1–280) containing the TBD *in vitro* (Fig. 5c), which was completely inhibited by XAV939 treatment (Fig. 5c). Using an antibody that specifically recognizes the PAR modification, we also observed that exogenously expressed GFP–axin 1 was PARsylated in cells (Fig. 5d). In addition, the PARsylation signal was strongly reduced in the presence of XAV939 (Fig. 5d), indicating that axin PARsylation is mediated by TNKS *in vivo*. To enhance the detection of endogenous axin 2 ubiquitination and PARsylation, we pretreated cells with XAV939 to increase endogenous axin 2 levels. Axin 2 was rapidly degraded within 1 h after XAV939 was washed off (Fig. 5e). As expected, treatment of MG132 blocked axin 2 degradation and strongly increased its polyubiquitination (Fig. 5e and Supplementary Fig. 7b). Treatment with XAV939 and MG132, however, completely blocked the ubiquitination of axin 2 (Fig. 5e and Supplementary Fig. 7b). Interestingly, an anti-PAR antibody reactive signal that co-migrated with axin 2 was detected when cells were treated with MG132 alone and disappeared when cells were also treated with XAV939 (Fig. 5e). Together, these findings are consistent with the hypothesis that TNKS promotes the ubiquitination and degradation of axin, which may be mediated, at least in part, through the direct PARsylation of axin.

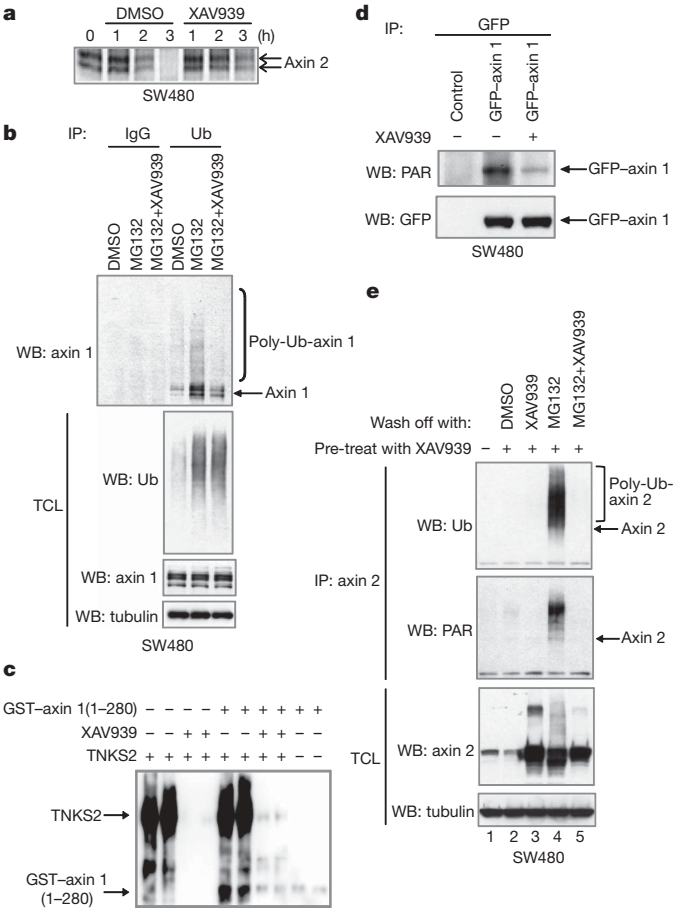


Figure 5 | XAV939 stabilizes axin protein and inhibits ubiquitination of axin. **a**, XAV939 (1 μM) stabilizes axin 2 in SW480 cells as indicated by a pulse-chase analysis. **b**, Ubiquitination of axin is inhibited by XAV939 (1 μM). The position of axin 1 is labelled with an arrow. Slow-migrating polyubiquitinated axin 1 conjugates are indicated. **c**, *In vitro* PARsylation of GST–axin 1 (amino acids 1–280) by TNKS2. **d**, XAV939 (1 μM, 6 h) inhibits *in vivo* PARsylation of axin 1. **e**, Post-translational modification of axin 2 in a compound wash-off experiment. SW480 cells pretreated with XAV939 were washed and incubated with medium supplemented with the indicated compound for 1 h, and analysed by the immunoprecipitation assay. The position of axin 2 is indicated by the arrows. IP, immunoprecipitation; TCL, total cell lysates.

XAV939 inhibits growth of DLD-1 cancer cells

Because XAV939 inhibited β-catenin signalling even in APC-deficient cells, we examined whether this compound could inhibit the proliferation of APC-deficient colorectal cancer cells using β-catenin-dependent DLD-1 cells and β-catenin-independent RKO cells (Supplementary Fig. 8a). Under low serum growth conditions, XAV939, but not the inactive analogue LDW643, significantly inhibited colony formation of DLD-1 cells (Fig. 6a). Importantly, XAV939 did not affect colony formation in RKO cells (Fig. 6b).

TNKS1 was proposed to be required for the resolution of sister telomere association or assembly of bipolar spindles, and *TNKS1* knockdown was reported to cause strong mitotic arrest[22,23]. However, using XAV939 treatment or individual/combinatorial *TNKS1/TNKS2* siRNA knockdown, we did not observe any overt mitotic arrest phenotype (Supplementary Fig. 8b and data not shown), demonstrating that XAV939 does not inhibit the proliferation of DLD-1 cells through an antimitotic function.

If the antiproliferative effect of XAV939 on DLD-1 cells was instead mediated by an increase in axin protein levels, knockdown of *AXIN1/2* may rescue the growth inhibitory effects of compound treatment. Indeed, siRNA-mediated depletion of axin 1 and axin 2 proteins completely abolished the antiproliferative effect of XAV939 (Fig. 6c), indicating that the growth inhibitory effects of XAV939 in DLD-1 cells are due to axin-dependent inhibition of Wnt signalling.

Discussion

Despite the fact that the Wnt pathway is an attractive target for anticancer therapy, there are few druggable targets in this pathway to generate small molecule Wnt inhibitors. Using a chemical genetics approach, we have discovered tankyrases as novel targets for Wnt inhibition and described a novel mechanism to promote β-catenin degradation through inhibition of tankyrases and stabilization of axin (Supplementary Fig. 9). This finding may pave the way for mechanism-based treatments of Wnt-dependent cancers.

Our data indicate that tankyrases are physiological regulators of axin protein homeostasis and Wnt signalling. Although the precise mechanistic details remain unclear, tankyrases seem to promote the ubiquitination of axin, possibly through direct PARsylation of axin. Notably, TNKS-mediated PARsylation seems to also stimulate ubiquitin-mediated proteolysis of TRF1 and TNKS itself[20,21]. These findings indicate that coupling of these biochemical processes may be a more general regulatory mechanism, akin to the widespread occurrence of phosphodegrons that couple phosphorylation to ubiquitination[24,25].

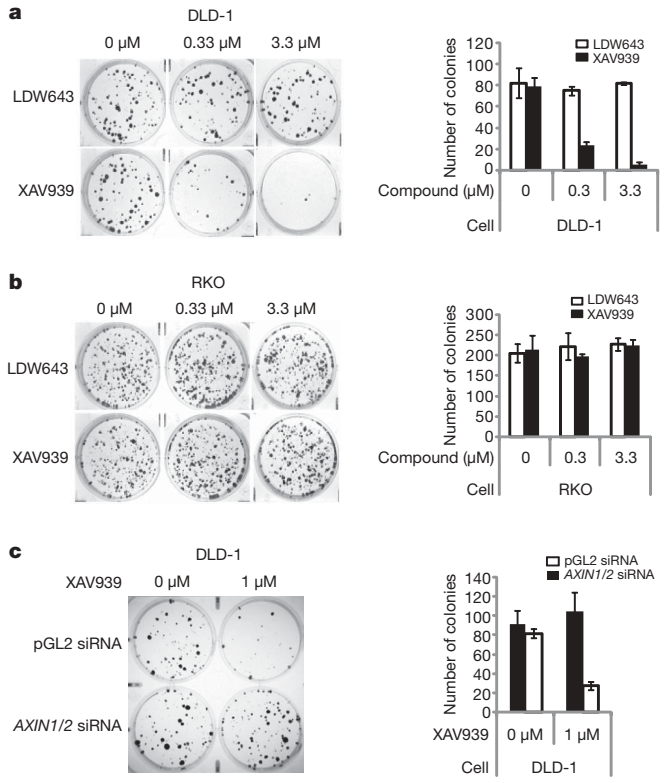


Figure 6 | XAV939 inhibits DLD-1 colony formation in an axin-dependent manner. a, **b**, XAV939 inhibits colony formation of DLD-1, but not RKO, cells. DLD-1 and RKO cells were seeded and grown in 0.5% serum medium containing the indicated compounds ($n = 3$). The number of colonies in each well was counted. **c**, The effect of XAV939 on DLD-1 colony formation is axin-dependent. DLD-1 cells were transfected with siRNAs against *AXIN1* and *AXIN2* or control pGL2 siRNA, and tested in the colony formation assay ($n = 3$).

Tnks1$^{-/-}$ and *Tnks2*$^{-/-}$ mice are largely normal and double knockout of *Tnks1/2* causes early embryonic lethality, indicating their redundancy in mouse development[26,27]. Conditional knockout of both *Tnks1* and *Tnks2* will be important to understand the role of tankyrases in regulation of Wnt signalling in tissue homeostasis and disease.

Axin is a concentration-limiting factor in the β-catenin degradation complex[7,8]. Our study further highlights the importance of axin as a key regulatory node in the Wnt signalling cascade. We speculate that axin may function more generally as a signal 'integrator' and that other proteins/pathways may modulate Wnt pathway activity by impinging on this key regulatory node. Further investigation of the molecular mechanisms that regulate axin protein and their potential deregulation in disease conditions is likely to provide additional avenues for treating Wnt-dependent cancers and other Wnt-related diseases.

METHODS SUMMARY

Compound affinity purification and mass spectrometry analysis. Compound affinity purification, mass spectrometry and data analysis were performed essentially as previously described[28]. A derivatized, bioactive analogue LDW639 with a primary amine group was coupled to NHS (*N*-hydroxysuccinimide)-activated Sepharose 4 beads. HEK293 cells were homogenized in lysis buffer (50 mM Tris-HCl, pH 7.5, 5% glycerol, 1.5 mM $MgCl_2$, 150 mM NaCl, 20 mM NaF, 1 mM Na_3VO_4, 1 mM dithiothreitol, 5 μM calyculin A, 0.8% Igepal-CA630 and a protease inhibitor cocktail). Cell lysates were pre-incubated with 20 μM XAV939, LDW643 or DMSO for 30 min, and incubated with LDW639-matrix for another 60 min. After extensive washes, bound material was eluted and resolved by SDS–polyacrylamide gel electrophoresis. Gel lanes were cut into slices across the full separation range and subjected to in-gel tryptic digestion, followed by iTRAQ labelling (Applied Biosystems). Peptides extracted from DMSO-treated samples were labelled with iTRAQ reagent 116 and combined with extracts of compound-treated samples labelled with iTRAQ reagents 114 and 115, respectively. Liquid chromatography–tandem mass spectrometry was done using an Eksigent 1D+ HPLC system coupled to a LTQ Orbitrap mass spectrometer (Thermo-Finnigan). Tandem mass spectra were generated using pulsed-Q dissociation, enabling detection of iTRAQ reporter ions[29]. Peptide mass and fragmentation data were used to query an in-house curated version of the IPI database using Mascot (Matrix Science). Protein identifications were validated using a decoy database. iTRAQ reporter ion-based quantification was performed with software developed in-house.

Received 7 May; accepted 30 July 2009.
Published online 16 September 2009.

1. Clevers, H. Wnt/β-catenin signaling in development and disease. *Cell* **127,** 469–480 (2006).
2. Polakis, P. The many ways of Wnt in cancer. *Curr. Opin. Genet. Dev.* **17,** 45–51 (2007).
3. Barker, N. & Clevers, H. Mining the Wnt pathway for cancer therapeutics. *Nature Rev. Drug Discov.* **5,** 997–1014 (2006).
4. Miyaki, M. *et al.* Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumors. *Cancer Res.* **54,** 3011–3020 (1994).
5. Mori, Y. *et al.* Somatic mutations of the *APC* gene in colorectal tumors: mutation cluster region in the *APC* gene. *Hum. Mol. Genet.* **1,** 229–233 (1992).
6. Powell, S. M. *et al.* APC mutations occur early during colorectal tumorigenesis. *Nature* **359,** 235–237 (1992).
7. Salic, A., Lee, E., Mayer, L. & Kirschner, M. W. Control of β-catenin stability: reconstitution of the cytoplasmic steps of the Wnt pathway in *Xenopus* egg extracts. *Mol. Cell* **5,** 523–532 (2000).
8. Lee, E., Salic, A., Kruger, R., Heinrich, R. & Kirschner, M. W. The roles of APC and axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.* **1,** E10 (2003).
9. Behrens, J. *et al.* Functional interaction of an axin homolog, conductin, with β-catenin, APC, and GSK3β. *Science* **280,** 596–599 (1998).
10. Kishida, M. *et al.* Axin prevents Wnt-3a-induced accumulation of β-catenin. *Oncogene* **18,** 979–985 (1999).
11. Hart, M. J., de los Santos, R., Albert, I. N., Rubinfeld, B. & Polakis, P. Downregulation of β-catenin by human axin and its association with the APC tumor suppressor, β-catenin and GSK3β. *Curr. Biol.* **8,** 573–581 (1998).
12. Leung, J. Y. *et al.* Activation of *AXIN2* expression by β-catenin-T cell factor. A feedback repressor pathway regulating Wnt signaling. *J. Biol. Chem.* **277,** 21657–21665 (2002).
13. Willert, K., Shibamoto, S. & Nusse, R. Wnt-induced dephosphorylation of axin releases β-catenin from the axin complex. *Genes Dev.* **13,** 1768–1773 (1999).
14. Donawho, C. K. *et al.* ABT-888, an orally active poly(ADP-ribose) polymerase inhibitor that potentiates DNA-damaging agents in preclinical tumor models. *Clin. Cancer Res.* **13,** 2728–2737 (2007).
15. Poss, K. D., Shen, J. & Keating, M. T. Induction of *lef1* during zebrafish fin regeneration. *Dev. Dyn.* **219,** 282–286 (2000).
16. Stoick-Cooper, C. L. *et al.* Distinct Wnt signaling pathways have opposing roles in appendage regeneration. *Development* **134,** 479–489 (2007).
17. Chen, B. *et al.* Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer. *Nature Chem. Biol.* **5,** 100–107 (2009).
18. Hsiao, S. J. & Smith, S. Tankyrase function at telomeres, spindle poles, and beyond. *Biochimie* **90,** 83–92 (2008).
19. Sbodio, J. I., Lodish, H. F. & Chi, N. W. Tankyrase-2 oligomerizes with tankyrase-1 and binds to both TRF1 (telomere-repeat-binding factor 1) and IRAP (insulin-responsive aminopeptidase). *Biochem. J.* **361,** 451–459 (2002).
20. Yeh, T.-Y. *et al.* Tankyrase recruitment to the lateral membrane in polarized epithelial cells: regulation by cell–cell contact and protein poly(ADP-ribosyl)ation. *Biochem. J.* **399,** 415–425 (2006).
21. Chang, W., Dynek, J. N. & Smith, S. TRF1 is degraded by ubiquitin-mediated proteolysis after release from telomeres. *Genes Dev.* **17,** 1328–1333 (2003).
22. Chang, P., Coughlin, M. & Mitchison, T. J. Tankyrase-1 polymerization of poly(ADP-ribose) is required for spindle structure and function. *Nature Cell Biol.* **7,** 1133–1139 (2005).
23. Dynek, J. N. & Smith, S. Resolution of sister telomere association is required for progression through mitosis. *Science* **304,** 97–100 (2004).
24. Winston, J. T. *et al.* The SCF$^{β\text{-}TRCP}$-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IκBα and β-catenin and stimulates IκBα ubiquitination *in vitro*. *Genes Dev.* **13,** 270–283 (1999).
25. Koepp, D. M. *et al.* Phosphorylation-dependent ubiquitination of cyclin E by the SCF^{Fbw7} ubiquitin ligase. *Science* **294,** 173–177 (2001).
26. Hsiao, S. J., Poitras, M. F., Cook, B. D., Liu, Y. & Smith, S. Tankyrase 2 poly(ADP-ribose) polymerase domain-deleted mice exhibit growth defects but have normal telomere length and capping. *Mol. Cell. Biol.* **26,** 2044–2054 (2006).
27. Chiang, Y. J. *et al.* Tankyrase 1 and tankyrase 2 are essential but redundant for mouse embryonic development. *PLoS ONE* **3,** e2639 (2008).
28. Bantscheff, M. *et al.* Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. *Nature Biotechnol.* **25,** 1035–1044 (2007).
29. Bantscheff, M. *et al.* Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Mol. Cell. Proteomics* **7,** 1702–1713 (2008).

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

Acknowledgements We thank D. Patel, F. Harbinski, J. Leighton-Davies, R. de Beaumont, X. Xiang, K. Bean, C. Xin, S. Zhao, B. Zhang and M. Xu for technical assistance, G. Wussler, H. Urquiza and W. Dai for zebrafish maintenance, and I. Cornella Taracido, S. Cleaver, A. Hernandez and Y. Ben-Neriah for comments and advice. In addition we are indebted to B. Kuster, J. Rick, M. Raida and A. Scholten for continued support and discussion.

Author Contributions S.-M.A.H., A.C., F.St., G.A.M., E.W., V.M., S.F., C.Lu, D.C., M.W.K., C.Le., P.M.F., J.A.T., T.B., J.A.P., A.B. and F.C. conceived and designed the study. S.-M.A.H., Y.M.M., S.L., A.C., F.St., G.A.M., O.C., E.W., Y.Z., S.W., M.H., X.S., C.W., C.M., A.F., R.T., F.Se., W.S., H.C., M.Sh., C.R., M.Sc., J.S., S.G., A.B. and F.C. designed and implemented experiments. S.-M.A.H., F.St., A.B., Y.M.M. and F.C. 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 F.C. (feng.cong@novartis.com).

ARTICLES

Inhibitors selective for mycobacterial versus human proteasomes

Gang Lin[1]*, Dongyang Li[3]*, Luiz Pedro Sorio de Carvalho[1], Haiteng Deng[4], Hui Tao[2], Guillaume Vogt[1], Kangyun Wu[1], Jean Schneider[1], Tamutenda Chidawanyika[1], J. David Warren[2], Huilin Li[3,5] & Carl Nathan[1]

Many anti-infectives inhibit the synthesis of bacterial proteins, but none selectively inhibits their degradation. Most anti-infectives kill replicating pathogens, but few preferentially kill pathogens that have been forced into a non-replicating state by conditions in the host. To explore these alternative approaches we sought selective inhibitors of the proteasome of *Mycobacterium tuberculosis*. Given that the proteasome structure is extensively conserved, it is not surprising that inhibitors of all chemical classes tested have blocked both eukaryotic and prokaryotic proteasomes, and no inhibitor has proved substantially more potent on proteasomes of pathogens than of their hosts. Here we show that certain oxathiazol-2-one compounds kill non-replicating *M. tuberculosis* and act as selective suicide-substrate inhibitors of the *M. tuberculosis* proteasome by cyclocarbonylating its active site threonine. Major conformational changes protect the inhibitor-enzyme intermediate from hydrolysis, allowing formation of an oxazolidin-2-one and preventing regeneration of active protease. Residues outside the active site whose hydrogen bonds stabilize the critical loop before and after it moves are extensively non-conserved. This may account for the ability of oxathiazol-2-one compounds to inhibit the mycobacterial proteasome potently and irreversibly while largely sparing the human homologue.

Proteasomes are structurally conserved from archaea to eukaryotes and are essential in eukaryotes[1]. The cytotoxicity of proteasome inhibitors has been exploited for cancer therapy[2] and suggested for the treatment of infections by eukaryotic pathogens, such as plasmodia and trypanosomes[3]. Unfortunately, the inherent toxicity of proteasome inhibitors is a drawback in the treatment of curable infections. Inhibitors that act with comparable potency on human proteasomes and proteasomes of infectious agents have not entered clinical practice as anti-infectives.

The only known bacterial pathogens with proteasomes are mycobacteria[4]. The tuberculosis pandemic has been declared a global health emergency as the growing resistance of *M. tuberculosis* to antibiotics[5] coincides with the spread of risk factors such as HIV/AIDS and diabetes[6,7]. Proteasomes degrade proteins that serve in signalling during adaptation, have become irreparably oxidized or are scavenged during starvation. The proteasome of *M. tuberculosis* is required for degradation of certain proteins[8,9], for it to survive nitroxidative stress *in vitro*[4] and for it to persist in mice[10]. The latter observation validates the *M. tuberculosis* proteasome as a drug target. However, all known proteasome inhibitors tested have inhibited mammalian proteasomes more potently than those of *M. tuberculosis*, including peptidyl epoxyketones, peptidyl aldehydes, γ-lactam-β-lactones such as salinosporamide A (NPI0052) (ref. 11 and G. Lin *et al.*, unpublished results) and the peptidyl boronate bortezomib (Velcade), which is in clinical use[2]. An effort to exploit the substrate preferences of the *M. tuberculosis* proteasome led to bortezomib analogues with varying amino acids at P1, the most selective of which inhibited the *M. tuberculosis* proteasome only eightfold more potently than a mammalian proteasome[12].

Nonetheless, biochemical and structural differences between proteasomes from *M. tuberculosis* and mammals encouraged us to seek species-selective proteasome inhibitors. In eukaryotic proteasomes the seven types of β-subunit forming the two heteroheptameric inner rings of the core particle[13] include three proteases with distinct specificities for oligopeptides. The *M. tuberculosis* proteasome has one type of β-subunit[14,15], but it is active against diverse benzyloxycarbonyl-capped tripeptides[14,15], perhaps because the side chains lining the active site of the *M. tuberculosis* proteasome have physicochemical properties that are a composite of those contributing to the three distinct active sites in eukaryotes[14,15]. Until now, residues so distant from the active site as to have no role in binding of substrate were not considered germane to the species selectivity of inhibitors, including those listed above, that react with the catalytic hydroxyl of the amino-terminal threonine (Thr 1).

Mycobactericidal, non-cytotoxic proteasome inhibitors

We expressed *M. tuberculosis* proteasome from the genes encoding proteasome components A (for the α-chain) and B (for the β-chain) (PrcBA) with and without the N-terminal octapeptide deleted from the α-chains. The latter 'open-gate' mutant (*M. tuberculosis* 20SOG) is thought to mimic a physiological mechanism for gate opening and has a higher specific activity than the wild type without a change in substrate preference[12,14,15]. By following the hydrolysis of Suc-LLVY-7-amido-4-methylcoumarin (AMC) by the *M. tuberculosis* 20SOG mutant, we screened 20,000 compounds (see Supplementary Information) and identified two inhibitors, 5-(5-methyl-2-(methylthio)thiophen-3-yl)-1,3,4-oxathiazol-2-one (GL5) and 5-(2-methyl-3-nitrothiophen-2-yl)-1,3,4-oxathiazol-2-one (HT1171), as well as an inactive congener, 5-(5-methyl-2-(methylsulphonyl)thiophen-3-yl)-1,3,4-oxathiazol-2-one (GL6) (Fig. 1a).

Like bortezomib, oxathiazol-2-one compounds were able to cross the cell wall of *M. tuberculosis*, insofar as GL5 and HT1171 inhibited proteasome activity on treatment of *Mycobacterium bovis* var. bacille

[1]Department of Microbiology and Immunology, Weill Cornell Medical College, [2]Milstein Chemistry Core Facility and Department of Biochemistry and Structural Biology, Weill Cornell Medical College, New York, New York 10065, USA. [3]Biology Department, Brookhaven National Laboratory, Upton, New York 11973-5000, USA. [4]Proteomics Resource Center, The Rockefeller University, New York, New York 10065, USA. [5]Department of Biochemistry and Cell Biology, Stony Brook University, Stony Brook, New York 11794, USA.
*These authors contributed equally to this work.

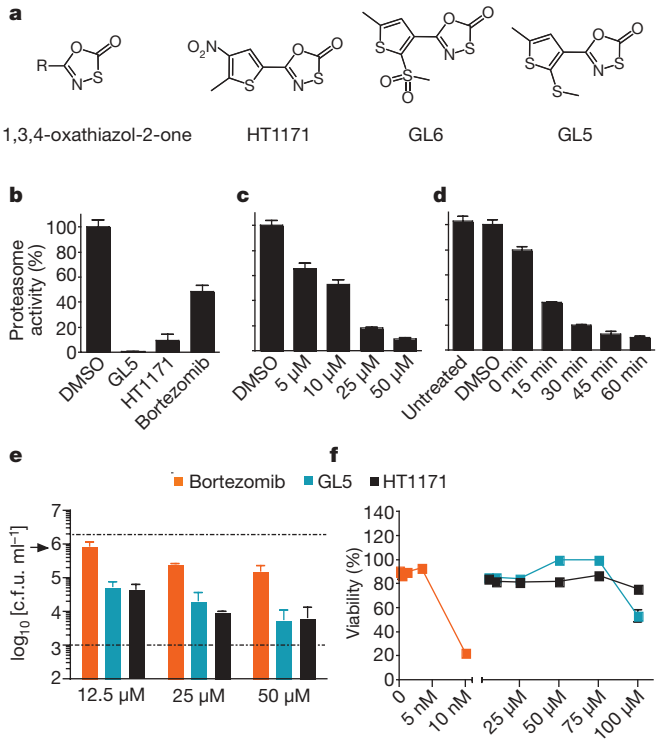


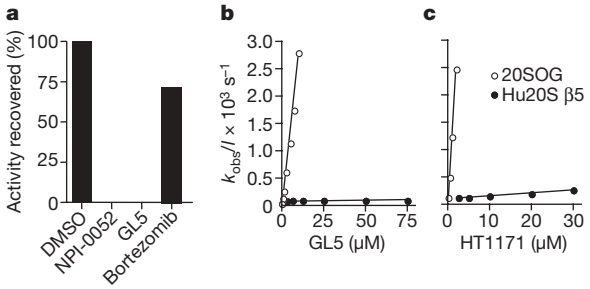
Figure 1 | Oxathiazol-2-one compounds inhibit mycobacterial proteasomes and kill non-replicating *M. tuberculosis*. **a**, Compound structures. **b**, Inhibition of proteasomes in intact BCG (A_{580} 0.6–1) by GL5, HT1171 or bortezomib (each 50 μM) after 4 h. Proteasome activity in lysates was tested with Ac-YQW-AMC (50 μM) as substrate. **c**, Concentration–response for GL5 after 1 h exposure as in **b**. Control, DMSO (<1% v/v) alone. **d**, Time course for effect of GL5 (50 μM) as in **b**, except that washing began at the times indicated. 'Untreated' and 'DMSO' cells were handled in the same manner but received nothing or DMSO at time 0 and were lysed at 60 min. 'T0' cells were treated and washed immediately. **e**, Killing of *M. tuberculosis* Erdman in Sauton's medium with time 0 addition of 50 μM DETA-NO, a nitric oxide donor ($t_{1/2}$, ~20 h), mimicking the nitroxidative stress that limits the replication of *M. tuberculosis* in mice[22]. Upper dashed line, c.f.u. ml^{-1} after exposure to DMSO control. Lower dashed line, limit of detection. Arrow, initial inoculum. **f**, Monkey kidney epithelial cells (Vero76) were incubated with compounds for 4 days before 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) assay for viability; microscopy gave concordant results. Data are means ± s.d. of triplicates in one of at least two experiments. Some error bars fall within the symbols.

Calmette–Guérin (BCG; Fig. 1b). At 50 μM, GL5 and HT1171 inhibited ~90% of mycobacterial proteasome activity, whereas bortezomib (50 μM) inhibited 52%. Exposure of BCG to 25 μM GL5 for 4 h led to >80% reduction of proteasome activity (Fig. 1c), whereas as little as ~15 min exposure to GL5 at 50 μM was sufficient to reduce activity by ~50% (Fig. 1d).

Moreover, GL5 killed BCG alone and in synergy with sub-bacteriostatic fluxes of nitric oxide arising from the decomposition of 2,2-(hydroxynitrosohydrazino)-bis-ethanamine (DETA-NO) (Supplementary Fig. 1). GL5 and HT1171 also dose-dependently killed 1.5–2.5 $\log_{10}$ *M. tuberculosis* over 4 days in synergy with sufficient nitric oxide to induce a pathophysiologically relevant state of bacterial non-replication[16] (Fig. 1e). Bortezomib was less mycobactericidal (Fig. 1e), and it was toxic to monkey epithelial cells (Fig. 1f) and human macrophages (Supplementary Fig. 2). In contrast, GL5 and HT1171 showed no apparent toxicity to mammalian cells (Fig. 1f and Supplementary Fig. 2) at concentrations up to 75 μM, 3,000-fold greater than those at which bortezomib killed the epithelial cells. The oxathiazol-2-one compounds exerted no antibacterial activity against *Mycobacterium avium intracellulare*, *Staphylococcus aureus*, *Salmonella enterica* var. Typhimurium or *Pseudomonas aeruginosa* (data not shown). Although some oxathiazol-2-one compounds were reported to react with thiols[17], those studied here did not inhibit the thiol-dependent cathepsin B (data not shown). Moreover, 11 of 23 oxathiazol-2-one compounds tested were <5% reactive with glutathione; the others reacted to a limited degree (Supplementary Table 1). Thus, at a functional level, the oxathiazol-2-one compounds tested here seem to be relatively selective and nontoxic, although they may have additional targets.

Selective inhibition of mycobacterial proteasomes

The different impact of oxathiazol-2-one compounds on *M. tuberculosis* and mammalian cells prompted us to ask if these compounds differentially inhibit isolated mycobacterial and human proteasomes. In dialysis (Fig. 2a) and kinetic studies (Fig. 2b, c), oxathiazol-2-one compounds inhibited *M. tuberculosis* proteasomes irreversibly (Fig. 2a), whereas inhibition of human proteasome β1, β2 and β5 sites was so minimal (Fig. 2b, c and Supplementary Table 1) as to preclude definition of a mode of inhibition. After establishing that oxathiazol-2-one compounds spontaneously hydrolyse in tissue culture medium to the corresponding amide with a half time ($t_{1/2}$) ranging from 7 to 180 min (Supplementary Table 2), we used partition ratios (the ratios of rate constants for catalysis and inactivation) to assess their relative potency[18]. By this measure, GL5 and HT1171 were >1,000-fold more effective against *M. tuberculosis* proteasomes than human proteasomes (Fig. 2d and Supplementary Fig. 3a, b). Inhibition was competitive with the substrate benzyloxycarbonyl-valyl-leucyl-arginyl-AMC (Z-VLR-AMC) (Supplementary Fig. 3c, d), indicating that the inhibitor binds at or near the active site, and was time-dependent (Supplementary Fig. 4). Potencies of all 21 oxathiazol-2-one compounds tested against wild-type *M. tuberculosis* proteasome correlated closely ($r^2 = 0.82$) with their potencies against the open-gate form (Supplementary Fig. 5). Oxathiazol-2-one-treated *M. tuberculosis* proteasomes lost the ability to degrade not only oligopeptides, but also a protein substrate, β-casein (Supplementary Fig. 6). In contrast, oxathiazol-2-one compounds were inactive or very weak inhibitors of trypsin (half-maximal inhibitory concentration (IC_{50}) > 50 μM), cathepsin B (IC_{50} > 50 μM), matrix metalloproteinase-2 (IC_{50} > 100 μM)



d

	k_{obs}/I M^{-1} s^{-1}				Partition ratio	
		Hu20S				
Compound	20SOG	β5	β2	β1	20SOG	Hu20S
GL5	376.4	0.4	NI	0.03	11.5	16,170
HT1171	2,134	10.1	6.9	2.8	8.9	8,963

Figure 2 | Kinetic analysis of inactivation of *M. tuberculosis* 20SOG and human proteasomes (Hu20S) by oxathiazol-2-one compounds. **a**, *M. tuberculosis* 20SOG (0.23 nM) was pre-treated with DMSO, GL5 (5 μM), salinosporamide A (NPI-0052, 10 nM) or bortezomib (2.7 μM) for 3 h and assayed, then dialysed against assay buffer overnight and assayed again. **b**, **c**, Plots of k_{obs} as function of inhibitor concentration for GL5 (**b**) and HT1171 (**c**). Values for k_{obs}, derived from the data in Supplementary Fig. 4, were plotted against inhibitor concentration *I*. **d**, Kinetic parameters and partition ratios of GL5 and HT1171, determined as in Supplementary Figs 3 and 4. K_{obs}/I is second-order rate constant of inactivation. Partition ratios for human proteasomes refer to the β5 subunit. Standard errors were <10%. NI, no inhibition.

and mycobacterial *blaC*-encoded β-lactamase, a serine protease-like enzyme ($IC_{50} > 100\,\mu M$). Although GL5 reversibly inhibited α-chymotrypsin with K_i 64 nM, antichymotryptic potency was minimal in 13 other oxathiazol-2-one compounds tested (Supplementary Table 4). Supplementary Table 1 summarizes biochemical and mycobactericidal properties of 24 oxathiazol-2-one compounds, including 21 we synthesized as described in the Supplementary Information.

Competitive, irreversible, mechanism-based inhibition indicated that the oxathiazol-2-one compounds inactivate the *M. tuberculosis* proteasome by covalent attack on the active site Thr 1. To test this, we trypsinized *M. tuberculosis* proteasomes that had been treated or not with HT1171. Results were identical for both wild-type and open-gate forms. Liquid chromatography–tandem mass spectrometry (LC–MS/MS) identified peptides representing 98% of the β-subunit. Only one peptide ion was unique to the *M. tuberculosis* proteasome β-subunit treated with HT1171 (Fig. 3a). Its mass was 26 Da higher than that of the N-terminal heptapeptide (TTIVALK) identified only in the untreated samples (Fig. 3a and Supplementary Fig. 7a, b), indicating the addition of a carbonyl at the expense of two hydrogen atoms ((1)→(2) in Fig. 3a). MS/MS analysis of this modified peptide (Supplementary Fig. 7a) indicated that the modification was on one of the first two residues. To determine whether the N-terminal residue was modified, the tryptic peptides from untreated and treated *M. tuberculosis* proteasome β-subunits were subjected to reductive glutaraldehydation, which modifies amino groups of proteins and peptides[19]. Subjecting the trypsinized oligopeptides to reductive glutaraldehydation without prior exposure to oxathiazol-2-one increased the mass of the N-terminal heptapeptide by 136.13 Da, consistent with modification of two primary amines, those of Thr 1 and Lys 7 (Fig. 3a, (1)→(3) and Supplementary Fig. 7c). In contrast, applying the same procedure to proteasomes that had been pre-treated with HT1171 increased the mass of the same peptide by only 68.06 Da (Fig. 3a, (2)→(4) and Supplementary Fig. 7d), indicating that one primary amine was no longer available for reductive alkylation. MS/MS analysis (Supplementary Fig. 7d) of this mono-alkylated peptide showed that only the lysine was modified, confirming that the N-terminal Thr was modified on HT1171 treatment. The mass spectrometric results were the same when HT1171 was replaced by GL5 or GL3 and when the proteasome was wild type or open-gate (not shown). These findings support the mechanism of inhibition shown in Fig. 3b: the attack of the oxathiazol-2-one by the OH of Thr 1 forms a carbonated or carbonothioated enzyme intermediate on Thr 1 that can either undergo hydrolysis to reactivate the enzyme, or donate a carbonyl to Thr 1 that links its α-amino and γ-hydroxyl groups to form an oxazolidin-2-one, a chemically stable moiety[20], consonant with the ability of oxathiazol-2-one compounds to cyclocarbonylate 1,2-aminoalcohols[21].

Structural basis of species selectivity

To determine the basis for species selectivity, we solved four crystal structures: wild-type *M. tuberculosis* proteasome following exposure to GL1 at 2.4 Å resolution and to HT1171 at 2.5 Å resolution, and the open-gate variant (20SOG) alone at 2.5 Å resolution or following exposure to HT1171 at 2.9 Å resolution. N-terminal octapeptide deletion in the α-subunit of 20SOG did not alter the overall structure of the *M. tuberculosis* proteasome (Supplementary Fig. 8a). Furthermore, wild type and open-gate proteasomes underwent the same conformational changes (described later) on inhibitor treatment (Supplementary Fig. 8b). The three structures (Fig. 4, Supplementary Figs 8, 9 and Supplementary Table 5) each confirmed that oxathiazol-2-one compounds cyclocarbonylate Thr 1. Not only the oxazolidin-2-one ring, but also its protruding methyl group and carbonyl oxygen, were

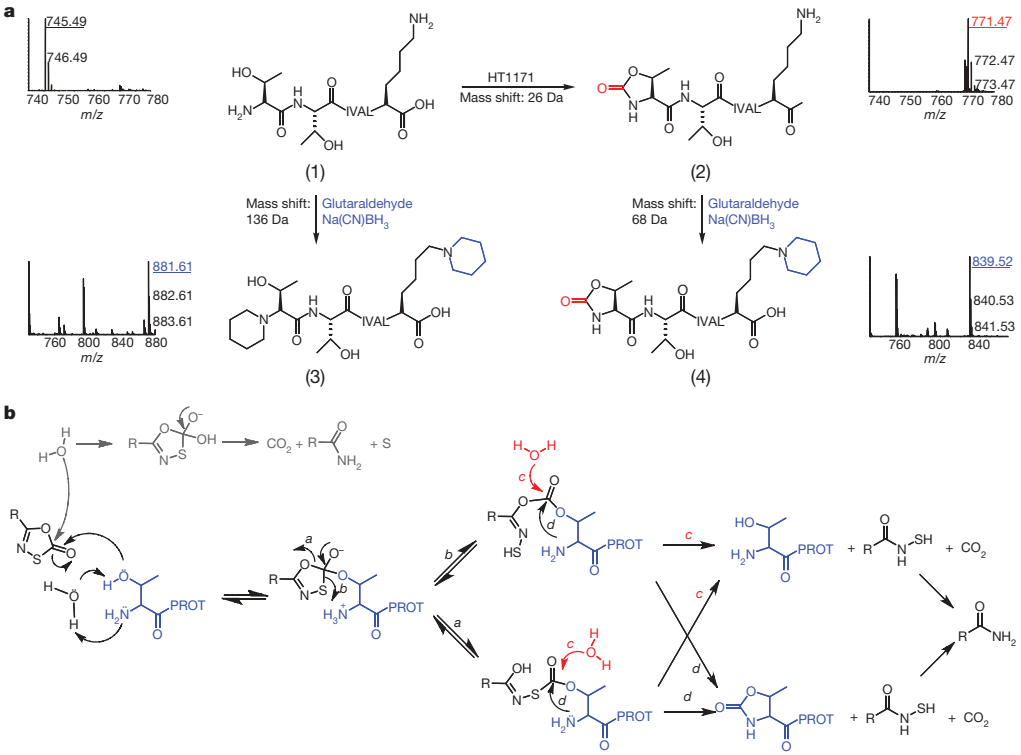


Figure 3 | LC–MS/MS identification of the modified N terminus of the *M. tuberculosis* proteasome treated with oxathiazol-2-one compounds. **a**, Mass spectra of tryptic N-terminal heptapeptides from *M. tuberculosis* proteasomes that were untreated (1), HT1171-treated (2), treated with glutaraldehyde/$Na(CN)BH_3$ after trypsin digestion (3), or treated with glutaraldehyde/$Na(CN)BH_3$ after HT1171 treatment and trypsin digestion (4). All ions were confirmed by MS/MS analysis (Supplementary Fig. 7a–d). Reaction equations illustrate proposed modification of active site Thr 1 by oxathiazol-2-one ((1)→(2)), and modification of primary amino groups at Thr 1 and Lys 7 with glutaraldehyde and $Na(CN)BH_3$ ((1)→(3) and (2)→(4)). **b**, Proposed mechanism of proteasome inactivation by oxathiazol-2-one. Paths marked by *a*, *b* and *d* lead to irreversible inhibition. In paths marked by *c*, hydrolysis of the inhibitor-enzyme intermediate allows the proteasome to degrade the oxathiazol-2-one without losing activity.

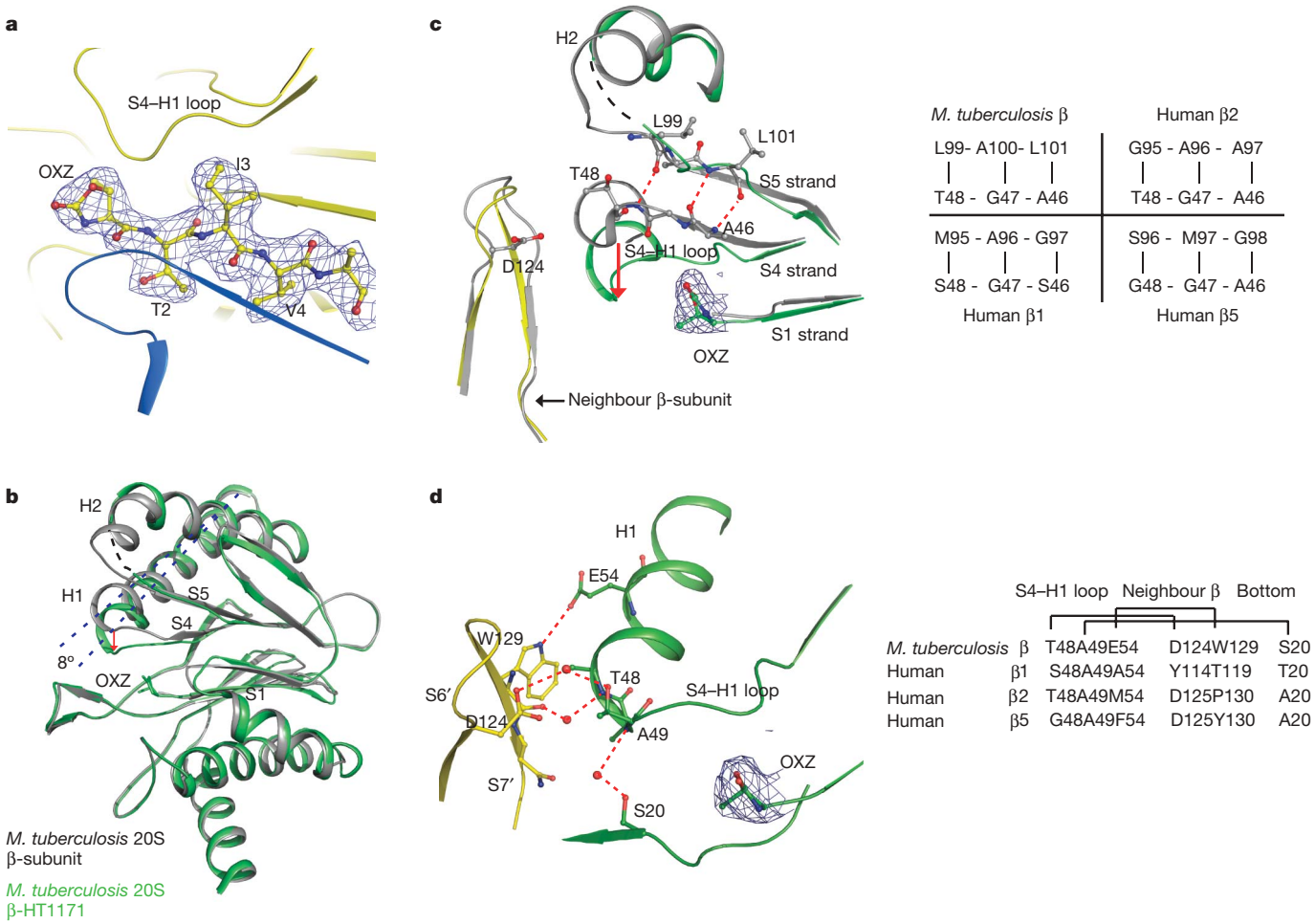


Figure 4 | Crystal structure of the full-length *M. tuberculosis* 20S proteasome after exposure to HT1171 shows cyclocarbonylation of active site Thr 1 and conformational changes in the β-subunit. **a**, $2F_o-F_c$ electron density map contoured at 1.2σ and superimposed on the crystal structure of HT1171-treated proteasome at the active site in the β-subunit. Map was calculated by omitting oxazolidin-2-one and the N-terminal four amino acids from the crystal structure. **b**, Superposition of HT1171-modified β-subunit in green with native β-subunit in grey (Protein Data Bank 2FHG). OXZ labels oxazolidin-2-one ring on Thr 1. The conformational changes can be approximately described by an 8° tilt of H1 (dashed line) and a downward shift of the S4–H1 loop (red arrow). **c**, Active site structure of HT1171-modified β-subunit in green, in comparison with native β-subunit structure in grey. In the native structure, Ala 46-Gly 47-Thr 48 is part of the S4 strand forming a β-sheet with the S5 strand. In the HT1171-treated structure, Ala 46-Gly 47-Thr 48 loses contacts with S5 and converts to the S4–H1 loop. Panel at right specifies amino acid pairs that form S4–S5 β-sheet interaction in this region in *M. tuberculosis* β-chain before oxathiazol-2-one treatment and compares these with corresponding sequences in the proteolytically active human proteasome β-chains. **d**, Active site structure of HT1171-treated *M. tuberculosis* 20S oriented to view hydrogen bonds stabilizing the new position of S4–H1 loop in the *M. tuberculosis* β-chain after oxathiazol-2-one treatment. S4–H1 loop as upper surface of the constricted substrate-binding pocket is stabilized by a hydrogen bond between Glu 54 and Trp 129 of the neighbouring β-subunit and by three pairs of water-mediated hydrogen bonds with the neighbouring β-subunit and with Ser 20 at the lower substrate-binding surface of the same subunit. Panel at right compares the hydrogen bonding of the residues of *M. tuberculosis* with human counterparts. The $2F_o-F_c$ electron density map contoured at 1.2σ is superimposed at the oxazolidin-2-one ring site in **c** and **d**. Thr 1 modification and β-subunit conformational changes on exposure to GL1 or HT1171 in the wild-type and the open-gate *M. tuberculosis* 20S are virtually the same.

resolved in the electron density (Fig. 4). The oxazolidin-2-one ring is stabilized by a hydrogen-bond network involving Ala 180, Ser 141 and Asn 24 of the neighbouring β-subunit, and a water molecule in the substrate cavity (Supplementary Fig. 9). Use of HT1171 and GL1 in the crystallographic studies and HT1171, GL5 and GL3 in the mass spectroscopic studies brought to four the number of oxathiazol-2-one compounds for which the same suicide-substrate inhibition mechanism was confirmed.

Surprisingly, the substrate-binding pocket of the *M. tuberculosis* proteasome underwent a major conformational change on cyclocarbonylation of Thr 1 by HT1171 or GL1. Such a change is unprecedented among the dozens of crystal structures of proteasomes in complex with inhibitors[11]. An ~8° downward tilt of the H1 helix in the β-subunit moved the N-terminal end of H1 (Ala 49–Phe 55) downward by as much as 4.2 Å (Fig. 4b). The H1 shift brought with it a 3-amino-acid stretch (Ala 46–Thr 48) of the S4 β-strand, converting this segment into a short loop (S4–H1). As a consequence, another short loop (Met 95-Gln 96-Gly 97) between H2 and S5 lost stabilizing contacts with the shifted components and became disordered, as illustrated by a black dashed curve in Fig. 4b. The S4–H1 loop region comprises the upper surface of the substrate-binding pocket[11,14]. Its downward shift constricted the pocket to the point that it could not accommodate a peptide substrate, affording an additional mechanism of inhibition over and above incorporation of the active site hydroxyl into an oxazolidin-2-one.

In the native *M. tuberculosis* β-subunit, the three S4–H1 loop amino acids Ala 46, Gly 47 and Thr 48 are in β-strand configuration and form a β-sheet interaction with Leu 101, Ala 100 and Leu 99, respectively, in the S5 β-strand (Fig. 4c). None of these residue pairs is conserved in the human proteasome β5 subunit and only one of them (Gly 47–Ala 96) in the human β1 and β2 subunits. One direct hydrogen bond and three pairs of water-mediated hydrogen bonds

stabilized the new position of the S4–H1 loop (Fig. 4d): between Glu 54 and Trp 129 of the neighbouring β-subunit (3.0 Å); between Thr 48, a water molecule and Asp 124 in the S6–S7 loop of the neighbouring β-subunit (2.8 and 3.1 Å, respectively); between Ala 50, a water molecule and the carbonyl O of Trp 129 in the short S7 β-strand of the neighbouring β-subunit (3.0 and 3.1 Å, respectively); and between Ala 49, a water molecule and Ser 20 in the same β-subunit (3.2 and 3.1 Å, respectively). This last pair of hydrogen bonds crosslinked the upper and the lower substrate-binding surfaces, sealing the entrance of the substrate pocket. The residue pairs involved in the post-shift hydrogen bonds are not well conserved in the human proteasome (Fig. 4d). We speculate that the upper substrate-binding surfaces (S4–H1 loops) in the three catalytic β-subunits of the human proteasome might have difficulty breaking off from their corresponding β-sheet cores on initial modification of Thr 1 by oxathiazol-2-one compounds.

Thus, selectivity of oxathiazol-2-one compounds seems to be imparted in three ways: by the presence of a 1,2-aminoalcohol at the active site of the target, which, among enzymes, is likely to be limited to the N-terminal Thr hydrolase family; by the ability of the inhibitor to bind rapidly (before its spontaneous decay) and precisely adjacent to the 1,2-aminoalcohol; and by the degree to which the amino group of the 1,2-aminoalcohol has better access to the inhibitor's carbonyl (now attached to the alcohol) than water has. The protein landscape near the 1,2-aminoalcohol can thus determine species selectivity, both in how it binds the R group on the oxathiazol-2-one and in the conformations it adopts, that may either permit or limit access of water to the intermediate formed during reaction of the oxathiazol-2-one with the active site. Detailed understanding of the sequence of steps by which oxathiazol-2-one compounds cause a major conformational shift in the substrate-binding domain could guide the design of the next generation of inhibitors selective for the *M. tuberculosis* proteasome over the human proteasome.

Discussion

Non-replicating *M. tuberculosis* displays 'phenotypic tolerance', that is, relative resistance to conventional anti-infectives, imposing the need to treat tuberculosis longer than almost any other infectious disease. Prolonged treatment leads to interruption of therapy and emergence of hereditable drug resistance. Hence agents are needed that can kill *M. tuberculosis* when its replication is halted by conditions encountered in the host, such as those imposed by inducible nitric oxide synthase[22]. Oxathiazol-2-one compounds identified here phenocopy genetic deletion of the *M. tuberculosis* proteasome[10] by killing *M. tuberculosis* rendered non-replicative by exposure to sublethal nitric oxide. Along with thioxothiazolidines that inhibit the dihydrolipoamide acyltransferase of *M. tuberculosis*[16], oxathiazol-2-one compounds are only the second class of compounds, to our knowledge, that are selectively bactericidal for a non-replicating pathogen.

Certain oxathiazol-2-one compounds are selective in other ways as well. They inhibit the *M. tuberculosis* proteasome, but not the human proteasome N-terminal threonine β1, β2 or β5 proteases, cysteine proteases, serine proteases, or metalloproteases. They do not kill monkey epithelial cells, human macrophages or any bacteria we tested other than mycobacteria. The oxathiazol-2-one compounds react with proteasomes at their active site; however, residues distant from the active site, with which oxathiazol-2-one compounds do not interact, seem to impart species selectivity.

Inhibition of protein synthesis at the stages of transcription (for example, by rifamycins) or translation (for example, by aminoglycosides and capreomycin) are among the best validated antibiotic strategies[23]. It may prove synergistic to interfere at the same time with bacterial protein degradation. The ability of a brief exposure to oxathiazol-2-one compounds to inhibit *M. tuberculosis* proteasomes permanently is a potential advantage. *M. tuberculosis* may have difficulty replacing irreversibly inactivated proteasomes, not only when the protein synthesis of *M. tuberculosis* is markedly diminished in the non-replicative state[24], but also when its protein synthesis is impaired by antibiotics.

Much of biology can be viewed as the interplay of two principles: conservation and diversity. The conservation of proteasomes across vast evolutionary distances is striking. The present work demonstrates that functionally exploitable diversity exists between the proteasomes of *M. tuberculosis* and its obligate human host.

METHODS SUMMARY

High-throughput screening was performed in 384-well plates with 33 μM test compounds. To test inhibition of proteasomes in intact mycobacteria, treated cells were centrifuged, washed twice and lysed mechanically, and aliquots (10 μg protein) of the supernatants assayed with Suc-LLVY-AMC (50 μM). Killing of DETANO-treated, non-replicating *M. tuberculosis* Erdman was tested in Sauton's medium, pH 7.4. After 4 days, samples were plated for enumeration of colony-forming units (c.f.u.) 3 weeks later. Monkey kidney cells were cultured in Dulbecco's Eagle's medium with 10% fetal bovine serum for cytotoxicity assessment after 48 h by tetrazolium reduction and microscopy. Recombinant *M. tuberculosis* proteasomes were purified as reported[15]. Human proteasomes were activated with PA28 (Boston Biochem). Kinetic studies were conducted in a Molecular Devices fluorescent plate reader with AMC-derivatized peptide substrates according to which proteasomal protease was being studied (*M. tuberculosis* or human β1, β2 or β5). Values of k_{obs} were derived from the fit of data to equation (1) (ref. 25) in Prism (GraphPad Software)

$$P = v_s t + \frac{(v_0 - v_s)}{k_{obs}}[1 - e^{(-k_{obs} t)}] \quad (1)$$

where P = product, v_s = steady-state (final) velocity, t = time, v_0 = initial velocity and k_{obs} = apparent first-order rate constant for interconversion between v_0 and v_s.

LC–MS/MS analysis used Thermo LTQ Orbitrap and Applied Biosystems QSTAR mass spectrometers. Matrix-assisted laser desorption/ionization–time of flight (MALDI–TOF) analysis was performed on a PerSeptive DE-STR instrument. Crystallization of the *M. tuberculosis* proteasome, diffraction data collection and structure solution and refinement are included along with other details in the Methods.

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

Received 9 June; accepted 4 August 2009.
Published online 16 September 2009.

1. Baumeister, W., Walz, J., Zuhl, F. & Seemuller, E. The proteasome: paradigm of a self-compartmentalizing protease. *Cell* **92,** 367–380 (1998).
2. Kropff, M. *et al.* Bortezomib in combination with intermediate-dose dexamethasone and continuous low-dose oral cyclophosphamide for relapsed multiple myeloma. *Br. J. Haematol.* **138,** 330–337 (2007).
3. Glenn, R. J. *et al.* Trypanocidal effect of α′, β′-epoxyketones indicates that trypanosomes are particularly sensitive to inhibitors of proteasome trypsin-like activity. *Int. J. Antimicrob. Agents* **24,** 286–289 (2004).
4. Darwin, K. H. *et al.* The proteasome of *Mycobacterium tuberculosis* is required for resistance to nitric oxide. *Science* **302,** 1963–1966 (2003).
5. Raviglione, M. C. & Smith, I. M. XDR tuberculosis–implications for global public health. *N. Engl. J. Med.* **356,** 656–659 (2007).
6. Corbett, E. L. *et al.* The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Arch. Intern. Med.* **163,** 1009–1021 (2003).
7. Restrepo, B. I. Convergence of the tuberculosis and diabetes epidemics: renewal of old acquaintances. *Clin. Infect. Dis.* **45,** 436–438 (2007).
8. Pearce, M. J. *et al.* Ubiquitin-like protein involved in the proteasome pathway of *Mycobacterium tuberculosis*. *Science* **322,** 1104–1107 (2008).
9. Burns, K. E. *et al.* Proteasomal protein degradation in mycobacteria is dependent upon a prokaryotic ubiquitin-like protein. *J. Biol. Chem.* **284,** 3069–3075 (2009).
10. Gandotra, S. *et al. In vivo* gene silencing identifies the *Mycobacterium tuberculosis* proteasome as essential for the bacteria to persist in mice. *Nature Med.* **13,** 1515–1520 (2007).
11. Borissenko, L. & Groll, M. 20S proteasome and its inhibitors: crystallographic knowledge for drug development. *Chem. Rev.* **107,** 687–717 (2007).
12. Lin, G. *et al.* Distinct specificities of *Mycobacterium tuberculosis* and mammalian proteasomes for *N*-acetyl tripeptide substrates. *J. Biol. Chem.* **283,** 34423–34431 (2008).
13. Kisselev, A. F. & Goldberg, A. L. Proteasome inhibitors: from research tools to drug candidates. *Chem. Biol.* **8,** 739–758 (2001).
14. Hu, G. *et al.* Structure of the *Mycobacterium tuberculosis* proteasome and mechanism of inhibition by a peptidyl boronate. *Mol. Microbiol.* **59,** 1417–1428 (2006).
15. Lin, G. *et al. Mycobacterium tuberculosis prcBA* genes encode a gated proteasome with broad oligopeptide specificity. *Mol. Microbiol.* **59,** 1405–1416 (2006).
16. Bryk, R. *et al.* Selective killing of nonreplicating mycobacteria. *Cell Host Microbe* **3,** 137–145 (2008).

17. Huth, J. R. *et al.* Toxicological evaluation of thiol-reactive compounds identified using a La assay to detect reactive molecules by nuclear magnetic resonance. *Chem. Res. Toxicol.* **20,** 1752–1759 (2007).
18. Copp, L. J. in *Enzyme Kinetics: A Modern Approach* (ed. Marangoni, A. G.) Ch. 13 158–173 (Wiley, 2003).
19. Russo, A., Chandramouli, N., Zhang, L. & Deng, H. T. Reductive glutaraldehydation of amine groups for identification of protein N-termini. *J. Proteome Res.* **7,** 4178–4182 (2008).
20. Lu, L. Q. *et al.* A new entry to cascade organocatalysis: reactions of stable sulfur ylides and nitroolefins sequentially catalyzed by thiourea and DMAP. *J. Am. Chem. Soc.* **130,** 6946–6948 (2008).
21. Rajca, A., Grobelny, D., Witek, S. & Zbirovsky, M. 5-Aryl-2-oxo-1,2,4-oxathiazoles as cyclocarbonylating agents for 2-aminoalcohols and 1,2-diamines. *Synthesis* **12,** 1032–1033 (1983).
22. MacMicking, J. D. *et al.* Identification of nitric oxide synthase as a protective locus against tuberculosis. *Proc. Natl Acad. Sci. USA* **94,** 5243–5248 (1997).
23. Walsh, C. Where will new antibiotics come from? *Nature Rev. Microbiol.* **1,** 65–70 (2003).
24. Hu, Y. M. *et al.* Protein synthesis is shutdown in dormant *Mycobacterium tuberculosis* and is reversed by oxygen or heat shock. *FEMS Microbiol. Lett.* **158,** 139–145 (1998).
25. Copeland, R. A. in *Enzymes: a Practical Introduction to Structure, Mechanism, and Data Analysis* Chs 9–10, 305–349 (Wiley, 2000).
26. Otwinoswki, Z. & Minor, W. in *Methods in Enzymology* (eds Carter, C. W. Jr & Sweet, R. M.) 307–326 (Academic, 1997).
27. McCoy, A. J. *et al. Phaser* crystallographic software. *J. Appl. Cryst.* **40,** 658–674 (2007).
28. Brünger, A. T. *et al. Crystallography & NMR system*: a new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54,** 905–921 (1998).
29. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).
30. Adams, P. D. *et al. PHENIX*: building new software for automated crystallographic structure determination. *Acta Crystallogr. D* **58,** 1948–1954 (2002).

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

Acknowledgements S. Eswaramoorthy helped with crystallography software, C. Lipinski, C. Walsh and M. Fischbach proposed reaction mechanisms, and C. Karan assisted with screening. S. Ehrt and S. Gandotra performed some bactericidal assays, C. Tsu and L. Dick donated a fluorimeter and J. Blanchard provided BlaC. Supported by NIH PO1-AI056293, NIH R01AI070285 and the Milstein Program in Chemical Biology of Infectious Diseases. X-ray diffraction data were collected at beamline X6A, X25 and X29 in the National Synchrotron Light Source, a facility supported by the US DOE and NIH. The Department of Microbiology and Immunology is supported by the William Randolph Hearst Foundation.

Author Contributions G.L. purified recombinant proteasome, conducted the screen, designed new oxathiazol-2-one compounds and performed most of the assays. D.L. and H.L. purified and crystallized recombinant proteasomes and solved their structures. L.P.S.C. helped analyse kinetics. H.D. performed mass spectrometry. H.T. synthesized oxathiazol-2-one compounds under the supervision of J.D.W.; G.V. studied human macrophages. K.W. conducted studies with viable *M. tuberculosis*. J.S. and T.C. performed kinetic, bactericidal and cytotoxicity experiments. C.N. organized the effort and helped design and interpret experiments. C.N., G.L. and H.L. wrote the paper.

Author Information Coordinates have been deposited in the Protein Data Bank under ID codes 3H6F, 3H6I, 3HFA and 3HF9. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.N. (cnathan@med.cornell.edu), G.L. (gal2005@med.cornell.edu) and H.L. (hli@bnl.gov).

METHODS

High-throughput screen. Compounds from ChemDiv, ChemBridge, Spectrum, Prestwick Chemicals and Cerep were dissolved in DMSO and robotically dispensed to Falcon Microtest 384-well plates, with DMSO as vehicle control. After a 45-min preincubation with *M. tuberculosis* 20SOG (10 μl) at 37 °C, reaction buffer containing substrate (5 μl) was added. Final concentrations were as follows: test compounds, 33 μM; Suc-LLVY-AMC, 50 μM; *M. tuberculosis* 20SOG, 1 nM; HEPES, 20 mM; EDTA, 0.5 mM; pH 7.5. Plates were placed on an orbital shaker in a humidified incubator at 37 °C for 45 min. Fluorescence was recorded at excitation 360 nm, emission 460 nm. Z' values were >0.5. We screened 20,400 compounds and picked 72 (0.35%) that gave inhibition ≥65%. Of these, inhibition was confirmed with 26 compounds; concentration-dependent activity was most potent for HT1171 and GL5. A search of the library with the core structure (1,3,4-oxathiazol-2-one) identified GL6. GL1 and GL3 were identified in commercial collections using SCIFINDER. GL1, GL3 and GL5 were purchased from TimTec and GL6 was purchased from ChemDiv. HT1171 was resynthesized and the other compounds listed in Supplementary Table 1 were synthesized at the Abby and Howard P. Milstein Chemistry Core Facility of Weill Cornell Medical College.

Mycobactericidal activity. Initial experiments showed that GL5 killed *M. bovis* var. BCG in synergy with nitric oxide, which was provided from the decomposition of 2,2-(hydroxynitrosohydrazino)-bis-ethanamine (DETA-NO) at 50 μM (Supplementary Fig. 1). In the experiment illustrated in Supplementary Fig. 1, DETA-NO alone reduced but did not completely prevent replication of BCG, whereas in the experiments with *M. tuberculosis* illustrated in Fig. 1e, DETA-NO prevented replication without reducing viability as assessed by plating. *M. tuberculosis* Erdman was cultivated in Sauton's medium pH 7.4 with 0.4% L-asparagine, 0.2% glycerol and 0.02% Tween 80. Mid-log phase cultures (A_{580} 0.8–1.0) were diluted to A_{580} 0.05–0.1 and quantified by c.f.u. Mycobacteria were incubated with bortezomib, GL5 or HT1171 at indicated concentrations in the presence of DETA-NO (50 μM) in 96-well plates in 200 μl for 4 days, then serially diluted in PBS with 0.02% Tween 80, pH 7.2 and plated for c.f.u. determination on Middlebrook 7H11 agar plates with 10% Middlebrook OADC enrichment (Fig. 1e).

Impact of oxathiazol-2-one compounds on eukaryotic cells. Vero76 African green monkey kidney cells (ATCC CRL-1587) were cultured in DMEM containing 4.5 g l^{-1} glucose, 0.584 g l^{-1} L-glutamine, 1 mM pyruvate, 100 IU ml^{-1} penicillin, 100 μg ml^{-1} streptomycin and 10% fetal bovine serum. Fresh medium was supplied containing test compounds at 25 or 50 μM. At indicated intervals of 1–25 h, this medium was removed and fresh medium added without test compounds. After exposure (48 h) to GL5, HT1171 or bortezomib at indicated concentrations, viability was assessed by the CellTiter96 Aqueous Assay (Promega) for tetrazolium reduction according to the manufacturer's protocol. Human monocyte-derived macrophages were allowed to differentiate *in vitro* for 2 weeks. GL5 or bortezomib were added in DMSO at indicated concentrations (<1%). Photomicrographs were taken 1 week later.

Titration of *M. tuberculosis* and human proteasome by GL5 or HT1171. GL5 or HT1171 at concentration 0.2–50 μM were incubated with 3 nM *M. tuberculosis* 20SOG or 0.5 nM Hu20S (from human red blood cells, purchased from Boston Biochem), 2 nM PA28 (eukaryotic proteasome activator, Boston Biochem) in assay buffer overnight (19 h). The remaining activity of proteasomes was assayed with 50 μM Suc-LLVY-AMC at 37 °C. The plot of v_i/v_0 versus I/E_t was used to estimate the partition ratios[18]. Here v_i is the velocity of enzymatic reaction in the presence of inhibitor, v_0 is the velocity of enzymatic reaction in the absence of inhibitor, E_t is the total concentration of the active sites of the enzymes. In this experiment, E_t of *M. tuberculosis* 20SOG is 42 nM counting the concentration of active sites (14 active sites for each *M. tuberculosis* proteasome core particle), and E_t of Hu20S is 1 nM counting the concentration of active sites (two β5 active sites for each human proteasome core particle). Because GL5 did not completely inhibit the Hu20S activity in our experiment (Supplementary Fig. 3a, b), its partition ratio as listed in Fig. 2b was significantly underestimated. Standard errors were less than 10%.

Determination of mode of inhibition. Assays were conducted on a Spectra MAX Gemini plate-reader from Molecular Devices. Assay buffer containing 1 nM *M. tuberculosis* 20SOG was added to the GL5 (5 μM, final concentration) and Z-VLR-AMC (indicated concentrations ranging from 15 to 60 μM). The reaction progress was recorded by monitoring fluorescence at 460 nm (λ_{ex} = 360 nm) for 60 min at 37 °C (Supplementary Fig. 3c). Values of k_{obs} were derived from the fit of data to equation (1) (ref. 25) in Prism (GraphPad Software), and then were plotted against substrate concentration (Supplementary Fig. 3d). The diminution in k_{obs} with increasing concentration of substrate indicated that GL5 was a competitive inhibitor. The line was the fit of data to equation (2) (ref. 25), where k_{inact} is the maximal rate of enzyme inactivation, I is the concentration of inhibitor, K_i is the dissociation constant for the inhibitor, S is the concentration of substrate and K_m is the substrate concentration at which the reaction reaches half-maximal velocity.

$$P = v_s t + \frac{(v_0 - v_s)}{k_{obs}}[1 - e^{(-k_{obs} t)}] \quad (1)$$

$$k_{obs} = \frac{k_{inact} I}{K_i(1 + S/K_m) + I} \quad (2)$$

Kinetics. Kinetic measurements were performed at 37 °C on a 96-well plate assay on a Spectra MAX Gemini plate-reader from Molecular Devices with 1.5 nM *M. tuberculosis* 20SOG, 5 nM *M. tuberculosis* 20SWT or 0.5 nM Hu20S plus 2 nM PA28 in 20 mM HEPES, 0.5 mM EDTA, pH 7.5, and 25 μM Suc-LLVY-AMC (for Hu20S β5 activity), 25 μM Z-LLE-AMC (for Hu20S β1 activity) or 25 μM Z-VLR-AMC (for h20S β2 activity). The release of AMC from substrate cleavage was monitored (λ_{ex} = 360 nm, λ_{em} = 460 nm) for 120 min. The k_{obs} values were determined by fitting the raw data to equation (1). The slopes of the plots of k_{obs} versus I gave an apparent value of k_{obs}/I, also referred to as k_{inact}/K_i for an irreversible inhibitor, which was then corrected by equation (3) to compensate for the effect of substrate competition, where *app* is the apparent value at different inhibitor concentrations (Supplementary Figs 4, 5 and Supplementary Table 1).

$$k_{obs}/I = (k_{obs}/I)^{app}(1 + S/K_m) \quad (3)$$

Inhibition of degradation of β-casein by *M. tuberculosis* 20SOG. *M. tuberculosis* 20SOG (final concentration 12 nM) was pre-incubated in the assay buffer with 10 μM GL5/GL6 at 37 °C for 1 h. β-Casein (Sigma) was added at a final concentration of 0.2 mg ml^{-1}. The mixtures were incubated at room temperature. Aliquots were withdrawn at 1, 2, 3 and 4 h, mixed with SDS– PAGE loading buffer and analysed by SDS–PAGE (Supplementary Fig. 6).

Stability of oxathiazol-2-one compounds in RPMI medium. Oxathiazol-2-one compounds in DMSO (final concentration 75 μM) were added to RPMI medium, which mimics human extracellular fluid, using pyrazinamide (75 μM) as an internal control. The solutions were incubated at 37 °C and aliquots were withdrawn at designed time points for analysis by ultra performance liquid chromatography/MS (Waters). Normalized data were fitted to an exponential decay equation to yield the half time for each compound.

LC–MS/MS analysis. *M. tuberculosis* 20SOG or *M. tuberculosis* 20SWT (415 μg ml^{-1}; 8.2 μM active sites) in 20 mM HEPES, 0.5 mM EDTA, pH 7.5, was incubated with 500 μM HT1171, GL5 or GL3 at room temperature until the activity assay demonstrated that inactivation was complete. A control sample was incubated for the same time with an equivalent volume of DMSO. The samples were then run on a SDS–PAGE gel to separate α- and β-subunits. The gel bands corresponding to untreated and inhibitor-treated PrcB were excised from the gel, reduced with 10 mM dithiothreitol, alkylated with 55 mM iodoacetamide and digested with sequencing grade modified trypsin (Promega) in ammonium bicarbonate buffer at 37 °C overnight. The digestion products were analysed by LC–MS/MS and LC–MS with Thermo LTQ Orbitrap and Applied Biosystems QSTAR mass spectrometers, respectively. One tenth of the digestion products for each sample was also analysed by MALDI–TOF with a PerSeptive MALDI–TOF DE-STR mass spectrometer. For LC–MS/MS analysis, each digestion product was separated by gradient elution with a Dionex capillary/nano-HPLC system that is directly interfaced with the mass spectrometer. MS/MS data were searched using the MASCOT search engine for identifying proteins and modifications (Fig. 3a and Supplementary Fig. 7a–d).

For in-gel modification of primary amine groups, gel slices of untreated and treated proteins were incubated with 500 mM sodium cyanoborohydride and 2.5% glutaraldehyde at 37 °C for 1 h. Reactions were stopped by addition of 1 M Tris-HCl. Note: sodium cyanoborohydride is toxic, glutaraldehyde is carcinogenic and mutagenic. All reactions were carried out in a fume hood.

Protease selectivity of oxathiazol-2-one compounds. Although the oxathiazol-2-one GL5 displayed marked selectivity for the *M. tuberculosis* proteasome over the three human proteasome subunits, the question remained whether oxathiazol-2-one compounds might inhibit proteases of other classes, given that many heterocyclic compounds are irreversible protease inhibitors. Therefore, we tested oxathiazol-2-one compounds on bovine spleen cathepsin B, a cysteine protease, human matrix metalloproteinase 2 (gelatinase), a metalloprotease, and bovine α-chymotrypsin, a serine protease. Most oxathiazol-2-one compounds were not active or weak inhibitors of Cat B and MMP-2 with IC_{50} values greater than 50 μM.

Inhibition of α-chymotrypsin by GL5 and HT1171. GL6, which was inactive on the *M. tuberculosis* proteasome, inhibited α-chymotrypsin more potently than GL5, with IC_{50} values of 0.1 and 0.4 μM, respectively. We further determined that

inhibition of α-chymotrypsin by GL5 was not time-dependent (data not shown), indicating that it was reversible. Double-reciprocal analysis of inhibition at different concentrations of GL5 versus different concentrations of substrate demonstrated that GL5 was a competitive inhibitor of α-chymotrypsin, with $K_i = 64$ nM (not shown). However, other oxathiazol-2-one compounds showed only weak activity against bovine α-chymotrypsin (Supplementary Table 4). For example, the IC_{50} for HT1171 was 18.9 μM.

Glutathione reactivity of oxathiazol-2-one compounds. Oxathiazol-2-one compounds (10 μM) were spotted in a 96-well plate in triplicate. *N*-ethylmaleimide (NEM) (10 μM) and DMSO were used as positive and negative controls, respectively. An identical 96-well plate was prepared to serve as the blank. The working buffer (100 mM potassium phosphate, 2 mM EDTA, pH 7.0) with glutathione (25 μM) was added to each well on one plate and working buffer without glutathione was added to the blank plate. The plates were incubated for 30 min at room temperature, at which point 5,5′-dithio-bis(2-nitrobenzoic acid) (DTNB; 100 μM) was added to each well. The plates were shaken and absorbance read at 405 nm. Percentage of reacted glutathione was calculated as % GSH reacted = (negative control − sample)/(negative control − positive control) × 100.

Purification and crystallization of the *M. tuberculosis* 20S proteasome. The wild-type *M. tuberculosis* proteasome was purified as described[14]. In our current work, we improved the sample preparation and crystallization conditions in order to obtain *M. tuberculosis* 20S proteasome crystals that diffracted to better than 3 Å resolution[14]. Briefly, the purified and concentrated *M. tuberculosis* proteasome was dialysed against 10 mM HEPES buffer at pH 7.5 for 12 h (with 0.2 M NaCl in the case of *M. tuberculosis* 20SOG). Then the sample was set up in crystallization plates with the sitting drop method at 4 °C. The crystallization droplet contained 3 μl of 20S proteasome at a concentration of 10 mg ml^{-1}, mixed with 3 μl of well solution containing 60 mM Na citrate (pH 5.8) and 13% PEG 6000 as the precipitant. The crystals with the approximate dimensions of 200 × 60 × 50 μm grew after 5–7 days. For cryo-crystallography, the crystal solution was replaced in several concentration steps with a cryo-protecting solution containing the original mother liquid and 35% dimethylformamide. For inhibitor soaking, the crystals were moved into the cryo-protecting solution containing 1 mM HT1171 or 1 mM GL1, and incubated at 4 °C for 10 to 14 h before the crystals were flash frozen.

Diffraction data collection and structure solution. X-ray diffraction data were collected at National Synchrotron Light Source in Brookhaven National Laboratory. The diffraction data were processed and scaled with the HKL2000 package[26]. The crystals of the *M. tuberculosis* 20S proteasome soaked with various inhibitors belonged to the same space group *P*21 (See Supplementary Table 5 for the statistics of the crystals). For crystals with unit cells different from our previously reported structure, molecular replacement was carried out using the program PHASER[27] with the published structure as search model (Protein Data Bank ID 2FGH)[14] and the top solution was subjected to rigid body refinement in CNS[28]. For crystals with the same unit cell parameters as our published structure, the native wild-type 20S structure was used directly to calculate the difference density maps.

Structure refinement. The following is the refinement procedure for the structure of wild-type 20S soaked with HT1171; the structure of GL1 was refined in the same way. The two loop regions, Thr 48 to Val 53 and Gly 89 to Leu 101, were rebuilt according to the electron density in program Coot[29]. The refinement was carried out using CNS with simulated annealing protocol. Fourteen non-crystallographic symmetry (NCS) restraints were applied during the refinements. A composite omit map was calculated to verify the conformation of the S4–H1 loop region. Several rounds of manual rebuilding and refinement improved the *R* factor to a value of 24.2% for R_{work} and 25.6% for R_{free}. At this point, the oxazolidin-2-one structure covalently linked to Thr 1 was built into the model. The topology and parameter files were generated for oxazolidin-2-one-modified Thr 1 with idealized geometry. The energy minimization and individual B-factor refinement were carried out with NCS restraints in CNS. Water molecules were then added. Certain residues had different conformations in different NCS-related subunits, and these side chains were fixed. The structure was further refined with relaxed NCS restraints. For the open-gated structures, further refinement of the model against the crystallographic data was achieved using the refinement program of PHENIX[30], which included anisotropic scaling, bulk solvent correction, rigid body refinement, simulated annealing, NCS and TLS (Translation/Libration/Screw, determined by TLSMD server) refinement. In the final models, >99% of the residues are in the core and allowed regions of the Ramachandran plot. Refinement and model statistics for all structures are given in Supplementary Table 5.

LETTERS

Universality of galactic surface densities within one dark halo scale-length

Gianfranco Gentile[1,2], Benoit Famaey[2,3,4], HongSheng Zhao[5,6] & Paolo Salucci[7]

It was recently discovered that the mean dark matter surface density within one dark halo scale-length (the radius within which the volume density profile of dark matter remains approximately flat) is constant across a wide range of galaxies[1]. This scaling relation holds for galaxies spanning a luminosity range of 14 magnitudes and the whole Hubble sequence[1–3]. Here we report that the luminous matter surface density is also constant within one scale-length of the dark halo. This means that the gravitational acceleration generated by the luminous component in galaxies is always the same at this radius. Although the total luminous-to-dark matter ratio is not constant, within one halo scale-length it is constant. Our finding can be interpreted as a close correlation between the enclosed surface densities of luminous and dark matter in galaxies[4].

As a result of mass modelling using a Burkert dark matter halo[5], the density profile of which is defined by a constant central density ρ_0, and a scale-length (or core radius, at which the local dark matter volume density reaches a quarter of its central value) r_0, the product $\rho_0 r_0$ has been found to be the same for all galaxies[1]: $141^{+82}_{-52}\ M_\odot \mathrm{pc}^{-2}$, where $M_\odot \mathrm{pc}^{-2}$ means solar masses per square parsec. The mean dark matter surface density within r_0, $\langle\Sigma\rangle_{0,\mathrm{dark}}$, is then also constant when defined in terms of $M_{<r0}$, the enclosed mass between $r=0$ and $r=r_0$: $\langle\Sigma\rangle_{0,\ \mathrm{dark}} = M_{<\ r_0}/\pi r_0^2 \approx 0.51\rho_0 r_0 = 72^{+42}_{-27}\ M_\odot \mathrm{pc}^{-2}$. We note that the mean surface density as defined above is also proportional to the gravitational acceleration generated by the dark matter halo at r_0, which is also universal: $g_{\mathrm{dark}}(r_0) = G\pi\langle\Sigma\rangle_{0,\ \mathrm{dark}} = 3.2^{+1.8}_{-1.2}\ 10^{-9}\mathrm{cm\ s}^{-2}$.

We sought to understand the intriguing universality[1–3] of the mean dark matter surface density within r_0 (and therefore of the gravitational acceleration due to dark matter at r_0). Hence, we searched for the dependency with galaxy magnitude of $g_{\mathrm{bary}}(r_0)$, the gravitational acceleration due to baryons (that is, luminous matter) at r_0, and therefore also of $\langle\Sigma\rangle_{0,\mathrm{bary}} = g_{\mathrm{bary}}(r_0)/G\pi$, the baryonic mean surface density within r_0. In Fig. 1 we plot $g_{\mathrm{bary}}(r_0)$ and $\langle\Sigma\rangle_{0,\mathrm{bary}}$ as a function of galaxy magnitude. As in the case of dark matter, these values are universal: $g_{\mathrm{bary}}(r_0) = 5.7^{+3.8}_{-2.8}\ 10^{-10}$ cm s^{-2}. We note that, owing to deviation from spherical symmetry, the error bars on $\langle\Sigma\rangle_{0,\mathrm{bary}}$ should be increased by at most 0.1 dex (decimal exponent, which is the unit of the logarithmic scale) in Fig. 1, but not the error bars on $g_{\mathrm{bary}}(r_0)$.

This is therefore a new scaling relation relating baryonic parameters to dark halo parameters. Together with the universality of the dark matter surface density within r_0 (ref. 1), it means that the baryonic-to-dark matter ratio is also universal within r_0. Our sample contains both small dark-matter-dominated dwarf spheroidals and large late-type spirals, so this finding is not linked with a possible universality of the total baryonic-to-dark matter ratio within the sample. Neither does this result mean that the baryonic surface density of galaxies is constant, a misconception that used to be known as Freeman's law.

Because the central surface density of baryons actually varies by about four orders of magnitudes within the range of luminosities spanned in this study[1], our finding implies that the larger core radii r_0 of larger and more luminous galaxies compensate precisely for their larger baryonic surface densities to keep the same mean baryonic surface density within r_0. Whereas the universality of the dark matter surface density[1] within r_0 implies that the central dark matter density ρ_0 is precisely anti-correlated with the core radius, this new result means that, for a galaxy of a given size, the central baryonic surface density is actually correlated with the core radius. A large central luminous density thus implies a large core radius, and in turn a small central dark matter density. This precise balance must be the result of some unknown, fine-tuned process in galaxy formation, because it is a priori difficult to envisage how such relations between dark and baryonic galaxy parameters can be achieved across galaxies that have experienced significantly different evolutionary histories, including numbers of mergers, baryon cooling or feedback from supernova-driven winds.

Another way of looking at this new relation is to say that the core radius of the dark matter halo is the radius beyond which the

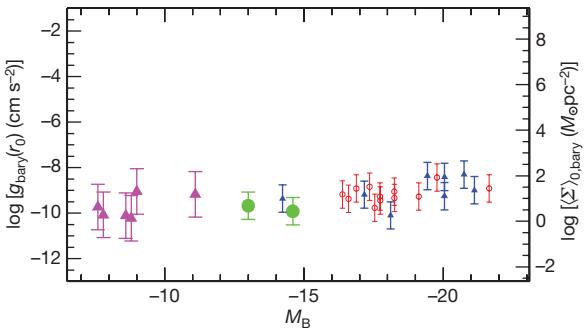


Figure 1 | Universality of the average surface density (and gravity) of baryons within the halo core radius. $\langle\Sigma\rangle_{0,\mathrm{bary}}$ and $g_{\mathrm{bary}}(r_0)$ are plotted as a function of the B-band absolute magnitude of the galaxies. From the original sample[1], we used the dwarf spheroidals data, NGC 3741 (ref. 14), DDO 47 (ref. 15), and the two samples of spiral galaxies[3,16], which all together span the whole magnitude range probed in the original sample[1]. We excluded the galaxies that had uncertain values of $g_{\mathrm{bary}}(r_0)$ owing to low-quality rotation curves, a low-quality fit, unrealistic stellar mass-to-light ratio ($\leq$0.1), a core radius much larger than the last measured point of the rotation curve, or poor sampling of the data. The error bars on the calculated $g_{\mathrm{bary}}(r_0)$ were based on a combination of the uncertainties on the stellar mass-to-light ratio and on the derived core radii (see Supplementary Information for details). The magenta triangles are the dwarf spheroidal galaxies[1], the left green point is NGC 3741 (ref. 14) and the right green point is DDO 47(ref. 15), and the empty red circles and filled blue triangles symbols are the Spano[3] and THINGS[16] spiral galaxies samples, respectively.

[1]Sterrenkundig Observatorium, Universiteit Gent, Krijgslaan 281, B-9000 Gent, Belgium. [2]Institut d'Astronomie et d'Astrophysique, Université Libre de Bruxelles, CP 226, Bvd du Triomphe, B-1050, Bruxelles, Belgium. [3]CNRS UMR 7550, Observatoire Astronomique, Université de Strasbourg, F-67000 Strasbourg, France. [4]AIfA, Universität Bonn, D-53121 Bonn, Germany. [5]SUPA, School of Physics and Astronomy, University of St Andrews, KY16 9SS, UK. [6]Leiden University, Sterrewacht and Instituut-Lorentz, Niels-Bohrweg 2, 2333 CA, Leiden, the Netherlands. [7]SISSA International School for Advanced Studies, via Beirut 4, I-34151, Trieste, Italy.

gravitational acceleration due to the baryons drops below a certain value, of the order of $6 \times 10^{-10}\,\mathrm{cm\,s^{-2}}$ (or $\log(g_{\mathrm{bary}}) \approx -9.2$, where g_{bary} is in centimetres per second squared). It is thus tempting to relate our finding to the mass discrepancy–acceleration relation[4,7,8], linking together the gravitational accelerations generated by the dark and luminous components at all radii in galaxies: other manifestations of this phenomenon are the universal rotation curve[9,10] (a parametrization of the rotation curves of spiral galaxies in which the rotation velocity depends only on radius and luminosity), or the radial Tully–Fisher[11] relation (a series of relations between luminosity and rotation velocity as a function of radius). Indeed, if the gravitational acceleration generated by the baryons is constant at r_0, it is a natural consequence of this mass discrepancy–acceleration relation that the baryonic-to-dark matter ratio should also be constant within this radius. Put another way, the mass discrepancy–acceleration relation means that the enclosed baryonic-to-dark matter ratio at any radius is uniquely correlated with the average surface density inside it[12].

To be more speculative, if we compare the effective dark matter profile predicted by the analytic form of the mass discrepancy–acceleration relation[7] with the enclosed mass profile of a Burkert halo[5] surrounding a point mass such that $g_{\mathrm{bary}} = 5.7 \times 10^{-10}\,\mathrm{cm\,s^{-2}}$ at the core radius, we find that the two profiles are very similar in shape. Moreover, this effective dark matter profile typically predicts a universal maximum gravitational acceleration generated by the dark halo[13], of the order of $3 \times 10^{-9}\,\mathrm{cm\,s^{-2}}$. For any Burkert halo the maximum acceleration is generated precisely at r_0: from there, it then follows naturally that the gravity due to dark matter at r_0 should be universal and of that order of magnitude[1,13]. Then, from this universal and maximal acceleration due to dark matter at r_0, the mass discrepancy–acceleration relation (by definition) predicts that there is a universal gravity due to baryons at this radius, and thus a universal mean luminous surface density within it.

Received 28 April; accepted 18 August 2009.

1. Donato, F. *et al.* A constant dark matter halo surface density in galaxies. *Mon. Not. R. Astron. Soc.* **397,** 1169–1176 (2009).
2. Kormendy, J. & Freeman, K. C. Scaling laws for dark matter halos in late-type and dwarf spheroidal galaxies. *IAU Symp. Astron. Soc. Pacif.* **220,** 377–395 (2004).
3. Spano, M. *et al.* GHASP: an Hα kinematic survey of spiral and irregular galaxies—V. Dark matter distribution in 36 nearby spiral galaxies. *Mon. Not. R. Astron. Soc.* **383,** 297–316 (2008).
4. McGaugh, S. S. The mass discrepancy-acceleration relation: disk mass and the dark matter distribution. *Astrophys. J.* **609,** 652–666 (2004).
5. Burkert, A. The structure of dark matter halos in dwarf galaxies. *Astrophys. J.* **447,** L25–L28 (1995).
6. McGaugh, S. S. Balance of dark and luminous mass in rotating galaxies. *Phys. Rev. Lett.* **95,** 171302 (2005).
7. Milgrom, M. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *Astrophys. J.* **270,** 365–370 (1983).
8. Famaey, B., Gentile, G., Bruneton, J.-P. & Zhao, H. S. Insight into the baryon-gravity relation in galaxies. *Phys. Rev. D* **75,** 063002 (2007).
9. Persic, M., Salucci, P. & Stel, F. The universal rotation curve of spiral galaxies—I. The dark matter connection. *Mon. Not. R. Astron. Soc.* **281,** 27–47 (1996).
10. Gentile, G. MOND and the universal rotation curve: similar phenomenologies. *Astrophys. J.* **684,** 1018–1025 (2008).
11. Yegorova, I. A. & Salucci, P. The radial Tully-Fisher relation for spiral galaxies—I. *Mon. Not. R. Astron. Soc.* **377,** 507–515 (2007).
12. Zwaan, M. A. *et al.* The Tully-Fisher relation for low surface brightness galaxies: implications for galaxy evolution. *Mon. Not. R. Astron. Soc.* **273,** L35–L38 (1995).
13. Milgrom, M. & Sanders, R. H. MOND predictions of halo phenomenology in disc galaxies. *Mon. Not. R. Astron. Soc.* **357,** 45–48 (2005).
14. Gentile, G., Salucci, P., Klein, U. & Granato, G. L. NGC 3741: the dark halo profile from the most extended rotation curve. *Mon. Not. R. Astron. Soc.* **375,** 199–212 (2007).
15. Gentile, G., Burkert, A., Salucci, P., Klein, U. & Walter, F. The dwarf galaxy DDO 47 as a dark matter laboratory: testing cusps hiding in triaxial halos. *Astrophys. J.* **634,** L145–L148 (2005).
16. de Blok, W. J. G. *et al.* High-resolution rotation curves and galaxy mass models from THINGS. *Astron. J.* **136,** 2648–2719 (2008).

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

Acknowledgements G.G. was supported by the FWO-Vlaanderen (Belgium), and B.F. was supported in part by the Alexander von Humboldt foundation (Germany), the FNRS (Belgium), and the CNRS (France). We thank A. Jorissen and L. Hill for their comments on the manuscript.

Author Contributions G.G. and B.F. contributed to analysing and interpreting the data, making the figures, and writing the paper. HS.Z. and P.S. contributed to interpreting the data.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.G. (gianfranco.gentile@ugent.be).

LETTERS

Plasmon lasers at deep subwavelength scale

Rupert F. Oulton[1]*, Volker J. Sorger[1]*, Thomas Zentgraf[1]*, Ren-Min Ma[3], Christopher Gladden[1], Lun Dai[3], Guy Bartal[1] & Xiang Zhang[1,2]

Laser science has been successful in producing increasingly high-powered, faster and smaller coherent light sources[1–9]. Examples of recent advances are microscopic lasers that can reach the diffraction limit, based on photonic crystals[3], metal-clad cavities[4] and nanowires[5–7]. However, such lasers are restricted, both in optical mode size and physical device dimension, to being larger than half the wavelength of the optical field, and it remains a key fundamental challenge to realize ultracompact lasers that can directly generate coherent optical fields at the nanometre scale, far beyond the diffraction limit[10,11]. A way of addressing this issue is to make use of surface plasmons[12,13], which are capable of tightly localizing light, but so far ohmic losses at optical frequencies have inhibited the realization of truly nanometre-scale lasers based on such approaches[14,15]. A recent theoretical work predicted that such losses could be significantly reduced while maintaining ultrasmall modes in a hybrid plasmonic waveguide[16]. Here we report the experimental demonstration of nanometre-scale plasmonic lasers, generating optical modes a hundred times smaller than the diffraction limit. We realize such lasers using a hybrid plasmonic waveguide consisting of a high-gain cadmium sulphide semiconductor nanowire, separated from a silver surface by a 5-nm-thick insulating gap. Direct measurements of the emission lifetime reveal a broad-band enhancement of the nanowire's exciton spontaneous emission rate by up to six times owing to the strong mode confinement[17] and the signature of apparently threshold-less lasing. Because plasmonic modes have no cutoff, we are able to demonstrate downscaling of the lateral dimensions of both the device and the optical mode. Plasmonic lasers thus offer the possibility of exploring extreme interactions between light and matter, opening up new avenues in the fields of active photonic circuits[18], bio-sensing[19] and quantum information technology[20].

Surface plasmon polaritons are the key to breaking down the diffraction limit of conventional optics because they allow the compact storage of optical energy in electron oscillations at the interfaces of metals and dielectrics[11–13]. Accessing subwavelength optical length scales introduces the prospect of compact optical devices with new functionalities by enhancing inherently weak physical processes, such as fluorescence and Raman scattering of single molecules[19] and nonlinear phenomena[21]. An optical source that couples electronic transitions directly to strongly localized optical modes is highly desirable because it would avoid the limitations of delivering light from a macroscopic external source to the nanometre scale, such as low coupling efficiency and difficulties in accessing individual optical modes[22].

Achieving stimulated amplification of surface plasmon polaritons at visible frequencies remains a challenge owing to the intrinsic ohmic losses of metals. This has driven recent research to examine stimulated surface plasmon polariton emission in systems that exhibit low loss, but only minimal confinement, which excludes such schemes from the rich new physics of nanometre-scale optics[14,15]. Recently, we have theoretically proposed a new approach hybridizing dielectric waveguiding with plasmonics, in which a semiconductor nanowire sits on top of a metallic surface, separated by a nanometre-scale insulating gap[16]. The coupling between the plasmonic and

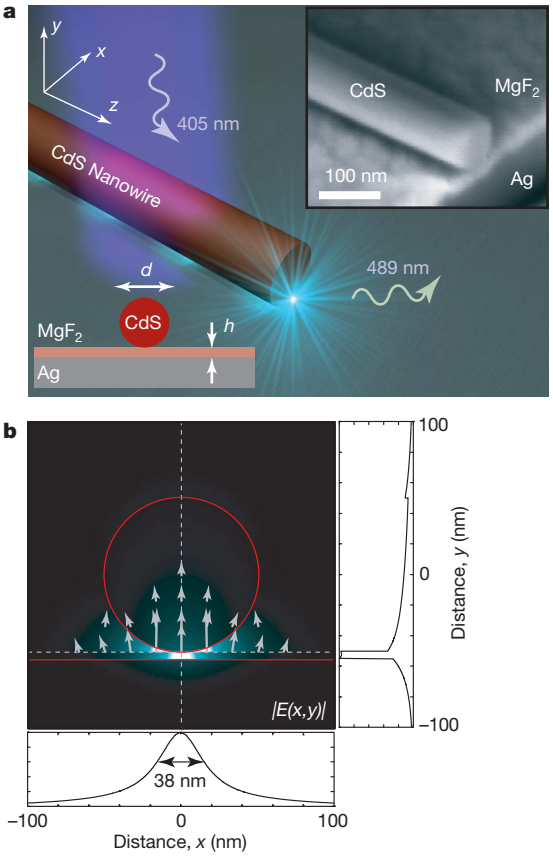


Figure 1 | The deep subwavelength plasmonic laser. a, The plasmonic laser consists of a CdS semiconductor nanowire on top of a silver substrate, separated by a nanometre-scale MgF_2 layer of thickness h. This structure supports a new type of plasmonic mode[16] the mode size of which can be a hundred times smaller than a diffraction-limited spot. The inset shows a scanning electron microscope image of a typical plasmonic laser, which has been sliced perpendicular to the nanowire's axis to show the underlying layers. **b**, The stimulated electric field distribution and direction $|E(x,y)|$ of a hybrid plasmonic mode at a wavelength of $\lambda = 489$ nm, corresponding to the CdS I_2 exciton line[25]. The cross-sectional field plots (along the broken lines in the field map) illustrate the strong overall confinement in the gap region between the nanowire and metal surface with sufficient modal overlap in the semiconductor to facilitate gain.

[1]NSF Nanoscale Science and Engineering Centre, 3112 Etcheverry Hall, University of California, Berkeley, California 94720, USA. [2]Materials Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA. [3]State Key Lab for Mesoscopic Physics and School of Physics, Peking University, Beijing 100871, China. *These authors contributed equally to this work.

waveguide modes across the gap enables energy storage in non-metallic regions. This hybridization allows surface plasmon polaritons to travel over larger distances with strong mode confinement[23] and the integration of a high-quality semiconductor gain material with the confinement mechanism itself.

In this Letter, we use the 'hybrid plasmonics' approach to show experimentally the laser action of surface plasmon polaritons with mode areas as small as $\lambda^2/400$. The truly nanometre-scale plasmonic laser devices consist of cadmium sulphide (CdS) nanowires[24] on a silver film, where the gap layer is magnesium fluoride (MgF_2, Fig. 1a). The close proximity of the semiconductor and metal interfaces concentrates light into an extremely small area as much as a hundred times smaller than a diffraction-limited spot[16] (Fig. 1b). To show the unique properties of hybridized plasmon modes, we compare the plasmonic lasers directly with CdS nanowire lasers on a quartz substrate, similar to typical nanowire lasers reported before[5–7]. Here, we will refer to these two devices as plasmonic and photonic lasers, respectively.

We optically pump these laser devices at a wavelength of 405 nm and measure emission from the dominant I_2 CdS exciton line at 489 nm (ref. 25). At moderate pump intensities (10–60 MW cm^{-2}), we observe the onset of amplified spontaneous emission peaks. These correspond to the longitudinal cavity modes that form when propagation losses are compensated by gain, allowing plasmonic modes to resonate between the reflective nanowire end-facets (see Supplementary Information and Fig. 2a). The large amount of gain needed to compensate both cavity and propagation losses produces a strong frequency pulling effect[26], making the Fabry–Perot modes much more closely spaced than expected for a passive cavity (Fig. 2c). A detailed discussion of this effect can be found in the Supplementary Information. The clear signature of multiple cavity mode resonances at higher pump powers demonstrates sufficient material gain to achieve full laser oscillation, shown by the nonlinear response of the integrated output power with increasing input intensity (Fig. 2b). The occurrence of these cavity resonances indicates that the laser mode's plasmonic coherence is determined by cavity feedback and not by its propagation distance.

Surmounting the limitations of conventional optics, not only do plasmonic lasers support nanometre-scale optical modes, their physical size can also be much smaller than conventional lasers, that is, plasmonic lasers operate under conditions where photonic modes cannot even exist[7]. Plasmonic lasers maintain strong confinement and optical mode-gain overlap over a broad range of nanowire diameters (Fig. 2e, f) with only a weak dependence on the nanowire diameter (see Supplementary Information). While hybrid modes do not experience mode cutoff, the threshold intensity increases for smaller nanowires owing to the reduction in the total gain volume.

Conversely, photonic lasers exhibit a strong dependence of the mode confinement on the nanowire diameter (Fig. 2g, h), resulting in a sharp increase in the threshold intensity at diameters near 150 nm owing to a poor overlap between the photonic mode and the gain material. Moreover, actual photonic lasers suffer mode cutoff because the leakage into the quartz substrate prevents lasing for nanowire diameters of less than 140 nm (ref. 7). That is, lasing of a nanowire with a diameter of just 52 nm is only feasible with hybrid plasmons and cannot occur in purely dielectric nanowires.

Plasmonic modes often exhibit highly polarized behaviour because the electric field normal to the metal surface binds most strongly to electronic surface charge. We have detected the signature of lasing plasmons from the polarization of scattered light from the nanowire end-facets (see Supplementary Information), which is in the same direction as the nanowire. Conversely, the polarization of scattered light from photonic lasers is perpendicular to the nanowire. This distinction provides a direct confirmation of the plasmonic mode.

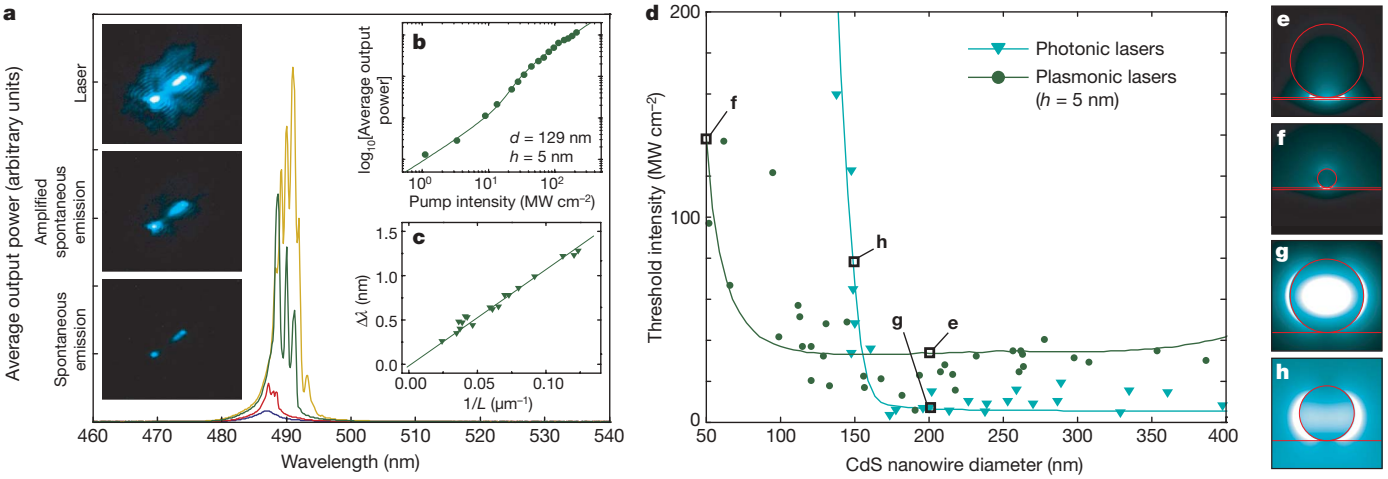


Figure 2 | Laser oscillation and threshold characteristics of plasmonic and photonic lasers. **a**, Laser oscillation of a plasmonic laser, $d = 129$ nm, $h = 5$ nm (longitudinal modes). The four spectra for different peak pump intensities exemplify the transition from spontaneous emission (21.25 MW cm^{-2}) via amplified spontaneous emission (32.50 MW cm^{-2}) to full laser oscillation (76.25 MW cm^{-2} and 131.25 MW cm^{-2}). **b**, The nonlinear response of the output power to the peak pump intensity. The relationship between mode spacing $\Delta\lambda$ and nanowire length L in **c** indicates a high group index of 11 due to the high material gain. The pictures on the left correspond to microscope images of a plasmon laser with $d = 66$ nm exhibiting spontaneous emission, amplified spontaneous emission and laser oscillation, where the scattered light output is from the end facets. **d**, Threshold intensity of plasmonic and photonic lasers versus nanowire diameter. The experimental data points correspond to the onset of amplified spontaneous emission, which occurs at slightly lower peak pump intensities compared to the threshold of gain saturation (see Supplementary Information). Amplified spontaneous emission in hybrid plasmonic modes occurs at moderate pump intensities of 10–60 MW cm^{-2} across a broad range of diameters. Plasmonic lasers maintain good mode confinement over a large range of diameters, as shown in **e** ($d = 200$ nm), and remain confined even for a 50-nm-diameter wire, as shown in **f**. Although the photonic lasers have similar threshold intensities around 10 MW cm^{-2} for all nanowires larger than 200 nm (shown in **g**), a sharp increase in the threshold occurs for diameters near 150 nm, owing to the loss of confinement within the nanowire and subsequent lack of overlap with the gain region[7] shown in **h**. The solid lines show a numerical fit to a simple rate equation model (see Supplementary Information). Below 140 nm, the photonic mode is cut off and lasing could not be observed at all. In contrast, plasmonic lasers maintain strong confinement and optical mode-gain overlap for diameters as small as 52 nm, a diameter for which a photonic mode does not even exist. The relatively small difference in the thresholds of plasmonic and photonic lasers can be attributed to high cavity mirror losses, which are of the same order of magnitude as plasmonic losses.

We find a strong increase of the spontaneous emission rate when the gap size between the nanowire and metal surface is decreased. Lifetime measurements reveal a Purcell factor of more than six for a gap width of 5 nm and nanowire diameters near 120 nm (Fig. 3), where hybrid plasmonic modes are most strongly localized[16]. This enhancement factor corresponds to a mode that is a hundred times smaller than the diffraction limit, which agrees well with our mode size calculations (see Supplementary Information). While the enhanced emission rate, or Purcell effect[17], is usually associated with micro-cavities[3,4], we observe a broad-band Purcell effect arising from mode confinement alone without a cavity[20].

We next examine the physics underlying the gain mechanism in the plasmonic lasers, which combines exciton dynamics, the modification of spontaneous emission[17] and mode competition (see Supplementary Information). Although photo-generated excitons have intrinsic lifetimes of up to 400 ps (ref. 25), they recombine faster at the edge of the nanowire near the gap region owing to strong optical confinement mediated by the hybrid plasmon mode (Fig. 3). The exciton diffusion length in bulk CdS is about a micrometre[27], which is much larger than the nanowire diameter. Therefore, the distribution of exciton recombination quickly adjusts itself to match the hybrid mode's profile (see Fig. 1b). The fast diffusion and enhanced emission rate into the hybrid plasmonic mode lead to preferential plasmon generation. In this way, the proportion of light that couples into the laser mode—known as the spontaneous emission factor β—can be high (Fig. 3b)[28]. The measured emission rates and a simple emission model show that the β-factor of the plasmonic mode is as high as 80% for a gap width of 5 nm. For gap widths below 5 nm, the exciton recombination is too close to the metal surface, causing rapid non-radiative quenching[29] (see Supplementary Information). Although nanowires placed in direct contact with the metal surface show the highest spontaneous emission rates, these devices exhibit weak luminescence and do not lase. Our calculations support these observations, indicating that there is a sharp reduction in the β-factor below gap widths of 5 nm.

The laser threshold is commonly manifested as a 'kink' between two linear regimes of the output power versus pump intensity response (log–log scale). However, it is known that lasers with strong mode confinement do not necessarily exhibit such behaviour, so that the laser threshold may be obscured[28,30]. The plasmonic lasers exhibit strong mode confinement, so we do indeed observe this 'smearing' of the threshold pump intensity (Fig. 4). The photonic lasers that we have measured, on the other hand, show the distinctive kink in the

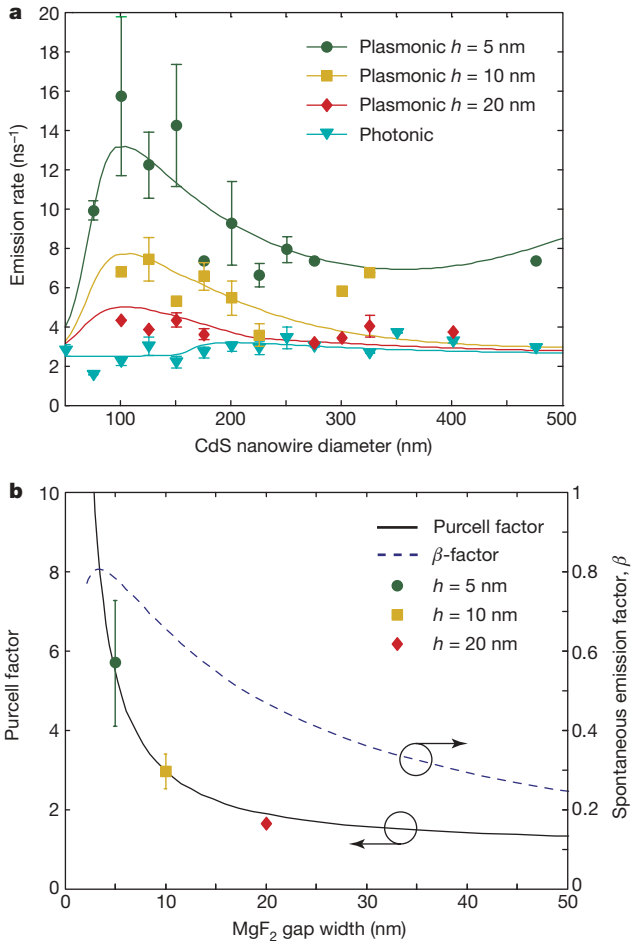


Figure 3 | The Purcell effect in plasmonic and photonic lasers. **a**, The emission rates of photonic and plasmonic nanowire lasers with different MgF_2 gap widths as a function of the nanowire diameter following the calculated trend. The optimal confinement condition of hybrid plasmonic modes is found near $d = 120$ nm, where the hybridization of nanowire and surface plasmon polariton modes is strongest, giving the highest emission rate[16]. The emission rates of nanowires placed in direct contact with the metal surface are given in the Supplementary Information. The error bars show standard deviation of emission rates from data collected within 25 nm nanowire diameter ranges. **b**, The Purcell factors determined from a numerical fit of the emission rate measurements to a simple emission model (see Supplementary Information). Near optimal confinement ($d = 120$ nm $\pm$ 20 nm), the average Purcell factor for devices with 5-nm gaps is more than six, which is considered high for a broad-band effect. We have also calculated the β-factor from the emission model fit by accounting for the possible emission pathways. The β-factor reaches a maximum of 80% for a gap width near 5 nm. At gap widths smaller than 5 nm, non-radiative quenching to the metal surface causes a sharp drop in the hybrid plasmon β-factor, which subsequently eliminates the possibility for lasing. The error bars show standard deviation of Purcell Factors for nanowire diameters, $d = [100, 140]$ nm.

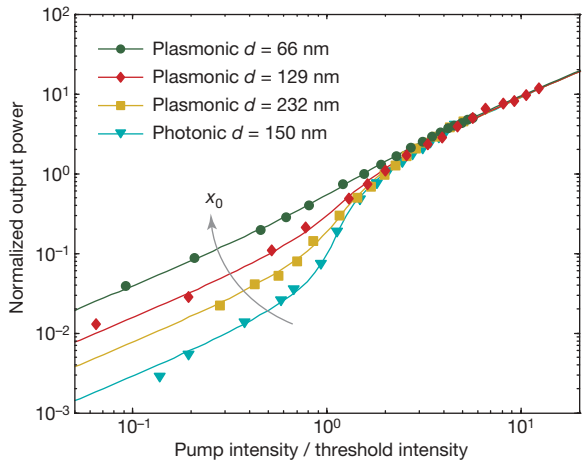


Figure 4 | The signature of threshold-less lasing due to high β-factor. The dependence of measured output power over pump intensity highlights clear differences in the physics underlying the plasmonic ($h = 5$ nm) and photonic lasers using a multi-mode lasing model[30]. In particular, our fitting parameter, x_0, is related to the gain saturation of individual longitudinal laser modes and their lateral mode area. A higher value of x_0 corresponds to a smaller mode area and a higher β-factor. Photonic lasers exhibit a clear transition between spontaneous emission and laser operation characterized by a change in the gradient of input peak pump intensity versus output power, corresponding to the laser threshold. For a 150-nm-diameter nanowire, the parameter $x_0 = 0.026$ is in agreement with a recent nanowire laser study[7]. Plasmonic lasers, however, show a strong dependence of x_0 on the nanowire diameter; a large multi-mode plasmonic laser ($d = 232$ nm) shows a somewhat smeared transition region ($x_0 = 0.074$), whereas smaller single-mode plasmonic lasers ($d = 129$ nm and $d = 66$ nm) have much less visible changes in gradient ($x_0 = 0.150$ and $x_0 = 0.380$ respectively). The large value of x_0 observed in the plasmonic lasers is associated with so-called threshold-less operation and is attributed to the strong mode confinement giving rise to a high β-factor. Nevertheless, the onset of amplified spontaneous emission peaks in all devices occurred at non-zero threshold intensities, as shown in Fig. 2d.

output power, which is in agreement with recently reported zinc oxide nanowire lasers[7]. Both plasmonic and photonic lasers in this work use the same gain material, so we conclude that the smeared response in output power arises from local electromagnetic field confinement. We therefore attribute this distinct behaviour to the increased β-factor of the hybrid plasmon mode. A high β-factor is often associated with low threshold laser operation where undesired emission modes are suppressed. Although the nearly linear response, shown in Fig. 4 for a nanowire diameter of 66 nm, may therefore give the impression of a lower threshold[28], in reality, the laser threshold occurs only once cavity losses are compensated (see Supplementary Information).

This demonstration of deep subwavelength plasmonic laser action at visible frequencies suggests new sources that may produce coherent light far below the diffraction limit. Extremely strong mode confinement, which is evident from the up to sixfold increase of the spontaneous emission rate, and the resulting high β-factor, are key aspects of the operation of deep subwavelength lasers. We have also shown that the advantage of plasmonic lasers is their ability to downscale the physical size of devices, as well as the optical modes they contain, unlike diffraction-limited lasers. Furthermore, the use of metals in plasmonics could provide a natural route towards electrical injection schemes that do not interfere with mode confinement[6]. The impact of plasmonic lasers on optoelectronics integration is potentially significant because the optical fields of these devices rival the smallest commercial transistor gate sizes and thereby reconcile the length scales of electronics and optics.

METHODS SUMMARY

The technical challenge of constructing the plasmonic lasers lies in ensuring good contact between the nanowire and the planar optical film. Low film roughness ($\sim$1 nm root mean square, r.m.s.) and accurate deposition allow a $<$2-nm gap width variation. Crystalline CdS nanowires exhibit extremely low surface roughness (see Supplementary Information), so the contact is limited by film roughness. Nanowires were grown using chemical vapour deposition of CdS powders on Si substrates by self-assembly from a 10-nm Au film seeding layer leading to random nanowire diameters of $d = 50$ nm–500 nm. CdS nanowires were deposited from solution by spin-coating onto pre-prepared films with varying MgF_2 thicknesses of 0, 5, 10 and 20 nm, along with control devices of nanowires on a quartz substrate.

A frequency-doubled, mode-locked Ti-sapphire laser (Spectra Physics) was used to pump the plasmonic and photonic lasers ($\lambda_{pump} = 405$ nm, repetition rate 80 MHz, pulse length 100 fs). An objective lens (20×, numerical aperture 0.4) was used to focus the pump beam to a 38-μm-diameter spot on the sample. All experiments were carried out at low temperature, $T < 10$ K, using a liquid-He-cooled cryostat (Janis Research). Individual spectra were recorded using a spectrometer with a resolution of 0.25 nm and a liquid-N_2-cooled charge-coupled device (CCD) (Princeton Instruments). The lifetime measurements were conducted under very low pump conditions to avoid heating and exciton–exciton scattering effects using time-correlated single photon counting (PicoQuant: PicoHarp 300, Micro Photon Devices APD, 40 ps timing resolution). A 490 ± 10 nm band pass filter was used to filter out ambient light and pass light from the I_2 CdS exciton line near 489 nm. An emission model was used to describe the enhanced lifetime data (Purcell effect), which relates to deep subwavelength mode confinement (see Supplementary Information).

Received 13 May; accepted 31 July 2009.
Published online 30 August 2009.

1. Gordon, J. P., Zeiger, H. J. & Townes, C. H. The maser—new type of microwave amplifier, frequency standard and spectrometer. *Phys. Rev.* **99,** 1264–1274 (1955).
2. Drescher, M. *et al.* X-ray pulses approaching the attosecond frontier. *Science* **291,** 1923–1927 (2001).
3. Altug, H., Englund, D. & Vučković, J. Ultrafast photonic crystal nanocavity laser. *Nature Phys.* **2,** 484–488 (2006).
4. Hill, M. T. *et al.* Lasing in metallic-coated nanocavities. *Nature Photon.* **1,** 589–594 (2007).
5. Johnson, J. C. *et al.* Single gallium nitride nanowire lasers. *Nature Mater.* **1,** 106–110 (2002).
6. Duan, X., Huang, Y., Agarwal, R. & Lieber, C. M. Single-nanowire electrically driven lasers. *Nature* **421,** 241–245 (2003).
7. Zimmler, M. A., Bao, J., Capasso, F., Müller, S. & Ronning, C. Laser action in nanowires: observation of the transition from amplified spontaneous emission to laser oscillation. *Appl. Phys. Lett.* **93,** 051101 (2008).
8. Astafiev, O. *et al.* Single artificial atom lasing. *Nature* **449,** 588–590 (2007).
9. Scalari, G. *et al.* THz and sub-THz quantum cascade lasers. *Laser Photon. Rev.* **3,** 45–66 (2009).
10. Bergman, D. J. & Stockman, M. I. Surface plasmon amplification by stimulated emission of radiation: quantum generation of coherent surface plasmons in nanosystems. *Phys. Rev. Lett.* **90,** 027402 (2003).
11. Zheludev, N. I., Prosvirnin, S. L., Papasimakis, N. & Fedotov, V. A. Lasing spaser. *Nature Photon.* **2,** 351–354 (2008).
12. Maier, S. A. *et al.* Local detection of electromagnetic energy transport below the diffraction limit in metal nanoparticle plasmon waveguides. *Nature Mater.* **2,** 229–232 (2003).
13. Pile, D. F. P. *et al.* Theoretical and experimental investigation of strongly localized plasmons on triangular metal wedges for subwavelength waveguiding. *Appl. Phys. Lett.* **87,** 061106 (2005).
14. Ambati, M. *et al.* Observation of stimulated emission of surface plasmon polaritons. *Nano Lett.* **8,** 3998–4001 (2008).
15. Noginov, M. A. *et al.* Stimulated emission of surface plasmon polaritons. *Phys. Rev. Lett.* **101,** 226806 (2008).
16. Oulton, R. F., Sorger, V. J., Genov, D. A., Pile, D. F. P. & Zhang, X. A hybrid plasmonic waveguide for sub-wavelength confinement and long-range propagation. *Nature Photon.* **2,** 495–500 (2008).
17. Purcell, E. M. Spontaneous emission probabilities at radio frequencies. *Phys. Rev.* **69,** 681 (1946).
18. Volkov, V. S., Devaux, E., Laluet, J.-Y., Ebbesen, T. W. & Bozhevolnyi, S. I. Channel plasmon subwavelength waveguide components including interferometers and ring resonators. *Nature* **440,** 508–511 (2006).
19. Anker, J. N. *et al.* Biosensing with plasmonic nanosensors. *Nature Mater.* **7,** 442–453 (2008).
20. Akimov, V. *et al.* Generation of single optical plasmons in metallic nanowires coupled to quantum dots. *Nature* **450,** 402–406 (2007).
21. Liu, Y., Bartal, G., Genov, D. A. & Zhang, X. Subwavelength discrete solitons in nonlinear metamaterials. *Phys. Rev. Lett.* **99,** 153901 (2007).
22. Stockman, M. I. Nanofocusing of optical energy in tapered plasmonic waveguides. *Phys. Rev. Lett.* **93,** 137404 (2004).
23. Oulton, R. F., Bartal, G., Pile, D. F. P. & Zhang, X. Confinement and propagation characteristics of subwavelength plasmonic modes. *N. J. Phys.* **10,** 105018 (2008).
24. Ma, R. M., Dai, L. & Qin, G. G. Enhancement-mode metal-semiconductor field-effect transistors based on single n-CdS nanowires. *Appl. Phys. Lett.* **90,** 093109 (2007).
25. Thomas, G. D. & Hopfield, J. J. Optical properties of bound exciton complexes in cadmium sulfide. *Phys. Rev.* **128,** 2135–2148 (1962).
26. Siegman, A. E. *Lasers* 466–472 (University Science Book, 1986).
27. Weber, C., Becker, U., Renner, R. & Klingshirn, C. Measurement of the diffusion-length of carriers and excitons in CdS using laser-induced transient gratings. *Z. Phys. B* **72,** 379–384 (1988).
28. Björk, G. & Yamamoto, Y. Analysis of semiconductor microcavity lasers using rate equations. *IEEE J. Quantum Electron.* **27,** 2386–2396 (1991).
29. Ford, G. W. & Weber, W. H. Electromagnetic interactions of molecules with metal surfaces. *Phys. Rep.* **113,** 195–287 (1984).
30. Casperson, L. W. Threshold characteristics of multimode laser oscillators. *J. Appl. Phys.* **12,** 5194–5201 (1975).

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

Acknowledgements We thank M. Ambati and D. Genov for discussions and the Lawrence Berkeley National Laboratory's Molecular Foundry for technical support. We acknowledge financial support from the US Air Force Office of Scientific Research (AFOSR) MURI programme under grant number FA9550-04-1-0434 and from the National Science Foundation Nano-scale Science and Engineering Center (NSF-NSEC) under award number CMMI-0751621. T.Z. acknowledges a fellowship from the Alexander von Humboldt Foundation. V.J.S. acknowledges a fellowship from the Intel Corporation. L.D. and R.-M.M. acknowledge the National Natural Science Foundation of China (award numbers 60576037 and 10774007) and the National Basic Research Program of China (grant numbers 2006CB921607 and 2007CB613402).

Author Contributions R.F.O. developed the device design and conducted theoretical simulations. V.J.S., T.Z. and R.F.O. performed the optical measurements. R.-M.M. and L.D. synthesized the CdS nanowires. V.J.S. and C.G. fabricated the devices. X.Z., G.B. and R.F.O. guided the theoretical and experimental investigations. R.F.O., V.J.S., T.Z., G.B. and X.Z. analysed data and 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 X.Z. (xiang@berkeley.edu).

LETTERS

X-ray observation of a transient hemiaminal trapped in a porous network

Takehide Kawamichi[1], Tsuyoshi Haneda[1], Masaki Kawano[1]† & Makoto Fujita[1,2]

X-ray crystallography is the method of choice for the direct structural analysis of crystalline compounds[1]. Extending its use to the *in situ* mapping of chemical transformations could provide valuable insights, as illustrated by time-resolved X-ray crystallography studies[2,3]; however, the transient nature of unstable reaction intermediates often poses a significant challenge. It has recently been demonstrated that standard chemical reactions can occur within the pores of porous coordination networks[4–6] and that the robust crystallinity of these networks facilitates *in situ* X-ray analysis of the adducts and products[7–11]. Here we show that such systems even enable X-ray observations of reaction intermediates that are usually transient and non-isolable. Our proof-of-concept demonstration examines the simple and ubiquitous reaction between an amine and an aldehyde, which normally form a very short-lived hemiaminal that then yields the Schiff-base product. The mechanism of this reaction has been exhaustively examined, but the hemiaminal intermediate has only rarely been observed[12–16]. We first determine the structure of a porous network with an aromatic amine embedded in it, then diffuse an aldehyde substrate into the material to transform the amine into a hemiaminal intermediate that is kinetically trapped and thus amenable to X-ray analysis, and finally raise the temperature of the system to obtain the imine product and determine its structure. These results establish that porous network materials provide a means of obtaining sequential X-ray-based snapshots of the structural transformations that occur during chemical reactions.

Figure 1 shows a simplified reaction profile for the two key steps involved in Schiff-base formation from an amine and an aldehyde, through an intermediate tetrahedral hemiaminal. The overall reaction is exothermic, but formation of the hemiaminal is endothermic and it is therefore short lived and typically observed only by spectroscopic methods in low-temperature matrices. Recently, however, Schiff-base formation in a molecular capsule was shown to be sufficiently retarded that the intermediate hemiaminal could be observed using NMR at ambient temperature[15,16]. Furthermore, in an exceptional case enabled by the use of an enzyme mutated to prevent reaction completion, X-ray crystallography revealed a hemiaminal intermediate bound by multiple hydrogen bonds in the enzyme pocket[17–19]. These examples illustrate that the free-energy surface for the reaction depends on whether it is performed in solution or in an environment deep within confined cavities where the hemiaminal intermediate can be kinetically stable. We reasoned that the pores of a crystalline network would also provide kinetic stabilization and hence enable direct observation by X-ray crystallography of trapped intermediates, providing structural and mechanistic information. Also, using porous materials, rather than the usually highly specific enzyme pockets, should provide a general method for observing reaction intermediates of many different organic transformations.

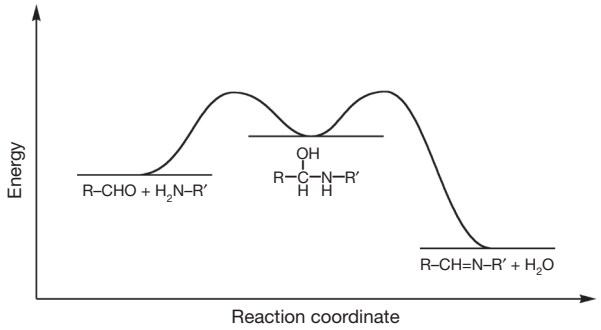


Figure 1 | A reaction profile of imine formation. Although the carbonyl addition of an amine to form a hemiaminal intermediate is endothermic, the overall reaction is exothermic. The precise reaction profile involves several protonated intermediates that are omitted here. R and R′ are typically aliphatic or aromatic groups.

The porous coordination network $\{[(ZnI_2)_3(\mathbf{2})_2(\mathbf{3})]\cdot x(\mathbf{G})\}_n$ (**1**; **G** = nitrobenzene, $x = 4$; n denotes a polymeric structure) that we used in this study was prepared by slowly diffusing layers of ZnI_2, 2,4,6-tris(4-pyridyl)-1,3,5-triazine (**2**) and 1-aminotriphenylene (**3**) in nitrobenzene[20]. As described previously, the structure of the resultant network has pores of two different shapes, A and B.

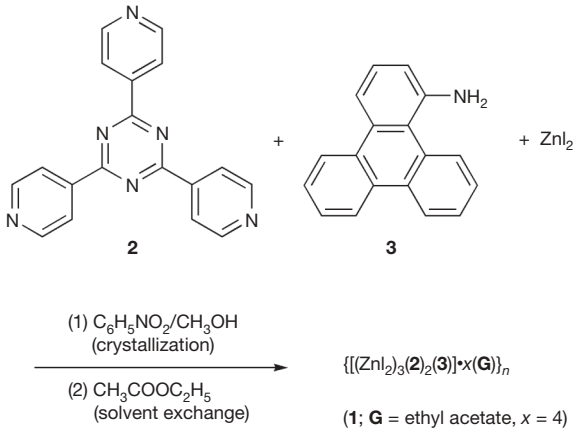


Figure 2 | Preparation of porous coordination network $\{[(ZnI_2)_3(\mathbf{2})_2(\mathbf{3})]\cdot x(\mathbf{G})\}_n$ (1; G = ethyl acetate, $x = \sim 4$). Network complex **1** (**G** = nitrobenzene, $x = 4$) was prepared according to established procedures[20]. Nitrobenzene solvent initially occupying the pores of **1** was replaced with ethyl acetate by immersing the crystals in ethyl acetate to give complex **1** (**G** = ethyl acetate, $x = \sim 4$).

[1]Department of Applied Chemistry, School of Engineering, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [2]Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. †Present address: Division of Advanced Materials Science, Pohang University of Science and Technology, San 31 Hyojadong, Pohang 790-784, South Korea.

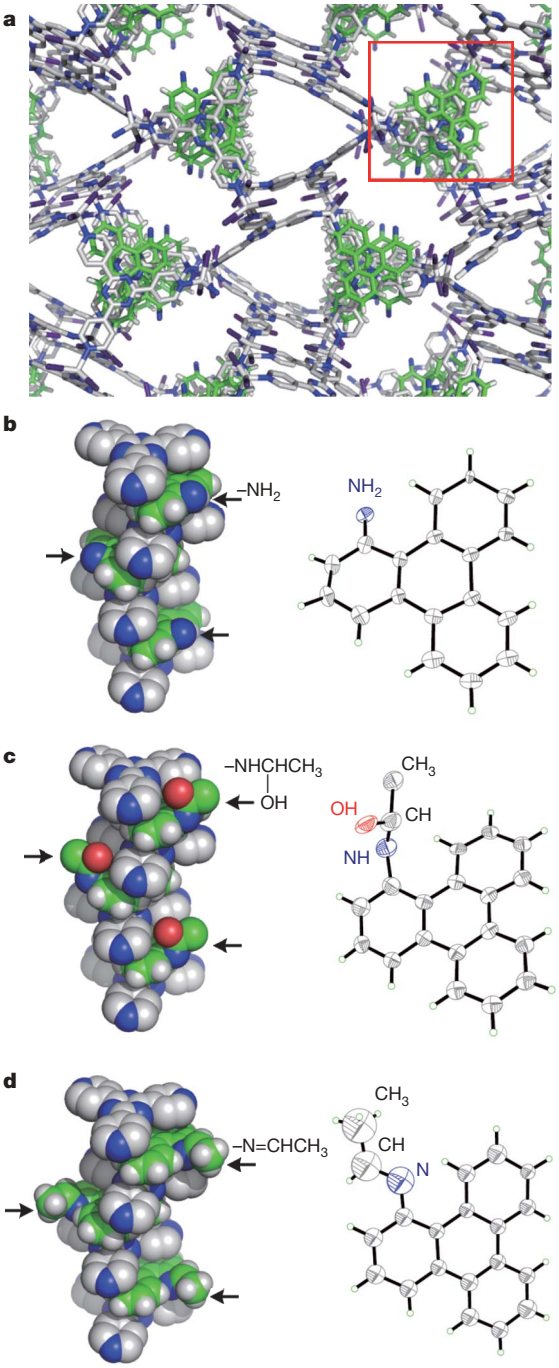


Figure 3 | Direct crystallographic observation of imine formation from acetaldehyde and embedded 1-aminotriphenylene within the pores of 1. a, Network structure of **1** (**G** = ethyl acetate, $x = 4$). In the boxed region, aromatic amine **3** stacks in an alternating donor–acceptor fashion with triazine ligand **2**. The pore is filled with solvent molecules (ethyl acetate), which are omitted for clarity. **b**, Left: view of the columnar stack of aromatic amine guest **3** and the triazine ligand; right: thermal-ellipsoid diagram of embedded guest **3**. **c**, The transient hemiaminal intermediate, formed *in situ* by treating a crystal of **1** with acetaldehyde at 215 K for 10 min. Left: view of the columnar stack of hemiaminal **4** and the triazine ligand **2**; right: embedded hemiaminal **4** (*R**-isomer, 18% occupancy (18% *S**-isomer not shown; *R** and *S** denote relative configurations)). **d**, The final product (imine **5**), which was obtained by warming hemiaminal **4** at 270 K for 30 min. Left: view of the columnar stack of imine **5** and triazine ligand **2**; right: embedded imine **5**. All thermal-ellipsoid diagrams are drawn at 30% probability. Colours: carbon, light grey in **2**, green in **3**, **4** and **5**; nitrogen, blue; oxygen, red; hydrogen, white. Hydrogen atoms of **2** are omitted for clarity. Arrows in **b**, **c** and **d** indicate the positions of $–NH_2$, $–NHCH(OH)CH_3$ and $–N{=}CHCH_3$ groups, respectively.

Aromatic cartridge **3** forms aromatic donor–acceptor stacks with ligand **2** and has its amino groups facing into type-A pores only. To allow rapid diffusion of reactants and thus enable chemical transformations in the pores, we replaced the nitrobenzene (which filled the pores of the network) with a non-aromatic solvent. Thus, crystals of **1** (**G** = nitrobenzene) were immersed in ethyl acetate for 24 h to produce solvent-exchanged **1** (**G** = ethyl acetate, $x = \sim 4$) (Fig. 2). Crystallinity persisted after solvent exchange and the structure of **1** (**G** = ethyl acetate) was solved by X-ray crystallography (Fig. 3a, b). At a temperature of 85 K, the amino groups face into type-A pores and are disordered over three positions (distributed 44%, 33% and 23%), which a previous study[7] showed to be in equilibrium through thermal rotation of **3** in the aromatic donor–acceptor stacks.

We examined the condensation of acetaldehyde with the amine groups of embedded **3** at low temperatures. A red single crystal of **1** (**G** = ethyl acetate) was mounted on a diffractometer, covered with a capillary and cooled to 215 K. An ethyl acetate solution of acetaldehyde was continuously introduced through the capillary at 215 K for 10 min, immersing the crystal in a gentle flow of the acetaldehyde (Fig. 4). We then cooled the crystal to 90 K to stop the reaction, and collected diffraction data. After data collection, we warmed the crystal to 270 K and allowed it to stand for 30 min to complete the condensation reaction; this was accompanied by a slight change in crystal colour from red to yellow. The crystal was then again cooled to 90 K and diffraction data collected. The reaction temperature and reaction time were optimized by using temperature-dependent infrared spectroscopy to observe ν(C=N) and ν(N–H) bands.

The X-ray structure obtained after flowing acetaldehyde over the crystal for 10 min at 215 K revealed the transient hemiaminal intermediate **4** and unreacted amine. Least-squares refinement converged to give yields of 36% of hemiaminal **4** and 64% of unreacted amine **3**. The tetrahedral sp^3 carbon of the hemiaminal was clearly observed (Fig. 3c), although the disorder is quite complex. However, synchrotron radiation is well suited to the analysis of minor products, and the refinement of the hemiaminal functional group converged successfully when we treated the electron density as a racemic disordered model.

As stated, the amino functional groups before reaction are disordered over three positions, sites a, b and c, with 44% at site a, 33% at site b and 23% at site c. Only the 33% of amino groups at site b react to form the hemiaminal intermediate, in agreement (within experimental error) with the observed yield of 36%. The amino groups at the other two disordered positions did not react, most likely because of steric hindrance. We note that neither reactant **3** nor intermediate **4** moves between sites a and c at 215 K, and that the modest yield of 36% is therefore reasonable and indicative of high reaction selectivity (Fig. 5).

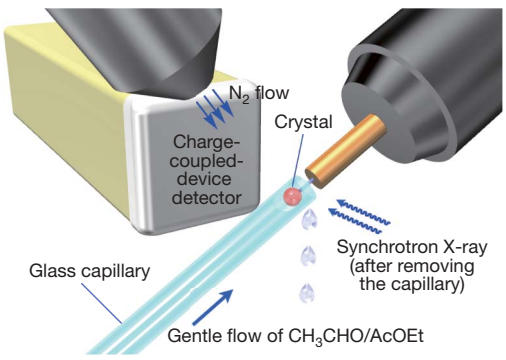


Figure 4 | Experimental device for the *in situ* hemiaminal formation in the crystal of 1. In a temperature-controlled cold nitrogen stream, a Pasteur pipette was set up in such a way that its capillary tip covered the crystal of **1** mounted on a diffractometer. At 215 K, an ethyl acetate solution of acetaldehyde was introduced onto the crystal through the capillary and continuously applied in such a way that the crystal was immersed in a gentle flow of the solution for 10 min. Then the pipette was removed and the crystal was cooled to 90 K before X-ray irradiation and data collection. Ac, acetyl; Et, ethyl.

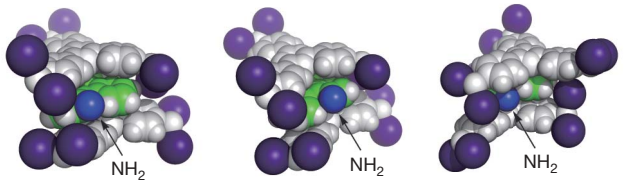


Figure 5 | Disorder of embedded amine 3 before the reaction. The guest amine **3** is disordered at sites a, b and c with respective occupancies of 44%, 33% and 22%. At the reaction temperature (215 K), guest **3** does not move among the three sites. The amino groups at sites a and c are sterically hindered and are not converted to hemiaminal upon treatment with acetaldehyde under the experimental conditions (215 K, 10 min). Only the amino group at site b was allowed to react with acetaldehyde to efficiently yield hemiaminal **4** (36% occupancy). The reactant **3** and the intermediate **4** at site b only are shown in Fig. 3. Colours: Zn and all atoms of **2**, light grey; iodine, purple; carbon in **3**, green; hydrogen in **3**, white; nitrogen in **3**, blue.

The hemiaminal intermediate is kinetically trapped at low temperatures, but simply warming the crystal facilitated its further conversion to an imine. The structure obtained after warming the crystal to 270 K and keeping it at that temperature for 30 min showed complete conversion of the hemiaminal to the final imine product (Fig. 3d). At the hemiaminal stage, the triphenylene core is not completely planar, owing to steric repulsions between the pendant N–H bond and the adjacent C–H bond. Steric repulsion is reduced in the final imine structure, however, where the triphenylene core is no longer distorted and adopts a planar conformation (Supplementary Information).

The chemical reactions in the pores were also monitored by carrying out single-crystal microscopic infrared spectroscopy under the same conditions. The temperature-dependent measurement showed excellent agreement with the X-ray results. When a crystal of **1** was immersed in acetaldehyde at 215 K, the intensity of the $-NH_2$ symmetry stretching mode of the amino group at 3,350 cm^{-1} decreased by ~38% and the –OH stretching vibrational mode attributed to hemiaminals appeared at around 3,520 cm^{-1}. The decreasing $-NH_2$ (amine) asymmetry stretching mode and the increasing –NH (hemiaminal) stretching mode overlap around 3,420 cm^{-1}, and absorption in this region thus remained almost unchanged. After completion of the reaction, the C=N vibrational mode of imine was observed at 1,650 cm^{-1} (Supplementary Information).

Our successful *in situ* observation by X-ray crystallography of a hemiaminal intermediate transiently produced during Schiff-base formation illustrates that reactions can occur smoothly within the pores of a coordination network, and that the network's robust crystallinity enables direct reaction monitoring through the collection of sequential X-ray 'snapshots'. We believe that porous coordination networks and the approach illustrated here offer a general method for acquiring sequential X-ray snapshots of chemical reactions, which will provide valuable direct mechanistic insights not obtainable by other means.

METHODS SUMMARY

Preparation of 1 (G = ethyl acetate). We immersed the as-synthesized crystals of **1** (**G** = nitrobenzene)[20] in ethyl acetate for 24 h at 300 K to produce guest-exchanged crystals of **1** (**G** = ethyl acetate). Elemental analysis of $\{[(ZnI_2)_3(\mathbf{2})_2(\mathbf{3})]\cdot(C_4H_8O_2)_4\}_n$ (percentage calculated, percentage found): C (38.60%, 38.45%), H (3.19%, 3.20%), N (8.36%, 8.08%).

***In situ* X-ray observation of the hemiaminal intermediate.** We placed a red single crystal of **1** (**G** = ethyl acetate) in a goniometer and immersed it in a gentle flow (1 ml h^{-1}) of acetaldehyde (ethyl acetate solution, 17.7 mM), which was supplied through a capillary with a syringe pump, at 215 K (see text and Fig. 4). After 10 min, we stopped the flow and cooled the crystal to 90 K to stop the reaction. Diffraction data were then rapidly collected. Crystallographic analysis showed that aromatic amine **3** in the crystal was 36% converted into hemiaminal intermediate **4** to give crystals with a formula of $\{[(ZnI_2)_3(\mathbf{2})_2(\mathbf{3})_{0.64}(\mathbf{4})_{0.36}]\cdot x(\mathbf{G}_1)\cdot y(\mathbf{G}_2)\}_n$ (**6**; $\mathbf{G}_1$ = ethyl acetate, $\mathbf{G}_2$ = acetaldehyde, $x = \sim 2.99$, $y = \sim 1.37$).

Conversion of hemiaminal intermediate into imine. After crystallographic analysis, we warmed the crystal containing intermediate **4** from 90 K to 270 K in a goniometer. After 30 min at 270 K, the crystal was again cooled to 90 K to stop the reaction and diffraction data were rapidly collected. Crystallographic analysis showed that hemiaminal **4** in the crystal was fully converted into Schiff base **5** to give crystals with a formula of $\{[(ZnI_2)_3(\mathbf{2})_2(\mathbf{3})_{0.44}(\mathbf{5})_{0.56}]\cdot x(\mathbf{G}_1)\cdot y(\mathbf{G}_2)\}_n$ (**7**; $\mathbf{G}_1$ = ethyl acetate, $\mathbf{G}_2$ = water, $x = \sim 2.97$, $y = \sim 0.56$). Details of the crystallographic analysis and crystal data are described in Supplementary Information.

Received 7 April; accepted 23 July 2009.

1. Ohashi, Y. in *Models, Mysteries and Magic of Molecules* (eds Boeyens, J. C. A. & Ogilvie, J. F.) 109–113 (Springer, 2008).
2. Coppens, P., Vorontsov, I. I., Graber, T., Gembicky, M. & Kovalevsky, A. Y. The structure of short-lived excited states of molecular complexes by time-resolved X-ray diffraction. *Acta Crystallogr. A* **61,** 162–172 (2005).
3. Moffat, K. Time-resolved biochemical crystallography: a mechanistic perspective. *Chem. Rev.* **101,** 1569–1581 (2001).
4. Batten, S. R. & Robson, R. Interpenetrating nets: ordered, periodic entanglement. *Angew. Chem. Int. Ed.* **37,** 1460–1494 (1998).
5. Eddaoudi, M. *et al.* Modular chemistry: secondary building units as a basis for the design of highly porous and robust metal-organic carboxylate frameworks. *Acc. Chem. Res.* **34,** 319–330 (2001).
6. Kitagawa, S., Kitaura, R. & Noro, S. Functional porous coordination polymers. *Angew. Chem. Int. Ed.* **43,** 2334–2375 (2004).
7. Haneda, T., Kawano, M., Kawamichi, T. & Fujita, M. Direct observation of the labile imine formation through single-crystal-to-single-crystal reactions in the pores of a porous coordination network. *J. Am. Chem. Soc.* **130,** 1578–1579 (2008).
8. Kawamichi, T., Kodama, T., Kawano, M. & Fujita, M. Single-crystalline molecular flasks: chemical transformation with bulky reagents in the pores of porous coordination networks. *Angew. Chem. Int. Ed.* **47,** 8030–8032 (2008).
9. Costa, J. S. *et al.* Chemical modification of a bridging ligand inside a metal–organic framework while maintaining the 3D structure. *Eur. J. Inorg. Chem.* 1551–1554 (2008).
10. Burrows, A. D., Frost, C. G., Mahon, M. F. & Richardson, C. Post-synthetic modification of tagged metal–organic frameworks. *Angew. Chem. Int. Ed.* **47,** 8482–8486 (2008).
11. Wang, Z. & Cohen, S. M. Postsynthetic modification of metal–organic frameworks. *Chem. Soc. Rev.* **38,** 1315–1329 (2009).
12. Forlani, L., Marianucci, E. & Todesco, P. E. ^{1}H nuclear magnetic resonance evidence for tetrahedral intermediates in the reactions between aromatic carbonyl groups and aliphatic amines. *J. Chem. Res. Synop.* 126–127 (1984).
13. Chudek, J. A., Foster, R. & Young, D. ^{13}C nuclear magnetic resonance studies of the products of reaction of acetaldehyde and of simple ketones in liquid ammonia, in hydrazine hydrate, and in some substituted hydrazine solutions. *J. Chem. Soc., Perkin Trans. 2* 1285–1289 (1985).
14. Evans, D. A., Borg, G. & Scheidt, K. A. Remarkably stable tetrahedral intermediates: carbinols from nucleophilic additions to N-acylpyrroles. *Angew. Chem. Int. Ed.* **41,** 3188–3191 (2002).
15. Iwasawa, T., Hooley, R. J. & Rebek, J. Jr. Stabilization of labile carbonyl addition intermediates by a synthetic receptor. *Science* **317,** 493–496 (2007).
16. Hooley, R. J., Iwasawa, T. & Rebek, J. Jr. detection of reactive tetrahedral intermediates in a deep cavity and with an introverted functionality. *J. Am. Chem. Soc.* **129,** 15330–15339 (2007).
17. Heine, A. *et al.* Observation of covalent intermediates in an enzyme mechanism at atomic resolution. *Science* **294,** 369–374 (2001).
18. Thorell, S., Schürmann, M., Georg, A., Sprenger, G. A. & Schneider, G. Crystal structure of decameric fructose-6-phosphate aldolase from *Escherichia coli* reveals inter-subunit helix swapping as a structural basis for assembly differences in the transaldolase family. *J. Mol. Biol.* **319,** 161–171 (2002).
19. Lorentzen, E., Siebers, B., Hensel, R. & Pohl, E. Mechanism of the Schiff base forming fructose-1,6-bisphosphate aldolase: structural analysis of reaction intermediates. *Biochemistry* **44,** 4222–4229 (2005).
20. Kawano, M., Kawamichi, T., Haneda, T., Kojima, T. & Fujita, M. The modular synthesis of functional porous coordination networks. *J. Am. Chem. Soc.* **129,** 15418–15419 (2007).

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

Acknowledgements This research was supported by the CREST project of the Japan Science and Technology Agency, of which M.F. is the principal investigator, and also in part by KAKENHI, the Japan Society for the Promotion of Science, the Global COE Program (Chemistry Innovation through Cooperation of Science and Engineering), MEXT, Japan, and Koei Chemical Co. Ltd. This work has been approved by the Photon Factory Program Advisory Committee.

Author Contributions M.K. and M.F. designed the project, analysed the results and wrote the manuscript. T.K. and T.H. performed the experimental work.

Author Information The X-ray crystallographic coordinates for structures reported in this paper have been deposited at the Cambridge Crystallographic Data Centre, under deposition numbers CCDC 725638 (**1**), CCDC 725639 (**6**) and CCDC 725640 (**7**). These data can be obtained free of charge from the Cambridge Crystallographic Data Centre (http://www.ccdc.cam.ac.uk/data_request/cif). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M.K. (mkawano@postech.ac.kr) or M.F. (mfujita@appchem.t.u-tokyo.ac.jp).

Vol 461|1 October 2009|doi:10.1038/nature08395

LETTERS

Remote triggering of fault-strength changes on the San Andreas fault at Parkfield

Taka'aki Taira[1], Paul G. Silver[1]‡, Fenglin Niu[2] & Robert M. Nadeau[3]

Fault strength is a fundamental property of seismogenic zones, and its temporal changes can increase or decrease the likelihood of failure and the ultimate triggering of seismic events. Although changes in fault strength have been suggested to explain various phenomena, such as the remote triggering of seismicity[1], there has been no means of actually monitoring this important property *in situ*. Here we argue that ~20 years of observation (1987–2008) of the Parkfield area at the San Andreas fault have revealed a means of monitoring fault strength. We have identified two occasions where long-term changes in fault strength have been most probably induced remotely by large seismic events, namely the 2004 magnitude (M) 9.1 Sumatra–Andaman earthquake and the earlier 1992 $M = 7.3$ Landers earthquake. In both cases, the change possessed two manifestations: temporal variations in the properties of seismic scatterers—probably reflecting the stress-induced migration of fluids—and systematic temporal variations in the characteristics of repeating-earthquake sequences that are most consistent with changes in fault strength. In the case of the 1992 Landers earthquake, a period of reduced strength probably triggered the 1993 Parkfield aseismic transient[2–5] as well as the accompanying cluster of four $M > 4$ earthquakes at Parkfield. The fault-strength changes produced by the distant 2004 Sumatra–Andaman earthquake are especially important, as they suggest that the very largest earthquakes may have a global influence on the strength of the Earth's fault systems. As such a perturbation would bring many fault zones closer to failure, it should lead to temporal clustering of global seismicity. This hypothesis seems to be supported by the unusually high number of $M \geq 8$ earthquakes occurring in the few years following the 2004 Sumatra–Andaman earthquake.

The time-varying properties of seismic scatterers have recently been used to probe stress-induced changes in the San Andreas fault zone near Parkfield in central California[4]. This temporal variability is most clearly observable through the use of repeating earthquakes that minimize variations in the source location and mechanism that could be mapped as apparent changes in the medium[4,6,7]. As a measure of temporal scatterer behaviour from one earthquake to the next, we made use of decorrelation index $D(t)$ derived from the cross-correlation of two seismograms (Supplementary Fig. 1 and Methods). Using particularly tightly clustered repeating microearthquakes[8,9] recorded by the High-Resolution Seismic Network (HRSN; Fig. 1), we have been able to track the behaviour of a group of time-dependent scatterers for a 22-year period (1987–2008); we will refer to this group of scatterers, first identified in ref. 4, as the target scatterer.

We find three excursions. The first excursion in $D(t)$ appears to have initiated between 10 September 1992 and 5 February 1993 (Supplementary Fig. 2). It reaches a peak in the mid 1990s and then gradually decays over the next decade (Fig. 2a). This variation closely tracks a well-documented transient in fault slip-rate along the San Andreas fault, as inferred from both geodetic measurements[2,3,5] and a change in seismic slip rate[9], $\dot{d} = M_0/T_r$, calculated from the recurrence interval, T_r, and seismic moment, M_0, of repeating earthquakes[9–11]. This excursion has been interpreted as reflecting stress-induced fluid migration associated with the 1993 Parkfield aseismic transient (ref. 4 and Supplementary Fig. 3).

The second excursion in target-scatterer properties is constrained to have initiated during the interval spanning the 28 September 2004

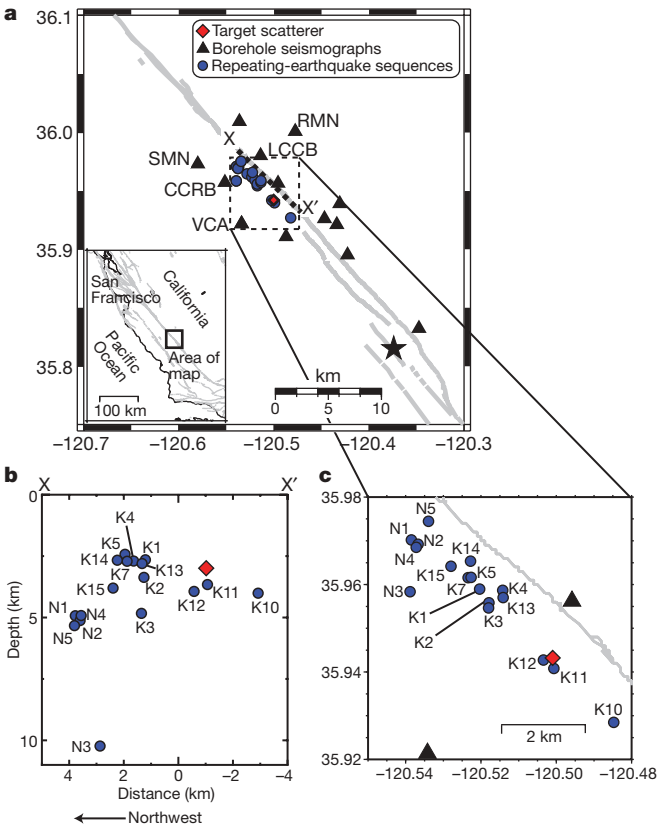


Figure 1 | Map of the Parkfield segment at the San Andreas fault. **a**, Locations of the target scatterer (red diamond), the HRSN borehole stations (triangles) and 17 repeating-earthquake sequences (circles) (ref. 8 and Methods). Labelled stations are those used in this study. The other stations either have a high noise level, are dominated by fault-zone guided waves, or do not respond to the scatterer defined by ref. 4. Star is the epicentre of the 2004 Parkfield earthquake. **b**, Vertical cross-section along X–X′ in **a**. **c**, Detailed map view.

[1]Department of Terrestrial Magnetism, Carnegie Institution of Washington, District of Columbia 20015, USA. [2]Department of Earth Science, Rice University, Houston, Texas 77005, USA. [3]Berkeley Seismological Laboratory, University of California, Berkeley, California 94720, USA.
‡Deceased.

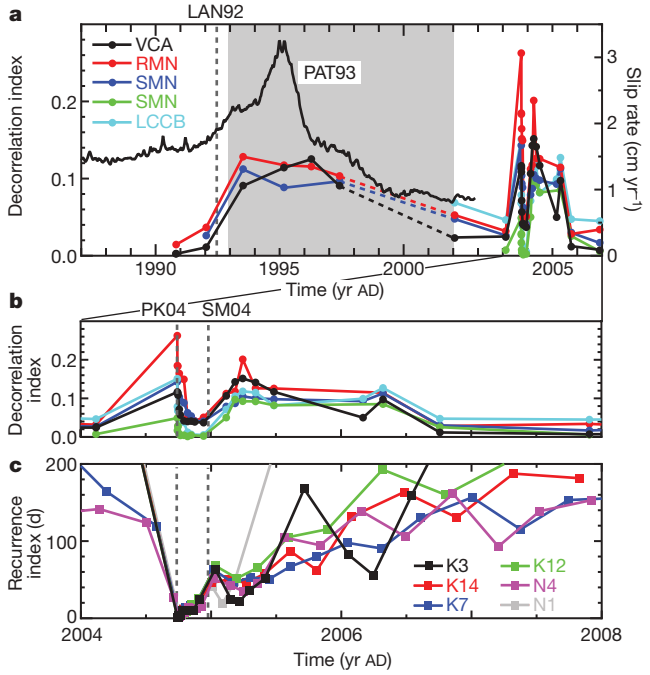


Figure 2 | Temporal changes in scatterer properties for a 22-year period. **a**, **b**, $D(t)$ reflecting target-scatter properties (Methods) for sequence K3 recorded by five stations during 1987–2008 (**a**) and 2004–08 (**b**). **c**, T_r for six sequences plotted at the midpoint between the origin times of the two events. Broken lines indicate HRSN was not functional (1998–2001). Solid line is geodetic inferred slip rate[5] that is concentrated in the section of the fault studied here (figure 5 in ref. 5). Vertical dashed-grey lines denote occurrence times of 1992 Landers (LAN92), 2004 Parkfield (PK04) and 2004 Sumatra–Andaman (SM04) earthquakes. PAT93 denotes the 1993 Parkfield aseismic transient (shaded area).

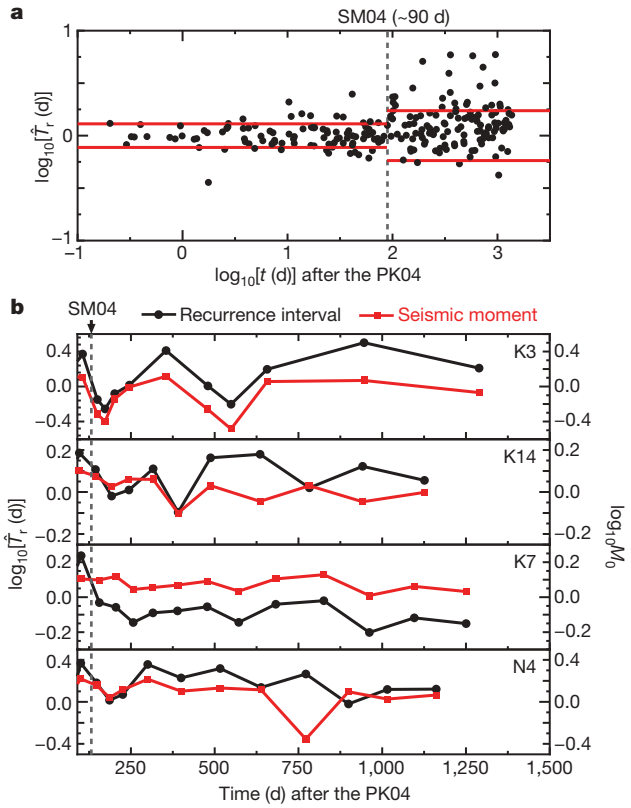


Figure 3 | Temporal variations in $\hat{T}_r$ following the 2004 Sumatra–Andaman earthquake. **a**, $\hat{T}_r$ for all available sequences. Red lines indicate standard deviations of $\hat{T}_r$. These variations are 0.11 and 0.24 for pre- and post-2004 Sumatra–Andaman earthquake periods, respectively. **b**, $\hat{T}_r$ (black circles) and M_0 (red squares) for four sequences. $\hat{T}_r$ and M_0 closely track each other, which suggests they reflect the time history of fault strength (see Supplementary Fig. 6). $\hat{T}_r$ is plotted at the midpoint between the origin times of the two events. M_0 of the second event in a pair is plotted at the same midpoint, and was normalized by its average value for the entire time period for each sequence.

$M_w = 6.0$ Parkfield earthquake[7] (from 24 days before to 2 hours after), and is thus consistent with an instantaneous coseismic change in $D(t)$ (Supplementary Fig. 4). The excursion then decays back to the pre-earthquake level after about 2–3 months (Fig. 2b). The observed structural changes are probably due to fluid redistribution around the fault resulting from a combination of both post-seismic stress relaxation[12] and direct fault-zone damage[13].

A third excursion in $D(t)$ occurred about three months after the 2004 Parkfield earthquake (Fig. 2b). The magnitude of change in $D(t)$ is comparable to that observed for the other two transients. The increase in $D(t)$ takes place over three months, after which time $D(t)$ decays slowly over a subsequent ~1 year period (Supplementary Fig. 5).

There are also changes in repeating-earthquake properties that accompany this third excursion in $D(t)$. Following the 2004 Parkfield earthquake, there is a characteristic increase in T_r, as is typically observed postseismically[14]. This trend, however, is interrupted roughly three months after the 2004 Parkfield earthquake (Fig. 2c). In 6 of the available 13 sequences, there is a systematic reduction in T_r that reaches a minimum about 6 months after the 2004 Parkfield earthquake. To explore this apparent disruption further, we removed the post-seismic effect produced by the 2004 Parkfield earthquake, assuming Omori's law, $T_r(t) = at^p$ (ref. 15 and Methods), and computed the residual recurrence interval, $\hat{T}_r(t) = T_r(t)/(at^p)$, where t is time after the earthquake, and a and p are constants to be estimated. Using all available sequences, we find that the logarithm of the residuals increases by roughly a factor of 2, beginning three months after the 2004 Parkfield earthquake (Fig. 3a). This increased variability suggests an additional perturbation to T_r and a temporal change in the mechanical properties of the fault.

An interesting feature of this variability is that there is a positive correlation between $\hat{T}_r$ and M_0 (Fig. 3b). Such a correlation is consistent with a slip-predictable model for earthquake occurrence[16], where the stress drop and T_r are both determined by the failure strength of the fault, for constant loading rate and constant minimum stress (Supplementary Fig. 6). Assuming rupture area is constant among members of a repeating-earthquake sequence[8], then fault slip will be proportional to M_0 which should in turn be proportional to stress drop.

The most dramatic correlation that we found is for sequence K3 (5 km depth; Fig. 3b). This sequence also has the least variability in event location, so that the assumption of constant fault area should be the most valid[8]. The correlation between $\hat{T}_r$ and M_0 suggests variations in fault strength over a factor of two, three months after the 2004 Parkfield earthquake. A reduction in $\hat{T}_r$ (and M_0) around 6 months after the 2004 Parkfield earthquake (Fig. 2c) indicates a temporary weakening of the fault (Fig. 3b).

We have localized the onset time of the third excursion in $D(t)$ using all of the repeating-earthquake sequences that displayed a change in $D(t)$ (Supplementary Fig. 7). We find that the excursion must have initiated between 21 December (sequence K2) and 26 December 2004 (sequence K5) (Supplementary Fig. 7b). The most dramatic tectonic event to occur within the 5-day time window is the 26 December 2004 $M_w = 9.1$ Sumatra–Andaman earthquake, whose origin time is 7 h before the end of the interval (Supplementary Fig. 7c). The timing strongly suggests that the dynamic stresses from this earthquake, estimated to be about 10 kPa (based on the amplitude of long period surface waves (>30 s) which are likely to have the strongest impact on fluid flow[17]), induced fluid flow that caused both a structural change in the fault zone region (that is, changes in $D(t)$) and, through variations

in pore pressure, consequent changes in the strength of the fault. It is now well documented that such dynamic stresses are capable of remotely triggering seismicity[18,19]. The present study suggests that these same dynamic stresses can actually produce long-term (~1 year for our study) changes in fault strength.

Given the behaviour of the third excursion, in particular the likely influence of a very large remote event, it is worthwhile revisiting the first excursion. Although initially interpreted as being produced by the 1993 Parkfield aseismic transient[4], similarities in the behaviour of $D(t)$ between the third and first excursions raise the possibility of remote triggering. Indeed, the largest observed dynamic stress change at Parkfield over the entire 22-year observing period is attributed to the 28 June 1992 $M_w = 7.3$ Landers earthquake, estimated to be 65 kPa, nearly an order of magnitude greater than for the 2004 Sumatra–Andaman earthquake. The Landers event triggered microseismicity throughout California and Nevada, both short-term (weeks)[18], as well as multiyear[20].

The excursion in $D(t)$ initiated a few months after the 1992 Landers earthquake. There were also changes in repeating-earthquake properties, namely a systematic reduction in M_0 (proportional to fault strength) from 6 of 17 available sequences following the 1992 Landers earthquake (Fig. 4). Moment reaches a minimum by the mid-1990s, during a time in which $D(t)$ and the geodetic slip rate both reach maxima[5]. All of these indicators subsequently returned to pre-1992 levels in advance of the 2004 Parkfield earthquake (Fig. 4). As with the third excursion, this overall behaviour suggests that the dynamic stresses of the 1992 Landers earthquake changed the structure and weakened the fault zone in the region of our study. In this interpretation, it is the 1992 Landers earthquake that triggered both the 1993 Parkfield aseismic transient as well as the accompanying cluster of four $M > 4$ earthquakes by weakening the fault[2,5].

Although there are many similarities between these two hypothesized episodes of remote triggering, there are also differences, which appear to be related to the occurrence of the 2004 Parkfield earthquake. First, the 1992 Landers earthquake apparently produced triggered slip on the San Andreas fault (both seismic and aseismic), whereas the 2004 Sumatra–Andaman earthquake did not. Second, for the 2004 Sumatra–Andaman earthquake, the onset of an observable change in $D(t)$ was essentially instantaneous, whereas it took months in the case of the 1992 Landers earthquake. Similarly, the corresponding changes

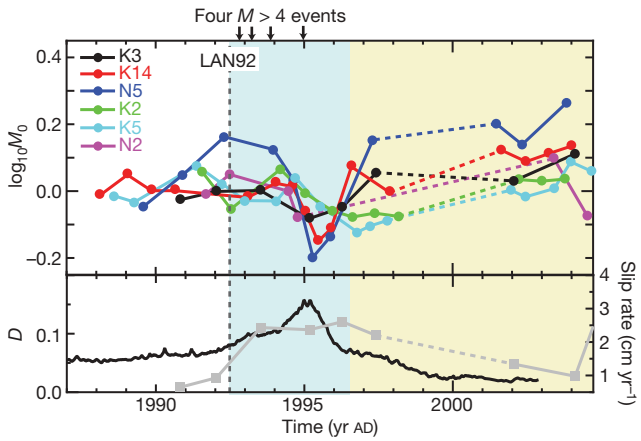


Figure 4 | Temporal changes in seismic moment for the post-Landers earthquake period. Top panel, reduction in M_0 is seen ~2–3 years after the 1992 Landers earthquake (LAN92, vertical dashed line), for 6 (of 17) sequences, suggesting temporary reduction in fault strength (see text). Arrows indicate four $M > 4$ earthquakes at Parkfield. M_0 for each sequence was normalized by average seismic moment from pre-1993 earthquakes. Bottom panel, geodetic slip rate[5] (solid line) and average $D(t)$ from sequence K3 (grey squares), using three stations (VCA, RMN and SMN). Blue (beige) field denotes inferred times of fault weakening (restrengthening). Broken lines indicate when HRSN was not functional.

in repeating-earthquake properties took a few months for the 2004 Sumatra–Andaman earthquake but a few years for the 1992 Landers earthquake. The third issue is sensitivity to dynamic stresses. There were three pre-Parkfield events in the 22-year observing period (1987–2008) that generated stresses roughly a factor of three greater than the 2004 Sumatra–Andaman earthquake (although a factor of 2 less than the 1992 Landers earthquake): the 3 November 2002 $M_w = 7.9$ Denali earthquake, the 16 October 1999 $M_w = 7.1$ Hector Mine earthquake, and the 22 December 2003 $M_w = 6.5$ San Simeon earthquake. They did not produce a detectable change in $D(t)$.

The 2004 Parkfield earthquake probably had two major effects on the fault zone. First, it damaged the fault zone by creating new fractures, as suggested by the dramatic post-seismic change in fault zone seismic properties[13]. Second, it relieved most of the stress stored in the fault zone. The timescale for fluid migration and resulting changes in fault strength should be a strong function of the fracture permeability of the medium. It is likely that Parkfield-earthquake-induced damage dramatically increased fracture permeability[21], thereby reducing the time constant for fluid migration after the earthquake. It also probably produced an increased sensitivity to dynamic stresses, thereby accounting for the positive response to Sumatra–Andaman dynamic stresses but not to the stronger 1999 Hector Mine, 2002 Denali and 2003 San Simeon dynamic stresses. Finally, the absence of accompanying slip for the 2004 Sumatra–Andaman earthquake can be explained by the low driving stress. Indeed, the hypothesized change in fault strength occurred after most of the Parkfield post-seismic deformation had already occurred.

Our inference that $D(t)$ is a proxy for fluid-induced changes in fault strength has several implications. First, it may be possible to continually monitor changes in fault strength, which is one of the most important parameters in assessing the seismogenic potential of a fault. Second, although clearly reflected in scatterer properties, such a fluid redistribution would be essentially undetectable by surface geodesy (in the absence of triggered slip), as the predicted signals are below the detection limits of currently available geodetic instruments (Supplementary Fig. 8 and Methods). The scatterers are thus measuring a fundamentally different physical property from that measured by geodesy. This would explain why the third excursion in $D(t)$ did not produce a detectable geodetic signal or a change in seismic slip rate, despite the change in scatterer properties and in fault strength inferred from repeating-earthquake systematics. Third, the effect on fault strength may be spatially heterogeneous, given that only about half of the repeating-earthquake sequences exhibited patterns consistent with fault weakening for both the Landers and Sumatra–Andaman earthquakes. This variability would be an expected consequence of fluid redistribution, which itself is likely to be spatially heterogeneous. Last, the long-range influence (~8,000 km) of the 2004 Sumatra–Andaman earthquake on this patch of the San Andreas fault suggests that many of the world's active faults were affected in the same way, thus bringing a significant fraction of them closer to failure. We speculate that such large events should produce a temporal clustering of global seismicity. This hypothesis appears to be supported by the unusually high number of $M \geq 8$ earthquakes occurring in the three years following the 2004 Sumatra–Andaman earthquake. No other large earthquake ($M \geq 8$) since 1900 was followed by as many for a comparable period.

METHODS SUMMARY

Cross-correlation analysis. We use two parameters derived from the cross-correlation of two seismograms, the lag time $\tau(t)$ and decorrelation index $D(t)$. We define $\tau(t)$ for a particular time window with centre (elapsed) time t as the lag that yields the maximum cross-correlation $C_{max}(t)$, and $D(t) = 1 - C_{max}(t)$ (ref. 4). We focus primarily on $D(t)$, because, unlike $\tau(t)$, it is insensitive to variations in the shallow near-station environment (Supplementary Fig. 1 and Methods).

Post-seismic effect on recurrence interval. We assume that T_r is governed by Omori's law, $T_r(t) = at^p$ (ref. 15), as a model of the expected increase in T_r due to a decay in fault-slip rate, where t is time after the earthquake and a and p are

constants to be estimated. Using data from the first three months of the post-Parkfield earthquake period and using all available sequences, we obtain a value of $p = 0.7$, comparable to what has been found in similar seismic studies[14], as well as to the geodetic data for this period[12].

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

Received 25 April; accepted 6 August 2009.

1. Hill, D. P. & Prejean, S. G. in *Treatise on Geophysics* (ed. Schubert, G.) Vol. 4 (ed. Kanamori, H.) 257–292 (Elsevier Science, 2007).
2. Langbein, J., Gwyther, R. L., Hart, R. H. G. & Gladwin, M. T. Slip-rate increase at Parkfield in 1993 detected by high-precision EDM and borehole tensor strainmeters. *Geophys. Res. Lett.* **26,** 2529–2532 (1999).
3. Gao, S., Silver, P. G. & Linde, A. T. Analysis of deformation data at Parkfield, California: detection of a long-term strain transient. *J. Geophys. Res.* **105** (B2), 2955–2967 (2000).
4. Niu, F., Silver, P. G., Nadeau, R. M. & McEvilly, T. V. Migration of seismic scatterers associated with the 1993 Parkfield aseismic transient event. *Nature* **426,** 544–548 (2003).
5. Murray, J. R. & Segall, P. Spatiotemporal evolution of a transient slip event on the San Andreas fault near Parkfield, California. *J. Geophys. Res.* **110,** B09407, doi:10.1029/2005JB003651 (2005).
6. Baisch, S. & Bokelmann, G. H. R. Seismic waveform attributes before and after the Loma Prieta earthquake: scattering change near the earthquake and temporal recovery. *J. Geophys. Res.* **106** (B8), 16323–16338 (2001).
7. Taira, T., Silver, P. G., Niu, F. & Nadeau, R. M. Detecting seismogenic stress evolution and constraining fault zone rheology in the San Andreas Fault following the 2004 Parkfield earthquake. *J. Geophys. Res.* **113,** B03303, doi:10.1029/2007JB005151 (2008).
8. Nadeau, R. M. & Johnson, L. R. Seismological studies at Parkfield VI: moment release rates and estimates of source parameters for small repeating earthquakes. *Bull. Seismol. Soc. Am.* **88,** 790–814 (1998).
9. Nadeau, R. M. & McEvilly, T. V. Fault slip rates at depth from recurrence intervals of repeating microearthquakes. *Science* **285,** 718–721 (1999).
10. Nadeau, R. M. & McEvilly, T. V. Periodic pulsing of characteristic microearthquakes on the San Andreas Fault. *Science* **303,** 220–222 (2004).
11. Bürgmann, R. *et al.* Earthquake potential along the Northern Hayward Fault, California. *Science* **289,** 1178–1182 (2000).
12. Langbein, J., Murray, J. R. & Snyder, H. A. Coseismic and initial postseismic deformation from the 2004 Parkfield, California, earthquake, observed by Global Positioning System, electronic distance meter, creepmeters, and borehole strainmeters. *Bull. Seismol. Soc. Am.* **96,** S304–S320 (2006).
13. Li, Y.-G., Chen, P., Cochran, E. S., Vidale, J. E. & Burdette, T. Seismic evidence for rock damage and healing on the San Andreas Fault associated with the 2004 M 6.0 Parkfield earthquake. *Bull. Seismol. Soc. Am.* **96,** S349–S363 (2006).
14. Peng, Z., Vidale, J. E., Marone, C. & Rubin, A. Systematic variations in recurrence interval and moment of repeating aftershocks. *Geophys. Res. Lett.* **32,** L15301, doi:10.1029/2005GL022626 (2005).
15. Schaff, D. P., Beroza, G. C. & Shaw, B. E. Postseismic response of repeating aftershocks. *Geophys. Res. Lett.* **25,** 4549–4552 (1998).
16. Shimazaki, K. & Nakata, T. Time-predictable recurrence model for large earthquakes. *Geophys. Res. Lett.* **7,** 279–282 (1980).
17. Brodsky, E. E. & Prejean, S. G. New constraints on mechanisms of remotely triggered seismicity at Long Valley Caldera. *J. Geophys. Res.* **110,** B04302, doi:10.1029/2004JB003211 (2005).
18. Hill, D. P. *et al.* Seismicity remotely triggered by the magnitude 7.3 Landers, California, earthquake. *Science* **260,** 1617–1623 (1993).
19. West, M., Sánchez, J. J. & McNutt, S. R. Periodically triggered seismicity at Mount Wrangell, Alaska, after the Sumatra earthquake. *Science* **308,** 1144–1146 (2005).
20. Gao, S. S., Silver, P. G., Linde, A. T. & Sacks, I. S. Annual modulation of triggered seismicity following the 1992 Landers earthquake in California. *Nature* **406,** 500–504 (2000).
21. Elkhoury, J. E., Brodsky, E. E. & Agnew, D. C. Seismic waves increase permeability. *Nature* **441,** 1135–1138 (2006).

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

Acknowledgements This work was supported by the National Science Foundation EAR-0337308, EAR-0408947, EAR-0409024, EAR-0453638, EAR-0537641, EAR-0544730 and DTM-2025-01, and by the Department of Terrestrial Magnetism, Carnegie Institution of Washington. The Parkfield High-Resolution Seismic Network is operated by University of California, Berkeley Seismological Laboratory with financial support from the US Geological Survey (USGS) through National Earthquake Hazards Reduction Program award 07HQAG0014. Seismic and geodetic data are archived at the Northern California Earthquake Data Center and USGS, respectively. We thank J. Murray-Moraleda for providing the slip rate data, and E. A. Roeloffs and A. T. Linde for discussions. This is Berkeley Seismological Laboratory contribution 09-14.

Author Contributions T.T., with P.G.S. and F.N, analysed data; R.M.N. identified and maintained the repeating earthquake catalogue; P.G.S. and T.T. wrote the manuscript with contributions from all authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.T. (taira@dtm.ciw.edu).

METHODS

Data set. The 17 repeating-earthquake sequences we have examined span a depth range of ~10 km and along-fault distance of ~8 km. Sequences K4, K10, K13 and K15 have been updated to 1998.

Decorrelation index and lag times from cross-correlation analysis. $\tau(t)$ is defined as the mean travel time perturbation of the lag times, τ_i, associated with the individual scattered phases in the time window, while $D(t)$ can be defined as $\omega^2 s_\tau^2/2$ where s_τ and ω are the standard deviation of τ_i and the characteristic frequency of the scattered phases, respectively[22]. Thus $\tau(t)$ and $D^{1/2}(t)$ are proportional to the mean and standard deviation of a weighted distribution of lag times, respectively. A constant velocity reduction in the background medium should yield a linear trend in $\tau(t)$ and $D(t) = 0$, given a short time window for calculating $\tau(t)$ and $D(t)$.

On the other hand, a previous numerical experiment[4] illustrates that isolated spikes in $\tau(t)$ and $D(t)$ can be explained by local changes in the location of a scatterer or by a velocity reduction around the scatterer. Our interest is in detecting discrete scatterers that permit the imaging of fault-zone properties at seismogenic depth, rather than changes in the background medium. We have therefore focused on detection of isolated spikes in $D(t)$ rather than $\tau(t)$ because the latter potentially contains this shallow contribution.

Measurement of decorrelation index. To measure $D(t)$, the chosen reference events for stations VCA, RMN and SMN were taken to be pre-1993, while those for stations CCRB and LCCB occurred ~50 days after the 2004 Parkfield earthquake. Elapsed times used for measuring $D(t)$ for stations VCA, RMN and SMN were the same as those for the previous study[4]. For stations CCRB and LCCB (unavailable in previous study), the first clear isolated peak in $D(t)$ in S-coda was chosen. A 0.5-s time window was used for calculating $D(t)$. As a way of assessing the data quality, we computed median $D(t)$ in a 1-s time window after the direct P wave. This gives us a measure of $D(t)$ in a part of the seismogram where we do not expect a contribution from the medium (ref. 4). On this basis we ranked the stations, averaged over all sequences, and found the stations VCA, CCRB and LCCB to be our best stations. We also used the same procedure to obtain a noise level for $D(t)$ for each sequence that may then be used to define a level of significance (for example, Supplementary Fig. 7b). Finally, we only use data where the median $D(t)$ (in a 2-s time window after the direct P wave) is less than 0.1.

Predicted geodetic response for a redistribution of fluid. Assume that we have a vertical fluid-filled fracture at 5 km depth with dimension one square kilometre in size and 10 mm in width, and that fluid migrates 500 m laterally (that is, 500 m of the fracture closes at one end and opens at the other end). The resulting deformation signal can be expressed as the sum of two sources, both 1 km by 500 m in dimension, opposite in sign, and separated by one kilometre. Using an elastic half-space model[23] to calculate the surface deformation, we find that the resulting strain is very small, due to the near-cancellation of the two sources, and is estimated to be less than one nanostrain, and the corresponding displacement of the order of micrometres. Given that the time constant for this change is probably a few months, such an event is undetectable, geodetically.

22. Snieder, R., Grêt, A., Douma, H. & Scales, J. Coda wave interferometry for estimating nonlinear behavior in seismic velocity. *Science* **295**, 2253–2255 (2002).
23. Okada, Y. Internal deformation due to shear and tensile faults in a half-space. *Bull. Seismol. Soc. Am.* **82**, 1018–1040 (1992).

doi:10.1038/nature08322

LETTERS

A pre-*Archaeopteryx* troodontid theropod from China with long feathers on the metatarsus

Dongyu Hu[1], Lianhai Hou[1,2], Lijun Zhang[1,3] & Xing Xu[1,2]

The early evolution of the major groups of derived non-avialan theropods is still not well understood, mainly because of their poor fossil record in the Jurassic. A well-known result of this problem is the 'temporal paradox' argument that is sometimes made against the theropod hypothesis of avian origins[1]. Here we report on an exceptionally well-preserved small theropod specimen collected from the earliest Late Jurassic Tiaojishan Formation of western Liaoning, China[2]. The specimen is referable to the Troodontidae, which are among the theropods most closely related to birds. This new find refutes the 'temporal paradox'[1] and provides significant information on the temporal framework of theropod divergence. Furthermore, the extensive feathering of this specimen, particularly the attachment of long pennaceous feathers to the pes, sheds new light on the early evolution of feathers and demonstrates the complex distribution of skeletal and integumentary features close to the dinosaur–bird transition.

Anchiornis huxleyi has been recently described, based on an incomplete specimen, as a basal avialan filling the morphological gap between non-avian and avian dinosaurs[3]. A nearly complete, extensively feathered specimen (LPM-B00169, housed in Liaoning Paleontological Museum) referable to *Anchiornis huxleyi* (Supplementary Information) has now been recovered from the Tiaojishan Formation at the Daxishan locality, Jianchang County. The Tiaojishan Formation has traditionally been regarded as Middle Jurassic but was recently dated to between 161 and 151 Myr (ref. 2) (Fig. 1; Supplementary Information). It is therefore older than the *Archaeopteryx*-bearing strata near Solnhofen, Germany, which date to less than 150 Myr (ref. 4).

The new specimen exhibits a sub-triangular skull in lateral view and a large surangular foramen, as in other paravians and deinonychosaurs, respectively[5–7] (Figs 2 and 3a). It has a large maxillary fenestra separated from the antorbital fenestra by a narrow interfenestral bar, a dorsoventrally flattened internarial bar, a distinct, posteriorly widening groove on the labial surface of the dentary housing the neurovascular foramina, and closely packed premaxillary and dentary teeth in the symphyseal region, derived features shared with other troodontids[5,7–9]. Furthermore, *Anchiornis* resembles *Mei* in having a large external naris extending posteriorly well beyond the anterior border of the antorbital fossa[5], a longitudinal groove along the dorsomedial margin of the slender sub-orbital ramus of the jugal, a maxillary tooth row that approaches the preorbital bar posteriorly[5] and unserrated teeth (also seen in the troodontid *Byronosaurus*[8]). A large paraquadrate foramen is present, in contrast to other troodontids but in agreement with dromaeosaurs[6].

Postcranially, *Anchiornis* also possesses a few troodontid features, including relatively long and slender transverse processes on the dorsal vertebrae and anterior-most caudal vertebrae[8]. As in *Archaeopteryx*[4], the middle and posterior caudal vertebrae each bear a distinct groove on the lateral surface near the junction of the neural arch and centrum (Fig. 3b), and the coracoid bears a laterally located coracoid tubercle (Fig. 3c). The ischium is strongly curved posteriorly, with an obturator

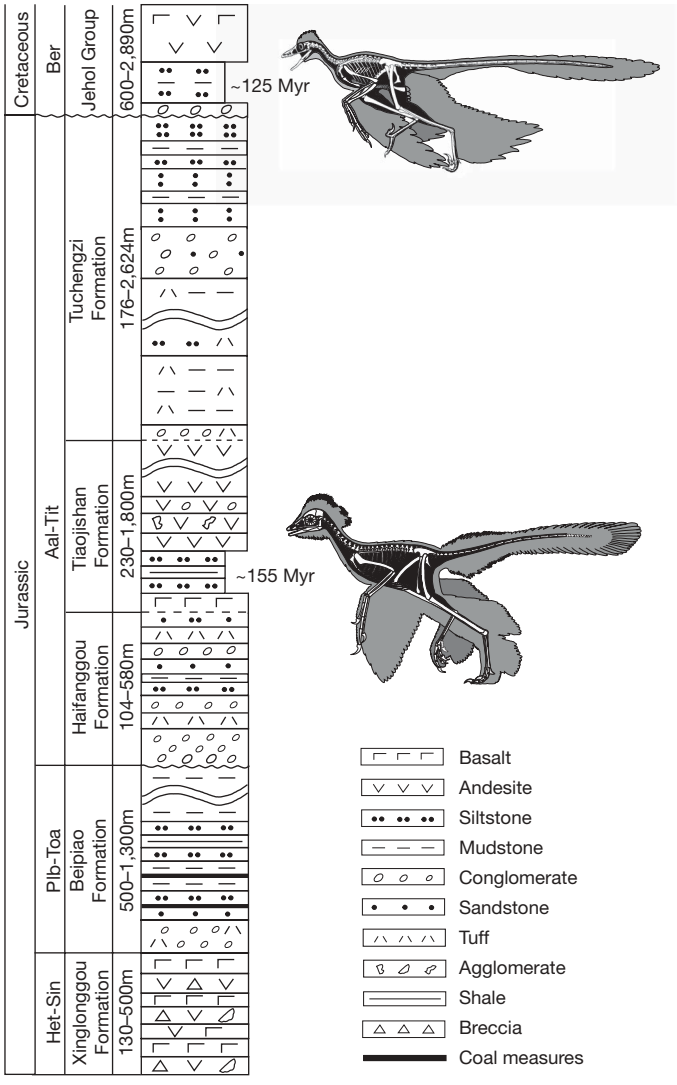


Figure 1 | Stratigraphic column of Jurassic and Lowest Cretaceous strata in western Liaoning, showing horizons from which feathered dinosaurs have been described. Two major horizons have produced such specimens: the Tiaojishan Formation has yielded *Anchiornis huxleyi* LPM-B00169 and dates to about 155 Myr, whereas the Jehol Group has yielded *Microraptor* and other feathered dinosaurs and dates to about 125 Myr.

[1]Paleontological Institute, Shenyang Normal University, 253 North Huanghe Street, Shenyang 110034, China. [2]Key Laboratory of Evolutionary Systematics of Vertebrates, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, 142 Xiwai Street, Beijing 100044, China. [3]Shenyang Institute of Geology and Mineral Resources, 25 Beiling Street, Shenyang 110032, China.

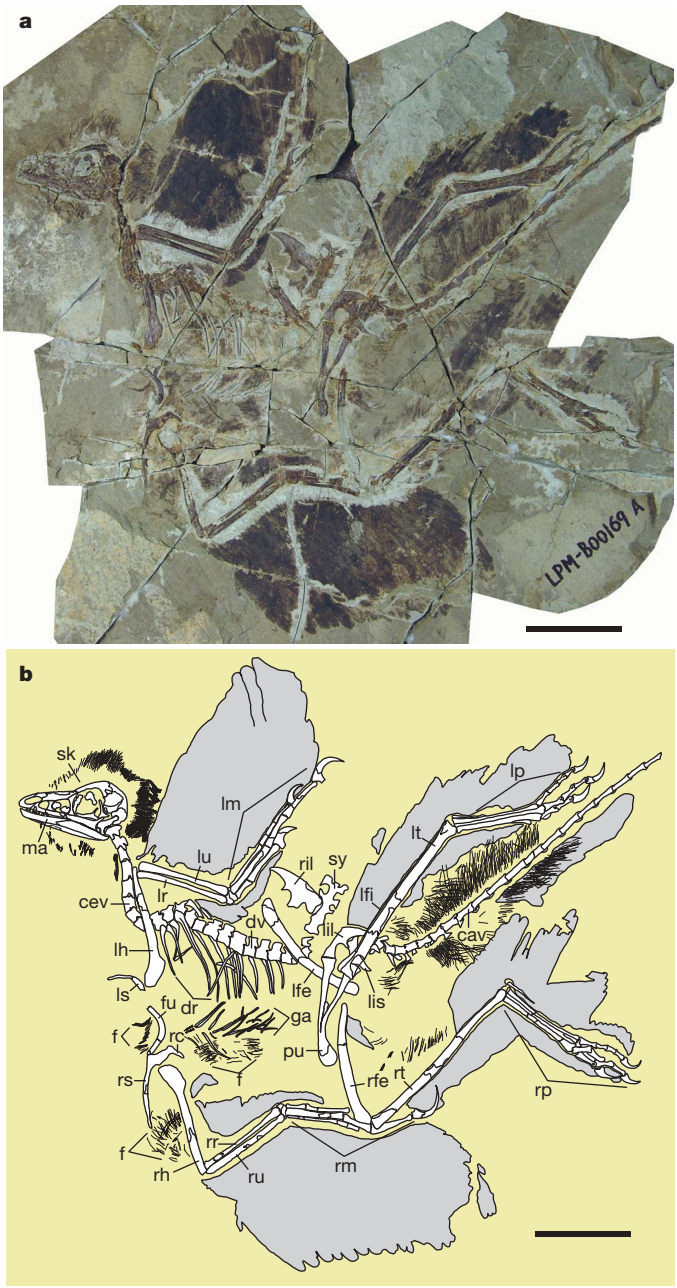


Figure 2 | *Anchiornis huxleyi* LPM-B00169. **a**, Photograph and **b**, line drawing of LPM-B00169 A (slab). Abbreviations: cav, caudal vertebra; cev, cervical vertebra; dr, dorsal rib; dv, dorsal vertebra; f, feather; fu, furcula; ga, gastralia; lfe, left femur; lfi, left fibula; lh, left humerus; lil, left ilium; lis, left ischium; lm, left manus; lp left pes; lr, left radius; ls, left scapula; lt, left tibiotarsus; lu, left ulna; ma, mandible; pu, pubis; rc, right coracoid; rfe, right femur; rh, right humerus; ril, right ilium; rm, right manus; rp, right pes; rr, right radius; rs, right scapula; rt, right tibiotarsus; ru, right ulna; sk, skull; sy, synsacrum. Scale bar, 5 cm.

process located close to the midpoint of the ischial shaft as in the dromaeosaurid *Buitreraptor*[10]. As in *Mei* and some dromaeosaurids[11], the distal articulation of metatarsal II is about as wide as that of metatarsal III.

There are about ten large pennaceous feathers attached to the forearm, 11 to the manus (Fig. 3d, e), 12–13 to the crus (Fig. 3f) and 10–11 to the metatarsus (Fig. 3g). The primary and secondary remiges are similar in size and morphology (Fig. 3j–l). The longest remiges, located near the distal end of the forearm and the proximal end of the manus, are about 150% of the humeral length. Because the longest wing feathers are close to the wrist, the broadest part of the

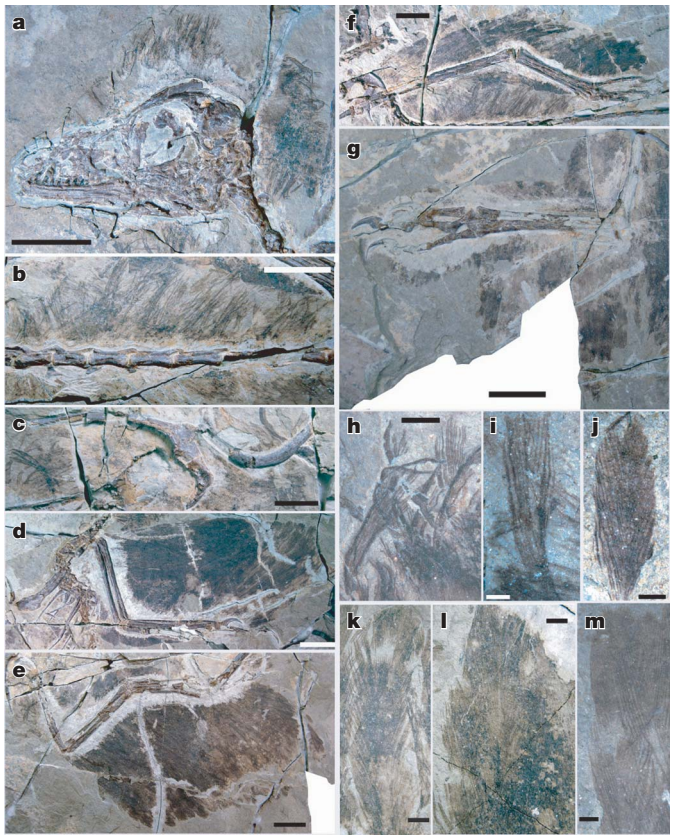


Figure 3 | Selected skeletal elements and associated feathers of LPM-B00169. **a**, Skull and mandible. **b**, Middle caudal vertebrae. **c**, Furcula and right scapulocoracoid. **d**, Left forelimb. **e**, Right forelimb. **f**, Left hindlimb. **g**, Right pes. **h**, Feathers near skull. **i**, Feathers near pectoral region. **j**, Distalmost primary remiges. **k**, Middle primary remiges. **l**, Secondary primary remiges. **m**, Pennaceous feathers attached to metatarsus. Arrows indicate the boundary of the soft tissue. Scale bars: **a**, **b**, **d**–**g**, 2 cm; **c**, 1 cm; **h**, **k**, **l**, 2 mm in; **i**, **j**, **m**, 1 mm.

wing is relatively proximal, in contrast to the basal dromaeosaurid *Microraptor* and basal avians such as *Archaeopteryx*[12]. The distal primary remiges resemble those of *Microraptor* and basal avians[4,12,13] in having curved rachides, but the remiges are relatively small, with thin rachides, symmetrical vanes, and blunt ends. In *Microraptor* and basal avians[12], the primary remiges are not only significantly longer than the secondary ones but also possess prominent rachides, asymmetrical vanes and somewhat pointed tips. Large pennaceous feathers also cover nearly the whole length of the lower leg of *Anchiornis* (Fig. 3m). The crural feathers are longer than the metatarsal feathers, which are nearly perpendicular to the long axis of the metatarsus. Much shorter feathers also attach to the pedal phalanges, apart from the unguals. These feathers appear to be pennaceous, a condition previously unknown in any fossil taxon.

Two types of plumulaceous feather can be distinguished, corresponding to types previously identified in the dromaeosaurid *Sinornithosaurus*[14]. Feathers of the first type are each composed of a bundle of filaments that are joined together proximally and remain nearly parallel as they pass distally (Fig. 3h). Feathers of the second type each consist of a series of filaments branching laterally from a slender central rachial filament (Fig. 3i).

Anchiornis huxleyi is posited as a basal troodontid by our phylogenetic analysis (Fig. 4) but also possesses several salient features of the Avialae and Dromaeosauridae, further blurring distinctions among the three major paravian groups (Supplementary Information). In particular, the forelimbs of *Anchiornis* are proportionally much longer than those of other troodontids, but similar in length to those of basal avialans and dromaeosaurids. *Anchiornis huxleyi* even exceeds *Microraptor* in the length of the forelimbs[3], a feature often

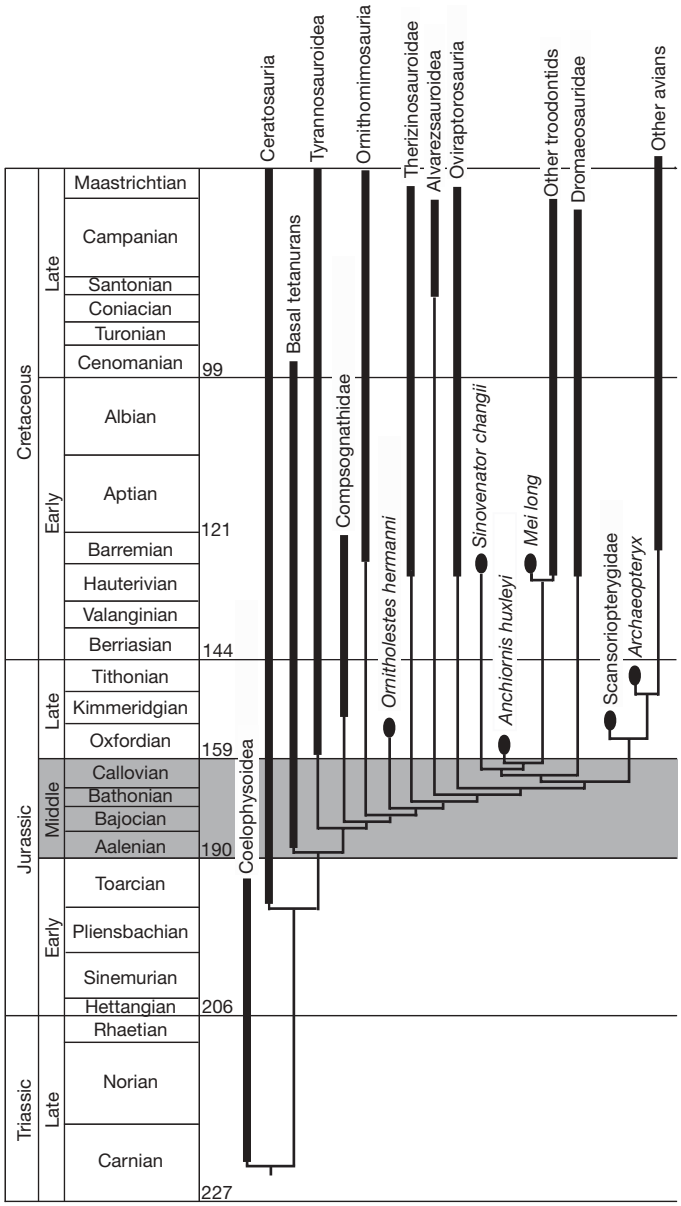


Figure 4 | Temporally calibrated phylogeny of the Theropoda. (See Supplementary Information for details.) The known temporal durations (solid bars) of the major theropod groups, based on well-corroborated fossil occurrences, indicate that the major tetanuran theropod groups diverged rapidly in the Middle Jurassic to earliest Late Jurassic.

regarded as an indicator of aerial capability. However, the remiges of *Anchiornis* are not obviously flight-adapted as in *Microraptor* and basal avians[12]. Furthermore, the extreme elongation of the lower legs in *Anchiornis huxleyi* suggests strong cursorial capability[3], but their long and extensive plumage is not common among cursorial animals: in cursorial birds and mammals, the lower legs tend to show reduction or even complete loss of the feathers or hair, respectively. Similarly, the recently reported Jurassic basal avialan *Epidexipteryx hui* closely resembles basal birds in various skeletal features usually interpreted as being related to aerial capability, but lacks pennaceous feathers that would facilitate aerial locomotion[15]. Collectively, these data point to a complex pattern of morphological evolution at the base of the Paraves.

The distribution of large pennaceous feathers, both phylogenetic and anatomical, implies that these structures first evolved on the distal portions of the forelimbs and tail in maniraptoran theropods and only spread proximally at a subsequent stage in theropod evolution[12]. Interestingly, such a distal-first pattern seems also to apply to the evolution of pennaceous feathers on the hindlimbs of paravians. Large pennaceous feathers are now known to occur on the lower leg and particularly the metatarsus of at least one basal member of each of the three major paravian groups, namely the basal troodontid *Anchiornis*, the basal avialan *Pedopenna* and the basal dromaeosaurid *Microraptor*[13,16]. Furthermore, many basal avians have proportionally large pennaceous feathers on the lower leg[13,17], which are reduced in more derived birds. This suggests that large pennaceous feathers first evolved distally on the hindlimbs, as on the forelimbs and tail. This distal-first development led to a four-winged condition at the base of the Paraves. Whereas the large feathers of the forewing developed further in subsequent avian evolution, the large hindwing feathers were reduced and even lost[12]. This suggests that extensive feathering of the pes was a critical modification in the transition to birds and that the pedal scales of extant birds might be secondarily derived structures, a possibility also supported by some developmental studies[18].

Extensive feathering of the pes is also seen in some modern birds, and serves an insulating or protective function[19]. In most cases the feathers are not organized into a coherent planar surface as in *Microraptor*, *Pedopenna* and *Anchiornis*[20], which indicates that the pedal feathers of these fossil taxa may have differed from those of extant birds in having an aerodynamic function. This would imply that a four-winged condition played a role in the origin of avian flight, as suggested by previous studies[12,17], although this conclusion is not universally accepted[21]. However, the significant differences noted above between the large pedal feathers of *Anchiornis* and those of *Microraptor* suggest that these feathers might have been less aerodynamically effective in *Anchiornis* than in *Microraptor*.

The poor fossil record of derived non-avialan theropods has hindered understanding of early coelurosaurian theropod evolution, and has created an apparent discrepancy between stratigraphy and phylogeny that has sometimes been used to argue against the theropod hypothesis of bird origins[1]. Although derived theropods including troodontids have been reported in the Jurassic, all have been based on fragmentary material[22–25]. *Anchiornis huxleyi* is probably Oxfordian in age, and unquestionably represents the oldest troodontid reported so far[7]. The presence of a troodontid in the earliest Late Jurassic indicates that all groups of derived theropods had originated by this time. A calibrated theropod phylogeny based only on well-corroborated fossil occurrences suggests that all major tetanuran groups, including Aves, might have originated and diversified rapidly in the Middle to earliest Late Jurassic (Fig. 4). This rapid divergence event would have coincided with documented palaeogeographical changes that took place around the same time. Alternatively, a calibrated theropod phylogeny incorporating fragmentary material suggests that tetanurans have a much longer evolutionary history, and that great potential exists for discovering derived theropod fossils, even in the Triassic[24].

Received 29 April; accepted 27 July 2009.
Published online 24 September 2009.

1. Feduccia, A. Birds are dinosaurs: simple answer to a complex problem. *Auk* **119,** 1187–1201 (2002).
2. Xu, K. *et al. Jurassic System in the North of China (VII): The Stratigraphic Region of Northeast China* (Petroleum Industry Press, 2003).
3. Xu, X. *et al.* A new feathered maniraptoran dinosaur fossil that fills a morphological gap in avian origin. *Chin. Sci. Bull.* **54,** 430–435 (2009).
4. Wellnhofer, P. *Archaeopteryx—Der urvogel von Solnhofen* (Friedrich Pfeil, 2008).
5. Xu, X. & Norell, M. A. A new troodontid from China with avian-like sleeping posture. *Nature* **431,** 838–841 (2004).
6. Norell, M. A. & Makovicky, P. J. in *The Dinosauria* 2nd edn (eds Weishampel, D. B., Dodson, P. & Osmolska, H.) 196–209 (Univ. California Press, 2004).
7. Makovicky, P. J. & Norell, M. A. in *The Dinosauria* 2nd edn (eds Weishampel, D. B., Dodson, P., Osmolska, H.) 184–195 (Univ. California Press, 2004).
8. Makovicky, P. J., Norell, M. A., Clark, J. M. & Rowe, T. Osteology and relationships of *Byronosaurus jaffei* (Theropda: Troodontidae). *Am. Mus. Novit.* **3402,** 1–32 (2003).
9. Currie, P. J. Bird-like characteristics of the jaws and teeth of troodontid theropods (Dinosauria: Saurischia). *J. Vert. Paleontol.* **7,** 72–81 (1987).
10. Makovicky, P. J. Apesteguía, S. & Agnolín, F. L. The earliest dromaeosaurid theropod from South America. *Nature* **437,** 1007–1011 (2005).
11. Xu, X. & Wang, X.-L. A new dromaeosaur (Dinosauria: Theropoda) from the Early Cretaceous Yixian Formation of western Liaoning. *Vert. PalAsia.* **42,** 111–119 (2004).

12. Xu, X. *et al.* Four-winged dinosaurs from China. *Nature* **421,** 335–340 (2003).
13. Zhang, F.-C. & Zhou, Z.-H. Leg feathers in an Early Cretaceous bird. *Nature* **431,** 925 (2004).
14. Xu, X., Zhou, Z.-H. & Prum, R. O. Branched integumental structures in *Sinornithosaurus* and the origin of feathers. *Nature* **410,** 200–204 (2001).
15. Zhang, F.-C., Zhou, Z.-H., Xu, X., Wang, X.-L. & Sullivan, C. A bizarre Jurassic maniraptoran from China with elongate ribbon-like feathers. *Nature* **455,** 1105–1108 (2008).
16. Xu, X. & Zhang, F.-C. A new maniraptoran with long metatarsalian feathers from China. *Naturwissenschaften* **92,** 173–177 (2005).
17. Longrich, N. R. Structure and function of hindlimb feathers in *Archaeopteryx lithographica. Paleobiology* **32,** 417–431 (2006).
18. Sawyer, R. H. & Knapp, L. W. Avian skin development and the evolutionary origin of feathers. *J. Exp. Zool.* **298B,** 57–72 (2003).
19. Weidensaul, S. *Raptors—The Birds of Prey* (Lyons & Burford, 1995).
20. Xu, X. *et al.* Origin of flight: could four-winged dinosaurs fly? *Nature* **438,** E3–E4 (2005).
21. Padian, K. Four-winged dinosaurs, bird precursors, or neither? *Bioscience* **53,** 450–452 (2003).
22. Jensen, J. A. & Padian, K. Small pterosaurs and dinosaurs from the Uncompahgre Fauna (Brushy Basin Member, Morrison Formation: ?Tithonian), Late Jurassic, western Colorado. *J. Paleontol.* **63,** 364–373 (1989).
23. Chure, D. J. *Koparion douglassi*, a new dinosaur from the Morrison Formation (Upper Jurassic) of Dinosaur National Monument: the oldest troodontid (Theropoda: Maniraptora). *Brigham Young Univ. Geol. Stud.* **40,** 11–15 (1994).
24. Xu, X., Zhao, X.-J. & Clark, J. M. A new therizinosaur from the Lower Jurassic Lufeng Formation of Yunnan, China. *J. Vert. Paleontol.* **21,** 477–483 (2001).
25. Evans, S. E. & Milner, A. R. in *In the Shadow of the Dinosaurs* (eds Fraser, N. C. & Sues, H.-D.) 303–321 (Cambridge Univ. Press, 1985).

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

Acknowledgements We thank S. Zheng for help with the fieldwork, T. Yu for preparing the specimen, C. Sullivan for editing and commenting on the manuscript, G. Sun, P. Cheng and F. Jin for discussions and comments, and J. Huang for Supplementary Fig. 4i. The fieldwork was supported by grants from the Education Bureau of Liaoning Province (numbers 20060805 and 2008S214) and the Special Fund of Shenyang Normal University. This study was also supported by grants to X.X. from the Chinese Academy of Sciences, the National Natural Science Foundation of China, and Major Basic Research Projects of the Ministry of Science and Technology, China.

Author Contributions D.-Y.H and X.X. designed the project. D.-Y.H., X.X., L.-H.H and L.-J.Z. performed the research. X.X. and D.-Y.H. 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 X.X. (xingxu@vip.sina.com) or D.-Y.H. (hudongyu@synu.edu.cn).

LETTERS

Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms

Daisuke Hattori[1], Yi Chen[1], Benjamin J. Matthews[2], Lukasz Salwinski[3], Chiara Sabatti[4], Wesley B. Grueber[5] & S. Lawrence Zipursky[1]

***Down Syndrome cell adhesion molecule* (*Dscam*) genes encode neuronal cell recognition proteins of the immunoglobulin superfamily[1,2]. In *Drosophila*, *Dscam1* generates 19,008 different ectodomains by alternative splicing of three exon clusters, each encoding half or a complete variable immunoglobulin domain[3]. Identical isoforms bind to each other, but rarely to isoforms differing at any one of the variable immunoglobulin domains[4,5]. Binding between isoforms on opposing membranes promotes repulsion[6]. Isoform diversity provides the molecular basis for neurite self-avoidance[6–11]. Self-avoidance refers to the tendency of branches from the same neuron (self-branches) to selectively avoid one another[12]. To ensure that repulsion is restricted to self-branches, different neurons express different sets of isoforms in a biased stochastic fashion[7,13]. Genetic studies demonstrated that Dscam1 diversity has a profound role in wiring the fly brain[11]. Here we show how many isoforms are required to provide an identification system that prevents non-self branches from inappropriately recognizing each other. Using homologous recombination, we generated mutant animals encoding 12, 24, 576 and 1,152 potential isoforms. Mutant animals with deletions encoding 4,752 and 14,256 isoforms[14] were also analysed. Branching phenotypes were assessed in three classes of neurons. Branching patterns improved as the potential number of isoforms increased, and this was independent of the identity of the isoforms. Although branching defects in animals with 1,152 potential isoforms remained substantial, animals with 4,752 isoforms were indistinguishable from wild-type controls. Mathematical modelling studies were consistent with the experimental results that thousands of isoforms are necessary to ensure acquisition of unique Dscam1 identities in many neurons. We conclude that thousands of isoforms are essential to provide neurons with a robust discrimination mechanism to distinguish between self and non-self during self-avoidance.**

Dscam1 diversity provides the molecular basis of neurite self-avoidance (for review, see ref. 1). Neurite self-avoidance facilitates uniform coverage of target fields by dendritic and axonal branches, while allowing the branches of different neurons to share the same field[12]. Self-avoidance also has an important role in branch segregation during axon guidance[7,8,11]. Self-avoidance requires that self-branches repel one another, whereas branches of different neurons do not repel each other. Each neuron expresses a unique combination of Dscam1 isoforms comprising some tens of isoforms in a largely stochastic manner[7,13]. Thus, isoform-specific homophilic recognition is restricted to self-branches. Binding elicits neurite repulsion[6]. That Dscam1 diversity underlies the phenomenon of self-avoidance is well documented in the formation of dendritic receptive fields formed by dendritic arborization neurons in the larval body wall[6,9,10] and in the segregation of axon branches of mushroom body neurons in the central brain[7,11,14,15]. Although self-avoidance demands that neurons express different sets of isoforms, the identity of the isoforms they express is not important[6,7,11,14] (see later).

How many isoforms are necessary for each neuron within a group to discriminate between self and non-self? To address this question, we analysed mutants with reduced Dscam1 diversity (Fig. 1a, b). Previously, deletion mutants were generated that removed subsets of exon 4 variants (for example, $Dscam1^{\Delta 4.1–4.3}$ and $Dscam1^{\Delta 4.4–4.12}$; Fig. 1b)[14]. The largest deletion ($Dscam1^{\Delta 4.4–4.12}$) removes nine out of twelve exon 4 variants and, thus, encodes 4,752 potential isoforms. To reduce Dscam1 diversity further, a series of knock-in mutants encoding between 12 and 1,152 isoforms were generated using homologous recombination (Fig. 1b and Supplementary Fig. 1). These include three different $Dscam1^{576\text{-isoform}}$ alleles, each of which contains only one exon 9 variant. Thus, each allele encodes a distinct and non-overlapping set of 576 potential isoforms. Furthermore, three $Dscam1^{12\text{-isoform}}$ alleles were generated, each of which contains one combination of exon 6 and 9 variants, thus encoding 12 potential isoforms. By combining two different $Dscam1^{576\text{-isoform}}$ alleles, animals with 1,152 potential isoforms were generated (referred to as $Dscam1^{1152\text{-isoform}}$). In a similar manner, animals with 24 potential isoforms were generated by combining two different $Dscam1^{12\text{-isoform}}$ alleles (referred to as $Dscam1^{24\text{-isoform}}$). Replacement of the genomic regions was confirmed by sequencing ~10 kilobases (kb) spanning the modified segment. Sequence analysis of complementary DNAs isolated from these alleles showed that the fixed exons were expressed with different combinations of the remaining variable exons (Supplementary Fig. 2a). Protein levels and distributions were similar to wild type, as assessed by western blots and immunostaining, respectively (Supplementary Fig. 2b, c). In summary, these alleles encode reduced numbers of isoforms expressed in a similar pattern and level as in wild type.

We first sought to determine the number of isoforms necessary for patterning dendritic arborization neurons. These neurons elaborate dendrites in a two-dimensional pattern over the larval body wall. Self-dendrites seldom overlap. In contrast, dendrites of different dendritic arborization neurons in the same territory overlap extensively[16]. Removal of all Dscam1 isoforms results in increased overlap and fasciculation of self-dendrites (Fig. 1c), whereas the overexpression of a single Dscam1 isoform in dendritic arborization neurons with normally overlapping dendrites leads to ectopic avoidance between them[6,9,10]. In all mutant animals encoding reduced diversity,

[1]Department of Biological Chemistry, Howard Hughes Medical Institute, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA. [2]Center for Neurobiology and Behavior, Columbia University Medical Center, New York, New York 10032, USA. [3]UCLA-DOE Institute for Genomics and Proteomics, Howard Hughes Medical Institute, Molecular Biology Institute, Los Angeles, California 90095, USA. [4]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California 90095, USA. [5]Department of Neuroscience and Department of Physiology and Cellular Biophysics, Columbia University Medical Center, New York, New York 10032, USA.

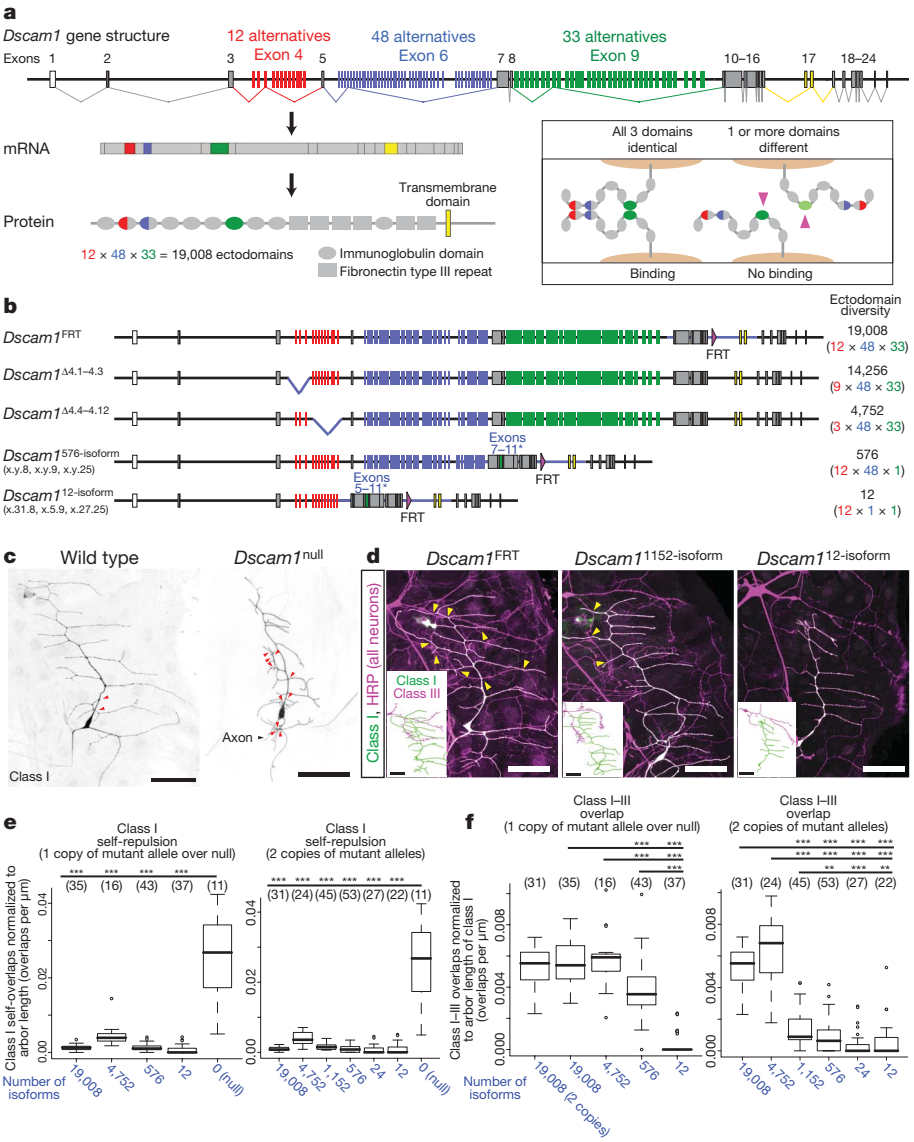


Figure 1 | Thousands of isoforms are required for dendrites of dendritic arborization neurons to distinguish between self and non-self. **a**, Alternative splicing of *Dscam1*. The inset shows a schematic of Dscam1 isoform-specific homophilic binding. The crystal structure shows that variable domains pair in an antiparallel configuration[23,24]. Binding requires matching of variable domains; one mismatch (right, pink arrowheads) disrupts binding. **b**, *Dscam1* alleles with reduced diversity. Alternative exons were replaced with cDNAs encoding exons 7–11 and 5–11 to generate $Dscam1^{576\text{-isoform}}$ and $Dscam1^{12\text{-isoform}}$ alleles, respectively (asterisks). Sequence modified is indicated with blue line. $Dscam1^{FRT}$ is a control for $Dscam1^{576\text{-isoform}}$ and $Dscam1^{12\text{-isoform}}$ knock-in alleles, and contains an FRT site as a result of homologous recombination[11]. Three different alleles were generated for each mutant. x.y.z refers to combinations of exon 4, 6 and 9, respectively, encoded by alleles. $Dscam1^{\Delta 4.1–4.3}$ and $Dscam1^{\Delta 4.4–4.12}$ were reported previously[14]. **c**, Self-repulsion of dendritic arborization neurons requires *Dscam1*. Red arrowheads, self-crossing of class I (vpda) dendrites. **d**, Representative images of dendritic arborization neuron dendrites in different *Dscam1* mutants. See Supplementary Fig. 3 for larger images. All neurons were visualized with anti-HRP antibody (magenta), and class I (vpda) neurons were labelled with GFP (green; appears white because they overlap with magenta). Class III (v′pda) neurons were traced from cell bodies[16] (insets). Yellow arrowheads, crossing between class I and III dendrites. Scale bars, 50 μm. **e**, Self-repulsion is intact in all mutants with reduced diversity. Numbers in parenthesis denote class I neurons analysed for each genotype. Data in boxplot format (thick line, median; box, 25–75% quartiles; error bar, data within 1.5× quartile range; circles, outliers). $***P < 0.001$; Student's *t*-test. **f**, Increasing the number of available isoforms increases overlaps between class I and class III dendrites. Numbers in parenthesis, class I/III pairs analysed. Data in boxplot format. $**P < 0.01$, $***P < 0.001$; Wilcoxon rank test. Thousands of isoforms are required for normal degree of overlap. For genotypes, see Methods.

self-dendrites of class I dendritic arborization neurons avoided each other as in wild-type controls (Fig. 1d, e). In contrast to wild type, however, dendrites of class I and class III neurons, which normally overlap, showed nearly complete avoidance in mutants encoding 12 and 24 isoforms (Fig. 1d, f and Supplementary Fig. 3). This indicates that ectopic repulsion occurred between class I and III dendrites in these mutants.

We next tested the effects of increasing the pool of isoforms to hundreds and thousands on the overlap between class I and III dendrites. We observed increased overlaps as the number of potential isoforms increased (Fig. 1f and Supplementary Fig. 3). Class I and III dendrites overlapped substantially more in $Dscam1^{576\text{-isoform}}/Dscam1^{null}$ mutants than in $Dscam1^{12\text{-isoform}}/Dscam1^{null}$ mutants, although still less than in wild type. Homozygous mutants with two copies of $Dscam1^{576\text{-isoform}}$ had fewer dendritic overlaps than the $Dscam1^{576\text{-isoform}}/Dscam1^{null}$ mutants. Thus, the number of overlaps between the dendrites of different neurons is sensitive to the overall level of Dscam1 protein with this level of diversity. $Dscam1^{1152\text{-isoform}}$ mutants showed a modest increase in overlaps compared to that seen in $Dscam1^{576\text{-isoform}}$ mutants. In contrast, mutant animals with 4,752 potential isoforms (that is, $Dscam1^{\Delta 4.4–4.12}$) had overlaps that were indistinguishable from the wild-type control. A linear regression analysis showed a strong correlation between the number of available isoforms and the degree of overlap between class I and III dendrites ($R^2 = 0.66$, *t*-test: $P < 2 \times 10^{-16}$). These findings suggest that a substantial fraction of isoforms is shared between class I and III neurons when tens or hundreds of isoforms are available. With thousands of isoforms available, however, different neurons express sufficiently different populations of Dscam1 isoforms to prevent ectopic repulsion. In summary, a pool of thousands of isoforms is necessary for each dendritic arborization neuron to reliably discriminate between self and non-self neurites.

We next assessed the extent of diversity required for mushroom body development. The mushroom body comprises some 2,500 neurons and most axons form two branches projecting into two orthogonal paths, which result in the formation of two distinct lobes[17]. Dscam1-mediated self-avoidance is required for self-branches to segregate with high fidelity (Fig. 2a)[7,11]. Mushroom bodies in mutants with only a single ectodomain encoded at the *Dscam1* locus ($Dscam1^{single}$) invariably form only one lobe, indicating that mushroom body axons do not segregate appropriately[11]. In contrast, self-branches of a single mushroom body neuron expressing a single

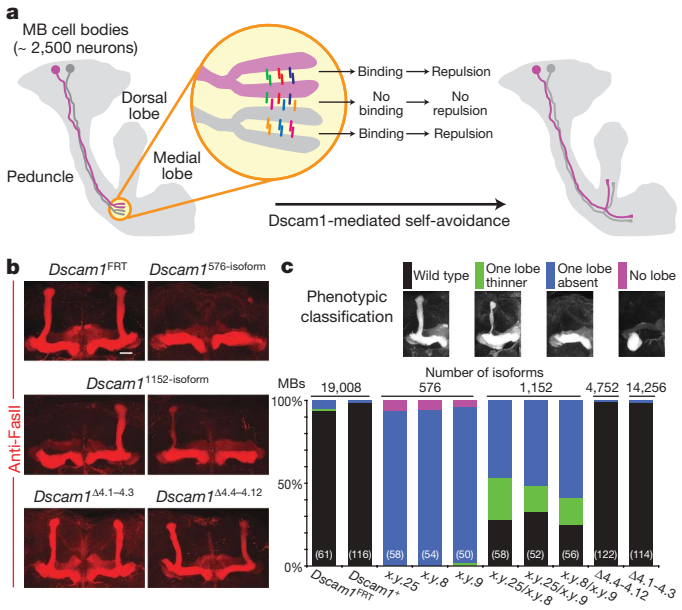


Figure 2 | Thousands of isoforms are required for mushroom body lobe development. a, Schematic of mushroom body (MB) development. The call-out circle represents a model of selective recognition between self-branches mediated by Dscam1 isoform-specific homophilic binding. Dscam1 isoforms are shown as coloured bars. **b**, Mushroom body lobe morphology in mutant animals, visualized with monoclonal antibody 1D4 (anti-FasII, red). Scale bar, 20 μm. **c**, Quantification of the mushroom body lobe phenotypes. The numbers in the parenthesis represent the number of mushroom bodies (that is, the number of brain hemispheres) examined for each genotype. See Methods for genotypes.

isoform surrounded by wild-type neighbours in a genetically mosaic animal segregate with high fidelity[7,11,14]. Together, these data suggest that the failure to form two lobes in *Dscam1*^{single} mutants reflects inappropriate recognition between the branches of different neurons[11]. All *Dscam1*$^{576\text{-isoform}}$ mutant animals form only a single lobe, similar to that seen in *Dscam1*^{single} mutants (Fig. 2b, c). Increasing the potential pool of isoforms to 1,152 by combining two different alleles of *Dscam1*$^{576\text{-isoform}}$, partially improved mushroom body lobe segregation (Fig. 2b, c). In contrast, the morphology of mushroom bodies in mutants with 4,752 potential isoforms (that is, *Dscam1*$^{\Delta4.4–4.12}$) was indistinguishable from wild type (Fig. 2b, c). These results show that between 1,152 and 4,752 isoforms are required for mushroom body axon branches to distinguish between self and non-self, thereby supporting orderly axon branch segregation and the formation of the bi-lobed mushroom body structure.

In summary, thousands of isoforms are essential for generating sufficiently different Dscam1 cell-surface identities to support a robust discrimination mechanism for dendrites of dendritic arborization neurons and axon branches of mushroom body neurons to distinguish between self and non-self. This self-avoidance function of Dscam1 does not depend on the identity of the isoforms, as different mutants with the same degree of diversity showed similar spectra of phenotypes, and deletion mutants collectively removing all isoforms (*Dscam1*$^{\Delta4.1–4.3}$ and *Dscam1*$^{\Delta4.4–4.12}$) showed normal mushroom body lobe segregation and patterns of dendrites of dendritic arborization neurons (Fig. 2b, c, Supplementary Fig. 3 and ref. 6).

A fundamentally different role for Dscam1 diversity in regulating branching was proposed in a previous study[18], where the complex branching patterns of posterior scutellar mechanosensory neurons were analysed in two different, although partially overlapping, exon 4 deletion mutants with the same degree of diversity (11,088 potential isoforms). Differences in the branching patterns of posterior scutellar neurons between deletions and between deletions and control were described. These included an increase in the number of animals with branches that are present in a fraction of control animals, and the appearance of some ectopic branches unique to each mutant. These findings led to the proposal that, in contrast to self-avoidance, it is the identity of specific isoforms that has a role in regulating the formation of specific branches of posterior scutellar neurons. Here we revisit the role of Dscam1 diversity in branching of posterior scutellar neurons. In contrast to previous work[18], our data demonstrate that there is no requirement for specific isoforms, and that it is the number of isoforms that is crucial in regulating branching.

All *Dscam1*$^{576\text{-isoform}}$ mutant animals had marked defects in the axonal branching patterns of posterior scutellar neurons (Fig. 3a–c). The branching patterns were improved in *Dscam1*$^{1152\text{-isoform}}$ mutants. The spectra of phenotypes were indistinguishable between different *Dscam1*$^{576\text{-isoform}}$ mutants and between different *Dscam1*$^{1152\text{-isoform}}$ mutants encoding different sets of isoforms (Fig. 3c, d and Supplementary Figs 5 and 6). Thus, the improvement seen in *Dscam1*$^{1152\text{-isoform}}$ mutants derives solely from the increase in the number of potential isoforms, and not from acquisition of 'correct' isoforms.

That isoform identity is not essential for elaborating the branching pattern of posterior scutellar neurons was demonstrated by examining branching patterns in *Dscam1*$^{\Delta4.1–4.3}$ and *Dscam1*$^{\Delta4.4–4.12}$ mutants, which collectively remove all Dscam1 isoforms. A few branch segments exhibited nominally significant differences between different genotypes. However, these differences did not persist in the same mutants in different genetic backgrounds (Fig. 3c, e and Supplementary Figs 5 and 6), or in mutants with one copy of a deletion allele and one copy of a protein-null allele (Fig. 3f and Supplementary Figs 4–6). Furthermore, cluster analysis showed that the branching patterns of *Dscam1*$^{\Delta4.1–4.3}$ and *Dscam1*$^{\Delta4.4–4.12}$ mutants were indistinguishable from controls (Fig. 3g and Supplementary Fig. 6). Thus, it is the number of isoforms, not their identity, that is crucial in regulating the branching pattern of posterior scutellar neurons.

In summary, we conclude that the branching of posterior scutellar neurons requires between 1,152 and 4,752 isoforms, similar to that observed for dendritic arborization and mushroom body neurons, and that no specific isoforms are required for this process. Although the details of branch formation of posterior scutellar neurons in the developing thoracic ganglion are not known, the target areas of branches contain neurites of many neurons. Thus, it is likely that the branching defects observed in mutants with reduced diversity reflect the failure of branches to adequately discriminate between self and non-self neurites. That is, insufficient diversity leads to inappropriate repulsive interactions between neurites of different neurons. Notably, thousands of isoforms are required for normal patterning in all three systems analysed in this paper.

To address why so many isoforms are required to prevent interactions between self and non-self, we mathematically modelled the relationship between the number of potential isoforms and the number of distinct Dscam1 identities that can be generated. Although isoform expression in dendritic arborization and posterior scutellar neurons has not been analysed, previous molecular studies on mushroom body neurons, as well as other neurons, showed that each neuron expresses some tens of isoforms in a largely stochastic fashion[7,13]. Accordingly, we used a Monte Carlo simulation and binomial distribution to estimate the number of neurons that reliably acquire 'unique' Dscam1 identities (that is, arrays of isoforms expressed in a pair of neurons that do not trigger ectopic repulsion between them), given different numbers of potential isoforms (Fig. 4a). We found that the probabilities that a pair of neurons express the same isoform(s) were rather high (Supplementary Fig. 7a). This finding is analogous to the 'birthday problem' in which the likelihood of two individuals sharing the same birthday from a randomly chosen population is counter-intuitively high (for example, there is a 50% likelihood that at least two individuals share the same birthday in a group of 23 people). For example, there is 4.4% chance that a pair of neurons shares at least one isoform out of 30 messenger RNAs even with a pool of 20,000

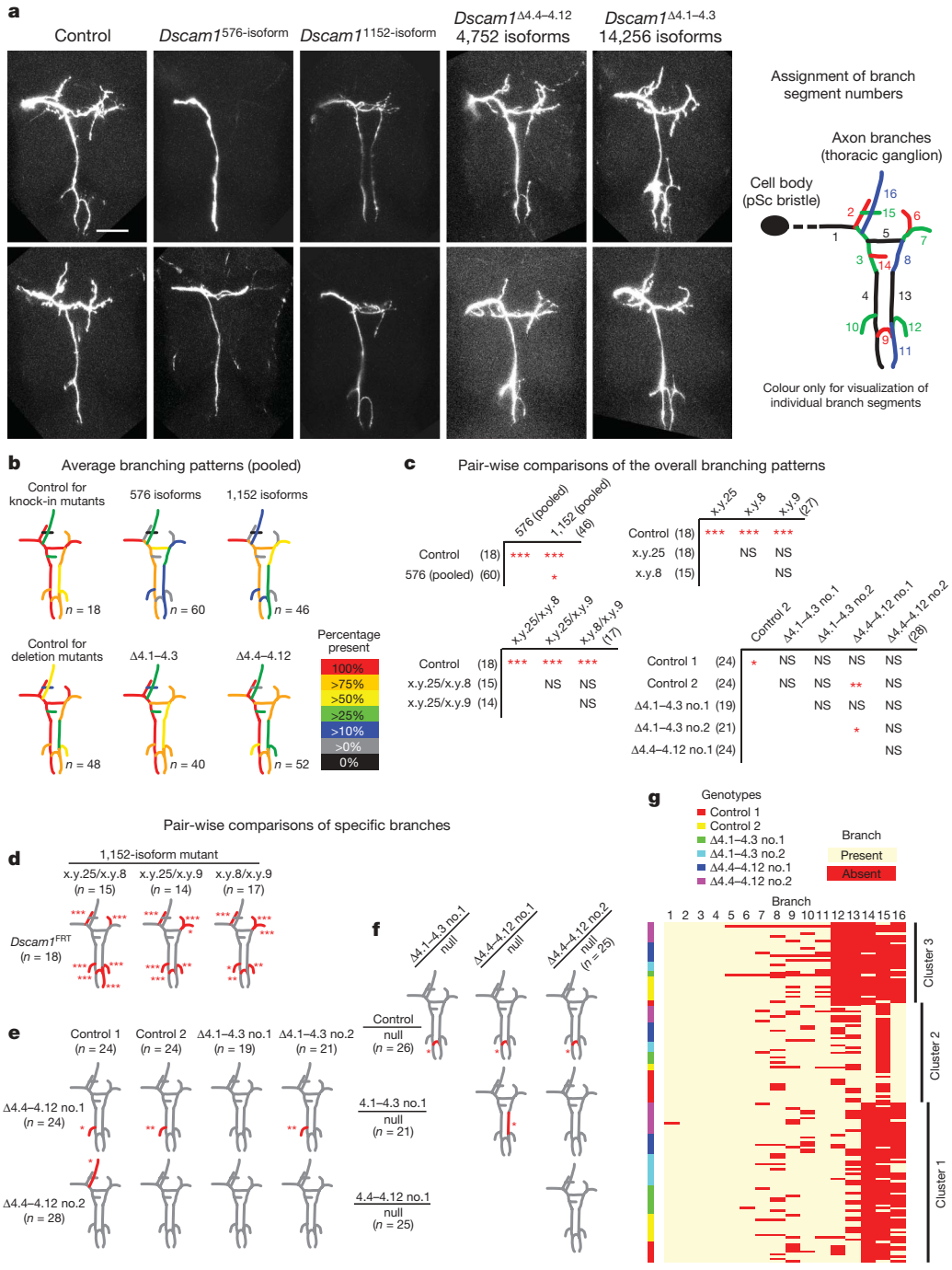
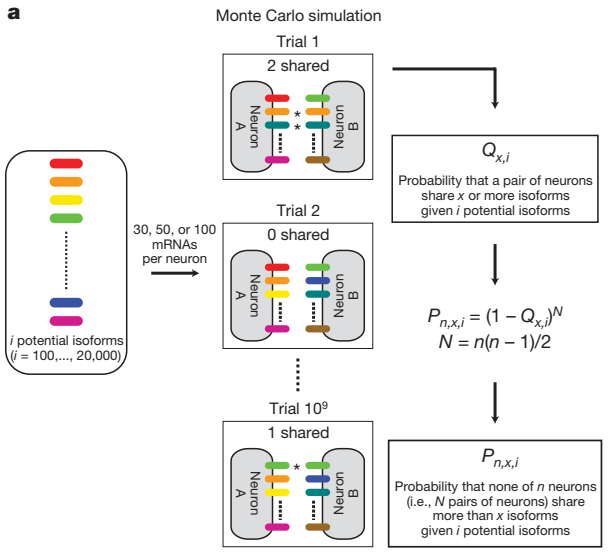


Figure 3 | Branching of posterior scutellar neurons requires thousands of isoforms independent of their identity. **a**, Representative images of posterior scutellar (pSc) neuron axon branch patterns in homozygous *Dscam1* mutants with different degrees of diversity (for *Dscam1* deletion mutants over $Dscam1^{null}$ alleles, see Supplementary Fig. 4). Scale bar, 50 μm. Right panel, branch segment numbers were assigned arbitrarily for scoring (scoring method, Supplementary Fig. 4). **b**, Average branch patterns for each genotype (for *Dscam1* deletion mutants over $Dscam1^{null}$ alleles, and raw data, see Supplementary Fig. 5). **c**, Pair-wise comparisons of overall branch patterns. *P* values were determined using permutation test with Bonferroni correction accounting for 30 different pair-wise comparisons (see Methods). Numbers in parenthesis indicate the number of animals analysed. See Supplementary Fig. 6 for raw statistical data. **d–f**, Branches that show difference in frequency between each pair are red. *P* values were calculated using Fisher's exact test with Bonferroni correction accounting for 16 different branch segments. **g**, The result of partitioning cluster analysis (homozygous deletion mutants). Each individual sample was clustered based on its branching pattern. Each of the three clusters contains controls, $Dscam1^{\Delta 4.1-4.3}$ mutants, and $Dscam1^{\Delta 4.4-4.12}$ mutants in similar ratios. See Supplementary Fig. 6 for results for knock-in mutants. NS, not significant; $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

isoforms. Thus, if one isoform shared between two neurons was sufficient to trigger repulsion, then fewer than three neurons would reliably acquire unique identities, even from a pool of 20,000 isoforms (Supplementary Fig. 7b). This indicates that some degree of overlap in the isoforms expressed between two neurons is probably permissive. When sharing some fraction of isoforms (for example, ~20%) between two neurons is permitted, a nonlinear increase in the number of neurons acquiring unique identities was observed as the pool of isoforms increases from hundreds to thousands, regardless of the number of isoforms expressed per neuron (Fig. 4b). Notably, thousands of isoforms are necessary to reliably generate tens of neurons or more with unique Dscam1 identities. Thus, this modelling study provides a mathematical explanation for why thousands of isoforms are necessary to promote repulsion between self-branches with minimal risk of ectopic recognition between non-self branches.

In conclusion, we demonstrate a profound requirement for thousands of Dscam1 isoforms in patterning neural circuits. The importance of this level of diversity is underscored by the observation that all arthropod species for which genome sequence is available encode thousands of different Dscam1 ectodomains (for example, ~3,000 isoforms in water flea[19], ~6,000 in honey bee[20], and up to 20,000 in different *Drosophila* species[20]). We show that around 5,000 isoforms are sufficient to pattern all three systems analysed in this paper. Further diversification of *Drosophila Dscam1* beyond 5,000 isoforms may reflect differential splicing mechanisms in different neuronal subclasses (for example, some neurons may show more extensive bias in splicing thus limiting the access to the full diversity), a function mediating recognition between neurons in other developmental contexts (for example, to promote adhesive interactions between synaptic partners) or functions elsewhere such as in the immune system[21]. Notably, our study argues that the demand for robust discrimination between self and non-self neurites underlying self-avoidance alone is sufficient to account for thousands of isoforms encoded at the *Dscam1* locus in *Drosophila*.



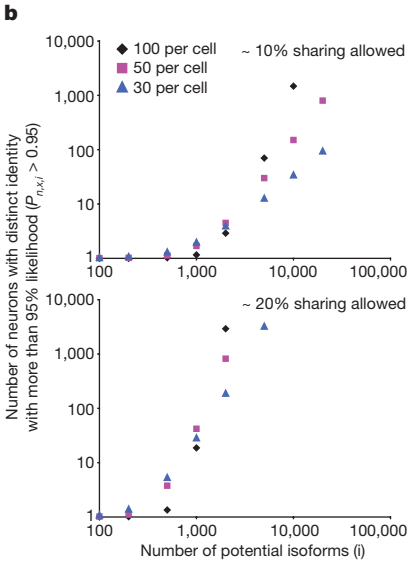


Figure 4 | Mathematical model showing that thousands of isoforms are required to give rise to many unique Dscam1 identities. **a**, A schematic representation of mathematical modelling used to determine number of neurons that acquire unique Dscam1 identities. Tables of $Q_{x,i}$ for each condition are shown in Supplementary Fig. 7a. **b**, The number of neurons that obtain unique identities at more than 95% likelihood with varying degrees of diversity. Both axes are in log scale. See Supplementary Fig. 7b for the condition in which no sharing of isoforms between neurons is allowed.

METHODS SUMMARY

Dscam1$^{576\text{-isoform}}$ and *Dscam1*$^{12\text{-isoform}}$ mutant alleles were generated essentially as described[11] (Supplementary Fig. 1). At least two independent recombinants were generated for each allele encoding distinct variants, and these were used to generate homozygous mutants to avoid effects from potential second site mutations (for example, *Dscam1*$^{x.y.25a}$/*Dscam1*$^{x.y.25b}$). Molecular characterization of these mutants was performed essentially as described[11]. Immunostaining was performed essentially as described[6,11]. The branching patterns of posterior scutellar mechanosensory neurons were analysed by DiI tracing as described[18,22]. The branching patterns of different exon 4 deletion mutants were scored blindly to avoid potential bias in scoring (for details of the scoring method, see Supplementary Fig. 4 and Methods). Pair-wise comparisons of branch occurrence between genotypes were performed by Fisher's exact test. Comparisons between genotypes that account for all branches were conducted using a permutation-based test. Because a large number of comparisons are carried out in this study, we used the Bonferroni criteria to correct for multiple testing.

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

Received 15 June; accepted 14 August 2009.

1. Hattori, D., Millard, S. S., Wojtowicz, W. M. & Zipursky, S. L. Dscam-mediated cell recognition regulates neural circuit formation. *Annu. Rev. Cell Dev. Biol.* **24,** 597–620 (2008).
2. Schmucker, D. & Chen, B. Dscam and DSCAM: complex genes in simple animals, complex animals yet simple genes. *Genes Dev.* **23,** 147–156 (2009).
3. Schmucker, D. *et al. Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. *Cell* **101,** 671–684 (2000).
4. Wojtowicz, W. M., Flanagan, J. J., Millard, S. S., Zipursky, S. L. & Clemens, J. C. Alternative splicing of *Drosophila* Dscam generates axon guidance receptors that exhibit isoform-specific homophilic binding. *Cell* **118,** 619–633 (2004).
5. Wojtowicz, W. M. *et al.* A vast repertoire of Dscam binding specificities arises from modular interactions of variable Ig domains. *Cell* **130,** 1134–1145 (2007).
6. Matthews, B. J. *et al.* Dendrite self-avoidance is controlled by Dscam. *Cell* **129,** 593–604 (2007).
7. Zhan, X. L. *et al.* Analysis of Dscam diversity in regulating axon guidance in *Drosophila* mushroom bodies. *Neuron* **43,** 673–686 (2004).
8. Zhu, H. *et al.* Dendritic patterning by Dscam and synaptic partner matching in the *Drosophila* antennal lobe. *Nature Neurosci.* **9,** 349–355 (2006).
9. Hughes, M. E. *et al.* Homophilic Dscam interactions control complex dendrite morphogenesis. *Neuron* **54,** 417–427 (2007).
10. Soba, P. *et al. Drosophila* sensory neurons require Dscam for dendritic self-avoidance and proper dendritic field organization. *Neuron* **54,** 403–416 (2007).
11. Hattori, D. *et al.* Dscam diversity is essential for neuronal wiring and self-recognition. *Nature* **449,** 223–227 (2007).
12. Kramer, A. P. & Stent, G. S. Developmental arborization of sensory neurons in the leech *Haementeria ghilianii*. II. Experimentally induced variations in the branching pattern. *J. Neurosci.* **5,** 768–775 (1985).
13. Neves, G., Zucker, J., Daly, M. & Chess, A. Stochastic yet biased expression of multiple Dscam splice variants by individual cells. *Nature Genet.* **36,** 240–246 (2004).
14. Wang, J. *et al.* Transmembrane/juxtamembrane domain-dependent Dscam distribution and function during mushroom body neuronal morphogenesis. *Neuron* **43,** 663–672 (2004).
15. Wang, J., Zugates, C. T., Liang, I. H., Lee, C. H. & Lee, T. *Drosophila* Dscam is required for divergent segregation of sister branches and suppresses ectopic bifurcation of axons. *Neuron* **33,** 559–571 (2002).
16. Grueber, W. B., Jan, L. Y. & Jan, Y. N. Tiling of the *Drosophila* epidermis by multidendritic sensory neurons. *Development* **129,** 2867–2878 (2002).
17. Lee, T., Lee, A. & Luo, L. Development of the *Drosophila* mushroom bodies: sequential generation of three distinct types of neurons from a neuroblast. *Development* **126,** 4065–4076 (1999).
18. Chen, B. E. *et al.* The molecular diversity of Dscam is functionally required for neuronal wiring specificity in *Drosophila*. *Cell* **125,** 607–620 (2006).
19. Brites, D. *et al.* The Dscam homologue of the crustacean Daphnia is diversified by alternative splicing like in insects. *Mol. Biol. Evol.* **25,** 1429–1439 (2008).
20. Graveley, B. R. *et al.* The organization and evolution of the dipteran and hymenopteran *Down syndrome cell adhesion molecule* (*Dscam*) genes. *RNA* **10,** 1499–1506 (2004).
21. Watson, F. L. *et al.* Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science* **309,** 1874–1878 (2005).
22. Grillenzoni, N., van Helden, J., Dambly-Chaudiere, C. & Ghysen, A. The iroquois complex controls the somatotopy of *Drosophila* notum mechanosensory projections. *Development* **125,** 3563–3569 (1998).
23. Meijers, R. *et al.* Structural basis of Dscam isoform specificity. *Nature* **449,** 487–491 (2007).
24. Sawaya, M. R. *et al.* A double S shape provides the structural basis for the extraordinary binding specificity of Dscam isoforms. *Cell* **134,** 1007–1018 (2008).

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

Acknowledgements We thank T. Lee and B. Dickson for fly stocks, B. Dickson and E. Demir for advice on homologous recombination procedures, S. Miura for staining of embryos, and N. Nocera and other members of the Zipursky laboratory for help in screening homologous recombinants. We thank D. Eisenberg for helpful discussion on mathematical modelling. We thank R. Axel, A. Nern, W. Wojtowicz and other members of the Zipursky and Grueber laboratories for discussion and comments on the manuscript. This work was supported by a grant from the NIH (S.L.Z.), and NRSA from NIH/NINDS (B.J.M.).W.B.G. is a Searle Scholar, McKnight Scholar and Klingenstein Fellow. S.L.Z. is an investigator of the Howard Hughes Medical Institute.

Author Contributions D.H. generated knock-in mutants. B.J.M. and W.B.G. analysed dendritic arborization neuron patterning. D.H. and Y.C. performed analyses of mushroom body and posterior scutellar neurons. L.S. and D.H. conducted mathematical modelling. C.S. performed statistical analysis on the branching of posterior scutellar neuron and linear regression analysis on dendritic arborization neuron data. D.H. and S.L.Z. designed the study 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 S.L.Z. (lzipursky@mednet.ucla.edu).

METHODS

Generation of *Dscam1*$^{576\text{-isoform}}$ and *Dscam1*$^{12\text{-isoform}}$ alleles. *Dscam1*$^{576\text{-isoform}}$ and *Dscam1*$^{12\text{-isoform}}$ alleles were generated through a two-step gene targeting strategy previously described[11]. Approximately 4-kb homology arms were PCR-amplified from the reference strain for the *Drosophila* genome sequencing project[25]. These genomic PCR products were combined with cDNA encoding a single exon 9 variant (for *Dscam1*$^{576\text{-isoform}}$) or a single exon 6 and a single exon 9 variant (for *Dscam1*$^{12\text{-isoform}}$) by overlap-extension PCR and cloned into a custom vector. These constructs were introduced into flies by a standard transgenic technique. At least two independent inserts of these constructs were used for gene targeting crosses. At least two independent *Dscam1*$^{5'\text{HR}}$ alleles were used to generate each knock-in allele and independent knock-in alleles were used to make homozygous animals (for example, *Dscam1*$^{\text{x.y.25a}}$/*Dscam1*$^{\text{x.y.25b}}$) to avoid effects of potential second site mutations.

Molecular characterization of *Dscam1*$^{576\text{-isoform}}$ and *Dscam1*$^{12\text{-isoform}}$ alleles. The molecular characterization of the *Dscam1*$^{576\text{-isoform}}$ and *Dscam1*$^{12\text{-isoform}}$ alleles was performed essentially as described[11]. cDNA sequencing was performed from RNA extracted from animals with one copy of mutant *Dscam1* allele over a deficiency allele, *Dscam1*$^{\text{Df6055}}$.

Genetics. Flies with the following genotypes were used in this study. *Dscam1*[21] and *Dscam1*[23] are protein-null alleles of *Dscam1*. For dendritic arborization neuron analysis: (1) *Dscam1*$^{\text{FRT}}$/*Dscam1*[23]; *221-Gal4,UAS-mCD8GFP*/+, (2) *Dscam1*[21]/*Dscam1*[23]; *221-Gal4,UAS-mCD8GFP*/+, (3) *Dscam1*$^{12\text{-isoform}}$/*Dscam1*[23]; *221-Gal4,UAS-mCD8GFP*/+, (4) *Dscam1*$^{576\text{-isoform}}$/*Dscam1*[23]; *221-Gal4,UAS-mCD8GFP*/+, (5) *Dscam1*$^{\Delta 4.4-4.12}$/*Dscam1*[23]; *221-Gal4,UAS-mCD8GFP*/+, (6) *Dscam1*$^{\text{FRT}}$; *221-Gal4,UAS-mCD8GFP*/+, (7) *Dscam1*$^{12\text{-isoform}}$; *221-Gal4,UAS-mCD8GFP*/+, (8) *Dscam1*$^{576\text{-isoform}}$; *221-Gal4,UAS-mCD8GFP*/+, (9) *Dscam1*$^{\text{x.27.25}}$/*Dscam1*$^{\text{x.31.8}}$; *221-Gal4,UAS-mCD8GFP*/+, (10) *Dscam1*$^{\text{x.27.25}}$/*Dscam1*$^{\text{x.5.9}}$; *221-Gal4,UAS-mCD8GFP*/+, (11) *Dscam1*$^{\text{x.y.25}}$/*Dscam1*$^{\text{x.y.8}}$; *221-Gal4,UAS-mCD8GFP*/+, (12) *Dscam1*$^{\text{x.y.25}}$/*Dscam1*$^{\text{x.y.9}}$; *221-Gal4,UAS-mCD8GFP*/+, (13) *Dscam1*$^{\text{x.y.8}}$/*Dscam1*$^{\text{x.y.9}}$; *221-Gal4,UAS-mCD8GFP*/+, and (14) *Dscam1*$^{\Delta 4.4-4.12}$; *221-Gal4,UAS-mCD8GFP*/+.

For mushroom body neuron analysis: (1) *Dscam1*$^{\text{FRT}}$, (2) *FRTG13,Dscam1*$^{+}$/*Dscam1*$^{\text{FRT}}$, (3) *FRTG13,Dscam1*$^{+}$/*Dscam1*$^{+}$, (4) *Dscam1*$^{576\text{-isoform}}$, (5) *Dscam1*$^{\text{x.y.25}}$/*Dscam1*$^{\text{x.y.8}}$, (6) *Dscam1*$^{\text{x.y.25}}$/*Dscam1*$^{\text{x.y.9}}$, (7) *Dscam1*$^{\text{x.y.8}}$/*Dscam1*$^{\text{x.y.9}}$, (8) FRTG13,*Dscam1*$^{\Delta 4.4-4.12}$/*Dscam1*$^{\Delta 4.4-4.12}$, (9) *Dscam1*$^{\Delta 4.4-4.12}$, (10) FRTG13,*Dscam1*$^{\Delta 4.1-4.3}$*/*FRTG13,Dscam1*$^{\Delta 4.1-4.3}$, and (11) *FRTG13,Dscam1*$^{\Delta 4.1-4.3}$.

For posterior scutellar neuron analysis: (1) *Dscam1*$^{\text{FRT}}$/*Dscam1*$^{+}$, (2) control 1: *FRTG13,Dscam1*$^{+}$/*Dscam1*$^{+}$, (3) control 2: *FRTG13,Dscam1*$^{+}$/*Dscam1*$^{\text{FRT}}$, (4) *Dscam1*$^{576\text{-isoform}}$, (5) *Dscam1*$^{\text{x.y.25}}$/*Dscam1*$^{\text{x.y.8}}$, (6) *Dscam1*$^{\text{x.y.25}}$/*Dscam1*$^{\text{x.y.9}}$, (7) *Dscam1*$^{\text{x.y.8}}$/*Dscam1*$^{\text{x.y.9}}$, (8) Δ4.4–4.12 no.1: *FRTG13,Dscam1*$^{\Delta 4.4-4.12}$/*Dscam1*$^{\Delta 4.4-4.12}$, (9) Δ4.4–4.12 no.2: *Dscam1*$^{\Delta 4.4-4.12}$, (10) Δ4.1–4.3 no.1: *FRTG13,Dscam1*$^{\Delta 4.1-4.3}$*/*FRTG13,Dscam1*$^{\Delta 4.1-4.3}$, (11) Δ4.1–4.3 no.2: *FRTG13,Dscam1*$^{\Delta 4.1-4.3}$, (12) *FRTG13,Dscam1*$^{+}$/*Dscam1*[21], (13) Δ4.4–4.12 no.1 (heterozygous): *FRTG13,Dscam1*$^{\Delta 4.4-4.12}$/*Dscam1*[21], (14) Δ4.4–4.12 no.2 (heterozygous): *Dscam1*$^{\Delta 4.4-4.12}$/*Dscam1*[21], and (15) *FRTG13,Dscam1*$^{\Delta 4.1-4.3}$*/*Dscam1*[21].

Two chromosomes containing the Δ4.1–4.3 mutation were kept in two different genetic backgrounds (one with a balancer and one without balancers). Asterisk indicates the chromosome that was maintained with a balancer.

Immunostaining. Immunostaining for the embryo and the mushroom body was performed essentially as described[11]. Dendritic arborization neuron analysis was performed essentially as described[6]. The antibodies were used at the following dilutions for immunostaining: monoclonal antibody 1D4 (anti-FasII) (1:10), monoclonal antibody 11G4 (anti-Dscam1) (1:2), and rabbit anti-GFP (Molecular Probes, 1:1,000). The following secondary antibodies were used (at 1:200 dilution): Alexa488-goat-anti-rabbit IgG, Alexa568-goat-anti-mouse IgG, and Cy5-conjugated anti-HRP (Molecular Probes).

Mechanosensory neuron analysis. Branching patterns of posterior scutellar mechanosensory neurons were visualized essentially as described previously[18,22]. In brief, head and abdominal segments were separated, and thoracic segments were fixed in 3.7% formaldehyde solution in 0.2 M sodium carbonate buffer (pH 9.5) overnight at 4 °C. DiI crystals were dissolved in 100% ethanol at 10 mg ml^{-1} concentration, and a paste of DiI was made by mixing this stock solution with 0.2 M sodium carbonate buffer (pH 9.5). Paste was injected into the socket of a posterior scutellar mechanosensory bristle in the scutellar segment of the thorax and DiI was allowed to transfer at room temperature for 3 days in a humidified chamber in 0.2 M sodium carbonate buffer with 0.05% sodium azide. Preparations were dissected further to isolate thoracic ganglia and imaged immediately using LSM510 confocal microscopy. All individual experiments were conducted with appropriate control genotypes for 'wild type' (that is, alleles with the *Dscam1* locus intact in different backgrounds). DiI injection and imaging were done blindly of the genotypes in all experiments involving deletion mutants. Each deletion was confirmed by PCR and sequencing. Only images with high signal to noise ratio containing one labelled posterior scutellar neuron were used for branching pattern analysis. All images of deletion mutants and wild-type controls were coded and scored blindly, independently by D.H. and Y.C., to avoid potential bias in scoring.

Statistical analysis of branching patterns of posterior scutellar neurons. The details of statistical analysis on branching patterns of posterior scutellar neurons are available on request. In brief, branching patterns were compared in a pairwise fashion between individual genotypes using Fisher's exact test for each branch segment. To test differences in the overall branching patterns between two genotypes, we used the minimum P value of the Fisher's exact test across each branch. The significance of these statistics is evaluated using permutation and results are reported as 'Permutation P values'. Both when looking at branch-specific and global tests, one must account for the large number of comparisons (that is, 16 branches and many genotype pairs). We suggest using a Bonferroni correction. In general, given the large number of tests carried out for this paper, a simplistic approximation would suggest that only tests with $P < 0.001$ should be considered significant. To perform cluster analysis, we use Partition Around Medoids[26], a partitioning algorithm. Similarity between neurons was evaluated using Gower's distance, which is well suited for the 0–1 values indicating the presence or absence of branches. Silhouette is a statistical measure of separation between clusters, which takes on values between 0 and 1. Silhouette values larger than 0.5 indicate reasonably reliable cluster structures.

Stochastic modelling. The details of stochastic modelling are available on request. In brief, a stochastic simulation algorithm was used to generate a table of probabilities ($Q_{x,i}$) that a pair of neurons share x or more isoforms when each neuron stochastically chooses 30, 50 or 100 *Dscam1* mRNAs from different numbers of potential isoforms (i). Then the probabilities that none of N pairs of neurons share x or more isoforms is represented by: $P_{n,x,i} = (1 - Q_{x,i})^N$. The number of neurons (n) contained in N pairs of neurons was calculated by solving: $N = n(n-1)/2$. The number of neurons that obtain unique identities at more than 95% likelihood was calculated by solving: $P_{n,x,i} \geq 0.95$.

25. Adams, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).
26. Kaufman, L. & Rousseeuw, P. J. *Finding Groups in Data: an Introduction to Cluster Analysis* (Wiley, 1990).

LETTERS

Direct reprogramming of human neural stem cells by *OCT4*

Jeong Beom Kim[1], Boris Greber[1], Marcos J. Araúzo-Bravo[1], Johann Meyer[2], Kook In Park[3], Holm Zaehres[1] & Hans R. Schöler[1]

Induced pluripotent stem (iPS) cells have been generated from mouse and human somatic cells by ectopic expression of four transcription factors (*OCT4* (also called *POU5F1*), *SOX2*, *c-Myc* and *KLF4*)[1–7]. We previously reported that *Oct4* alone is sufficient to reprogram directly adult mouse neural stem cells to iPS cells[8]. Here we report the generation of one-factor human iPS cells from human fetal neural stem cells (one-factor (1F) human NiPS cells) by ectopic expression of *OCT4* alone. One-factor human NiPS cells resemble human embryonic stem cells in global gene expression profiles, epigenetic status, as well as pluripotency *in vitro* and *in vivo*. These findings demonstrate that the transcription factor *OCT4* is sufficient to reprogram human neural stem cells to pluripotency. One-factor iPS cell generation will advance the field further towards understanding reprogramming and generating patient-specific pluripotent stem cells.

Induced pluripotent stem (iPS) cells have been generated from mouse and human somatic cells by overexpression of defined factors[1–7,9,10]. Mouse iPS cells are capable of differentiating into all three germ layers and have developmental potency of contributing to chimaeric mice and the germ line[2,3,9]. Human iPS cells are also very similar to human embryonic stem (ES) cells with regard to their differentiation potential *in vitro* and *in vivo*, which holds promise for research aimed at drug development and regenerative medicine to produce 'patient-specific iPS cells'. However, clinical applications of human iPS cells require avoiding viral transgenes, especially the oncogenes *c-Myc* and *KLF4*, from the host genome, as reactivation of the *c-myc* oncogene in mice frequently causes tumours in chimaeras and offspring derived from iPS cells[2]. It has become possible to generate iPS cells from various mouse and human somatic cells without the oncogenes *c-Myc* or *KLF4*[11–14]. Mouse neural stem cells (NSCs), which endogenously express *Sox2*, *Klf4* and *c-myc*, can be reprogrammed by two factors (*Oct4* and *Klf4*)[15] or one factor (*Oct4*)[8].

We continued to investigate human iPS cell generation from human NSCs to determine whether the single transcription factor *OCT4* alone, which is sufficient to reprogram mouse NSCs[8,15], can also reprogram human NSCs. We initially expanded human NSCs derived from human fetal brain tissue in serum-free NSC medium as previously described (Fig. 1a)[16,17]. Human NSCs were infected with retroviruses carrying human *OCT4* and *KLF4* (two factor, 2F) or *OCT4* (1F) and maintained in NSC medium for up to 7 days. Day 8 after infection, the cells were re-plated on to feeder layers in human ES cell medium containing 10 ng ml^{-1} bFGF and mouse embryonic fibroblast (MEF)-conditioned medium (CM) in a 1:1 ratio. Within 10–11 weeks after infection, we identified a colony resembling human ES cells in the *OCT4*-infected cultures. Notably, reprogrammed human NSCs did form a neural rosette in the centre of a colony (Fig. 1b). The colony grew larger, displaying typical human ES-cell-like morphology within another 5–6 days but the unreprogrammed neural rosettes still remained in the centre of the colony (Fig. 1c). Human NSCs may be reprogrammed to the pluripotent state through neural ectoderm. The neural rosettes were removed from the human ES-cell-like colony and the rest of the colony was mechanically divided and re-plated several times on to new feeder layers (passage 1 and 10, Fig. 1d, e).

We successfully established two clones out of three human ES-cell-like colonies by picking from *OCT4*-infected human NSCs (1F human NiPS clone A and C, reprogramming efficiency 0.004%), and three clones out of five human ES-cell-like colonies in *OCT4* and *KLF4*-infected human NSCs (2F human NiPS clones A, B and C, reprogramming efficiency 0.006%) within 7–8 weeks after infection. We next investigated the expression of pluripotency markers from non-transduced NSCs to work out the possibility that prolonged human ES cell culture conditions can induce pluripotency of human NSCs. These control human NSCs differentiated mostly into astrocytes and neurons (Supplementary Fig. 1a). In addition, polymerase chain reaction with reverse transcription (RT–PCR) analysis against pluripotency and neural markers demonstrated detection of neural markers, but absence of pluripotency marker expression (Supplementary Fig. 1b).

All iPS cell lines could be expanded in human ES cell culture conditions. The 1F human NiPS cells were morphologically similar to human ES cells and stained positive for alkaline phosphatase (Fig. 1e, f). Immunofluorescence staining confirmed that 2F and 1F human NiPS cells uniformly expressed human ES cell markers, including OCT4, SSEA4, TRA-1-60 and TRA-1-81 (Fig. 1g and Supplementary Fig. 2a). These results demonstrate that human iPS cells can be generated from human NSCs by *OCT4* and *KLF4* as well as by *OCT4* alone.

Next, we tested messenger RNA expression levels of pluripotency marker genes and known reprogramming factors in these iPS cells using quantitative RT–PCR (qRT–PCR). 2F and 1F human NiPS cells expressed human ES cell markers, including endogenous *OCT4* and *KLF4*, at similar levels as H9 and H1 human ES cells, whereas parental human NSCs showed endogenous expression of *SOX2* and *c-Myc* (Fig. 2a). Genotyping PCR demonstrated that 1F human NiPS cell clones only carried the *OCT4* transgene without contamination of transgenic *KLF4*, and 2F human NiPS cell clones contained the *OCT4* and *KLF4* transgenes in their genomes (Supplementary Fig. 3). We also confirmed that expression of transgenic *OCT4* or *KLF4* was significantly silenced in 2F and 1F human NiPS cell clones, except *OCT4* expression of 2F human NiPS cell clone B (Supplementary Fig. 4). Southern blot analysis confirmed three integrations of the *OCT4* transgene in 1F human NiPS cell clones and 2F human NiPS

[1]Max Planck Institute for Molecular Biomedicine, Department of Cell and Developmental Biology, Röntgenstrasse 20, 48149 Münster, NRW, Germany. [2]Hannover Medical School, Department of Experimental Hematology, Carl-Neuberg-Strasse 1, 30625 Hannover, Germany. [3]Department of Pediatrics, BK21 Project for Medical Sciences, Yonsei University College of Medicine, 250 Seongsanno, Seodaemoon-gu, Seoul 120-752, Korea.

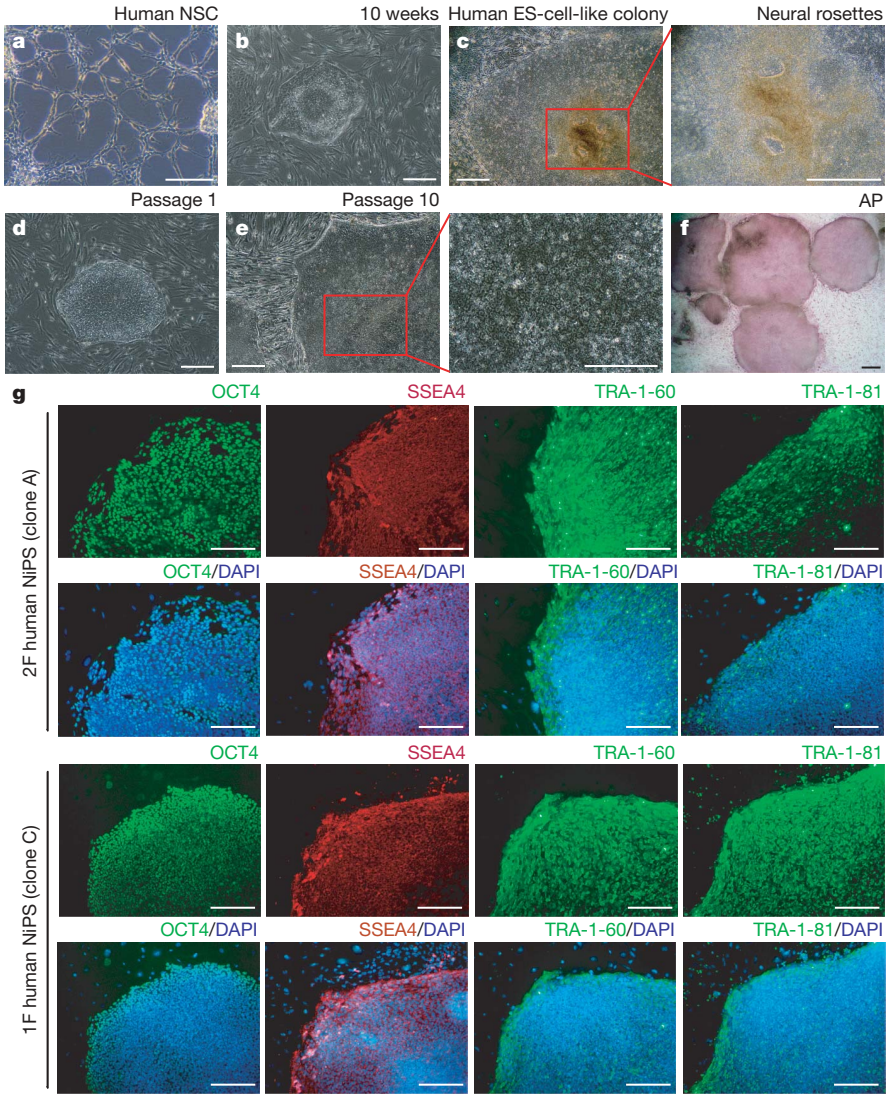


Figure 1 | 1F human NSC-derived iPS cell colony formation and cell line characterization. **a**, Morphology of human NSCs grown in NSC medium. **b**, Colony formation of human *OCT4*-infected cells 10 weeks after infection. **c**, The colony shows human ES-cell-like morphology but unreprogrammed neural rosettes remain in its centre. **d**, Typical human ES-cell-like iPS colony growing on a feeder layer after mechanical isolation at passage 1 (1F human NiPS clone C). **e**, High magnification of iPS colony at passage 10. **f**, 1F human NiPS colonies were stained for alkaline phosphatase (AP). Scale bars, 250 μm. **g**, Immunocytochemical analysis of pluripotency and surface markers (OCT4, SSEA4, TRA-1-60 and TRA-1-81) in 2F human NiPS (clone A) and 1F human NiPS (clone C) cells. Nuclei are stained with DAPI (blue). Scale bars, 250 μm.

cell clones (Supplementary Fig. 5). To exclude the possibility that iPS cell clones arose from human ES cell contamination in the laboratory, DNA fingerprinting analysis confirmed the genetic identity between human NiPS cells and donor human NSCs (Supplementary Table 1). To evaluate chromosomal abnormalities of human NiPS cells, we performed cytogenetic analysis by metaphase spread. 1F and 2F human NiPS cells maintained normal karyotype (46, XY) during the whole reprogramming period (Supplementary Fig. 6).

To confirm epigenetic remodelling in reprogrammed cells, we performed bisulphite sequencing analysis to determine the degree of DNA methylation of the *OCT4* and *NANOG* promoters (Fig. 2b). Similar to human ES cells, both promoter regions were demethylated in 2F and 1F human NiPS cells relative to the donor human NSCs. Taken together, a single transcription factor, *OCT4*, can reprogram human NSCs into iPS cells that are very similar to human ES cells at the molecular level.

Next, we tested *in vitro* pluripotency of 2F and 1F human NiPS cells by embryoid body (EB) differentiation and directed differentiation. During EB differentiation, human NiPS cells readily differentiated into vesicular structures with a variety of cell types through hanging drops including endoderm (AFP), mesoderm (α-SMA) and ectoderm (TUJ1) (Fig. 3a and Supplementary Fig. 2b). To enhance the differentiation towards all three germ layers, human NiPS cells were cultured by a directed differentiation protocol (see Methods). We then additionally confirmed the expression of markers for all three germ layers by qRT–PCR analysis (Fig. 3b).

To evaluate *in vivo* pluripotency of these human iPS cells, they were subcutaneously transplanted into severe combined immunodeficient (SCID) mice. Six-to-eight weeks after injection, 2F and 1F human NiPS cells gave rise to teratomas containing derivatives of all three germ layers, including respiratory epithelium, skeletal muscle, cartilage and neural epithelium (Fig. 4a and Supplementary Fig. 7). These results indicate that, like human ES cells, 2F and 1F human NiPS cells are pluripotent *in vitro* and *in vivo*.

Finally, we performed global gene expression analysis on human NSCs, 1F and 2F human NiPS cells derived from human NSCs, as well as H9 and H1 human ES cells using oligonucleotide microarrays. A heat map shows that the global 1F and 2F human NiPS cell expression profiles were overall similar to those in human ES cells, and different from parental human NSCs (Fig. 4b, left panel). Consistent with this, hierarchical clustering analysis showed that 1F and 2F human NiPS cells clustered closely with human ES cells and were distinct from

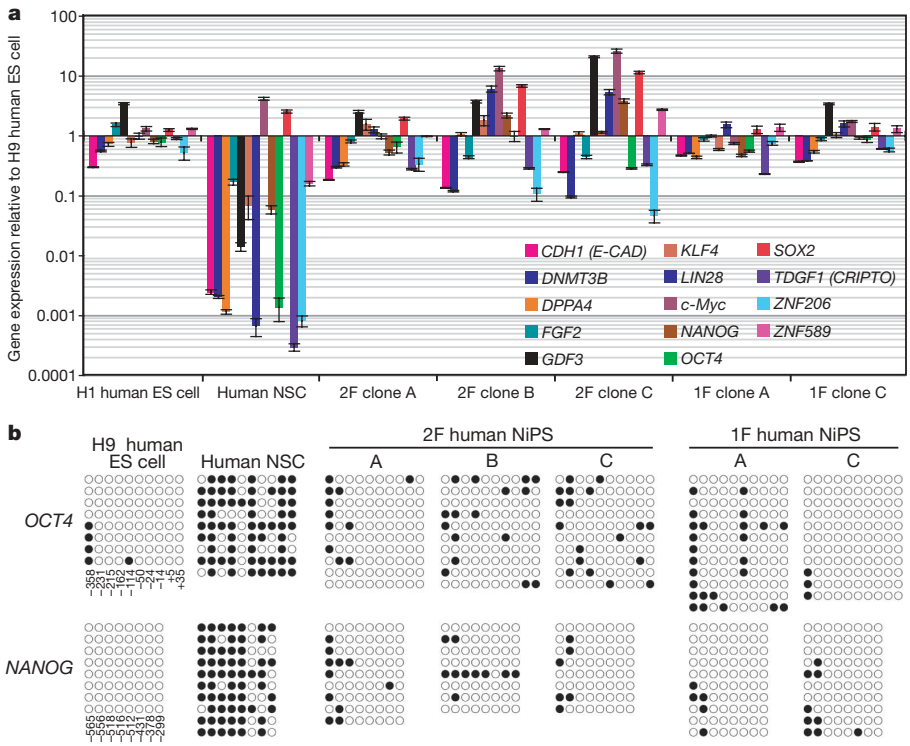


Figure 2 | Expression levels of pluripotent markers and DNA methylation analysis in human NSC-derived iPS cells. a, Quantitative PCR analysis for pluripotent markers in H1 human ES cells, human NSCs, 2F human NiPS clones (A, B and C) and 1F human NiPS clones (A and C). The expression levels are relative to H9 human ES cells using primers specific for endogenous transcripts, and on the logarithmic scale. Transcript levels were normalized to β-actin expression. Error bars indicate standard errors from duplicate samples ($n = 2$). **b**, Bisulphite sequencing analysis of *OCT4* and *NANOG* promoter regions in H9 human ES cells, human NSCs, 2F human NiPS clones (A, B and C) and 1F human NiPS clones (A and C). Each row of circles for a given amplicon represents the methylation status of CpG dinucleotides in one bacterial clone for that region. Open circles represent unmethylated CpGs; closed circles represent methylated CpGs. Bottom numbers of each column indicate CpG locations relative to the translational start site.

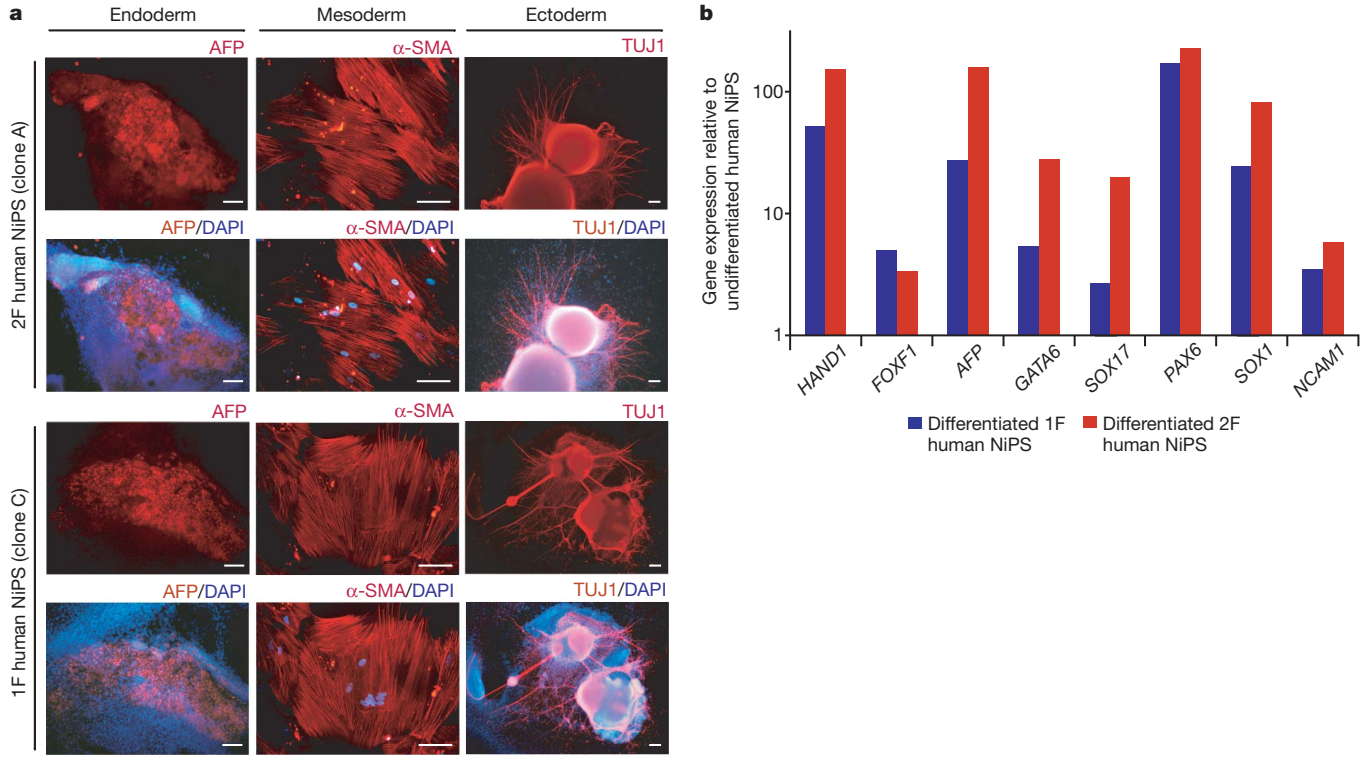


Figure 3 | *In vitro* differentiation of human NSC-derived iPS cells into all three germ layers. a, Immunofluorescence analysis of 2F human NiPS (clone A) and 1F human NiPS (clone C) derived cells of all three germ layers after EB differentiation: endoderm (α-fetoprotein; AFP), mesoderm (α-smooth muscle actin; α-SMA) and ectoderm (β-tublin lllb; TUJ1). Nuclei are stained with DAPI (blue). Scale bars, 100 μm. **b**, Quantitative PCR analyses of all three germ layer markers from differentiated 2F human NiPS (clone A) and 1F human NiPS (clone C) cells after directed differentiation: mesoderm (*HAND1*, *FOXF1*), endoderm (*AFP*, *GATA6*, *SOX17*) and ectoderm (*PAX6*, *SOX1*, *NCAM1*). Data denote β-actin-normalized fold changes relative to undifferentiated parental human NiPS cells.

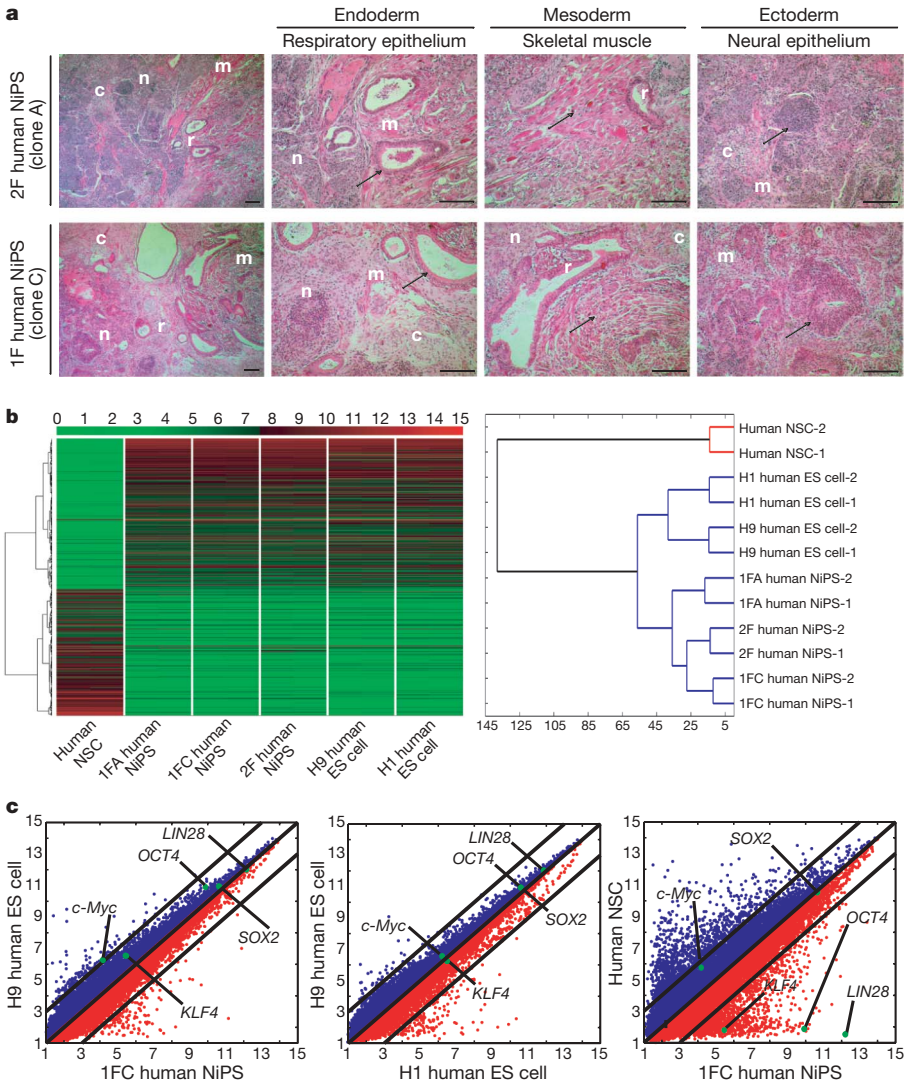


Figure 4 | *In vivo* pluripotency and global gene expression profile of human NSC-derived iPS cells. **a**, Teratoma formation after 6–8 weeks transplantation of 2F human NiPS (clone A) and 1F human NiPS (clone C) cells into SCID mice. Teratomas were sectioned and stained with haematoxylin and eosin at 6–8 weeks. Histological section of identified cells representing all three germ layers: endoderm (respiratory epithelium, r), mesoderm (skeletal muscle, m; cartilage, c) and ectoderm (neural epithelium, n). Enlargements of sections showing tissues are indicated by arrows. Scale bars, 100 μm. **b**, Heat map (left panel) and hierarchical cluster analysis (right panel) of global gene expression from human NSCs, 1F human NiPS (clone A and clone C), 2F human NiPS (clone A), H9 human ES cells and H1 human ES cells. A colour bar (top) indicates the colour code gene expression in $\log_2$ scale. The abscissa numbers in the hierarchical cluster correspond with the standardized Euclidean distance. **c**, Scatter plots comparing global gene expression profiles between H9 human ES cells and 1F human NiPS cells (clone C) (left panel), H9 human ES cells and H1 human ES cells (middle panel), and human NSCs and 1F human NiPS cells (clone C) (right panel). The black lines indicate twofold differences in gene expression levels between the paired cell populations. The transcript expression levels are on the $\log_2$ scale.

parental human NSCs (Fig. 4b, right panel). Scatter plot analysis demonstrated that the global expression differences between 1F human NiPS cells (clone C) and human ES cells were in the same range as those between two different human ES cell lines (H9 versus H1) (Fig. 4c). 1F (clone A) and 2F (clone A) human NiPS cells also showed high similarity with H9 and H1 human ES cells (Supplementary Fig. 8). Hence, global gene expression profiles revealed a high degree of similarity between 1F and 2F human NiPS cells and human ES cells.

Our study demonstrates the generation of one-factor (*OCT4*) and two-factor (*OCT4* and *KLF4*) human iPS cells that resemble human ES cells at the molecular level and with respect to their differentiation potential. The present study extends our findings on the feasibility of reprogramming from mouse to human NSCs with OCT4 alone. It demonstrates that endogenous expression of the reprogramming factors can complement exogenously added factors[8,15]. The feasibility to reprogram directly NSCs by *OCT4* alone might reflect their higher similarity in transcriptional profiles to ES cells than to other stem cells like haematopoietic stem cells or to their differentiated counterparts[18].

So far, small molecule approaches can be used to compensate for viral transduction of some of the four factors when the other factors are transduced[19]. Direct delivery of recombinant proteins also facilitated generation of human iPS cells[20]. Future studies will show if direct reprogramming is possible with small molecules or OCT4 recombinant protein alone. Knowing that specifically *Oct4* needs to be activated in NSCs might facilitate such an approach. It will be interesting to extend this study to human NSCs derived from other sources, such as dental pulp[21,22], as well as to other stem-cell types.

METHODS SUMMARY

Human NSCs were derived from fetal tissue and transduced with retroviral vectors expressing human OCT4 and OCT4 plus KLF4 as previously described[8,15,16,23]. The cells were re-plated on mouse feeder layers under human ES cell culture conditions until human ES-cell-like colonies could be mechanically isolated and further cultivated. Human iPS cells were characterized by

immunostaining, karyotyping, qRT–PCR, gene expression profiling by microarray, bisulphite sequencing, Southern blot analysis and short tandem repeat analysis. *In vitro* differentiation by EB formation and teratoma formation in SCID mice was used to measure the developmental potential of the human iPS cells.

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

Received 2 April; accepted 21 August 2009.
Published online 28 August 2009.

1. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126,** 663–676 (2006).
2. Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448,** 313–317 (2007).
3. Wernig, M. *et al. In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448,** 318–324 (2007).
4. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131,** 861–872 (2007).
5. Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318,** 1917–1920 (2007).
6. Park, I. H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451,** 141–146 (2008).
7. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl Acad. Sci. USA* **105,** 2883–2888 (2008).
8. Kim, J. B. *et al.* Oct4-induced pluripotency in adult neural stem cells. *Cell* **136,** 411–419 (2009).
9. Maherali, N. *et al.* Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell* **1,** 55–70 (2007).
10. Dimos, J. T. *et al.* Induced pluripotent stem cells generated from patients with ALS can be differentiated into motor neurons. *Science* **321,** 1218–1221 (2008).
11. Nakagawa, M. *et al.* Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotechnol.* **26,** 101–106 (2008).
12. Aoi, T. *et al.* Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science* **321,** 699–702 (2008).
13. Wernig, M., Meissner, A., Cassady, J. P. & Jaenisch, R. c-Myc is dispensable for direct reprogramming of mouse fibroblasts. *Cell Stem Cell* **2,** 10–12 (2008).
14. Huangfu, D. *et al.* Induction of pluripotent stem cells from primary human fibroblasts with only *Oct4* and *Sox2*. *Nature Biotechnol.* **26,** 1269–1275 (2008).
15. Kim, J. B. *et al.* Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature* **454,** 646–650 (2008).
16. Park, K. I., Teng, Y. D. & Snyder, E. Y. The injured brain interacts reciprocally with neural stem cells supported by scaffolds to reconstitute lost tissue. *Nature Biotechnol.* **20,** 1111–1117 (2002).
17. Kim, H. T. *et al.* Human neurospheres derived from the fetal central nervous system are regionally and temporally specified but are not committed. *Exp. Neurol.* **199,** 222–235 (2006).
18. Ramalho-Santos, M., Yoon, S., Matsuzaki, Y., Mulligan, R. C. & Melton, D. A. ''Stemness'': transcriptional profiling of embryonic and adult stem cells. *Science* **298,** 597–600 (2002).
19. Shi, Y. *et al.* A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell* **2,** 525–528 (2008).
20. Kim, D. *et al.* Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell* **4,** 472–476 (2009).
21. Seo, B. M. *et al.* Investigation of multipotent postnatal stem cells from human periodontal ligament. *Lancet* **364,** 149–155 (2004).
22. Widera, D. *et al.* Adult palatum as a novel source of neural-crest related stem cells. *Stem Cells* **27,** 1899–1910 (2009).
23. Zaehres, H. & Daley, G. Q. Transgene expression and RNA interference in embryonic stem cells. *Methods Enzymol.* **420,** 49–64 (2006).

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

Acknowledgements We thank B. Schäfer for histology, T. Kitamura for the pMX retroviral vector and A. Malapetsas and S. Kölsch for editing. This work has been supported, in part, by a grant from the German Federal Ministry of Education and Research BMBF 01GN0811 and in part by a grant by the Deutsche Forschungsgemeinschaft DFG SCHO 340/4-1.

Author Contributions J.B.K., project design, generation and characterization of iPS cells, preparation of manuscript; B.G., M.J.A.-B., J.M., characterization of iPS cells; K.I.P., providing human NSCs; H.Z., project design, generation of iPS cells, preparation of manuscript; H.R.S., project design, preparation of manuscript.

Author Information The microarray data sets are available from GEO (Gene Expression Omnibus) under accession number GSE15355. Human NiPS cells are in the process of being deposited in the UK Stem Cell Bank. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to H.R.S. (office@mpi-muenster.mpg.de).

METHODS

Cell culture. Human fetal tissue from cadavers at 13 weeks of gestation was obtained with full patient consent and the approval of the research ethics committee of Yonsei University College of Medicine, Seoul, Korea. The methods of acquisition conformed to NIH and Korean Government guidelines. Human NSCs were derived from the telencephalon (HFT13) as previously described[17]. Briefly, telencephalon tissue was freshly dissected, dissociated in 0.1% trypsin for 30 min and seeded into 10 cm plates at a density of 200,000 cells ml^{-1} in NSC medium. These cells were cultured in NSC medium as previously described[16,17]. Human ES and iPS cells were maintained on mitomycin-C-treated CF1 mouse feeder layers (Millipore) in human ES cell medium, which contains knockout DMEM (Invitrogen) supplemented with 20% knockout serum replacement (Invitrogen), 1 mM L-glutamine, 1% non-essential amino acids, 0.1 mM β-mercaptoethanol, penicillin/streptomycin and 10 ng ml^{-1} human basic fibroblast growth factor (bFGF) (Invitrogen) as previously described[24].

Induction of 1F and 2F human NiPS cells. The pMX-based retroviral vectors encoding the human cDNAs of *OCT4* and *KLF4* were co-transfected with packaging-defective helper plasmids into 293 cells using Fugene transfection reagent (Roche) to produce vesicular stomatitis virus (VSV) G protein pseudotyped virus as previously described[4,23]. Viral supernatants were collected and concentrated by ultracentrifugation 48 h after transfection to infect human NSCs. For generation of iPS cells, human NSCs were seeded at a density of 5×10^4 cells per 6-well plate and incubated with virus-containing supernatants for *OCT4* or *OCT4* and *KLF4* supplemented with 6 μg ml^{-1} protamine sulphate (Sigma) for 24 h. The next day, the medium was replaced with fresh NSC medium at day 1 after infection and maintained up to day 7 after infection. Cells were further cultured in human ES cell medium from day 8 after infection. The iPS colonies were mechanically isolated at 2 months or 2.5 months after infection and were subsequently re-plated and maintained on CF1 mouse feeder layers (Millipore) in human ES cell medium.

qRT–PCR. Total RNA was isolated from bulk cell culture samples or hand-picked undifferentiated colonies using RNeasy columns (Qiagen) with on-column DNA digestion. cDNA was produced using oligo-dT_{15} priming and M-MLV reverse transcriptase (USB) according to the manufacturer's instructions at 42 °C for 1 h. About 50 ng of total RNA equivalent was typically used as template in 20 μl SYBR green PCR reactions (40 cycles of 15 s, 95 °C/60 s, 60 °C on Applied Biosystems 7300 instrumentation) that additionally contained 0.375 μM of each primer and 10 μl of SYBR green PCR mix (ABI). All primers used were confirmed to amplify the predicted product at close-to-optimal efficiency without side products. Primer sequences are given in Supplementary Table 2. Relative expression levels were calculated using the comparative C_t method, based on biological control samples and two housekeeping genes for normalization. Error bars reflect standard errors arising from biological replicates (marker gene expression data) or from using independent housekeeping genes for normalization (transgene silencing data).

Global gene expression analysis. For transcriptome profiling, 400 ng of total DNA-free RNA was used as input for labelled cRNA synthesis (Illumina TotalPrep RNA Amplification Kit, Ambion) following the manufacturer's instructions (IVT: 10 h). Quality-checked cRNA samples were hybridized as biological or technical duplicates for 18 h onto HumanRef-8 v3 expression BeadChips (Illumina), washed, stained and scanned following guidelines using materials/instrumentation supplied/suggested by the manufacturer. The microarray data are available from the GEO (Gene Expression Omnibus) website under accession number GSE15355.

Microarray data processing. The bead intensities were mapped to gene information using BeadStudio 3.2 (Illumina), background correction was performed using the Affymetrix Robust Multi-array Analysis (RMA) background correction model, variance stabilization was performed using the $\log_2$ scaling, and gene expression normalization was calculated with the quantile method implemented in the lumi package of R-Bioconductor. Data post-processing and graphics were performed with in-house developed functions in Matlab. Hierarchical clusters of genes and samples were performed with a standardized Euclidean metric and the Ward's linkage method.

Bisulphite sequencing. Genomic DNA was isolated from bulk cell culture samples or hand-picked undifferentiated colonies using DNeasy columns (Qiagen). A total of 300 ng was used as input for bisulphite conversion (EpiTect Bisulfite Kit, Qiagen). A total of 50 ng of converted DNA was used as a template for conventional nested PCRs amplifying 467- and 336-bp regions of the *OCT4* and *NANOG* promoters, respectively. Primers were specific for conversion of the sense DNA strand and are given in Supplementary Table 2. Purified PCRs were TA-cloned into pCR2.1-TOPO (Invitrogen). Insert sequences of randomly picked clones were analysed using the BiQ Analyser program, following its quality-check-based suggestions to drop individual clones if appropriate. Data from one CpG site at position +20 relative to the *OCT4* translation start codon is not shown as it was uninformative.

Teratoma formation. Human NiPS cells and human NSCs ($3–5 \times 10^6$ cells per mouse) were injected subcutaneously into the dorsal flank of SCID mice. Teratomas were fixed in 4% PFA overnight and embedded in paraffin after 6–8 weeks injection. Sections were stained with haematoxylin and eosin dyes.

Alkaline phosphatase and immunofluorescence staining. Alkaline phosphatase staining was performed with the ES Cell Characterization Kit (Chemicon) according to the manufacturer's protocol. Immunofluorescence staining was performed using the following primary antibodies: AFP (Sigma, 1:100), α-SMA (Sigma, 1:50), TUJ1 (Chemicon, 1:500), OCT4 (Santa Cruz, 1:200), SSEA4 (Chemicon, 1:200), TRA-1-60 (Chemicon, 1:200) and TRA-1-81 (Chemicon, 1:200).

***In vitro* differentiation of human iPS cells.** For immunocytochemistry, embryoid bodies (EBs) were generated from iPS cells with the hanging-drop method in MEF-conditioned medium. After 5 days, EBs were transferred to gelatin-coated plates and subsequently cultured for another 14 days in knockout DMEM (Invitrogen) supplemented with 20% FBS, 1 mM L-glutamine, 1% non-essential amino acids, 0.1 mM β-mercaptoethanol, and penicillin/streptomycin. For qRT–PCR, iPS colonies were mechanically isolated and re-plated on Matrigel-coated plates in MEF-conditioned medium. After 2 days, medium was replaced with medium for each of the three germ layers. Endodermal differentiation: RPMI1640 medium supplemented with 2% FBS, 100 ng ml^{-1} activin A (R&D Systems), L-glutamine and penicillin/streptomycin for 3 weeks[14]. For mesodermal differentiation: knockout DMEM supplemented with 100 μM ascorbic acid (Sigma), 20% FBS, 1 mM L-glutamine, 1% non-essential amino acids, 0.1 mM β-mercaptoethanol and penicillin/streptomycin for 3 weeks[25]. For ectodermal differentiation: the cells were maintained in N2B27 medium for 7 days and the medium replaced with N2 medium supplemented with 10 ng ml^{-1} bFGF2 (peprotech), 100 ng ml^{-1} Sonic Hedgehog (R&D Systems), 10 ng ml^{-1} PDFG (R&D Systems), L-glutamine and penicillin/streptomycin for 2 weeks. The medium was changed every other day. Primer sequences are given in Supplementary Table 2.

24. Thomson, J. A. *et al.* Embryonic stem cell lines derived from human blastocysts. *Science* **282**, 1145–1147 (1998).
25. Aasen, T. *et al.* Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nature Biotechnol.* **26**, 1276–1284 (2008).

LETTERS

Discovery of Atg5/Atg7-independent alternative macroautophagy

Yuya Nishida[1,2]*, Satoko Arakawa[1]*, Kenji Fujitani[1], Hirofumi Yamaguchi[1], Takeshi Mizuta[1], Toku Kanaseki[1], Masaaki Komatsu[4], Kinya Otsu[3], Yoshihide Tsujimoto[2] & Shigeomi Shimizu[1]

Macroautophagy is a process that leads to the bulk degradation of subcellular constituents by producing autophagosomes/autolysosomes[1–3]. It is believed that *Atg5* (ref. 4) and *Atg7* (ref. 5) are essential genes for mammalian macroautophagy. Here we show, however, that mouse cells lacking Atg5 or Atg7 can still form autophagosomes/autolysosomes and perform autophagy-mediated protein degradation when subjected to certain stressors. Although lipidation of the microtubule-associated protein light chain 3 (LC3, also known as Map1lc3a) to form LC3-II is generally considered to be a good indicator of macroautophagy[6], it did not occur during the Atg5/Atg7-independent alternative process of macroautophagy. We also found that this alternative process of macroautophagy was regulated by several autophagic proteins, including Unc-51-like kinase 1 (Ulk1) and beclin 1. Unlike conventional macroautophagy, autophagosomes seemed to be generated in a Rab9-dependent manner by the fusion of isolation membranes with vesicles derived from the *trans*-Golgi and late endosomes. *In vivo*, Atg5-independent alternative macroautophagy was detected in several embryonic tissues. It also had a function in clearing mitochondria during erythroid maturation. These results indicate that mammalian macroautophagy can occur through at least two different pathways: an Atg5/Atg7-dependent conventional pathway and an Atg5/Atg7-independent alternative pathway.

The biological functions of macroautophagy have been examined by generating mice lacking several *Atg* genes, including *Atg5* (ref. 4), beclin 1 (*Becn1*)[7,8] and *Atg7* (ref. 5), considered essential for macroautophagy; constitutive and starvation-induced autophagy is impaired in such mice. However, because $Atg5^{-/-}$ or $Atg7^{-/-}$ mice remain healthy until the perinatal period[4,5], an alternative mechanism of bulk protein degradation may compensate for the lack of Atg5/Atg7-dependent macroautophagy.

To investigate possible alternative protein degradation mechanisms in $Atg5^{-/-}$ mice, we obtained embryonic fibroblasts from *Atg5* knockout mice ($Atg5^{-/-}$ MEFs) and their control littermates (wild-type (WT) MEFs) (Supplementary Fig. 2a). Treatment with rapamycin induced macroautophagy in WT MEFs but not in $Atg5^{-/-}$ MEFs (Supplementary Fig. 2b), as reported previously[9]. We investigated autophagy induced by cytotoxic stressors (such as etoposide). After exposure to etoposide, each cell detached and underwent apoptosis in the same proportion (Supplementary Fig. 3). To exclude apoptotic consequences, only attached MEFs were examined. Etoposide resulted in the formation of numerous autophagic vacuoles in WT MEFs (Supplementary Fig. 2c), and an equivalent size and number of autophagic vacuoles also appeared in $Atg5^{-/-}$ MEFs (Fig. 1a (electron microscopy, EM), b and Supplementary Fig. 2c). The autophagic area in each cell increased equivalently over time during exposure to etoposide in WT and $Atg5^{-/-}$ MEFs (Fig. 1c). Typical autophagic structures, including double-membrane structures, autophagosomes, amphisomes[10,11] and autolysosomes, were observed in etoposide-treated $Atg5^{-/-}$ MEFs (Fig. 1d and Supplementary Fig. 4a). The presence of multilamellar bodies in autolysosomes indicated autophagic degradation of subcellular constituents (Fig. 1d). These autolysosomal structures were merged with the immunofluorescence dots of Lamp2, a lysosomal protein (Fig. 1a and Supplementary Fig. 5). There were numerous double-membrane structures and few autolysosomes in $Atg5^{-/-}$ MEFs after treatment with etoposide for 12 h; the situation was reversed at 18 h (Supplementary Fig. 4b), indicating that the autophagic process progressed over time after exposure to etoposide. Similar structures were observed when cells were examined by electron microscopy after quick freezing and freeze-substitution, a technique that achieves superior preservation of cellular architecture (Fig. 1e and Supplementary Fig. 4a). Similar results were observed when apoptosis was inhibited by the pan-caspase inhibitor zVAD-fmk (see later). Furthermore, similar macroautophagy was observed in etoposide-treated $Atg5^{-/-}$ MEFs from a different $Atg5^{-/-}$ mouse (Supplementary Fig. 6a, b) and in staurosporine-treated $Atg5^{-/-}$ MEFs and $Atg5^{-/-}$ thymocytes (Supplementary Fig. 6c, d). This suggests that cells possess the Atg5-independent macroautophagy system, designated 'alternative macroautophagy'.

To confirm the induction of alternative macroautophagy in etoposide-treated $Atg5^{-/-}$ MEFs, we examined the effect of bafilomycin A1, which prevents the fusion of autophagosomes with lysosomes[12]. Exposure to bafilomycin A1 was expected to increase and decrease the number of autophagosomes and autolysosomes, respectively. We obtained the expected results in WT MEFs and $Atg5^{-/-}$ MEFs, supporting the notion of Atg5-independent induction of macroautophagy (Fig. 1f and Supplementary Fig. 7a). Moreover, 3-methyladenine, a phosphatidylinositol-3-OH kinase (PI(3)K) inhibitor that blocks Atg5-dependent autophagosome formation[13], significantly suppressed autophagosome formation in etoposide-treated $Atg5^{-/-}$ MEFs (Fig. 1g and Supplementary Fig. 7b), suggesting the involvement of PI(3)K in Atg5-independent macroautophagy.

We next assessed whether and to what extent protein degradation occurs in etoposide-treated $Atg5^{-/-}$ MEFs. Etoposide induced protein degradation and had a similar effect on WT and $Atg5^{-/-}$ MEFs (Fig. 1h), which is consistent with the morphological analysis (Fig. 1c). Suppression of protein degradation by bafilomycin A1 and a lysosomal protease inhibitor cocktail suggested the involvement of lysosomal proteases (Fig. 1i). Although lysosomal proteases are also involved in macroautophagy-independent proteolysis, including chaperone-mediated autophagy[14], alternative macroautophagy may possibly have functioned in etoposide-induced proteolysis, because protein degradation was suppressed by 3-methyladenine (Fig. 1i),

[1]Department of Pathological Cell Biology, Medical Research Institute, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan. [2]Department of Medical Genetics, [3]Department of Cardiovascular Medicine, Osaka University Medical School, 2-2 Yamadaoka, Suita, Osaka 565-0871, Japan. [4]Laboratory of Frontier Science, Tokyo Metropolitan Institute of Medical Science, Bunkyo-ku, Tokyo 113-8613, Japan.
*These authors contributed equally to this work.

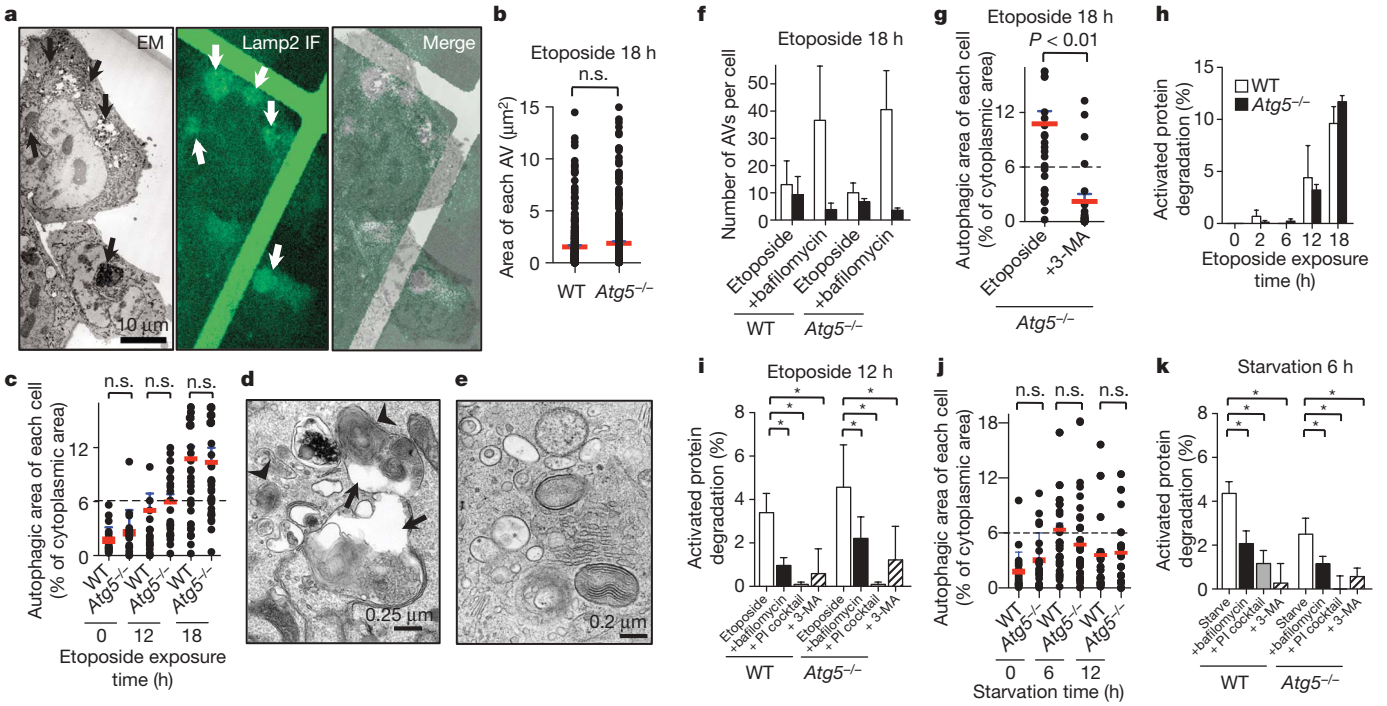


Figure 1 | Induction of macroautophagy in *Atg5*$^{-/-}$ MEFs by etoposide and starvation. **a**, Etoposide-treated *Atg5*$^{-/-}$ MEFs were assessed by EM (left) and by Lamp2 immunofluorescence (IF; middle). Right: reverse Lamp2 immunofluorescence image merged with the electron microscopy image. Arrows indicate Lamp2-positive autolysosomes. Magnified photos are provided in Supplementary Fig. 5. **b**, The size of each autophagic vacuole (AV) ($n = 30$ cells) in MEFs treated with etoposide. The total number of AVs in 30 cells was 483 in WT cells and 494 in *Atg5*$^{-/-}$ cells. **c**, Percentage autophagic area ($n \geq 20$ cells) in MEFs treated with etoposide. **d**, **e**, Representative macroautophagy in etoposide-treated *Atg5*$^{-/-}$ MEFs. Conventional (**d**) and quick freezing and freeze-substitution (**e**) techniques were used. The autolysosomes (arrows) contain multilamellar bodies (arrowheads). **f**, The number of autophagosomes (white columns) and autolysosomes (black columns) in MEFs ($n = 10$ cells) treated with etoposide. **g**, Reduction of percentage autophagic area by 3-methyladenine (3-MA; $n = 22$ cells). **h**, Time-course analysis of long-lived protein degradation ($n = 4$). **i**, Inhibition of long-lived protein degradation by bafilomycin A1, protease inhibitor (PI) cocktail and 3-MA (*$P < 0.05$; $n = 4$). **j**, **k**, Similar experiments to those in **c** and **i** were performed by starvation. Red and blue lines in **b**, **c**, **g** and **j** indicate means and s.e.m., respectively. The dashed line in **c**, **g** and **j** indicates the autophagic cell border (6%). Black error bars indicate s.d.

which blocked macroautophagy (Fig. 1g) but not chaperone-mediated autophagy[14]. zVAD-fmk was used in these experiments to avoid the influence of apoptosis-related proteolysis; its addition did not influence etoposide-induced macroautophagy (Supplementary Fig. 8).

We investigated starvation-induced Atg5-independent macroautophagy. Although a lack of autophagosomes in *Atg5*$^{-/-}$ cells has been reported for embryonic stem cells[15], we found mature autophagic vacuoles in starved *Atg5*$^{-/-}$ MEFs (Supplementary Fig. 9a). The extent of this macroautophagy was slightly less than that in starved WT MEFs (Fig. 1j, k and Supplementary Fig. 9b, c). Protein degradation in starved *Atg5*$^{-/-}$ MEFs was inhibited by bafilomycin A1, protease inhibitor cocktail and 3-methyladenine (Fig. 1k), indicating that starvation-induced proteolysis was also mediated by alternative macroautophagy in *Atg5*$^{-/-}$ MEFs. Mammalian macroautophagy can therefore occur through two pathways—Atg5-dependent conventional and Atg5-independent alternative pathways—possibly activated by different stimuli.

We examined whether autophagic modification of LC3 occurs during Atg5-independent alternative macroautophagy. LC3-II, a lipid conjugate form, and a punctate fluorescence pattern of green fluorescent protein (GFP)-tagged LC3 were observed in WT MEFs but not in *Atg5*$^{-/-}$ MEFs during treatment with etoposide (Fig. 2a–c), suggesting an absence of LC3 modification in alternative macroautophagy. This was confirmed by observing GFP–LC3 fluorescence followed by electron microscopy analysis in the same *Atg5*$^{-/-}$ MEFs. Autophagic vacuoles were generated in etoposide-treated *Atg5*$^{-/-}$ MEFs (Fig. 2d, EM); GFP–LC3 was distributed homogeneously throughout these cells (Fig. 2d, GFP–LC3). Mammalian cells have two other yeast Atg8 homologues, Gabarap and Gabarapl2 (ref. 16). Examination with yellow fluorescent protein (YFP)-tagged Gabarap and YFP–Gabarapl2 yielded results similar to those for GFP–LC3 (Supplementary Fig. 10). WT and *Atg5*$^{-/-}$ MEFs showed similar macroautophagy levels after exposure to etoposide (Fig. 1c), even though both macroautophagy types were activated in WT MEFs (Supplementary Fig. 11) and only the alternative form was activated in *Atg5*$^{-/-}$ MEFs; however, this may have resulted from the mutual regulation between conventional and alternative macroautophagy.

Immunostaining was performed for Lamp2 to assess the extent of alternative macroautophagy in etoposide-treated *Atg5*$^{-/-}$ MEFs, because the fluorescence pattern changed from diffuse to punctate during autolysosome generation (Fig. 2e). The following evidence supports the validity of this assay: first, the Lamp2 fluorescent dots coincided with autolysosomes (Fig. 1a); second, the number of *Atg5*$^{-/-}$ MEFs with fluorescent dots matched that of autophagic cells assessed by electron microscopy (Supplementary Fig. 12a); and third, bafilomycin A1 decreased the number of cells with dots (Supplementary Fig. 12b). We searched for molecules involved in alternative macroautophagy by using the Lamp2 immunostaining assay. Because many DNA damage-induced events are under transcriptional regulation, we compared the gene expression profiles of healthy and etoposide-treated *Atg5*$^{-/-}$ MEFs. The autophagic molecule Ulk1 (ref. 17) was upregulated in etoposide-treated *Atg5*$^{-/-}$ MEFs (Fig. 2f) and confirmed by quantitative PCR with reverse transcription (qRT–PCR) (Fig. 2g) and western blotting (Fig. 2h). Furthermore, silencing both Ulk1 and Fip200 (also known as Rb1cc1)—a component of the Ulk1 kinase complex[18]—decreased the number of autophagic cells (Fig. 2i, j and Supplementary Fig. 13), indicating that the Ulk1 complex is important in alternative macroautophagy.

Because Ulk1 is involved in alternative macroautophagy, some conventional autophagy-related molecules might function in the alternative process. Conventional macroautophagy machinery includes the Ulk1 complex[18], the PI(3)K complex, the Atg9 systems

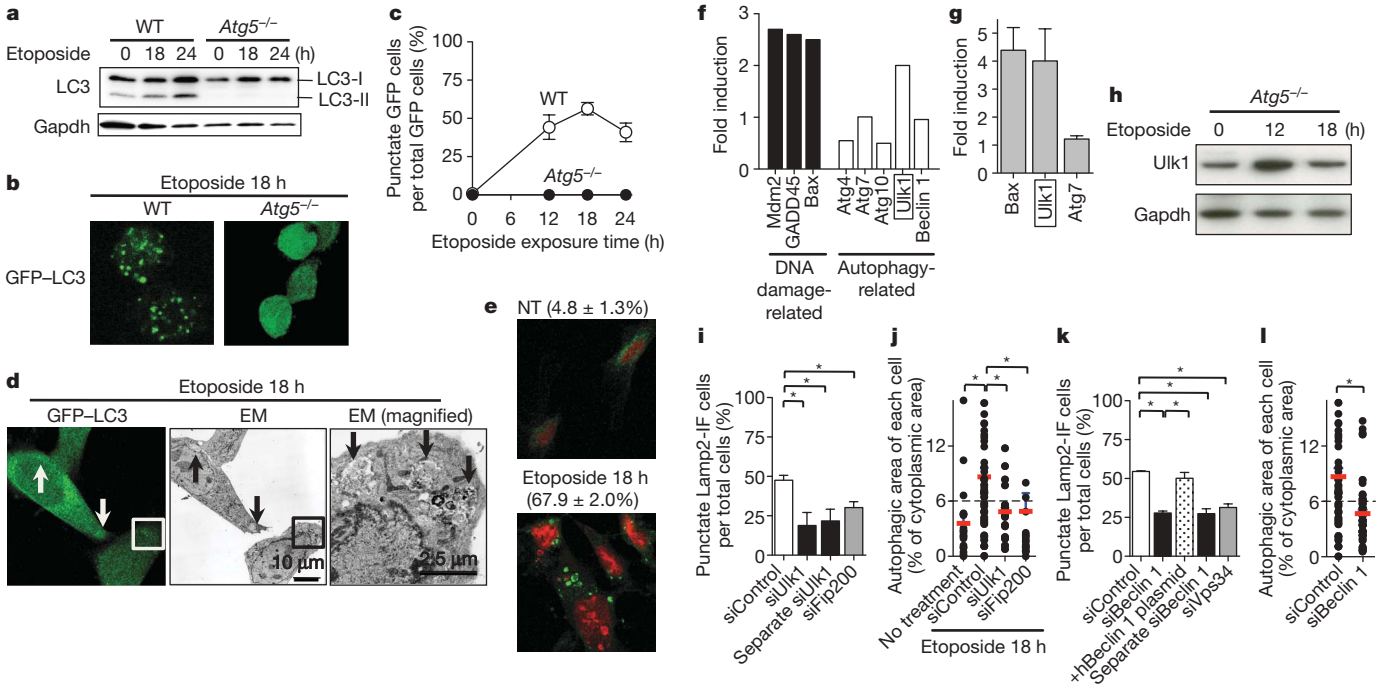


Figure 2 | Involvement of the Ulk1 and PI(3)K complexes in etoposide-induced alternative macroautophagy. **a**, The production of LC3-II. Gapdh was used as a loading control. **b**, **c**, Punctate GFP–LC3 fluorescence in WT MEFs but not in *Atg5*$^{-/-}$ MEFs. **b**, Representative photographs. **c**, Graphical presentation of the percentages of WT MEFs and *Atg5*$^{-/-}$ MEFs with punctate GFP–LC3 fluorescence. **d**, The same GFP–LC3-transfected *Atg5*$^{-/-}$ MEFs were assessed by IF and by EM. The magnified photo is the area indicated by the squares. Arrows indicate AVs. **e**, *Atg5*$^{-/-}$ MEFs were incubated without (NT) or with etoposide and immunostained with an anti-Lamp2 antibody (green). Nuclei were counterstained (red). The percentages of cells with punctate Lamp2 immunostaining are shown (means ± s.d., $n = 4$). **f**, Microarray analysis of the *Atg5*$^{-/-}$ MEFs during treatment with etoposide. **g**, **h**, Induction of Ulk1 in the etoposide-treated *Atg5*$^{-/-}$ MEFs was assessed by qPCR (**g**) and by immunoblotting (**h**). **i**–**l**, *Atg5*$^{-/-}$ MEFs were transfected with the siRNAs related to the Ulk1 complex (**i**, **j**) and the PI(3)K complex (**k**, **l**) and a beclin 1 plasmid (**k**) for 24 h and treated with etoposide for 18 h. Induction of macroautophagy was assessed by Lamp2 IF (**i**, **k**) and by EM (**j**, **l**). In **i** and **k**, $^*P < 0.05$ ($n = 4$); in **j** and **l**, $^*P < 0.01$ ($n = 35$ cells each). hBeclin 1, human beclin 1. Red and blue lines in **j** and **l** indicate means and s.e.m., respectively. The dashed line in **j** and **l** indicates the autophagic cell border. Black error bars indicate s.d. ($n = 4$). Original magnification (**b**, **e**), × 200.

and two Atg7-mediated ubiquitin-like protein systems (Atg8 and Atg12 is conjugated with phosphatidylethanolamine and Atg5, respectively, and Atg16 interacts with Atg5)[2]. When *Atg7*$^{-/-}$ MEFs were treated with etoposide, macroautophagy was induced without generating LC3-II (Supplementary Fig. 14). Moreover, silencing of Atg7, Atg12 or Atg16 did not suppress alternative macroautophagy in *Atg5*$^{-/-}$ MEFs (Supplementary Fig. 15). Together with the data on Atg5 and Atg8 homologues, these experiments indicated that the ubiquitin-like protein systems are not required for alternative macroautophagy. In addition, silencing of Atg9 did not show any influence (Supplementary Fig. 15). In contrast, silencing of beclin 1 decreased the number of autophagic cells (Fig. 2k, l and Supplementary Fig. 16). The involvement of beclin 1 was further confirmed by the restoration of macroautophagy after overexpression of human beclin 1 in beclin 1-silenced *Atg5*$^{-/-}$ MEFs (Fig. 2k). Silencing of Vps34, another component of the PI(3)K complex, decreased the number of autophagic cells (Fig. 2k and Supplementary Fig. 16), indicating that the complex is involved in etoposide-induced alternative macroautophagy; this is consistent with the data for 3-methyladenine (Fig. 1g, i). Ulk1, Fip200, beclin 1 and Vps34 (but not Atg7, Atg12, Atg16 or Atg9) were involved in alternative macroautophagy induced by starvation (Supplementary Fig. 17a) and staurosporine (Supplementary Fig. 17b).

The following observations were also noted: first, almost all autophagic vacuoles were localized near the Golgi apparatus (Fig. 3a, b); second, some isolation membranes extended from the Golgi membranes (Fig. 3a, inset); third, closure of some isolation membranes occurred by fusion with vesicles with thicker membranes[19], presumably derived from the *trans*-Golgi (*trans*-Golgi cisternae and *trans*-Golgi network) (Fig. 3c, d); and fourth, inhibition of Golgi-derived membrane by brefeldin A resulted in the inhibition of alternative but not conventional macroautophagy (Supplementary Fig. 18). We therefore examined the involvement of the *trans*-Golgi or endosomes in the extension and closure of isolation membranes. First, we introduced genes encoding GFP-tagged organelle marker proteins into *Atg5*$^{-/-}$ MEFs and analysed the co-localization with Lamp2-positive vacuoles. Etoposide induced the redistribution of a fraction of mannose 6-phosphate receptors (a marker for the *trans*-Golgi and late endosomes[20]); it was co-localized with Lamp2-positive autolysosomes (Fig. 3e and Supplementary Fig. 19a). Similar results were observed with TGN38 (a marker for the *trans*-Golgi network[20]) (Supplementary Fig. 19b) and syntaxin 7 (a marker for late endosomes) (Fig. 3e and Supplementary Fig. 19a) but not calnexin (a marker for the endoplasmic reticulum) (Supplementary Fig. 19c), confirming the involvement of the *trans*-Golgi or late endosomes in alternative macroautophagy. Because Rab9 is involved in trafficking from late endosomes to the *trans*-Golgi[20], we suspected its role in alternative macroautophagy. As expected, GFP–Rab9 was co-localized with Lamp2-positive autolysosomes in etoposide-treated *Atg5*$^{-/-}$ MEFs (Fig. 3f and Supplementary Fig. 20a). This co-localization was slightly increased with GFP–Rab9(Q66L), a GTP-preferring active Rab9 form[20], and decreased with GFP–Rab9(S21N), a GDP-preferring dominant-negative Rab9 form[20] (Supplementary Fig. 20b). Furthermore, Rab9 silencing (Supplementary Fig. 20c) decreased etoposide-induced alternative macroautophagy (Fig. 3g, h and Supplementary Fig. 20d, e), indicating that Rab9 is involved in Atg5-independent macroautophagy. Rab9 silencing decreased the number of autophagic vacuoles but induced the accumulation of isolation membranes (Fig. 3i). Because many isolation membranes were normally generated from the early exposure to etoposide (Fig. 3i), the effect of Rab9 silencing was not due merely to a slow progression of macroautophagy but also to inhibition of autophagosome maturation. This accumulation of

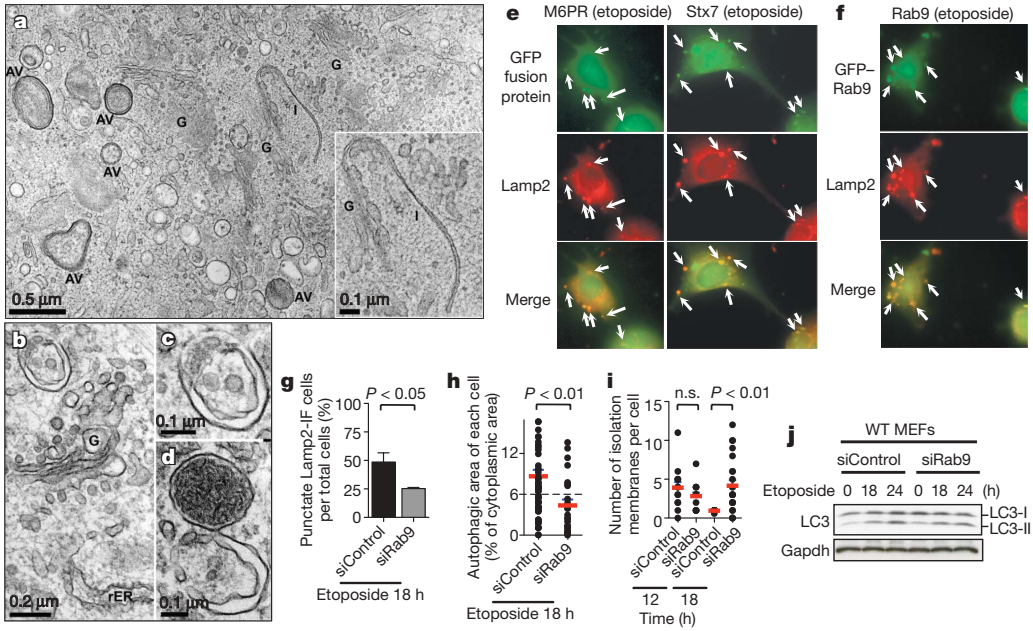


Figure 3 | Involvement of Rab9 in alternative autophagosome generation. **a–d**, Electron micrographs of etoposide-treated $Atg5^{-/-}$ MEFs. **a**, AVs were observed near the Golgi apparatus (G). Inset, the isolation membrane (I) was extended from the Golgi stack. **b**, Developing autophagosome. rER, rough endoplasmic reticulum. **c**, **d**, Isolation membrane fusing with vesicles containing thick membrane (**d**) and a complete autophagosome (**c**) is generated. **e**, **f**, Co-localization of Lamp2 with GFP–M6PR, GFP–Stx7 or GFP–Rab9 in etoposide-treated $Atg5^{-/-}$ MEFs. Arrows indicate co-localized dots. Original magnification, × 200. **g–i**, $Atg5^{-/-}$ MEFs were transfected with the indicated siRNAs for 24 h and treated with etoposide. Induction of macroautophagy was assessed by Lamp2 IF (**g**) and by EM (**h**), and the number of isolation membranes was calculated (**i**). In **g**, error bars indicate s.d. ($n = 4$). In **h** and **i**, red and blue lines indicate means and s.e.m., respectively; $n = 35$ cells each. **j**, No effect of silencing of Rab9 (siRab9) on the production of LC3-II in etoposide-treated WT MEFs.

isolation membranes was not observed with Ulk1 silencing (Supplementary Fig. 13d) or beclin 1 silencing (Supplementary Fig. 16d). Rab9 silencing had no influence on conventional macroautophagy (Fig. 3j), indicating that Rab9 is required for alternative but not conventional macroautophagy.

Finally, we examined alternative macroautophagy in $Atg5^{-/-}$ embryos to understand its physiological relevance. In $Atg5^{-/-}$ fetal brains (Fig. 4a), livers (Fig. 4b–d) and hearts (Fig. 4e), we detected autophagic vacuoles at the same level as in WT fetal tissues (data not shown). We then investigated alternative macroautophagy during erythrocyte maturation. Erythrocytes undergo organelle clearance during terminal differentiation, and macroautophagy may function in this process[21,22]. In fact, ultrastructural analysis of the WT fetal liver (embryonic day (E)14.5), where haematopoiesis occurs, showed that autophagic vacuoles in reticulocytes engulfed and digested mitochondria (Fig. 4f). We observed a few more autophagic vacuoles in $Atg5^{-/-}$ reticulocytes (Fig. 4g–i). Examination of the circulating erythrocytes showed that although the number of autophagic vacuoles was decreased, a few mitochondria were still engulfed and digested in the vacuoles in both mice groups (Fig. 4k–n). Moreover, the number of persisting mitochondria in $Atg5^{-/-}$ reticulocytes and erythrocytes was the same as in WT cells of each type (Fig. 4j, o).

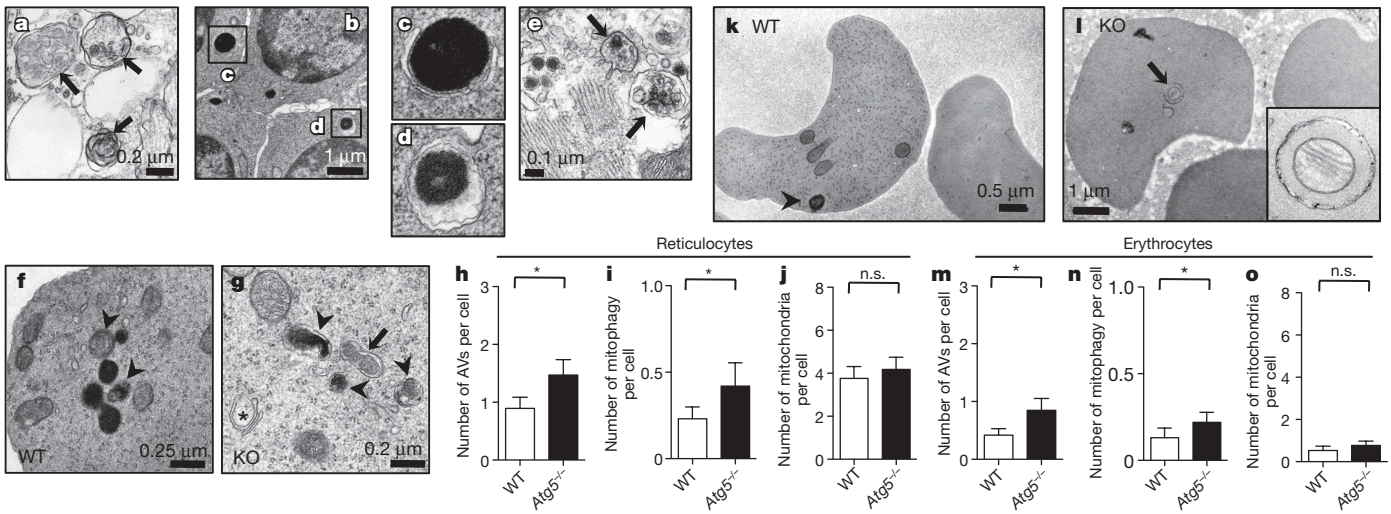


Figure 4 | Physiological roles of alternative macroautophagy. **a–e**, Typical autophagic structures in the tissues of an $Atg5^{-/-}$ embryo (E14.5): **a**, midbrain; **b–d**, liver; **c**, **d**, developing autophagosomes; **e**, heart. Arrows indicate the autophagic vacuoles. **f**, **g**, **k**, **l**, Electron micrographs of wild-type (WT) and $Atg5^{-/-}$ (KO) reticulocytes (**f**, **g**) and erythrocytes (**k**, **l**). An isolation membrane (asterisk), autophagosomes (arrows) and autolysosomes (arrowheads) are shown. Inset, engulfed mitochondria. **h–j**, **m–o**, The numbers of autophagic vacuoles (**h**, **m**) and autophagic vacuoles containing mitochondria (**i**, **n**) and persisting mitochondria (**j**, **o**) in reticulocytes (**h–j**) and erythrocytes (**m–o**). Error bars indicate s.d. (*$P < 0.05$; $n = 55$ cells).

This was confirmed when residual mitochondria were examined by MitoTracker Green in Ter119$^+$ erythroid cells (Supplementary Fig. 21a), which is consistent with previous findings[23]. Similar results were observed when different mice were analysed (Supplementary Fig. 21b). Furthermore, Ter119$^+$ CD71$^+$ and Ter119$^+$ CD71$^-$ erythroid cell populations were similar in WT and *Atg5*$^{-/-}$ mice, indicating that terminal differentiation proceeded equally (Supplementary Fig. 21c). Mitochondrial clearance from erythroid cells may be due to Atg5-independent alternative macroautophagy. This is consistent with observations that a lack of Ulk1 and the addition of 3-methyladenine results in disturbance in autophagic clearance of mitochondria[22,24]. Taken together, these results show that Ulk1-mediated alternative macroautophagy may function in the terminal differentiation of erythrocytes.

Our findings indicate that macroautophagy is more complex than previously realized. Conventional macroautophagy is crucial for basal and starvation-induced autophagy *in vitro*[4,5], and is required for neuronal protein aggregate clearance[25,26] and overcoming early neonatal starvation *in vivo*[4,5]. In contrast, we showed that alternative macroautophagy can be triggered by cellular stress *in vitro* and that it functions in the autophagic elimination of organelles during erythrocyte differentiation *in vivo*. Although both processes lead to the bulk degradation of cellular proteins, they may be activated by different stimuli in different cell types and may have different physiological functions. To understand macroautophagy better, it is important to classify autophagy-related molecules according to the type of macroautophagy in which they are involved.

METHODS SUMMARY

DNA transfection. MEFs were transfected with plasmid DNAs or short interfering RNAs (siRNAs) with the Amaxa electroporation system[27]. In some experiments, plasmids were introduced into MEFs by retroviral infections by using Plat-E cells.

Induction of autophagy. MEFs were treated with 10 μM etoposide for 18 h or starved for 6 h, unless otherwise indicated.

Electron microscopy. Cells and tissues were fixed by a conventional method or by a quick-freeze substitution method. Fixed samples were embedded in Epon 812, and thin sections were then cut and stained with uranyl acetate and lead citrate for observation under a Jeol-1010 electron microscope (Jeol) at 80 kV. The area of every autophagic vacuole and the total cytoplasmic area were calculated on the photographs by using a planimeter. For each cell, the autophagic area was calculated by expressing the total area of autophagic vacuoles as a percentage of the cytoplasmic area. In experiments combining fluorescence microscopy and electron microscopy, cells were cultured on coverslips with grids and fixed with paraformaldehyde and glutaraldehyde. In the Lamp2 experiment, cells were then immunostained with anti-Lamp2 monoclonal antibody. Subsequently, some sections were viewed under a confocal fluorescence microscope; the cells were fixed with OsO_4 and examined by electron microscopy. In the GFP–LC3 experiment, after GFP–LC3 fluorescence had been viewed, the same cells were fixed with OsO_4 and examined by electron microscopy.

Long-lived protein degradation assay. MEFs were starved or treated with etoposide with zVAD-fmk (100 μM), and the degradation of long-lived proteins was measured by a standard method[28]. The extent of protein degradation was calculated as the percentage of protein degradation in stress-induced cells minus the percentage of protein degradation in healthy cells.

Immunofluorescence analysis. Cells were fixed in 4% formaldehyde, permeabilized in 0.1% Triton X-100 and stained with anti-Lamp2 monoclonal antibody, followed by a FITC-conjugated secondary antibody.

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

Received 5 May; accepted 24 August 2009.

1. Mizushima, N., Levine, B., Cuervo, A. M. & Klionsky, D. J. Autophagy fights disease through cellular self-digestion. *Nature* **451,** 1069–1075 (2008).
2. Xie, Z. & Klionsky, D. J. Autophagosome formation: core machinery and adaptations. *Nature Cell Biol.* **9,** 1102–1109 (2007).
3. Mizushima, N., Ohsumi, Y. & Yoshimori, T. Autophagosome formation in mammalian cells. *Cell Struct. Funct.* **27,** 421–429 (2002).
4. Kuma, A. *et al.* The role of autophagy during the early neonatal starvation period. *Nature* **432,** 1032–1036 (2004).
5. Komatsu, M. *et al.* Impairment of starvation-induced and constitutive autophagy in Atg7-deficient mice. *J. Cell Biol.* **169,** 425–434 (2005).
6. Kabeya, Y. *et al.* LC3, a mammalian homologue of yeast Apg8p, is localized in autophagosome membranes after processing. *EMBO J.* **19,** 5720–5728 (2000).
7. Yue, Z., Jin, S., Yang, C., Levine, A. J. & Heintz, N. Beclin 1, an autophagy gene essential for early embryonic development, is a haploinsufficient tumor suppressor. *Proc. Natl Acad. Sci. USA* **100,** 15077–15082 (2003).
8. Qu, X. *et al.* Promotion of tumorigenesis by heterozygous disruption of the beclin 1 autophagy gene. *J. Clin. Invest.* **112,** 1809–1820 (2003).
9. Rosenbluth, J. M. & Pietenpol, J. A. mTOR regulates autophagy-associated genes downstream of p73. *Autophagy* **5,** 114–116 (2009).
10. Liou, W., Geuze, H. J., Geelen, M. J. H. & Slot, J. W. The autophagic and endocytic pathways converge at the nascent autophagic vacuoles. *J. Cell Biol.* **136,** 61–70 (1997).
11. Razi, M., Chan, E. Y. W. & Tooze, S. A. Early endosomes and endosomal coatomer are required for autophagy. *J. Cell Biol.* **185,** 305–321 (2009).
12. Yoshimori, T., Yamamoto, A., Moriyama, Y., Futai, M. & Tashiro, Y. Bafilomycin A1, a specific inhibitor of vacuolar-type H^+-ATPase, inhibits acidification and protein degradation in lysosomes of cultured cells. *J. Biol. Chem.* **266,** 17707–17712 (1991).
13. Bursch, W. The autophagosomal–lysosomal compartment in programmed cell death. *Cell Death Differ.* **8,** 569–581 (2001).
14. Cuervo, A. M. Autophagy: many paths to the same end. *Mol. Cell. Biochem.* **263,** 55–72 (2004).
15. Mizushima, N. *et al.* Dissection of autophagosome formation using Apg5-deficient mouse embryonic stem cells. *J. Cell Biol.* **152,** 657–668 (2001).
16. Kabeya, Y. *et al.* LC3, GABARAP and GATE16 localize to autophagosomal membrane depending on form-II formation. *J. Cell Sci.* **117,** 2805–2812 (2004).
17. Young, A. R. *et al.* Starvation and ULK1-dependent cycling of mammalian Atg9 between the TGN and endosomes. *J. Cell Sci.* **119,** 3888–3900 (2006).
18. Hara, T. *et al.* FIP200, a ULK-interacting protein, is required for autophagosome formation in mammalian cells. *J. Cell Biol.* **181,** 497–510 (2008).
19. Gagnon, E. Endoplasmic reticulum-mediated phagocytosis is a mechanism of entry into macrophages. *Cell* **110,** 119–131 (2002).
20. Riederer, M. A., Soldati, T., Shapiro, A. D., Lin, J. & Pfeffer, S. R. Lysosome biogenesis requires Rab9 function and receptor recycling from endosomes to the *trans*-Golgi network. *J. Cell Biol.* **125,** 573–582 (1994).
21. Fader, C. M. & Colombo, M. I. Multivesicular bodies and autophagy in erythrocyte maturation. *Autophagy* **2,** 122–125 (2006).
22. Sandoval, H. *et al.* Essential role for Nix in autophagic maturation of erythroid cells. *Nature* **454,** 232–235 (2008).
23. Matsui, M., Yamamoto, A., Kuma, A., Ohsumi, Y. & Mizushima, N. Organelle degradation during the lens and erythroid differentiation is independent of autophagy. *Biochem. Biophys. Res. Commun.* **339,** 485–489 (2006).
24. Kundu, M. *et al.* Ulk1 plays a critical role in the autophagic clearance of mitochondria and ribosomes during reticulocyte maturation. *Blood* **112,** 1493–1502 (2008).
25. Hara, T. *et al.* Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. *Nature* **441,** 885–889 (2006).
26. Komatsu, M. *et al.* Loss of autophagy in the central nervous system causes neurodegeneration in mice. *Nature* **441,** 880–884 (2006).
27. Shimizu, S. *et al.* Role of Bcl-2 family proteins in a non-apoptotic programmed cell death dependent on autophagy genes. *Nature Cell Biol.* **6,** 1221–1228 (2004).
28. Ogier-Denis, E., Houri, J. J., Bauvy, C. & Codogno, P. Guanine nucleotide exchange on heterotrimeric Gi3 protein controls autophagic sequestration in HT-29 cells. *J. Biol. Chem.* **271,** 28593–28600 (1996).

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

Acknowledgements We thank M. Narita and A. R. J. Young for critical reading of the manuscript; N. Mizushima for providing *Atg5*$^{+/-}$ mice and the expression plasmid of GFP–LC3; T. Yoshimori for the human beclin 1 expression plasmid; and T. Kitamura for providing Plat-E cells. This study was supported in part by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NIBIO), a grant for Creative Scientific Research, a grant for the 21st Century COE Program from the Japanese Ministry of Education, Science, Sports and Culture, a grant for Comprehensive Research on Aging and Health from the Japanese Ministry of Health, Labor and Welfare, and a grant for Solution-Oriented Research for Science and Technology (SORST) from the Japan Science and Technology Corporation. This study was also supported by grants from the Uehara Memorial Foundation, the Sagawa Foundation for Promotion of Cancer Research, the YASUDA Medical Foundation, the Astellas foundation for research on metabolic disorders, and the Foundation for Promotion of Cancer Research.

Author Contributions Y.N. performed the biochemical analyses. S.A. and T.K. performed the electron microscopy analyses. K.F. and H.Y. performed the Rab9 study. T.M. developed the Lamp2 immunofluorescence assay. M.K. provided the *Atg7*$^{-/-}$ cells. K.O. contributed data analysis. Y.T. supervised data interpretation. S.S. designed the research 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 S.S. (shimizu.pcb@mri.tmd.ac.jp).

METHODS

Antibodies and chemicals. Anti-Ulk1 (A7481) polyclonal antibodies were purchased from Sigma-Aldrich, and anti-Gapdh (6G7) monoclonal antibodies were purchased from BD Biosciences. Anti-LC3 (5F10) monoclonal antibodies were obtained from NanoTools, and anti-Lamp2 (GL2A7) monoclonal antibodies were obtained from Abcam. Lysosome protease inhibitor cocktail (containing 1.5 mM E-64, 2 mM leupeptin and 1 mM pepstatin A) was purchased from Nacalai Tesque. Etoposide and bafilomycin A1 were from Sigma-Aldrich, and other chemicals were purchased from Nacalai Tesque.

Cell culture and DNA transfection. MEFs generated from WT and *Atg5*$^{-/-}$ embryos at embryonic day 13.5 were immortalized with SV40 T antigen. MEFs were grown in modified DMEM[27]. Cells (10^6) were transfected with plasmid DNA using the Amaxa electroporation system (kit V, program U-20) in accordance with the manufacturer's instructions. The transfection efficiency was more than 75%, as assessed by co-transfection with DNA for GFP. The siRNA sequences used were as follows: mouse *Ulk1*, 5′-GGGUGGACACAUGCUAAUA-3′; separate mouse *Ulk1*, 5′-GGAAACUUCAGAUGUUUAA-3′; mouse *Fip200*, 5′-CCCAAAGAUUAUUCAACCA-3′; mouse beclin 1, 5′-GGUUUGGAAAGAUGCUUUA-3′; separate mouse beclin 1, 5′-GACAGUUUGGCACAAUCAA-3′; mouse *Vps34*, 5′-GAAAGAUAGUCAAACUUCA-3′; mouse *Rab9*, 5′-GCAGUACUAUUAAAGAUUATT-3′. We also used control siRNA (Dharmacon siGENOME Non-Targeting siRNA#1 D-001210-01-20; Thermo Scientific). Cells (10^6) were transfected with 10 μg of siRNA with the use of the Amaxa electroporation system.

The retroviral GFP fusion expression vectors pMSCV–GFP–MCS–Zeo and pMSCV–MCS–GFP–Zeo were constructed by replacing the HindIII–ClaI fragment of pMSCV puro (Clontech) with the Zeocin resistance gene and subsequently cloning the enhanced GFP (EGFP) sequence into the BglII site and EcoRI site, respectively. cDNA corresponding to mouse Rab9a and its mutants, mouse syntaxin7 and mouse M6PR, was subsequently cloned into these vectors to generate pMSCV–GFP–Rab9a–Zeo, pMSCV–GFP–Syntaxin7–Zeo, pMSCV–M6PR–GFP–Zeo, pMSCV–GFP–Rab9a^{S21N}–Zeo and pMSCV–GFP–Rab9a^{Q66L}–Zeo. Each EGFP fusion plasmid was introduced into the MEFs by retroviral infection with the use of Plat-E cells.

Quantitative RT–PCR and microarray analysis. *Atg5*$^{-/-}$ MEFs were incubated with or without 10 μM etoposide for 12 h, and microarray analysis and qRT–PCR were performed. For microarray analysis, total RNA was purified with TRIzol reagent (Invitrogen) and used for preparing cRNA by Message AmpII (Ambion). cRNA was hybridized to Mouse Whole Genome 4 × 44K microarrays (Agilent Technologies) in accordance with the manufacturer's instructions. Two biological replicates were performed for each set of experimental conditions. Data were analysed with Feature Extraction Software (v. 9.5.3.1) (Agilent). For qRT–PCR analysis, total RNA was purified on an RNeasy column (Qiagen) and used for mRNA purification by Turbo Capture8 (Qiagen). Synthesis of cDNA and PCR amplification were performed with the iScript One-step RT–PCR kit with SYBR Green (Bio-Rad). Quantitative determination was performed with the CFX96 real-time PCR system (Bio-Rad). All samples were normalized for *Gapdh* level.

Electron microscopy. Attached cells and tissues were fixed by a conventional method (1.5% paraformaldehyde and 3% glutaraldehyde in 0.1 M phosphate buffer, pH 7.3, followed by an aqueous solution of 1% OsO_4) or by a quick-freeze substitution method (cells were frozen in liquid nitrogen and then fixed with 1% OsO_4 in acetone at −80 °C). Fixed samples were embedded in Epon 812, and thin sections (70–80 nm) were then cut and stained with uranyl acetate and lead citrate for observation under a Jeol-1010 electron microscope (Jeol) at 80 kV.

Fixed adherent cells were sectioned up to 3 μm from the base. The extent of macroautophagy was assessed on electron micrographs that contained both the nucleus and cytoplasm of individual cells. The area of every autophagic vacuole and total cytoplasmic area were calculated on the enlarged photographs with the use of a planimeter (Planix). For each cell, the autophagic area was calculated by expressing the total area of autophagic vacuoles as a percentage of the cytoplasmic area, and cells with an autophagic area of more than 6% were defined as autophagic cells (6% was the upper limit in healthy cells). Macroautophagy was quantified in each sample for at least 20 cells and was confirmed by two additional independent experiments.

In experiments combining fluorescence microscopy and electron microscopy, MEFs cultured on coverslips with grids were treated with 10 μM etoposide for 18 h and fixed with glutaraldehyde and paraformaldehyde. In the Lamp2 experiment, cells were immunostained with anti-Lamp2 monoclonal antibody. Subsequently, three or four sections were acquired under confocal fluorescence microscopy and the cells were treated with OsO_4. Cells were identified by using the grid as a guide, and the same sections as those employed for fluorescence analysis were also examined by electron microscopy. In the GFP–LC3 experiment, after GFP–LC3 fluorescence had been viewed, the same cells were fixed with OsO_4 and examined by electron microscopy.

The number of autophagic vacuoles and mitochondria in erythrocytes or reticulocytes was counted in cells that had a long axis of more than 4.5 μm.

Long-lived protein degradation assay. MEFs were starved or treated with etoposide with zVAD-fmk (100 μM), and degradation of long-lived proteins was measured by a standard method[28]. In brief, cells were labelled for 20 h with a medium containing 0.2 μCi ml^{-1} L-[^{14}C]valine (GE Healthcare). After washing and incubation in a medium containing 10 mM unlabelled valine for 60 min, the medium was replaced with fresh medium and the incubation was continued for the indicated durations. The medium was then precipitated in 10% trichloroacetic acid (TCA), and TCA-soluble radioactivity was measured. Release of [^{14}C]valine was calculated from the radioactivity in the TCA-soluble supernatant as a percentage of the total cell radioactivity. The extent of protein degradation was calculated by subtracting [^{14}C]valine release by untreated cells from that by starved or etoposide-treated cells, as the percentage of protein degradation in stress-induced cells minus the percentage of protein degradation in healthy cells.

Staining of conventional macroautophagosomes. Cells stably expressing GFP–LC3 were treated with etoposide or subjected to starvation, and the fluorescence of attached cells was observed under an FV500 confocal fluorescence microscope (Olympus) after 18 h.

Immunofluorescence analysis. Cells grown in eight-well slide chambers were fixed in 4% formaldehyde, permeabilized in 0.1% Triton X-100 and stained with anti-Lamp2 monoclonal antibody, followed by a FITC-conjugated secondary antibody (Invitrogen). The coverslips were then mounted in mounting medium (Beckman Coulter) with propidium iodide and examined by fluorescence microscopy.

Statistical analysis. Statistical evaluation was performed with the non-paired *t*-test.

LETTERS

Membrane-bound Fas ligand only is essential for Fas-induced apoptosis

Lorraine A. O' Reilly[1], Lin Tai[1], Lily Lee[1], Elizabeth A. Kruse[1,2], Stephanie Grabow[1,2], W. Douglas Fairlie[1], Nicole M. Haynes[3], David M. Tarlinton[1], Jian-Guo Zhang[1], Gabrielle T. Belz[1], Mark J. Smyth[3], Philippe Bouillet[1], Lorraine Robb[1] & Andreas Strasser[1]

Fas ligand (FasL), an apoptosis-inducing member of the TNF cytokine family, and its receptor Fas are critical for the shutdown of chronic immune responses[1–3] and prevention of autoimmunity[4,5]. Accordingly, mutations in their genes cause severe lymphadenopathy and autoimmune disease in mice[6,7] and humans[8,9]. FasL function is regulated by deposition in the plasma membrane and metalloprotease-mediated shedding[10,11]. Here we generated gene-targeted mice that selectively lack either secreted FasL (sFasL) or membrane-bound FasL (mFasL) to resolve which of these forms is required for cell killing and to explore their hypothesized non-apoptotic activities. Mice lacking sFasL (*FasL*$^{\Delta s/\Delta s}$) appeared normal and their T cells readily killed target cells, whereas T cells lacking mFasL (*FasL*$^{\Delta m/\Delta m}$) could not kill cells through Fas activation. *FasL*$^{\Delta m/\Delta m}$ mice developed lymphadenopathy and hypergammaglobulinaemia, similar to *FasL*$^{gld/gld}$ mice, which express a mutant form of FasL that cannot bind Fas, but surprisingly, *FasL*$^{\Delta m/\Delta m}$ mice (on a C57BL/6 background) succumbed to systemic lupus erythematosus (SLE)-like autoimmune kidney destruction and histiocytic sarcoma, diseases that occur only rarely and much later in *FasL*$^{gld/gld}$ mice. These results demonstrate that mFasL is essential for cytotoxic activity and constitutes the guardian against lymphadenopathy, autoimmunity and cancer, whereas excess sFasL appears to promote autoimmunity and tumorigenesis through non-apoptotic activities.

Although Fas-induced apoptosis is thought to require extensive aggregation of pre-assembled Fas trimers[12–15], it has not been resolved whether mFasL, sFasL or both cause cell killing when expressed physiologically[4,5]. It is also debated whether either or both of these forms of FasL may have non-apoptotic activities, such as induction of inflammatory responses[5,16,17]. Indeed, sFasL is capable of activating the NF-κB pathway (Supplementary Fig. 1 and refs 18 and 19). We sought to determine the physiological functions of mFasL and sFasL by generating gene-targeted mice that cannot shed FasL but do express mFasL (*FasL*$^{\Delta s/\Delta s}$) and mice that lack mFasL but are capable of producing sFasL (*FasL*$^{\Delta m/\Delta m}$). Studies using FasL overexpression in transfected cell lines showed that the former can be accomplished by mutating the sequences in the *fasl* gene encoding the amino acids required for metalloprotease-mediated cleavage[12–14] (Fig. 1a, Supplementary Fig. 2a, b, d). Conversely, the latter can be achieved by replacing the sequences in the *fasl* gene encoding the transmembrane and intra-cellular regions of FasL with those encoding the signal peptide of the cytokine G-CSF[12–14] (Fig. 1a, Supplementary Fig. 2a, c, e).

To verify that the mutations had the intended consequences, we compared the expression and subcellular localization of FasL in mitogenically activated T lymphocytes from *FasL*$^{\Delta s/\Delta s}$, *FasL*$^{\Delta m/\Delta m}$ and wild-type (WT) mice. Immunofluorescent staining and confocal microscopic analysis of fixed cells showed that intracellular localization and levels of the FasL$^{\Delta s}$ and FasL$^{\Delta m}$ mutant proteins were comparable to those of WT FasL (Fig. 1b). Enzyme-linked immunosorbent assay (ELISA) demonstrated that mitogen-activated T cells from *FasL*$^{\Delta m/\Delta m}$ and WT mice contained substantial levels of FasL in their supernatants, whereas *FasL*$^{\Delta s/\Delta s}$ T cells had significantly less (Fig. 1c). FasL in cellular supernatants can be found as either sFasL, derived by metalloprotease-mediated cleavage, or mFasL present on vesicles that had been shed by cells[12]. mFasL on vesicles can efficiently trigger Fas-mediated apoptosis in cultured cells[12], although the physiological relevance of this remains unclear. Fast protein liquid chromatography (FPLC) and ultra-centrifugation revealed that, in contrast to FasL from supernatants of WT or *FasL*$^{\Delta m/\Delta m}$ T cells, a substantial fraction of FasL in supernatants of *FasL*$^{\Delta s/\Delta s}$ T cells resided in membranous fractions (Supplementary Fig. 3). Finally, immunofluorescent cell surface staining and fluorescence-activated cell sorting (FACS) analysis identified significantly higher levels of mFasL on activated T cells from *FasL*$^{\Delta s/\Delta s}$ mice than in WT T cells (Fig. 1d), consistent with the notion that metalloprotease-mediated cleavage reduces the levels of mFasL[11–14]. As we expected, no FasL was detected on the surface of *FasL*$^{\Delta m/\Delta m}$ T cells (Fig. 1e). These results verify that *FasL*$^{\Delta s/\Delta s}$ mice produce mFasL that cannot be shed by metalloproteases, whereas *FasL*$^{\Delta m/\Delta m}$ mice lack mFasL but produce sFasL.

FasL contributes to the killing of virus-infected and other target cells[4], so we wanted to determine which form is essential to this action. We used mitogen-activated T cells from WT or the mutant *fasl* knock-in mice as killers and the FasL-sensitive CH1 mouse B lymphoma cells (Supplementary Fig. 4a, b) as targets. *FasL*$^{\Delta s/\Delta s}$ T cells killed CH1 cells with significantly higher efficiency than did WT T cells (Fig. 2a). In contrast, *FasL*$^{\Delta m/\Delta m}$ T cells possessed very poor cytotoxic activity, comparable to the T cells of FasL-deficient *FasL*$^{gld/gld}$ mice (Fig. 2b). FasL neutralization inhibited the cytotoxicity of WT and *FasL*$^{\Delta s/\Delta s}$ T cells but did not further reduce the poor killing by *FasL*$^{\Delta m/\Delta m}$ or *FasL*$^{gld/gld}$ T cells (Supplementary Fig. 4c–e), demonstrating that only mFasL triggered a FasL/Fas-dependent apoptotic process.

Restimulation of activated T cells *in vitro* causes activation-induced cell death (AICD), which is largely dependent on FasL-mediated (paracrine and autocrine) Fas activation[4]. Stimulation with mitogenic antibodies to CD3 triggered AICD in *FasL*$^{\Delta s/\Delta s}$ T-cell blasts as efficiently as in WT T-cell blasts (Fig. 2c). In contrast, AICD was abnormally reduced in *FasL*$^{\Delta m/\Delta m}$ T cells, to a similar extent as in their *FasL*$^{gld/gld}$ counterparts (Fig. 2d). FasL neutralization significantly reduced AICD in WT and *FasL*$^{\Delta s/\Delta s}$ T cells but did not further diminish the already reduced killing ability of *FasL*$^{\Delta m/\Delta m}$ or *FasL*$^{gld/gld}$ T cells (Supplementary Fig. 5). This is consistent with

[1]The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3052, Australia. [2]Department of Medical Biology, The University of Melbourne, Parkville, Victoria 3010, Australia. [3]Peter MacCallum Cancer Centre, East Melbourne, Victoria 3002, Australia.

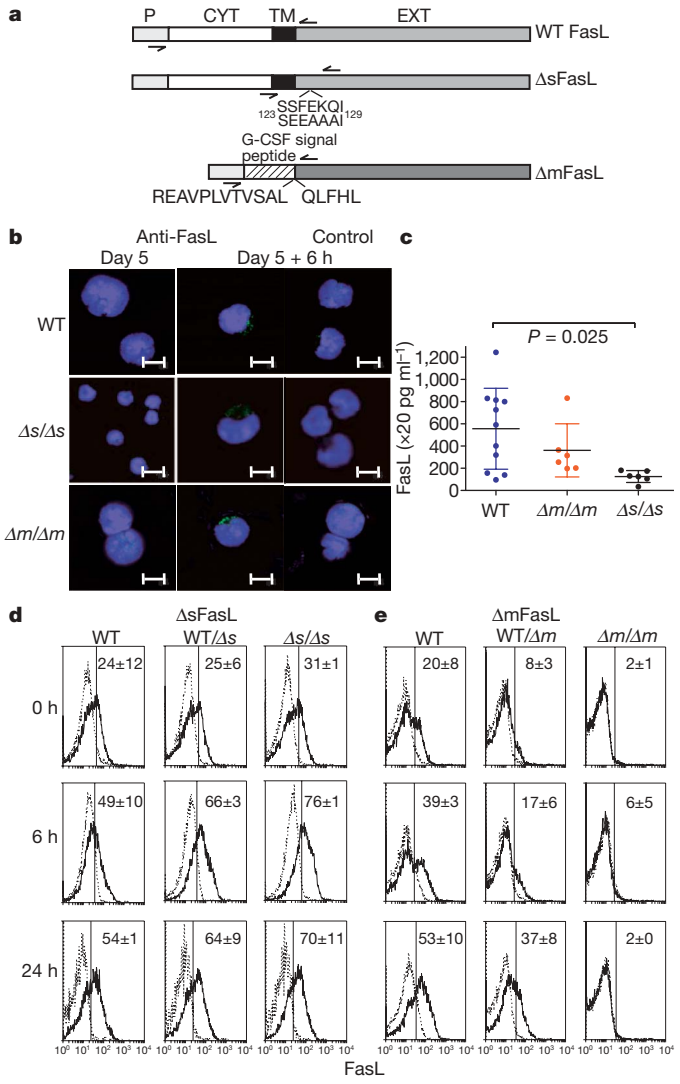


Figure 1 | Generation of mutant mice that specifically lack either sFasL or mFasL. **a**, Schematic diagram of the WT mouse *fasl* gene, the mutation for creating the $FasL^{\Delta s}$ allele and the mutation for creating the $FasL^{\Delta m}$ allele. To make $FasL^{\Delta s}$, the metalloproteinase recognition site (middle sequence) was altered to prevent FasL shedding. To make $FasL^{\Delta m}$, sequences encoding the human G-CSF signal sequence were fused in frame with those for the extracellular region of mouse FasL (lowest sequence) to preclude insertion of FasL into the plasma membrane but allowing secretion of FasL (sFasL). P, CYTO, TM and EXT represent the promoter, cytoplasmic, transmembrane and extracellular regions, respectively. The positions of polymerase chain reaction (PCR) primers used for genotyping of the gene-targeted mice are indicated. **b**, Immunofluorescent staining and confocal microscopy to demonstrate intracellular localization of FasL (green) in activated T cells from WT, $FasL^{\Delta s/\Delta s}$ and $FasL^{\Delta m/\Delta m}$ mice. DAPI (blue) was used to label nuclei. Scale bars represent 6 μm. **c**, ELISA to quantify the levels of FasL in the supernatants of activated T cells from WT, $FasL^{\Delta s/\Delta s}$ and $FasL^{\Delta m/\Delta m}$ mice. Each dot represents a single mouse and the bar indicates the mean ±s.d. $FasL^{\Delta s/\Delta s}$ T cells had significantly ($P < 0.03$) less FasL in their supernatants than WT T cells. **d**, **e**, Cell surface immunofluorescent staining and FACS analysis to measure expression of mFasL on activated T cells from WT, $FasL^{\Delta s/\Delta s}$ (**d**) and $FasL^{\Delta m/\Delta m}$ (**e**) mice. Gene-targeted mice that selectively lack sFasL are denoted ΔsFasL; those that lack mFasL are denoted ΔmFasL. Solid lines show staining with an anti-FasL antibody and dotted lines show staining with an isotype-matched control antibody. Values represent mean ±s.d. of $mFasL^+$ T cells from three independent experiments ($P < 0.05$ WT versus $FasL^{\Delta m/\Delta m}$ mice at 6 and 24 h; $P < 0.05$ WT versus $FasL^{\Delta s/\Delta s}$ mice at 6 h). For all experiments (**b**–**e**), T cells were stimulated with concanavalin A for 3 days, rested for 2 days in interleukin IL-2 (0 h) and then restimulated with phorbol 12-myristate 13-acetate (PMA) plus ionomycin for 6 or 24 h.

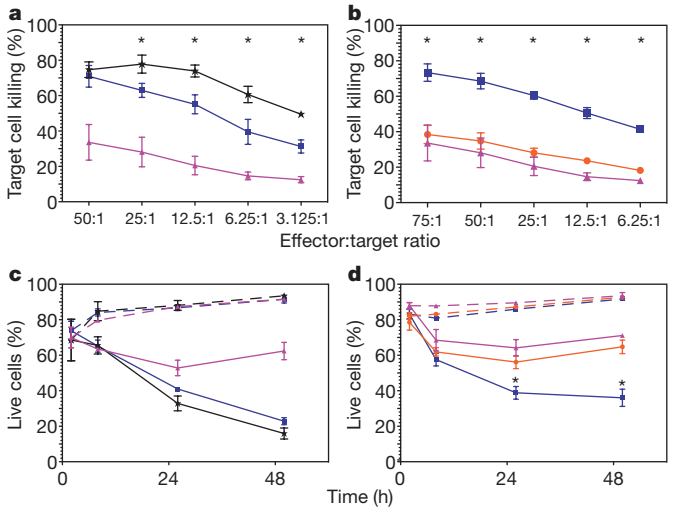


Figure 2 | mFasL but not sFasL is essential for target cell killing and AICD. **a**, **b**, Target cell killing. CH1 target cells (FasL-sensitive) were co-cultured with activated T cells from WT (blue squares), $FasL^{\Delta s/\Delta s}$ (black stars) or $FasL^{gld/gld}$ (purple triangles) mice (**a**) or from WT (blue squares), $FasL^{\Delta m/\Delta m}$ (red circles) or $FasL^{gld/gld}$ (purple triangles) mice (**b**) at the indicated effector:target ratio. The percentage of CH1 target cell killing was measured by FACS analysis after 24 h. **c**, **d**, AICD. Activated T cells from WT (blue squares), $FasL^{\Delta s/\Delta s}$ (black stars) or $FasL^{gld/gld}$ (purple triangles) mice (**c**) or from WT (blue squares) $FasL^{\Delta m/\Delta m}$ (red circles) or $FasL^{gld/gld}$ (purple triangles) mice (**d**) were either cultured in medium plus IL-2 (broken lines) or restimulated for 6, 24 or 48 h with monoclonal antibodies to CD3 (solid lines). Cell survival was measured by staining with the vital dye PI plus fluorescein-isothiocyanate (FITC)-coupled annexin V and FACS analysis. Values in graphs in **a** to **d** represent mean ±s.e.m. from three independent experiments. Asterisks indicate $P < 0.05$ for $FasL^{\Delta s/\Delta s}$ versus WT (**a**) and $FasL^{\Delta m/\Delta m}$ versus WT (**b**, **d**).

the notion that AICD involves FasL/Fas-dependent as well as FasL/Fas-independent mechanisms[4,20]. Collectively, these results demonstrate that mFasL but not sFasL is essential for Fas-induced killing of target cells and AICD, and they indicate that metalloprotease-mediated cleavage of mFasL reduces the cytotoxic activity of activated T cells. This is consistent with certain findings using transfected cell lines overexpressing FasL[12–14] and with the notion that apoptosis induction requires not only binding of FasL trimers to pre-assembled Fas trimers but more extensive aggregation of Fas trimers[12–15]. This contrasts with TNFα, which can kill $TNF\text{-}R1^+$ target cells in both its membrane-bound and its secreted form[4,5].

Defects in the AICD of mature T cells are thought to be the cause of the lymphadenopathy, hyper-gammaglobulinaemia and autoimmunity in mice and humans deficient in FasL (for example, $FasL^{gld/gld}$) or Fas (for example, $Fas^{lpr/lpr}$)[4]. As reported[21], by ~14 weeks of age $FasL^{gld/gld}$ mice (on an inbred C57BL/6 background) started to show signs of lymphadenopathy and splenomegaly (Fig. 3a, b), including accumulation of large numbers of 'unusual' $TCR\alpha/\beta^+CD4^-CD8^-B220^+$ T cells (Fig. 3c, d). This was accompanied by hyper-gammaglobulinaemia (Fig. 3e) with high titres of anti-nuclear auto-antibodies (Fig. 3f, g). $FasL^{\Delta m/\Delta m}$ mice developed lymphadenopathy, splenomegaly, hyper-gammaglobulinaemia and anti-nuclear auto-antibodies at a similar rate and to a similar extent as $FasL^{gld/gld}$ mice (Fig. 3). Remarkably, however, the titres of anti-DNA auto-antibodies were significantly higher in $FasL^{\Delta m/\Delta m}$ mice than in $FasL^{gld/gld}$ mice (67% versus 14% were symptomatic of SLE-like disease (>12 international units per millilitre, IU ml^{-1}), $P < 0.05$; Fig. 3h). In contrast, $FasL^{\Delta s/\Delta s}$ mice exhibited none of these abnormalities and had a normal lifespan (Supplementary Fig. 6). These results show that mFasL but not sFasL is essential for the killing of unwanted lymphocytes that is required to prevent lymphadenopathy, hyper-gammaglobulinaemia and accumulation of auto-antibodies.

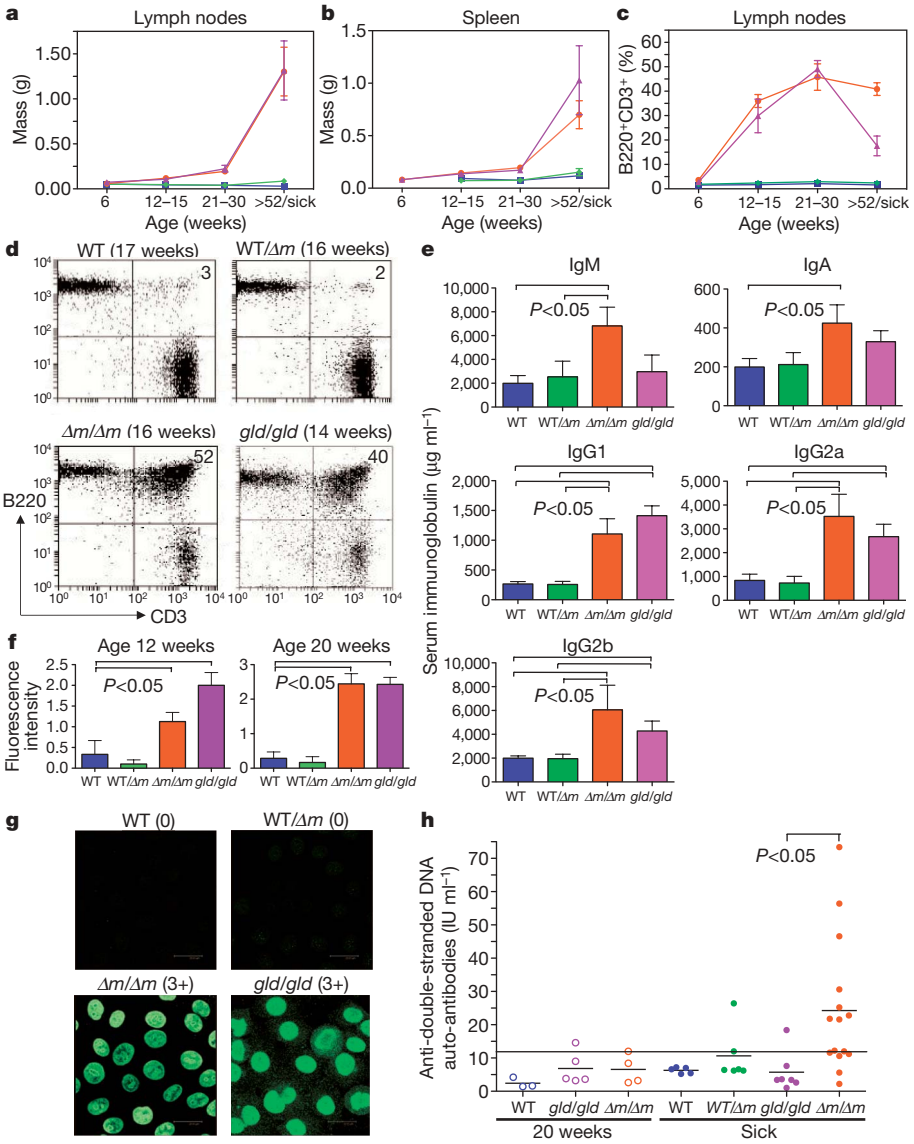
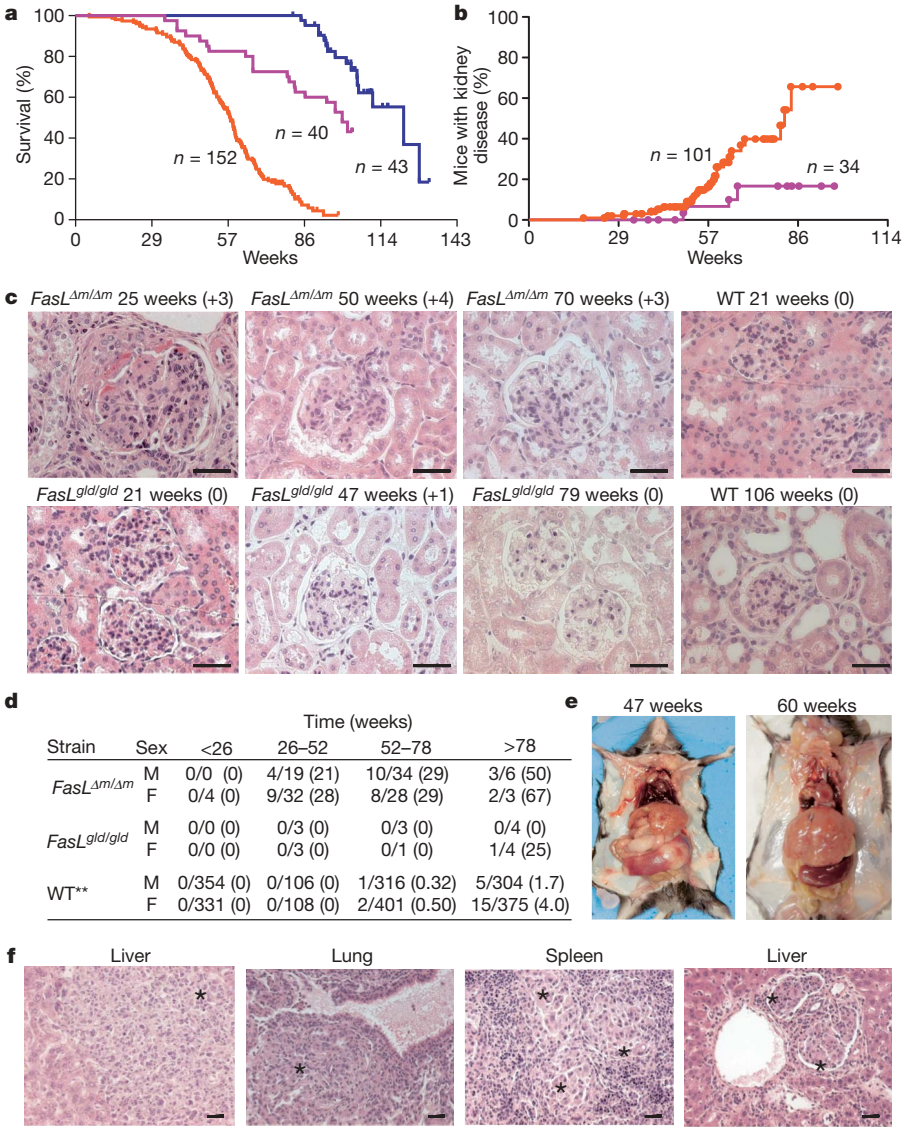


Figure 3 | mFasL but not sFasL is essential to prevent lymphadenopathy, splenomegaly and hyper-gammaglobulinaemia with anti-nuclear autoantibodies. **a**, **b**, Lymph nodes (axillary, brachial, inguinal and mesenteric) (**a**) and spleens (**b**) from WT (blue), $FasL^{gld/gld}$ (purple), $FasL^{\Delta m/WT}$ (green) or $FasL^{\Delta m/\Delta m}$ (red) mice of the ages indicated were weighed. '>52/sick' means that the mice were either over a year old or sick. **c**, **d**, The percentages of the 'unusual' $CD3^+B220^+$ T cells in the lymph nodes of these mice were determined by FACS analysis (numbers indicate percentage of cells in each quadrant). **e**, The levels of the indicated immunoglobulin isotypes in the sera of these mice (aged 20 weeks) were determined by ELISA, mean ±s.d. **f**, The levels of anti-nuclear auto-antibodies in the sera of these mice, aged 12 (left panel) or 20 weeks (right panel) were quantified by indirect immunofluorescence staining (1/100 serum dilution) of slides covered with human HEp-2 epithelial cells and scoring of brightness of fluorescence intensity on a scale of 0 (no fluorescence) to 3+ (maximal fluorescence). **g**, Examples of anti-nuclear auto-antibodies quantification in sera of mice of the indicated genotypes. Immunofluorescence intensity score is indicated in brackets. Values in graphs in **a** and **b** represent mean ± s.e.m. and in **c**, **e** and **f** they represent mean ± s.d. from a minimum of 10 up to 52 $FasL^{\Delta m/\Delta m}$ mice. **h**, The levels of anti-double-stranded DNA auto-antibodies in sera of mice of the indicated genotypes and ages were measured by ELISA (each dot represents the value for a single mouse). The bar at 12.5 IU ml^{-1} indicates the diagnostic level for SLE.

$FasL^{gld/gld}$ mice express a mutated $FasL^{gld}$ protein (in membrane-bound and secreted form) that is unable to bind to its receptor Fas[7], whereas the sFasL in the $FasL^{\Delta m/\Delta m}$ mice (like WT sFasL) can bind to Fas (Supplementary Fig. 1 and refs 12–14). If sFasL plays a role in inflammation, for example through NF-κB activation (Supplementary Fig. 1 and refs 18, 19), we would expect significant differences in morbidity and mortality between $FasL^{\Delta m/\Delta m}$ and $FasL^{gld/gld}$ mice. It was therefore remarkable that $FasL^{\Delta m/\Delta m}$ mice became sick significantly earlier than $FasL^{gld/gld}$ mice (Fig. 4a, median latency 58 versus 94 weeks; $P < 0.0001$). At autopsy 62% of terminally ill $FasL^{\Delta m/\Delta m}$ mice showed signs of SLE-like autoimmune disease, including cellular crescents, protein casts (Fig. 4b, c), deposition of immunoglobulin IgM as well as IgG and complement in renal glomeruli (Supplementary Fig. 7). In contrast, such pathologies were only rarely observed in $FasL^{gld/gld}$ mice, and then only at a considerably older age and with less severity. By 57 weeks about half of the $FasL^{\Delta m/\Delta m}$ mice but only ~15% of the $FasL^{gld/gld}$ mice had developed fatal SLE-like autoimmune kidney disease (Fig. 4b).

By 20 weeks 46% of $FasL^{\Delta m/\Delta m}$ mice presented with very high (>3,000 ng ml^{-1}) serum IgE levels and ~30% had developed severe dermatitis with lesions appearing on their ears and necks (Supplementary Fig. 8a, b). Although not previously reported, we observed this autoimmune pathology also in some $FasL^{gld/gld}$ mice, albeit at decidedly lower incidence (~10%) and later in life than in the $FasL^{\Delta m/\Delta m}$ mice. In $Fas^{lpr/lpr}$ mice lymphadenopathy and accumulation of $TCR\alpha/\beta^+CD4^-CD8^-B220^+$ T cells are accompanied by abnormally increased serum levels of pro-inflammatory cytokines, including TNFα, IL-6, IFNγ and FasL itself[22]. FasL was shown to activate NF-κB transcription factors and expression of pro-inflammatory cytokines and chemokines (Supplementary Fig. 1 and refs 5, 18, 19). We therefore hypothesized that $FasL^{\Delta m/\Delta m}$ mice may develop autoimmune disease more rapidly and at higher incidence than $FasL^{gld/gld}$ animals because only $FasL^{\Delta m/\Delta m}$ mice produce excess sFasL that can bind to its receptor Fas, which may then activate NF-κB and thereby drive production of pro-inflammatory cytokines (Supplementary Figs 8 and 9). Consistent with this idea, at 12 to 20 weeks $FasL^{\Delta m/\Delta m}$ mice contained substantial numbers of cells with high levels of nuclear (that is, active) p65/NF-kB in the spleen, liver (Supplementary Fig. 10) and kidneys (not shown) and high serum levels of TNFα (Supplementary Fig. 8; $P < 0.0001$ for TNFα). These abnormalities were significantly less prevalent in $FasL^{gld/gld}$ mice and were not detected in $FasL^{\Delta s/\Delta s}$ or WT animals (Supplementary Figs 8, 9 and 10).

Fas is expressed on hepatocytes and its activation causes apoptosis[23,24]. FasL can also be found in the liver, produced by infiltrating T lymphocytes or resident myeloid cells, and it has therefore been hypothesized that FasL/Fas-induced apoptosis prevents tumorigenesis in the liver[25]. Interestingly, a significant fraction (27% by 78 weeks) of both strains of



Strain	Sex	Time (weeks) <26	26–52	52–78	>78
$FasL^{\Delta m/\Delta m}$	M	0/0 (0)	4/19 (21)	10/34 (29)	3/6 (50)
	F	0/4 (0)	9/32 (28)	8/28 (29)	2/3 (67)
$FasL^{gld/gld}$	M	0/0 (0)	0/3 (0)	0/3 (0)	0/4 (0)
	F	0/0 (0)	0/3 (0)	0/1 (0)	1/4 (25)
WT**	M	0/354 (0)	0/106 (0)	1/316 (0.32)	5/304 (1.7)
	F	0/331 (0)	0/108 (0)	2/401 (0.50)	15/375 (4.0)

Figure 4 | $FasL^{\Delta m/\Delta m}$ mice die considerably earlier than $FasL^{gld/gld}$ mice owing to SLE-like fatal glomerulonephritis and histiocytic sarcoma. **a**, Kaplan–Meyer survival curves for control (blue line, WT and $FasL^{\Delta m/WT}$ combined), $FasL^{\Delta m/\Delta m}$ (red line) and $FasL^{gld/gld}$ mice (purple line). $P < 0.0001$ for control versus $FasL^{gld/gld}$; $P < 0.0001$ for control versus $FasL^{\Delta m/\Delta m}$; $P < 0.0001$ for $FasL^{\Delta m/\Delta m}$ versus $FasL^{gld/gld}$. **b**, Incidence of severe autoimmune kidney disease (glomerulonephritis score ≥ 3) in $FasL^{\Delta m/\Delta m}$ (red line) and $FasL^{gld/gld}$ (purple line) mice. $P < 0.0175$ for $FasL^{\Delta m/\Delta m}$ versus $FasL^{gld/gld}$. **c**, H&E-stained sections of kidneys from mice of the indicated genotypes and ages were examined for pathological changes, such as hypercellularity, cellular crescents, dilated tubules or sclerotic glomeruli and the percentage of total glomeruli affected scored on a scale of 0 to 4 (0 = normal, 1 = 0–25%, 2 = 25–50%, 3 = 50–75%, 4 = >75% affected; scores indicated in parentheses). **d**, Incidence of histiocytic sarcoma (percentage incidence in parentheses) in $FasL^{\Delta m/\Delta m}$, $FasL^{gld/gld}$ and control WT mice. $^{**}P = 0.03$ for $FasL^{\Delta m/\Delta m}$ versus $FasL^{gld/gld}$ (ref. 27). **e**, Pictures of two $FasL^{\Delta m/\Delta m}$ mice with histiocytic sarcoma (aged 47 and 60 weeks). Macroscopic lesions are readily seen in the livers and spleens. **f**, H&E-stained sections of liver, spleen and lung of histiocytic sarcoma-burdened $FasL^{\Delta m/\Delta m}$ mice aged 47 weeks (leftmost 'liver' panel) and 60 weeks ('lung', 'spleen' and rightmost 'liver' panel). Asterisks indicate histiocytic lesions. Scale bars in **c** and **f** represent 155 μm.

$FasL^{\Delta m/\Delta m}$ mice developed hepatic tumours with deposits in the spleen and lungs (Fig. 4d, e). Microscopically and by immuno-phenotype ($Mac\text{-}1^+Mac\text{-}2^+F4/80^+B220^-Thy\text{-}1^-CD3^-$) these tumours were characteristic of histiocytic sarcoma (Fig. 4e, f and Supplementary Fig. 11), being composed of oval cells with eosinophilic cytoplasm and elongated or folded nuclei[26]. These tumours were transplantable in C57BL/6 mice (Supplementary Fig. 11), confirming their malignant status. Histiocytic sarcoma was rarely seen in $FasL^{\Delta s/\Delta s}$ or $FasL^{gld/gld}$ animals and in C57BL/6 (WT) mice such tumours are observed only at very low frequency late in life (~5% >78 weeks[26,27]; Fig. 4d and Supplementary Fig. 6).

Our findings that mFasL but not sFasL is critical for AICD of T cells *in vitro* and for prevention of lymphadenopathy, hyper-gammaglobulinaemia and accumulation of auto-antibodies within the whole animal are consistent with the notion that repeated T-cell antigen receptor (TCR) stimulation kills chronically activated T cells that are specific for auto-antigens or persistent pathogens through FasL/Fas signalling, thereby preventing lymphadenopathy[1–4]. The observation that $FasL^{\Delta m/\Delta m}$ mice develop SLE-like glomerulonephritis and histiocytic sarcoma considerably earlier and with higher incidence than $FasL^{gld/gld}$ mice indicates that the high levels of sFasL produced in the $FasL^{\Delta m/\Delta m}$ mice may promote autoimmunity and tumorigenesis (because unlike sFasL from $FasL^{gld/gld}$ mice, sFasL from $FasL^{\Delta m/\Delta m}$ mice can engage its receptor Fas[12,13]). sFasL from $FasL^{\Delta m/\Delta m}$ mice may achieve this by triggering non-apoptotic signalling pathways, such as NF-κB-dependent inflammatory processes. Alternatively, differences between $FasL^{gld/gld}$ and $FasL^{\Delta m/\Delta m}$ mice may be due to the fact that retrograde signalling through FasL[28,29] can only occur in $FasL^{gld/gld}$ mice but not in $FasL^{\Delta m/\Delta m}$ mice, although upon challenge

with influenza virus *in vivo* or stimulation with suboptimal doses of anti-CD3 antibodies *in vitro*, $CD8^+$ T-cell responses were indistinguishable in WT, $FasL^{gld/gld}$ and $FasL^{\Delta m/\Delta m}$ mice (Supplementary Figs 12 and 13). It is theoretically also possible that $FasL^{\Delta m/\Delta m}$ mice die earlier than $FasL^{gld/gld}$ mice because the $FasL^{\Delta m}$ mutation causes complete loss of function whereas $FasL^{gld}$ represents a partial loss-of-function mutation. For two reasons this appears unlikely: (1) lymphadenopathy and hyper-gammaglobulinaemia occur in the two mutant strains ($FasL^{\Delta m/\Delta m}$ and $FasL^{gld/gld}$) with comparable kinetics and magnitude (Fig. 3), indicating that the two mutations do not differ markedly in their potency, and (2) histiocytic sarcoma has not been reported in FasL knockout mice[30].

We therefore hypothesize that tumorigenesis may be driven by a combination of loss of mFasL-mediated apoptosis of cells undergoing transformation and sFasL/Fas-induced non-apoptotic signals, perhaps NF-κB-mediated stimulation of cell proliferation, survival or inflammation within an elevated cytokine milieu.

METHODS SUMMARY

Generation of FasL mutant mice. The mouse *fasl* locus and known restriction sites were used to construct the targeting vector and diagnose homologous recombination in embryonic stem cells and gene-targeted ($FasL^{\Delta m/\Delta m}$, $FasL^{\Delta s/\Delta s}$) mice. Targeting knock-in vectors were made with the *loxP/pGKNeo/loxP* cassette cloned into the Pac1 site. Targeting constructs for the mutant FasL mice (Supplementary Fig. 2) were linearized and electroporated into C57BL/6-derived Bruce-4 embryonic stem cells.

Analysis of FasL mutant mice. All experiments with mice were performed according to the guidelines of the Animal Ethics committees of our institutions. Mice were killed at 6, 12 or 20 weeks for analysis and further cohorts were monitored daily for morbidity and killed when showing signs of illness. Tissues were fixed for microscopic analysis in 80% Histochoice (Amresco)/20% ethanol or 10% buffered formalin and embedded in paraffin; conventional histopathology was performed on haematoxylin plus eosin (H&E)-stained sections. For detailed methods of immunohistochemical staining, immunofluorescent staining, confocal microscopy, cell preparation, flow cytometric analysis, ELISA, AICD, chromatography, viral infection and target cell killing, T-cell proliferation assays and western blotting, see the online-only Methods.

Statistical analysis was performed using Student's *t*-test, log rank (Mantel–Cox) test for survival curves or one-way analysis of variance using Turkey's comparison test to compare multiple groups where appropriate.

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

Received 8 July; accepted 14 August 2009.

1. Hughes, P. D. *et al.* Apoptosis regulators Fas and Bim cooperate in shutdown of chronic immune responses and prevention of autoimmunity. *Immunity* **28,** 197–205 (2008).
2. Hutcheson, J. *et al.* Combined deficiency of proapoptotic regulators Bim and Fas results in the early onset of systemic autoimmunity. *Immunity* **28,** 206–217 (2008).
3. Weant, A. E. *et al.* Apoptosis regulators Bim and Fas function concurrently to control autoimmunity and CD8+ T cell contraction. *Immunity* **28,** 218–230 (2008).
4. Krammer, P. H. CD95's deadly mission in the immune system. *Nature* **407,** 789–795 (2000).
5. Peter, M. E. *et al.* The CD95 receptor: apoptosis revisited. *Cell* **129,** 447–450 (2007).
6. Watanabe-Fukunaga, R. *et al.* The cDNA structure, expression, and chromosomal assignment of the mouse Fas antigen. *J. Immunol.* **148,** 1274–1279 (1992).
7. Takahashi, T. *et al.* Generalized lymphoproliferative disease in mice, caused by a point mutation in the Fas ligand. *Cell* **76,** 969–976 (1994).
8. Rieux-Laucat, F. *et al.* Mutations in Fas associated with human lymphoproliferative syndrome and autoimmunity. *Science* **268,** 1347–1349 (1995).
9. Fisher, G. H. *et al.* Dominant interfering Fas gene mutations impair apoptosis in a human autoimmune lymphoproliferative syndrome. *Cell* **81,** 935–946 (1995).
10. Trambas, C. M. & Griffiths, G. M. Delivering the kiss of death. *Nature Immunol.* **4,** 399–403 (2003).
11. Schulte, M. *et al.* ADAM10 regulates FasL cell surface expression and modulates FasL-induced cytotoxicity and activation-induced cell death. *Cell Death Differ.* **14,** 1040–1049 (2007).
12. Schneider, P. *et al.* Conversion of membrane-bound Fas(CD95) ligand to its soluble form is associated with downregulation of its proapoptotic activity and loss of liver toxicity. *J. Exp. Med.* **187,** 1205–1213 (1998).
13. Tanaka, M., Itai, T., Adachi, M. & Nagata, S. Downregulation of Fas ligand by shedding. *Nature Med.* **4,** 31–36 (1998).
14. Hohlbaum, A. M., Moe, S. & Marshak-Rothstein, A. Opposing effects of transmembrane and soluble Fas ligand expression on inflammation and tumor cell survival. *J. Exp. Med.* **191,** 1209–1219 (2000).
15. Siegel, R. M. *et al.* Fas preassociation required for apoptosis signaling and dominant inhibition by pathogenic mutations. *Science* **288,** 2354–2357 (2000).
16. Chen, J.-J., Sun, Y. & Nabel, G. J. Regulation of the proinflammatory effects of Fas ligand (CD95L). *Science* **282,** 1714–1717 (1998).
17. Wilson, N. S., Dixit, V. & Ashkenazi, A. Death receptor signal transducers: nodes of coordination in immune signaling networks. *Nature Immunol.* **10,** 348–355 (2009).
18. Ahn, J. H. *et al.* Non-apoptotic signaling pathways activated by soluble Fas ligand in serum-starved human fibroblasts. Mitogen-activated protein kinases and NF-κB-dependent gene expression. *J. Biol. Chem.* **276,** 47100–47106 (2001).
19. Xiao, S., Jodo, S., Sung, S. S., Marshak-Rothstein, A. & Ju, S. T. A novel signaling mechanism for soluble CD95 ligand. Synergy with anti-CD95 monoclonal antibodies for apoptosis and NF-κB nuclear translocation. *J. Biol. Chem.* **277,** 50907–50913 (2002).
20. Oliveira, J. B. *et al.* NRAS mutation causes a human autoimmune lymphoproliferative syndrome. *Proc. Natl Acad. Sci. USA* **104,** 8953–8958 (2007).
21. Kono, D. H. & Theofilopoulos, A. N. Genetic susceptibility to spontaneous lupus in mice. *Curr. Dir. Autoimmun.* **1,** 72–98 (1999).
22. Watanabe, D., Suda, T., Hashimoto, H. & Nagata, S. Constitutive activation of the Fas ligand gene in mouse lymphoproliferative disorders. *EMBO J.* **14,** 12–18 (1995).
23. Ogasawara, J. *et al.* Lethal effect of the anti-Fas antibody in mice. *Nature* **364,** 806–809 (1993).
24. Huang, D. C. *et al.* Activation of Fas by FasL induces apoptosis by a mechanism that cannot be blocked by Bcl-2 or Bcl-x_L. *Proc. Natl Acad. Sci. USA* **96,** 14871–14876 (1999).
25. Strand, S. *et al.* Lymphocyte apoptosis induced by CD95 (APO-1/Fas) ligand-expressing tumor cells—a mechanism of immune evasion? *Nature Med.* **2,** 1361–1366 (1996).
26. Fredrickson, T. N. & Harris, A. W. *Atlas of Mouse Hematopathology* (Harwood Academic, 2000).
27. Frith, C. H. & Wiley, L. D. Morphologic classification and correlation of incidence of hyperplastic and neoplastic hematopoietic lesions in mice with age. *J. Gerontol.* **36,** 534–545 (1981).
28. Suzuki, I. & Fink, P. J. Maximal proliferation of cytotoxic T lymphocytes requires reverse signaling through Fas ligand. *J. Exp. Med.* **187,** 123–128 (1998).
29. Sun, M. & Fink, P. J. A new class of reverse signaling costimulators belongs to the TNF family. *J. Immunol.* **179,** 4307–4312 (2007).
30. Karray, S. *et al.* Complete loss of Fas ligand gene causes massive lymphoproliferation and early death, indicating a residual activity of gld allele. *J. Immunol.* **172,** 2118–2125 (2004).

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

Acknowledgements We thank Genentech, in particular A. Ashkenazi and S. Masters, for Fas-Fc fusion proteins. We also thank our colleagues at the Walter and Eliza Hall Institute of Medical Research (Melbourne), G. Siciliano, N. Iannarella, J. Coughlin for animal care; J. Corbin for automated blood analysis; B. Helbert and C. Young for genotyping; S. Mihajlovic, E. Tsui, A. Hasanein, V. Babo and K. Weston for histological sections; A. Light and K. O'Donnell for help with antibody measurements; and S. Drake for help with cytokine quantification. We also thank J. Melny (Royal Melbourne Hospital) for measuring anti-DNA antibody levels; A. Banerjee, S. Gerondakis and R. Gugasyan (Burnet Medical Research Institute, Prahran) for antibodies and advice; J. Silke and L. Wong (La Trobe University, Bundoora) for TNFα; P. Morgan for assistance with protein purification; and A. Silva and J. Sharkey for help with animal procedures (Peter MacCallum Cancer Centre). This work was supported in part by the NHMRC (programme numbers 461221 and 454569, a CJ Martin fellowship to N.M.H. and an RD Wright fellowship to L.A.O'R.). We also acknowledge support from the IRIISS (grant numbers 361646 and 257502 to P.B.), the Victorian State Government (an OIS grant), the Leukemia and Lymphoma Society (SCOR grant number 7015), the NIH (grant numbers CA043540-18 and CA80188-6), the JDRF/NHMRC, the Association for International Cancer Research, the Charles and Sylvia Viertel Charitable Foundation (to P.B.) and the Leukemia Research Foundation.

Author Contributions L.A.O'R. planned and performed most experiments and wrote the manuscript. L.T., L.L., E.A.K., S.G., W.D.F., N.M.H., D.M.T., J.-G.Z., G.T.B., M.J.S., P.B. and L.R. contributed to the planning and execution of the experiments and the writing of the manuscript. A.S. conceived the study, planned experiments and 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 A.S. (strasser@wehi.edu.au).

METHODS

Generation of FasL mutant mice. Knowledge of the mouse *fasl* locus and restriction sites was used to construct the targeting vectors and to set up Southern blotting and PCR strategies to diagnose homologous recombination in embryonic stem cells and gene-targeted $FasL^{\Delta m/\Delta m}$, $FasL^{\Delta s/\Delta s}$ mice. Gene targeting knock-in vectors were constructed with the *loxP/pGKNeo/loxP* cassette cloned into the Pac1 site. Targeting constructs (Supplementary Fig. 1) were linearized and electroporated into C57BL/6-derived Bruce-4 embryonic stem cells according to standard procedures. The *neo* cassette was removed by crossing heterozygous gene-targeted mice with *Cre* deleter mice[31]. For each of the mutations two or three independent strains were established on an inbred C57BL/6 background: $FasL^{\Delta s/\Delta s}$ strains del494, del510 and del511 and $FasL^{\Delta m/\Delta m}$ strains del478 and del480. Independent strains with the same mutation were found to have identical phenotypes.

Immunohistochemical and immunofluorescent staining, confocal microscopy. Immunohistochemical staining was performed as previously described[32] using antibodies to S100 (Neomarkers), CD3** (Abcam), B220 (RA3-6B2), Mac-2 (M3/84.6.34), F4/80*, lysozyme (Abcam), and NF-κB p65** (sc-109, Santa Cruz) using optimized antigen recovery methods (*proteinase-K treatment; **microwave retrieval). Secondary staining was performed with biotinylated goat anti-rabbit IgG antibodies (Vector) or rat IgG-isotype-specific antibodies (Pharmingen), followed by incubation with the biotin-avidin-diaminobenzine-based detection system (Vector-Elite standard). In all cases positive (spleen, lymph node, melanoma) and negative controls (isotype-matched control Abs or normal serum) were used to verify the reliability of the results. The results of the immunohistochemically stained sections were graded semiquantitatively as described[33]: + = 75–100% of the neoplastic cells positive; +/− = 50–75 of the neoplastic cells positive; −/+ = 25–50% of the neoplastic cells positive; rare = 10–25% of the neoplastic cells positive; − = <10% of the neoplastic cells positive.

To stain for immune complex deposits, kidneys were snap-frozen in isopentane, sectioned, acetone-fixed, blocked with PBS/2% FCS, followed by staining with FITC-coupled goat antibodies specific to mouse IgM, IgG or IgA (Southern Biotechnology), blocked with mouse serum and then stained with anti-C1q antibodies (Abcam) followed by incubation with biotinylated mouse anti-rat Igκ light chain monoclonal antibody (MAR 18.5), followed by staining with Cy5-conjugated avidin (Amersham Pharmacia Biotech) in PBS/2% FCS with DAPI (to stain nuclei).

Anti-nuclear auto-antibodies in sera of mice were detected by staining using 1/100 serum dilutions of slides coated with HEp-2 human epithelial cells according to the manufacturer's instructions (Immuno Concepts). Anti-nuclear auto-antibodies levels were semi-quantitated according to brightness of fluorescence intensity on a scale of 0 (no fluorescence) to 3+ (maximum fluorescence intensity).

To detect cytoplasmic FasL in activated T lymphocytes, cells were fixed with ice-cold methanol as previously described[34] and stained for FasL using MFL-3 (BD Pharmingen) or FLIM58 (MBL) monoclonal antibodies or, as a negative control, with a hamster IgG1/κ Ig isotype-matched control monoclonal antibodies (BD Biosciences). Secondary staining was performed with biotinylated goat anti-Armenian Hamster IgG antibodies (Jackson ImmunoResearch) and detection with FITC-conjugated streptavidin (Caltag). All immunofluorescent staining was visualized by confocal microscopy (Leica).

Cell preparation and flow cytometric analysis. Single cell suspensions were prepared from histiocytic sarcoma (tumour)-burdened livers or spleens by digestion for 20 min at room temperature (22 °C) with collagenase-DNAase followed by treatment with ethylene-diaminetetraacetic acid (EDTA) as described previously[35]. Aliquots of cell suspensions (2×10^6) were injected intravenously into each of three (non-irradiated) 6–8-week-old C57BL/6 mice. Histiocytic sarcoma cells were cultured in the high glucose version of Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 50 μM 2-mercaptoethanol, 13 μM folic acid, 100 μM L-asparagine plus L929 fibroblast supernatant (5–10%). Cultured histiocytic sarcoma cells were trypsinized, washed, resuspended in PBS and 1×10^6 live tumour cells injected intravenously into C57BL/6 mice.

Single cell suspensions of spleen, lymph nodes, peripheral blood and liver were stained with FITC-or R-PE-conjugated surface marker specific monoclonal antibodies (RB6-8C5: anti-Gr-1, MI/70: anti-Mac-1, M3/84.6.34: anti-Mac-2, Ter119: anti-erythroid cell surface marker, F4/80: anti-macrophage surface marker, T24.31.2: anti-Thy-1, GK1.5: anti-CD4, 53.6.72: anti-CD8, RA3-6B2: anti-CD45R-B220, PK136: anti-NK1.1 (BD Biosciences)) plus the vital dye PI (1 μg ml^{-1}) and analysed in a FACScan (Becton Dickinson). Surface Fas expression was detected by pre-blocking with anti-FcγR antibody (24G2) + 5% normal goat serum and staining with biotinylated anti-Fas monoclonal antibody Jo2 or a biotinylated isotype-matched control hamster IgG antibody (BD Biosciences) followed by incubation with R-PE-conjugated streptavidin (Caltag). Surface FasL expression was detected by pre-blocking with anti-FcγR antibody (24G2) + 5% normal rat serum and staining with FLAG-tagged FasL (Alexis) or FLAG-tagged $FasL^{\Delta m}$ plus biotinylated anti-FLAG antibody or a biotinylated Ig isotype-matched control antibody, followed by detection with R-PE-conjugated streptavidin (Caltag).

ELISA. Serum immunoglobulin concentrations were determined by ELISA as described previously[36], using sheep anti-mouse Ig antibodies (Silenus Laboratories) as a capture reagent, and developed with mouse Ig isotype-specific goat antibodies that had been conjugated with horseradish peroxidase (Southern Biotechnology). Purified myeloma proteins were used as standards (Sigma). Cytokine concentrations in sera were measured by ELISA using the appropriately matched pairs of antibodies specific to IL-6, IFNγ, TNFα (BD Pharmingen) or FasL (R&D Systems, FasL Duoset kit) following the manufacturers' instructions.

AICD and target cell killing. Lymph node or spleen cells were mitogenically stimulated for 3 days in culture with concanavalin A (2 μg ml^{-1}, Pharmacia) plus IL-2 (100 IU ml^{-1}), washed in medium and then cultured in IL-2 for a further 2 days. The resulting activated T lymphoblasts were then restimulated in the presence of IL-2 for 6 h with PMA (2 ng ml^{-1}) plus ionomycin (2 μg ml^{-1}) or with monoclonal antibodies to CD3 (145-2C11) and CD28 (37N51) (coated on culture plates at 10 μg ml^{-1} in PBS).

Activated T cells were washed and used as effector cells at effector:target ratios of 50:1 to 3.125:1. As target cells we used the FasL-sensitive mouse CH1 B lymphoma cells (2×10^4 cells in 200 μl in 96-well plates) that had been labelled 24 h earlier with CFSE (Molecular Probes) as described[37]. After 24 h of co-culture, CH1 cell survival was determined by staining with PI, followed by FACS analysis with $CFSE^+PI^-$ cells being considered as live CH1 target cells. Killing with FLAG-tagged FasL (10 ng ml^{-1}, Alexis) crosslinked with anti-FLAG M2 monoclonal antibody (1 μg ml^{-1}, Sigma) was used as positive control for CH1 cell killing.

For AICD, activated T lymphoblasts were prepared as described above. These cells were washed in medium and re-plated at 10^5 cells ml^{-1} on tissue culture plates that had been coated with anti-CD3 monoclonal antibody (145-2C11; 10 μg ml^{-1}) or cultured in IL-2 alone. T-cell viability was determined after 6, 12 or 24 h by staining with PI plus FITC-conjugated annexin V.

The effects of Fas–Fc fusion proteins (100 ng ml^{-1}; gift from A. Ashkenazi and S. Masters, Genentech) or an Ig isotype-matched mouse monoclonal antibody (Bcl-2-100) on inhibition of CH1 target cell killing or AICD were assessed by adding these reagents to the cultures 30 min before the start of the assay.

Size-exclusion chromatography. The supernatants of activated T lymphocytes were concentrated 25 times and 200 μl aliquots loaded onto a column (Superdex-200 HR10/30; Pharmacia). Proteins were eluted in PBS at 0.5 ml min^{-1} and fractions analysed by a FasL-specific ELISA (see above). The column was calibrated with standard proteins: thyroglobulin (669 kDa), ferritin (Fer, 440 kDa), bovine serum albumin (67 kDa), ovalbumin (45 kDa) and RNase A (RibA, 13.7 kDa).

Viral infections and measurements of anti-viral $CD8^+$ T-cell responses. Mice (6–8 weeks old) were inoculated intra-nasally with $10^{4.5}$ plaque-forming units of the HKx31 (H3N2, X31)[38] influenza virus. Virus stocks were grown in the allantoic cavity of 10-day-old embryonated hen's eggs and stored in aliquots at −80 °C. Anti-viral T-cell immune responses were enumerated by staining cells from spleens or lymph nodes with phycoerythrin (R-PE)-labelled MHC class I/peptide tetrameric complexes specific for the H-$2D^b$-restricted nucleoprotein ($D^bNP_{367–374}$) epitope of influenza virus[39,40] plus antibodies to CD8. To measure antigen-specific cytokine expression, $CD8^+$ T cells were stimulated in culture in the presence of 1 μM $D^bNP_{367–374}$ peptide and 1 μg ml^{-1} Brefeldin A (Epicentre Biotechnologies) for 5 h. Cells were then stained for CD8α, fixed and permeabilized using the Cytofix/Cytoperm reagent (BD Biosciences) and co-stained for intracellular IFNγ or TNFα. Intracellular granzyme B (Caltag) content of cells was measured after stimulation of the cells with 50 ng ml^{-1} PMA and 500 ng ml^{-1} ionomycin for 5 h (with 1 μg ml^{-1} Brefeldin A present for the last 2 h) and subsequent surface marker staining. Analysis was performed on an LSR II (BD Biosciences) and data analysed using Weasel flow cytometry analysis software.

Cultures of NP-specific $CD8^+$ T cells and ^{51}Cr release assay. Spleen cells from mice infected intra-nasally with HKx31 influenza virus 9 days earlier were cultured for 5 days with 2×10^7 γ-irradiated (1,000 rad) $NP_{367-374}$-coated C57BL/6 spleen cells plus 10 IU ml^{-1} recombinant IL-2. Cytotoxicity was assessed by conventional ^{51}Cr-release assay using as targets EL4 cells (H-2^b) pulsed with 1 μM $NP_{367–374}$ (ref. 41).

T-cell proliferation assay. Allogeneic $CD8^+$ T-cell responders were seeded on plates coated with suboptimal amounts of anti-CD3 antibodies (145-2C11) that were also coated with either murine Fas-Fc IgG fusion proteins or an Ig isotype-matched control mouse monoclonal antibody (Bcl-2-100) as previously

described[28]. [3H]TdR uptake, a measure of proliferation of the allogeneic $CD8^+$ T cells, was determined as described[28].

Production and purification of FLAG-tagged $FasL^{\Delta m}$ protein. Recombinant sFasL used in this study, FLAG-tagged $FasL^{\Delta m}$, was prepared from culture supernatants of 293T cells that had been transiently transfected using lipofectamine with an expression vector encoding FLAG-tagged $FasL^{\Delta m}$ 72 h earlier. Filtered supernatant was applied to a 2 ml anti-FLAG M2 antibody conjugated agarose column (Sigma). After extensive washing with PBS, bound proteins were eluted with 100 μg ml^{-1} FLAG peptide in PBS (Auspep). The fractions containing the recombinant proteins were concentrated to 0.5 ml using a Millipore Ultrafree-MC centrifugal filter unit with a regenerated cellulose membrane (molecular weight cut-off of 10 kDa) and subjected to size-exclusion chromatography for further purification using a pre-packed Superdex 200 column (GE Healthcare; 300 mm × 10 mm internal diameter) operated at 0.4 ml min^{-1} in PBS. The column was calibrated with Bio-Rad standard containing thyroglobulin dimer (670 kDa), bovine γ-globulin (158 kDa), chicken ovalbumin (44 kDa), equine myoglobin (17.5 kDa) and vitamin B12 (1.35 kDa). Fractions (0.4 ml) containing FLAG-tagged $sFasL^{\Delta m}$, as determined by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS–PAGE) and western blot analysis (using anti-FLAG monoclonal antibody), were pooled and the protein concentration quantified.

Cell culture and western blot analysis for NF-κB pathway activation. NIH-3T3 cells were seeded at 5×10^5 per flask in DMEM containing 10% FCS, allowed to adhere overnight, washed and then cultured for 24 h in DMEM containing only 0.25% FCS[18]. The T-cell hybridoma 3DO.54.3, which are resistant to FasL/Fas-induced apoptosis, were cultured in the high glucose version of DMEM supplemented with 10% fetal bovine serum, 50 μM 2-mercaptoethanol, 13 μM folic acid, 100 μM L-asparagine. Cell were stimulated for 1 h or 5 h with TNFα (10 ng ml^{-1}), FLAG-tagged FasL (100 ng ml^{-1}) (Alexis) plus anti-FLAG M2 antibodies (1–2 μg ml^{-1}; Sigma) or FLAG-tagged $mFasL^{\Delta m}$ (0.2 or 1 μg ml^{-1}; without addition of anti-FLAG antibodies). Cell lysates were prepared in RIPA lysis buffer, proteins were size-fractionated on polyacrylamide gels (Novex) and transferred to nitrocellulose membranes (Amersham Pharmacia). Non-specific binding of antibodies to membranes was blocked by incubation in PBS, 5% skimmed milk, 1% casein, 0.05% Tween-20. Membranes were then probed with the following antibodies: rabbit anti-phospho-IκBα (Ser32), rabbit anti-total IκBα (all from Cell Signaling), rat anti-caspase-8 (clone 10B12, Alexis), mouse anti-β-actin (Sigma) or mouse anti-FLAG M2 (Sigma). Bound antibodies were visualized with goat anti-rat IgG (Southern Biotechnology), sheep anti-rabbit Ig (Chemicon) or sheep anti-mouse IgG antibodies (Chemicon), all conjugated to HRP, followed by enhanced chemiluminescence (Amersham Pharmacia).

31. Schwenk, F., Baron, U. & Rajewsky, K. A *cre*-transgenic mouse strain for the ubiquitous deletion of *loxP*-flanked gene segments including deletion in germ cells. *Nucleic Acids Res.* **23,** 5080–5081 (1995).
32. Phillips, J. M., O'Reilly, L., Bland, C., Foulis, A. K. & Cooke, A. Patients with chronic pancreatitis have islet progenitor cells in their ducts, but reversal of overt diabetes in NOD mice by anti-CD3 shows no evidence for islet regeneration. *Diabetes* **56,** 634–640 (2007).
33. Pileri, S. A. *et al.* Tumours of histiocytes and accessory dendritic cells: an immunohistochemical approach to classification from the International Lymphoma Study Group based on 61 cases. *Histopathology* **41,** 1–29 (2002).
34. Blott, E. J., Bossi, G., Clark, R., Zvelebil, M. & Griffiths, G. M. Fas ligand is targeted to secretory lysosomes via a proline-rich domain in its cytoplasmic tail. *J. Cell Sci.* **114,** 2405–2416 (2001).
35. Naik, S., Vremec, D., Wu, L., O'Keeffe, M. & Shortman, K. CD8α+ mouse spleen dendritic cells do not originate from the CD8 α-dendritic cell subset. *Blood* **102,** 601–604 (2003).
36. Hibbs, M. L. *et al.* Multiple defects in the immune system of *Lyn*-deficient mice, culminating in autoimmune disease. *Cell* **83,** 301–311 (1995).
37. Lyons, A. B. & Parish, C. R. Determination of lymphocyte division by flow cytometry. *J. Immunol. Methods* **171,** 131–137 (1994).
38. Kilbourne, E. D. Future influenza vaccines and the use of genetic recombinants. *Bull. WHO* **41,** 643–645 (1969).
39. Flynn, K. J. *et al.* Virus-specific $CD8^+$ T cells in primary and secondary influenza pneumonia. *Immunity* **8,** 683–691 (1998).
40. Belz, G. T., Xie, W., Altman, J. D. & Doherty, P. C. A previously unrecognized H-2D(b)-restricted peptide prominent in the primary influenza A virus-specific CD8(+) T-cell response is much less apparent following secondary challenge. *J. Virol.* **74,** 3486–3493 (2000).
41. Belz, G. T., Xie, W. & Doherty, P. C. Diversity of epitope and cytokine profiles for primary and secondary influenza a virus-specific CD8+ T cell responses. *J. Immunol.* **166,** 4627–4633 (2001).

Vol 461|1 October 2009|doi:10.1038/nature08397

LETTERS

Cooperative binding of two acetylation marks on a histone tail by a single bromodomain

Jeanne Morinière[1,2], Sophie Rousseaux[3], Ulrich Steuerwald[4], Montserrat Soler-López[1,2]†, Sandrine Curtet[3], Anne-Laure Vitte[3], Jérôme Govin[3], Jonathan Gaucher[3], Karin Sadoul[3], Darren J. Hart[1,2], Jeroen Krijgsveld[4], Saadi Khochbin[3], Christoph W. Müller[4] & Carlo Petosa[5]

A key step in many chromatin-related processes is the recognition of histone post-translational modifications by effector modules such as bromodomains and chromo-like domains of the Royal family[1,2]. Whereas effector-mediated recognition of single post-translational modifications is well characterized[3], how the cell achieves combinatorial readout of histones bearing multiple modifications is poorly understood. One mechanism involves multivalent binding by linked effector modules[4]. For example, the tandem bromodomains of human TATA-binding protein-associated factor-1 (TAF1) bind better to a diacetylated histone H4 tail than to monoacetylated tails, a cooperative effect attributed to each bromodomain engaging one acetyl-lysine mark[5]. Here we report a distinct mechanism of combinatorial readout for the mouse TAF1 homologue Brdt, a testis-specific member of the BET protein family[6]. Brdt associates with hyperacetylated histone H4 (ref. 7) and is implicated in the marked chromatin remodelling that follows histone hyperacetylation during spermiogenesis, the stage of spermatogenesis in which post-meiotic germ cells mature into fully differentiated sperm[7–10]. Notably, we find that a single bromodomain (BD1) of Brdt is responsible for selectively recognizing histone H4 tails bearing two or more acetylation marks. The crystal structure of BD1 bound to a diacetylated H4 tail shows how two acetyl-lysine residues cooperate to interact with one binding pocket. Structure-based mutagenesis that reduces the selectivity of BD1 towards diacetylated tails destabilizes the association of Brdt with acetylated chromatin *in vivo*. Structural analysis suggests that other chromatin-associated proteins may be capable of a similar mode of ligand recognition, including yeast Bdf1, human TAF1 and human CBP/p300 (also known as CREBBP and EP300, respectively). Our findings describe a new mechanism for the combinatorial readout of histone modifications in which a single effector module engages two marks on a histone tail as a composite binding epitope.

Bromodomains and extraterminal (BET)-family proteins (including yeast Bdf1, *Drosophila* Fsh and mammalian Brd2, Brd3 and Brd4) contain two bromodomains (BD1 and BD2) spaced ~150 residues apart and followed by an 'extra terminal' (ET) domain[6]. These proteins recognize multi-acetylated histones H3 and H4 (refs 7, 11–14) and are distinguished from other bromodomain proteins by their retention on mitotic chromosomes, implicating BET proteins in epigenetic memory[15–17]. Brdt associates with, and has the unique ability to compact, hyperacetylated chromatin *in vitro* and *in vivo*[7]. Compaction activity is ATP-independent, is enhanced by a carboxy-terminal truncation that removes the ET domain, and requires integrity of BD1 (ref. 7). Deletion of BD1 results in male sterility in mice[8].

To investigate how Brdt recognizes acetylated chromatin, we assessed binding to histone H3 and H4 tail peptides by isothermal calorimetry (ITC), using a murine construct (δC-sBrdt) with enhanced remodelling activity[7] (Fig. 1a). δC-sBrdt specifically binds a tetra-acetylated H4 peptide (H4ac4) with a dissociation constant ($K_d = 11.4 \pm 2.2\ \mu M$) similar to that reported for Bdf1 (8.4 μM)[13] and for the TAF1 double bromodomain (5.3 μM)[5] (Supplementary Fig. 1a and Table 1). We next analysed binding by the isolated Brdt bromodomains, expecting each to engage a unique acetyl-lysine residue within H4. Surprisingly, BD2 did not bind tetra- or mono-acetylated H4 peptides, whereas BD1 bound H4ac4 nearly as well ($K_d = 28.0 \pm 6.1\ \mu M$) as δC-sBrdt, suggesting that the latter recognizes H4ac4 primarily through BD1 alone (Supplementary Fig. 1b and Table 1). Recognition does not seem to involve protein oligomerization, as both BD1 and δC-sBrdt are monomeric in solution (Supplementary Fig. 2).

BD1 failed to bind monoacetylated H4 tails, but recognized four diacetylated H4 peptides (out of six tested), highlighting the cooperative nature of ligand recognition (Fig. 1b and Table 1). The affinity was highest for H4 diacetylated on K5/K8, equalling that for tetra-acetylated H4, and 5–15 times lower for the three other diacetylated H4 peptides recognized. Likewise, BD1 failed to bind monoacetylated H3 peptides, but recognized tetra-acetylated H3 (albeit more weakly than H4ac4) (Table 1). This behaviour contrasts with that of BD2, which binds H3 peptides acetylated on K18 but does not discriminate mono-, di- and tetra-acetylated forms (Table 1). BD1 also binds an H2B peptide bearing a diacetylated motif (Kac-GS-Kac) resembling that of H4K5acK8ac (Kac-GG-Kac) (Supplementary Fig. 1c). However, the affinity ($K_d = 157 \pm 15\ \mu M$) is eight times lower than for H4K5acK8ac, indicating that sequence context, and not merely spacing between acetylation marks, is important for optimal binding.

To verify the physiological relevance of the high-affinity ligand of BD1 (H4K5acK8ac), we performed immunoblot analysis on fractionated mouse spermatogenic cells using antibodies raised against K5-, K8-, K12-, K16- or tetra-acetylated H4 tails, and additionally assessed H4 acetylation by immunofluorescence microscopy. In rat, chicken and other species, the wave of histone hyperacetylation associated with spermiogenesis is known to occur during spermatid elongation, just before histone replacement and chromatin compaction[9]. Consistent with this, we found that K16 acetylation was constitutive, whereas K5, K8 and K12 gained acetylation during post-meiotic stages in elongating spermatids (Fig. 1c, d). Tandem mass spectrometry confirmed

[1]European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, BP 181, 38042 Grenoble Cedex 9, France. [2]Unit of Virus Host-Cell Interactions, UMI 3265 Université Joseph Fourier-EMBL-CNRS, 6 rue Jules Horowitz, BP 181, 38042 Grenoble Cedex 9, France. [3]INSERM, U823; Université Joseph Fourier, Institut Albert Bonniot, F-38700 Grenoble, France. [4]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [5]Institut de Biologie Structurale Jean-Pierre Ebel, UMR 5075 CEA-CNRS-Université Joseph Fourier, 41 Jules Horowitz, 38027 Grenoble Cedex 1, France. †Present address: Institute for Research in Biomedicine (IRB Barcelona), Parc Científic de Barcelona, calle Baldiri Reixac 10–12, 08028 Barcelona, Spain.

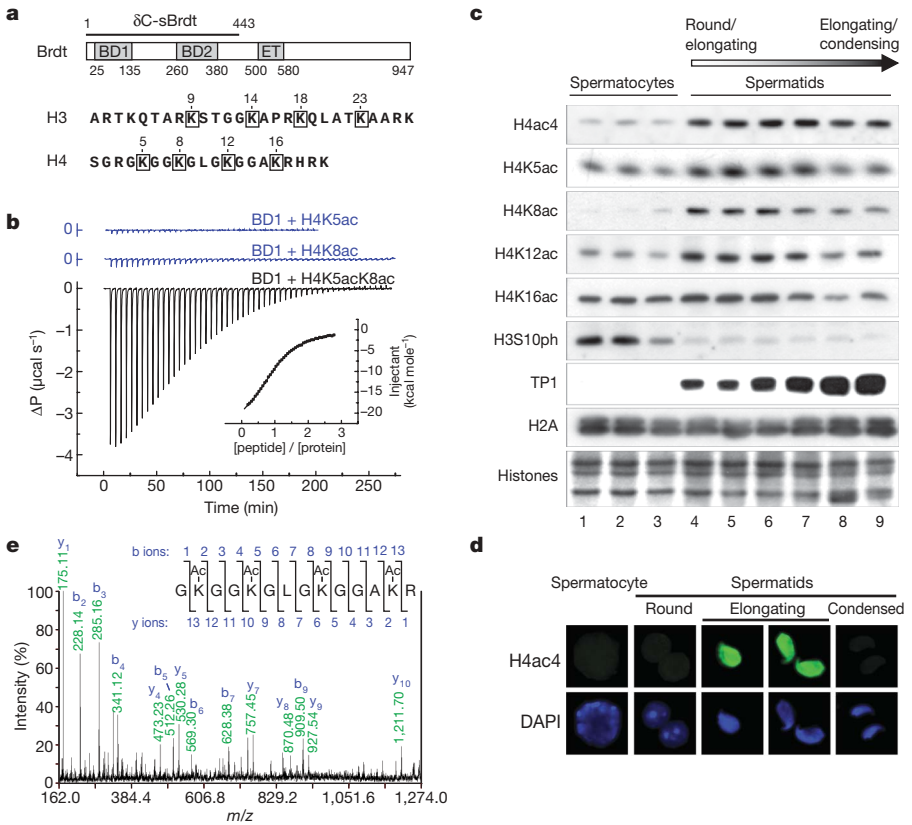


Figure 1 | Brdt recognizes a diacetylated H4 peptide through BD1. **a**, Brdt domain organization and H3 and H4 tail sequences. **b**, ITC profiles showing the binding of BD1 to H4 peptides monoacetylated on K5 or K8 or diacetylated on K5 and K8. Protein (150 μM) was titrated with 2 mM peptide. ΔP, differential power. **c**, Western blot showing the occurrence of H4 hyperacetylation in post-meiotic spermatogenic cells. Extracts prepared from spermatogenic cells as previously described[23] were immunoblotted with antibodies against the indicated species. Anti-H2A and anti-histone blots serve as loading controls. Histone H3 phosphorylated on Ser 10 (H3S10ph) and transition protein 1 (TP1) were used as specific markers for meiotic and post-meiotic cells, respectively[9]. **d**, Immunofluorescence microscopy. Mouse spermatogenic cells at the indicated stages of development were stained by DAPI (4,6-diamidino-2-phenylindole) and by an antibody against tetra-acetylated H4 tails. Original magnification, ×380. **e**, Tandem mass spectrometry (MS) analysis confirms the presence of histone H4 tetra-acetylated on lysines 5, 8, 12 and 16 in post-meiotic spermatids. Histone residues extracted from nuclei of murine round/elongated spermatids were separated by SDS–PAGE. The band corresponding to histone H4 was digested by trypsin and analysed by matrix-assisted laser desorption/ionization–tandem time of flight (MALDI–TOF/TOF; Applied BioSystems 4800) MS. A peak with mass M+1 = 1438.8125 corresponding to G(Kac^5)GG(Kac^8)GLG(Kac^{12})GGA(Kac^{16})R was detected and its identity confirmed by the fragmentation spectrum.

that histone H4 was indeed tetra-acetylated on lysines 5, 8, 12 and 16 in round/elongating spermatids (Fig. 1e). A semiquantitative analysis revealed that the level of tetra-acetylated H4 was more than fivefold higher in round/elongating spermatids than in spermatocytes, comparable to the increase induced in a somatic mouse cell line by treatment with the histone deacetylase inhibitor trichostatin A (TSA) (Supplementary Fig. 3). Importantly, the first defects detected in BD1 knockout mice also appear in post-meiotic elongating spermatids[8]. The selectivity of BD1 towards H4K5acK8ac *in vitro* fits neatly with these observations.

To understand how BD1 mediates cooperative binding, we solved the 2.4 Å crystal structure of BD1 bound to an H4K5acK8ac peptide. (Statistics are summarized in Supplementary Table 1.) BD1 exhibits the standard bromodomain fold[18]: a left-handed bundle of four helices (αZ, αA, αB and αC) with a peptide-binding pocket defined by a short BC and an elaborate ZA loop (Fig. 2a). BD1 aligns well with known bromodomain structures, and among ligand-bound structures most closely resembles that of yeast GCN5 (ref. 19) (yGCN5; Supplementary Table 2). BD1 recognizes H4 residues 4–10 in an orientation antiparallel to that of the yGCN5 ligand (Fig. 2b and Supplementary Figs 4 and 5a). Both acetyl-lysine side chains point into the BD1-binding pocket: K5ac extends perpendicularly from the peptide backbone, inserting deep into the pocket, while K8ac reaches diagonally towards K5ac. A water-mediated hydrogen bond between the K5ac acetyl and K8ac amide groups structurally links the two post-translational modifications (PTMs) (Fig. 2c, d). The Kac-GG-Kac tetrapeptide motif thus forms a composite binding epitope resembling a two-pronged plug, with the left prong bent towards the right.

BD1 recognition of the K5ac side chain is prototypical: compared to yGCN5, the hydrogen bond between K5ac and Asn 108, the hydrophobic contacts involving Val 55, Leu 62 and Cys 104, and the characteristic ring of five water molecules stabilized by ZA- and BC-loop carbonyl groups are all conserved (Fig. 2c, d). However, in yGCN5 a tyrosine (Tyr 413) packs against the entire length of the acetyl-lysine side chain, whereas in BD1 the corresponding isoleucine (Ile 114) contacts only the acetyl group (Supplementary Fig. 5b). As a result, the keyhole-like pocket that engages the acetyl-lysine side chain in yGCN5 is wider in BD1, accommodating both acetyl-lysines (Fig. 2b). Mutating residue Ile 114 either to tyrosine (as in yGCN5) or to tryptophan (for increased steric hindrance) abolished the binding of BD1 to H4K5acK8ac, confirming the importance of this residue in ligand recognition (Supplementary Fig. 5c and Table 1). The K8ac side chain is recognized by BD1 primarily through hydrophobic contacts: the acetyl moiety's methyl group by Leu 60, and the aliphatic lysyl moiety by residues Trp 49, Pro 50 and Met 117. A similarly oriented unmodified lysine would unfavourably bury its positive charge in this apolar environment, explaining why monoacetylated H4K5ac binds poorly. Additional specificity is due to H4 residue Leu 10, which interacts with three hydrophobic residues (Phe 47, Trp 49 and Val 116), and to extensive recognition of the peptide backbone

Table 1 | Summary of ITC data

Binding experiment	Lysines acetylated	K_d (μM)	*N**		
δC-sBrdt + H4				5 8 12 16	
(1)	None	Large†	ND		−
(2)	5/8/12/16	11.4 ± 2.2	1.01 ± 0.03		++
(3)	5/8	8.9 ± 1.2	1.03 ± 0.03		++
(4)	5/16	350 ± 45	1.01 ± 0.06		+
BD1 + H4					
(5)	None	Large	ND		−
(6)	5/8/12/16	28.0 ± 6.1	1.01 ± 0.02		++
(7)	5	Large	ND		−
(8)	8	Large	ND		−
(9)	12	Large	ND		−
(10)	16	Large	ND		−
(11)	5/8	21.9 ± 1.7	1.00 ± 0.03		++
(12)	8/12	193 ± 22	1.06 ± 0.01		+
(13)	12/16	117 ± 22	0.99 ± 0.04		+
(14)	5/12	340 ± 67	1.01 ± 0.04		+
(15)	8/16	Large	ND		−
(16)	5/16	Large	ND		−
BD1 + H2B				12 15	
(17)	12/15	157 ± 15	1.00 ± 0.04		+
BD1 + H3				9 14 18 23	
(18)	9/14/18/23	390 ± 36	1.00 ± 0.06		+
(19)	18	Large	ND		−
BD2 + H3					
(20)	None	Large	ND		−
(21)	9/14/18/23	214 ± 48	1.03 ± 0.11		+
(22)	9/14	Large	ND		−
(23)	14/18	217 ± 28	1.00 ± 0.05		+
(24)	18/23	176 ± 52	1.00 ± 0.03		+
(25)	9/18	360 ± 140	1.01 ± 0.03		+
(26)	14/23	Large	ND		−
(27)	14	Large	ND		−
(28)	18	251 ± 62	1.03 ± 0.11		+
(29)	23	Large	ND		−
BD2 + H4					
(30)	5/8/12/16	Large	ND		−
BD1-I114Y + H4					
(31)	5	Large	ND		−
(32)	5/8	Large	ND		−
BD1-I114W + H4					
(33)	5	Large	ND		−
(34)	5/8	Large	ND		−
BD1-mutBC + H4					
(35)	5	1,010 ± 50	1.00 ± 0.10		±
(36)	5/8	195 ± 23	1.00 ± 0.05		+
BD1-mutBC + H3					
(37)	18	345 ± 45	1.00 ± 0.08		+

Mean and s.d. were determined from experiments performed in triplicate. Horizontal lines separate experiments involving different protein constructs. At right, relative binding is summarized as strong (++), significant (+) or weak/non-detectable (−). ND, not determined.
* Binding stoichiometry.
† Poor or no binding detected (K_d > 1 mM).

by the BC and ZA loops, which make water-mediated and direct hydrogen bonds with K5ac, Gly 6, K8ac and Gly 9 (Fig. 2c–e). These backbone contacts would be partly disrupted in diacetylated ligands with three or more residues between acetylation marks, accounting for the lower affinity of such peptides (Table 1).

To gain additional insights we solved the 2.15 Å structure of BD2 bound to H3K18ac. BD2 interacts with H3 residues 16–22: recognition of K18ac matches that of K5ac by BD1, and further specificity is achieved through a hydrogen bond to the peptide main chain, a salt bridge to Arg 17 and hydrophobic contacts to Leu 20 (Supplementary Fig. 6). Like BD1, BD2 has a binding pocket wide enough to accommodate two acetyl-lysine side chains. However, aligning the BD1 and BD2 structures reveals a steric clash between the BC loop of BD2 and the superimposed H4K5acK8ac peptide (Supplementary Fig. 7), underscoring that pocket size does not uniquely determine diacetylated ligand recognition. We surmised that replacing the BC loop of BD1 with that of BD2 would result in a mutant (BD1-mutBC) with altered ligand selectivity. Indeed, BD1-mutBC binds to H4K5ac and H3K18ac slightly better, and to H4K5acK8ac nine times worse, than wild-type BD1, corresponding to a more than tenfold reduction in selectivity towards the diacetylated ligand (Table 1, compare rows 7, 11 and 19 with 35–37; see also Supplementary Fig. 8). We previously described two triple mutations within the ZA loops, P50A/F51A/V55A (mutPFV1) and P293A/F294A/V298A (mutPFV2), that abolish ligand binding by BD1 and BD2, respectively[7]. This loss of binding is rationalized by our crystal structures, which show that the mutated residues form direct and water-mediated contacts with the H4K5ac and H3K18ac side chains (Fig. 2c and Supplementary Fig. 6c).

To assess the functional significance of BD1-mediated H4K5acK8ac recognition, we introduced the above mutations into an haemagglutinin (HA)-tagged δC-sBrdt construct and assayed for acetylation-dependent chromatin compaction activity using TSA to induce chromatin hyperacetylation. As previously reported[7], wild-type δC-sBrdt and mutPFV2 both compacted chromatin in a TSA-dependent manner, whereas compaction was abolished for mutPFV1 (Supplementary Fig. 9, Hoechst panels). The I114Y and I114W mutants also abolished TSA-dependent compaction, whereas mutBC behaved like the wild type. These results indicate that (at least partial) ligand-binding activity by BD1 is required for compaction, whereas that by BD2 is dispensable (Fig. 3c). For further insights we introduced the mutations into a green fluorescent protein (GFP)-tagged δC-sBrdt construct and used fluorescence recovery after photobleaching (FRAP) to examine retention on hyperacetylated chromatin. In the absence of TSA, mutants and wild type behaved similarly: bleached regions rapidly recovered fluorescence, exhibiting a time to half-maximal recovery ($t_{1/2}$) of ~3 s (Fig. 3 and Supplementary Fig. 10). Addition of TSA greatly retarded recovery of the wild-type protein and of mutPFV2 ($t_{1/2}$ ~20 s), indicating significant retention on acetylated chromatin. In contrast, TSA had no effect on mutants with BD1 domains completely defective for ligand binding (mutPFV1, I114Y/W), and had an intermediate effect ($t_{1/2}$ ~8 s) for mutBC, whose BD1 domain has reduced affinity for H4K5acK8ac (Table 1). The recovery time thus mirrors the binding affinity of BD1 towards H4K5acK8ac (last two columns in Fig. 3c), highlighting the importance of this interaction for the stability of Brdt on acetylated chromatin *in vivo*.

Superimposing our BD1–ligand complex onto known bromodomain structures revealed that several bromodomains were sterically compatible with the superimposed K5ac-GG-K8ac tetrapeptide (Supplementary Fig. 11). This has possible implications for the following proteins.

First, other BET proteins. The BD1 and BD2 domains are highly conserved among BET proteins (for example, Brd2-BD1 is 70% identical to Brdt-BD1 and preserves all contacts with the K5ac-GG-K8ac tetrapeptide when structurally aligned with Brdt-BD1). Moreover, the ligand specificity of many BET proteins resembles that of Brdt-BD1: Brd2 and Brd4 both bind to tetra-acetylated H4 and to an H4 peptide diacetylated on K5/K12, but not on K8/K16 (refs 16, 20), and Bdf1 binds hyperacetylated H3 and H4 tails, associating more stably with the latter[13,14]. Thus, BD1-mediated diacetylated ligand recognition is probably a conserved feature of the BET family.

Second, TAF1. The TAF1 double bromodomain cooperatively binds H4 peptides diacetylated on K5/K12 or K8/K16, an ability putatively attributed to TAF1 simultaneously engaging one acetyl-lysine residue in each binding pocket[5]. As previously noted[19], however, such a binding mode would imply a large conformational change of TAF1 (Supplementary Fig. 12). In light of the present study, an alternative (or additional) binding mode in which TAF1 recognizes diacetylated ligands through a single bromodomain is conceivable.

Third, CBP/p300. The bromodomains of CBP and p300 bind the transcription factor MyoD by recognizing a motif acetylated either on Lys 99 or on both Lys 99 and Lys 102 (refs 21, 22). CBP discriminates poorly between mono- and diacetylated MyoD, whereas p300 prefers the diacetylated form[22]. Structural alignment with Brdt shows that both bromodomains (which share 97% sequence identity) are sterically compatible with the superimposed K5ac-GG-K8ac tetrapeptide,

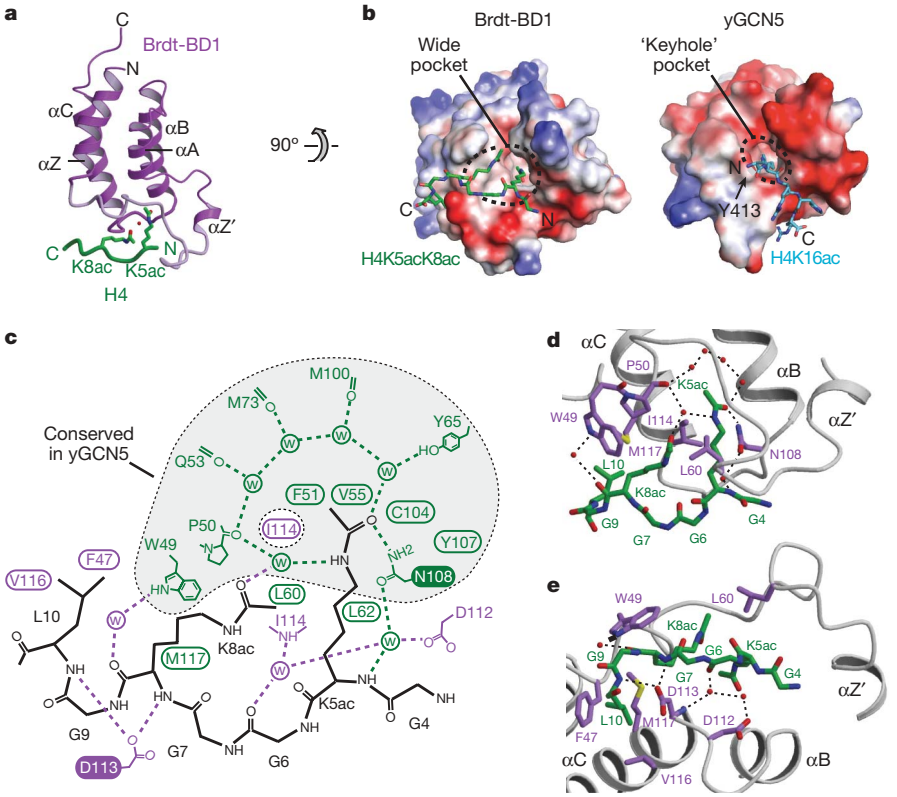
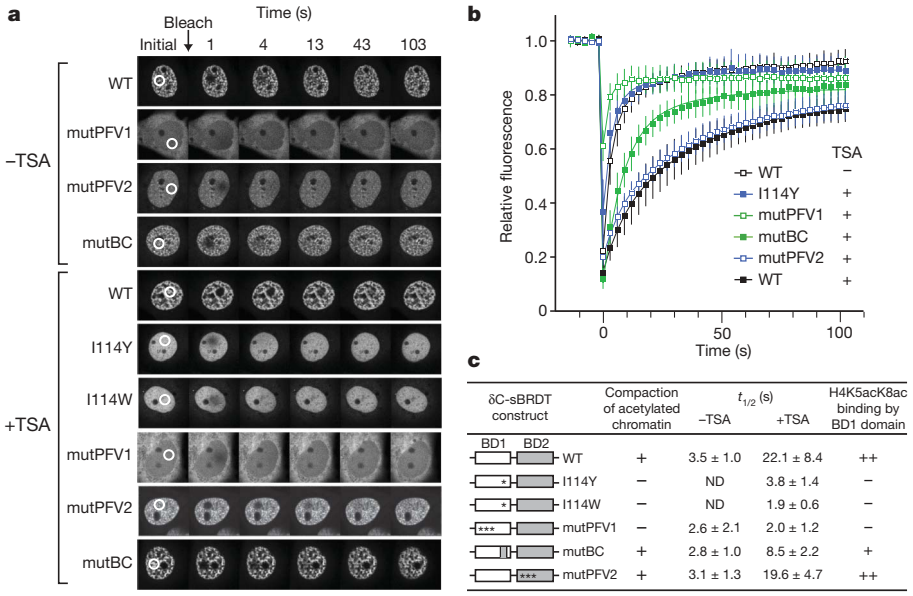


Figure 2 | Structure of the Brdt-BD1–H4K5acK8ac complex. **a**, Ribbon diagram. **b**, Molecular surface showing regions of positive (blue) and negative (red) electrostatic potential. The protein–peptide interface buries 498 Å^2 of accessible surface area, or 7% and 50% of the total protein and peptide accessible surface areas, respectively. The yGCN5 bromodomain is shown for comparison. **c**, Schematic summary of interactions between BD1 and H4K5acK8ac. Residues making van der Waals contacts are shown in an oval near the peptide atom they contact; residues that directly hydrogen bond with the peptide are in inverse font. Interactions conserved in Brdt-BD2 are in green, those unique to BD1 in magenta; those conserved in yGCN5 (PDB 1E6I) are shown within the shaded envelope. Dashed lines, hydrogen bonds; 'w', solvent molecule. **d**, **e**, Orthogonal views of the peptide-binding pocket. BD1 and H4 carbon atoms are in magenta and green, respectively.



c

δC-sBRDT construct		Compaction of acetylated chromatin	$t_{1/2}$ (s) –TSA	$t_{1/2}$ (s) +TSA	H4K5acK8ac binding by BD1 domain
BD1 BD2	WT	+	3.5 ± 1.0	22.1 ± 8.4	++
*	I114Y	–	ND	3.8 ± 1.4	–
*	I114W	–	ND	1.9 ± 0.6	–
***	mutPFV1	–	2.6 ± 2.1	2.0 ± 1.2	–
	mutBC	+	2.8 ± 1.0	8.5 ± 2.2	+
***	mutPFV2	+	3.1 ± 1.3	19.6 ± 4.7	++

Figure 3 | Stability of Brdt on acetylated chromatin mirrors BD1-mediated H4K5acK8ac recognition. **a**, FRAP experiment comparing the mobility of wild-type (WT) and mutated δC-sBrdt. White circles indicate bleached regions. FRAP analysis of I114Y and I114W was only performed in the presence of TSA, as in its absence these mutants aggregated (Supplementary Fig. 9). MutPVF1 shows a partial cytoplasmic localization, as previously observed[7]. Original magnification, ×350. **b**, Time course of fluorescence in bleached area, showing mean intensity and standard deviation from ten independent experiments. For clarity, plots are shown for only a subset of the experiments in **a**. **c**, Summary of compaction activity and $t_{1/2}$ values. Strength of H4K5acK8ac binding is from Table 1.

but that p300 can hypothetically form an extra hydrogen bond (Supplementary Fig. 13). The ability to rationalize the difference in selectivity suggests that CBP/p300 may use a mode of diacetylated ligand recognition similar to that by Brdt-BD1.

The present work describes an effector module that mediates positive cooperativity by engaging two histone PTMs in a common binding pocket. Unlike linked modules that are constrained by a minimal inter-domain distance, a common pocket is ideal for mediating cross-talk between PTMs located closely on a histone tail. Brdt-BD1 binds two PTMs of the same chemical type. Whether a single module can mediate agonistic interactions between heterotypic PTMs is a question for future study.

METHODS SUMMARY

Recombinant δC-sBrdt, BD1 and BD2 constructs were bacterially expressed and purified using standard techniques. BD1 and BD2 were co-crystallized with acetylated histone tail peptides by hanging drop vapour diffusion. Diffraction data were collected using synchrotron radiation. Crystal structures were solved by molecular replacement.

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

Received 19 April; accepted 13 August 2009.

1. Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447,** 407–412 (2007).
2. Kouzarides, T. Chromatin modifications and their function. *Cell* **128,** 693–705 (2007).
3. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Struct. Mol. Biol.* **14,** 1025–1040 (2007).
4. Ruthenburg, A. J., Li, H., Patel, D. J. & Allis, C. D. Multivalent engagement of chromatin modifications by linked binding modules. *Nature Rev. Mol. Cell Biol.* **8,** 983–994 (2007).
5. Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human $TAF_{II}250$ double bromodomain module. *Science* **288,** 1422–1425 (2000).
6. Florence, B. & Faller, D. V. You BET-cha: a novel family of transcriptional regulators. *Front. Biosci.* **6,** d1008–d1018 (2001).
7. Pivot-Pajot, C. *et al.* Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol. Cell. Biol.* **23,** 5354–5365 (2003).
8. Shang, E., Nickerson, H. D., Wen, D., Wang, X. & Wolgemuth, D. J. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development* **134,** 3507–3515 (2007).
9. Govin, J., Caron, C., Lestrat, C., Rousseaux, S. & Khochbin, S. The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur. J. Biochem.* **271,** 3459–3469 (2004).
10. Rousseaux, S. *et al.* Establishment of male-specific epigenetic information. *Gene* **345,** 139–153 (2005).
11. Crowley, T. E., Kaine, E. M., Yoshida, M., Nandi, A. & Wolgemuth, D. J. Reproductive cycle regulation of nuclear import, euchromatic localization, and association with components of Pol II mediator of a mammalian double-bromodomain protein. *Mol. Endocrinol.* **16,** 1727–1737 (2002).
12. Leroy, G., Rickards, B. & Flint, S. J. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Mol. Cell* **30,** 51–60 (2008).
13. Ladurner, A. G., Inouye, C., Jain, R. & Tjian, R. Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol. Cell* **11,** 365–376 (2003).
14. Matangkasombut, O. & Buratowski, S. Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol. Cell* **11,** 353–363 (2003).
15. Toyama, R., Rebbert, M. L., Dey, A., Ozato, K. & Dawid, I. B. Brd4 associates with mitotic chromosomes throughout early zebrafish embryogenesis. *Dev. Dyn.* **237,** 1636–1644 (2008).
16. Dey, A., Chitsaz, F., Abbasi, A., Misteli, T. & Ozato, K. The double bromodomain protein Brd4 binds to acetylated chromatin during interphase and mitosis. *Proc. Natl Acad. Sci. USA* **100,** 8758–8763 (2003).
17. Dey, A. *et al.* A bromodomain protein, MCAP, associates with mitotic chromosomes and affects G_2-to-M transition. *Mol. Cell. Biol.* **20,** 6537–6549 (2000).
18. Mujtaba, S., Zeng, L. & Zhou, M. M. Structure and acetyl-lysine recognition of the bromodomain. *Oncogene* **26,** 5521–5527 (2007).
19. Owen, D. J. *et al.* The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase Gcn5p. *EMBO J.* **19,** 6141–6149 (2000).
20. Kanno, T. *et al.* Selective recognition of acetylated histones by bromodomain proteins visualized in living cells. *Mol. Cell* **13,** 33–43 (2004).
21. Polesskaya, A. *et al.* Interaction between acetylated MyoD and the bromodomain of CBP and/or p300. *Mol. Cell. Biol.* **21,** 5312–5320 (2001).
22. Wei, L., Jamonnak, N., Choy, J., Wang, Z. & Zheng, W. Differential binding modes of the bromodomains of CREB-binding protein (CBP) and p300 with acetylated MyoD. *Biochem. Biophys. Res. Commun.* **368,** 279–284 (2008).
23. Govin, J. *et al.* Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J. Cell Biol.* **176,** 283–294 (2007).

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

Acknowledgements We thank the ESRF and EMBL staff for beamline assistance, the Partnership for Structural Biology (PSB) for access to technical platforms, M. Jamin for MALLS experiments, G. Natrajan and V. Rybin for help with ITC, C. Soucier for help with FRAP, D. Panne and C. Clapier for comments on the manuscript, and S. Cusack for critical support. Work in the S.K. laboratory was supported by ANR blanche 'EpiSperm' and 'Empreinte', INCa and 'ARECA' (ARC) research programmes. J.Ga. was supported by a Ph.D. fellowship from the Rhône-Alpes region. J.M. was supported by an 'E-STAR' fellowship funded by EU FP6. C.P. acknowledges support from the ANRS/Fondation de France AIJC and CNRS ATIP programmes.

Author Contributions M.S.-L., D.J.H., S.K. and C.W.M. initiated the study. S.K., C.P. and C.W.M. coordinated the entire project and specific author contributions. J.M., S.R., U.S., M.S.-L., A.-L.V. and D.J.H. prepared constructs. J.M., U.S. and M.S.-L. expressed, purified and crystallized BD1 and BD2. J.M. and U.S. performed and analysed ITC and biophysical experiments. J.M., M.S.-L. and C.P. measured diffraction data and solved the crystal structures. C.P. and C.W.M. analysed structural and biochemical data. J.Go. performed immunoblot analysis and immunofluorescence microscopy. S.C. and S.R. performed FRAP and chromatin compaction experiments. J.Ga. prepared histones for MS/MS analysis. K.S. and J.K. performed the MS/MS analysis. C.P. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Author Information Atomic coordinates and structure factors have been deposited with the Protein Data Bank under accession codes 2WP2 (BD1) and 2WP1 (BD2). Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.W.M. (christoph.mueller@embl.de).

METHODS

Protein purification and histone peptides. Murine δC-sBrdt (residues 2–442), BD2 (residues 257–383) and two BD1 constructs (residues 23–142 for ITC; residues 17–136 for crystallization) were cloned from mouse Brdt cDNA (GenBank accession number AF019085) as fusion constructs containing an N-terminal His tag and a TEV protease site. δC-sBrdt corresponds closely to the human sequences with GenBank accession numbers AAH05281, AAH17582, AAH62700 and AAH47900, but lacking their C-terminal 20-residue lysine-rich tail. BD1-mutBC was generated from BD1 (residues 23–142) by replacing residues 109–114 (KTGDDI) with residues 352–357 (PPDHEV). Expression of BD1 constructs was from a pETM11 vector in *Escherichia coli* strain BL21(DE3). Expression of BD2 and δC-sBrdt constructs was from a pHAR1201 vector (a pTriEx (Novagen) derivative with N-terminal His tag and TEV protease site) in *E. coli* strain BL21-AI (Invitrogen). Proteins were purified by two Ni-NTA (Qiagen) chromatography steps interspaced by TEV cleavage to remove the His tag, and followed by either Superdex 75 (BD1 and BD2) or Superdex 200 (δC-sBrdt) chromatography (GE Healthcare). Cell lysis and Ni-NTA steps were performed in 500 mM NaCl, 1 mM DTT and 50 mM HEPES (pH 7.5 for BD1 and BD2, pH 8.0 for δC-sBrdt). Gel filtration was in 100 mM NaCl, 10 mM HEPES (pH 7.5 or 8.0), and either 1 mM (for crystallization) or no DTT (for ITC). Protein concentrations were measured by ultraviolet absorbance at 280 nm using an extinction coefficient determined by amino acid analysis.

Peptides corresponding to mouse histone H4 residues 1–20, histone H2B residues 3–22, and either histone H3 residues 4–27 (for ITC) or H3 residues 14–23 (for crystallization) were chemically synthesized with the relevant lysine(s) acetylated. Lyophilized peptides were dissolved in the gel filtration buffer used to purify the bromodomain of interest. Peptide concentrations were determined by amino acid analysis.

Isothermal titration calorimetry. Experiments were performed with a MicroCal VP-ITC calorimeter at 25 °C. Peptides (1.9–5.6 mM) were injected in 5 μl aliquots every 300 s into a cell containing 1.7 ml of protein (50–160 μM). Data were analysed using ORIGIN software (MicroCal). Average values of K_d and N (binding stoichiometry) were determined from three or more experiments.

Crystallization. Hanging drop crystallization was performed at 20 °C by combining equal volumes of reservoir solution and one of three protein/peptide complexes, made by mixing either BD1 (residues 17–136) at 15 mg ml^{-1} with H4K5acK8ac in a 1:20 molar ratio; or BD2 at 25 mg ml^{-1} with H3K18ac in a 1:5 molar ratio. Crystallization conditions were 2.4 M $(NH_4)_2SO_4$, 0.1 M HEPES (pH 7), and 2.0 M $(NH_4)_2SO_4$, 2% PEG 400, 0.1 M HEPES (pH 7.5).

Structure determination. Crystals were cryoprotected with 30% glycerol and flash-cooled in liquid nitrogen. Diffraction data were collected at ESRF beamlines ID23-1 (λ = 0.954 Å; BD2) and ID23-2 (λ = 0.873 Å; BD1) on an ADSC Quantum 4 detector and processed with XDS[24] and programs of the CCP4 suite[25]. The BD2 structure was solved by molecular replacement with Phaser[26] using the yGCN5 bromodomain as search model. The BD1 structure was solved using the refined BD2 structure as search model. Model building was performed with Coot[27] and refinement with CNS[28].

FRAP. Cos7 cells grown on Labtek glass coverslips were transfected with 2 μg of wild-type or mutated GFP–δC-sBrdt using lipofectamine, treated with TSA (50 ng ml^{-1}) and incubated at 37 °C in 5% CO_2 for 16 h. FRAP analysis was performed on ten independent nuclei using a Confocor II Zeiss microscope equipped with a 488 nm laser and a LP505 filter. A circular region 4.4 μm in diameter was bleached for 0.7 s; fluorescence was measured in an area comprising over 50% of the bleached region.

24. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.* **26**, 795–800 (1993).
25. CCP4. The *CCP*4 suite: programs for protein crystallography. *Acta Crystallogr D* **50**, 760–763 (1994).
26. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D* **60**, 432–438 (2004).
27. Emsley, P. & Cowtan, K. *Coot*: model-building tools for molecular graphics. *Acta Crystallogr D* **60**, 2126–2132 (2004).
28. Brünger, A. T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**, 905–921 (1998).

LETTERS

Substrate interactions and promiscuity in a viral DNA packaging motor

K. Aathavan[1,2]*†, Adam T. Politzer[1,2]*, Ariel Kaplan[2,3,4]*, Jeffrey R. Moffitt[2,4]*, Yann R. Chemla[2,4]†, Shelley Grimes[5], Paul J. Jardine[5], Dwight L. Anderson[5,6] & Carlos Bustamante[1,2,3,4,7]

The ASCE (additional strand, conserved E) superfamily of proteins consists of structurally similar ATPases associated with diverse cellular activities involving metabolism and transport of proteins and nucleic acids in all forms of life[1]. A subset of these enzymes consists of multimeric ringed pumps responsible for DNA transport in processes including genome packaging in adenoviruses, herpesviruses, poxviruses and tailed bacteriophages[2]. Although their mechanism of mechanochemical conversion is beginning to be understood[3], little is known about how these motors engage their nucleic acid substrates. Questions remain as to whether the motors contact a single DNA element, such as a phosphate or a base, or whether contacts are distributed over several parts of the DNA. Furthermore, the role of these contacts in the mechanochemical cycle is unknown. Here we use the genome packaging motor of the *Bacillus subtilis* bacteriophage φ29 (ref. 4) to address these questions. The full mechanochemical cycle of the motor, in which the ATPase is a pentameric-ring[5] of gene product 16 (gp16), involves two phases—an ATP-loading dwell followed by a translocation burst of four 2.5-base-pair (bp) steps[6] triggered by hydrolysis product release[7]. By challenging the motor with a variety of modified DNA substrates, we show that during the dwell phase important contacts are made with adjacent phosphates every 10-bp on the 5′–3′ strand in the direction of packaging. As well as providing stable, long-lived contacts, these phosphate interactions also regulate the chemical cycle. In contrast, during the burst phase, we find that DNA translocation is driven against large forces by extensive contacts, some of which are not specific to the chemical moieties of DNA. Such promiscuous, nonspecific contacts may reflect common translocase–substrate interactions for both the nucleic acid and protein translocases of the ASCE superfamily[1].

To test the role of the phosphate backbone charge in motor–DNA interactions we inserted a 10-bp region of double-stranded methylphosphonate DNA (dsMeP) into the middle of an ~8-kilobase-pair (kb) native DNA molecule (Fig. 1a), and followed the packaging of this molecule by a single φ29 prohead–motor complex using optical tweezers. In MeP the charged oxygen on DNA is replaced with an uncharged isosteric methyl group while conserving the B-form structure of DNA[8,9] (Fig. 1b, inset). Thus, it is possible to determine the role of this chemical modification in a native geometric context. Figure 1b shows sample packaging traces under 5 pN of constant tension and saturating [ATP] (1 mM). Packaging proceeds normally until the motor encounters the inserted modification where it pauses, and then either successfully traverses the insert or completely dissociates. In stark contrast to related helicases, in which disruption of a single charge interaction[10–14] completely abolishes translocation, the packaging motor traverses 10 bp of neutral DNA with a probability of ~80% under a tension of 5 pN (Fig. 1c).

To rule out the possibility that the motor crosses the neutral insert by diffusive fluctuations as opposed to making direct contact with uncharged moieties, we took advantage of the strong force dependence of diffusive traversal times[15]. We found that there is only a twofold increase in pause duration with a 15 pN increase in force (Fig. 1c)—much less than the 10^5-fold increase predicted for diffusion across a 10-bp distance (Supplementary Discussion). Furthermore, lowering [ATP] increases the pause duration and decreases the traversal probability, providing further support for an active, ATP-dependent crossing mechanism. Thus, the motor actively traverses the insert by making contacts with elements other than the phosphate charge, albeit with reduced efficiency, indicating that native packaging involves both charge and non-charge contacts.

To determine whether phosphate charges from both strands are equally important, we created hybrid inserts in which only one strand contains the MeP backbone. We used a 30-bp insert to accentuate the effect of the uncharged section because the traversal probability of a 30-bp dsMeP insert at 5 pN is ~4%. At this force the traversal probability of the hybrid insert with MeP on the strand packaged from 3′–5′ is almost 90%, whereas that of the hybrid insert with MeP in the 5′–3′strand is reduced to 10% (Fig. 1d). This result clearly indicates that the most important phosphate interactions are made with the 5′–3′ strand in the direction of packaging. Such preferential interaction with a single DNA strand has been shown for the monomeric dsDNA translocase EcoR124I (ref. 16), although it is more surprising in this case of a ring-ATPase in which several subunits of the ring are simultaneously in close proximity to both strands.

Next, we addressed whether a critical length is involved in the interaction of the motor and its DNA substrate. Figure 1e shows that as we increase the length of the double-stranded neutral insert from 5 to 10 bp there is no statistically significant change in traversal probability, but the 1-bp increase from 10 to 11 bp results in a twofold reduction. Further increasing the length to 15 bp does not produce a similar change. The location of this discrete change in traversal probability at first seems to be inconsistent with the 10 bp of DNA packaged by the motor each full mechanochemical cycle[6], but these results are easily reconciled if the motor makes contact with two adjacent phosphates, with either contact being sufficient for packaging (Fig. 1f). The co-crystal structure of the related BPV helicase E1 with its single-stranded DNA (ssDNA) substrate reveals simultaneous

[1]Biophysics Graduate Group, [2]Jason L. Choy Laboratory of Single-Molecule Biophysics, [3]QB3 Institute, and [4]Department of Physics, University of California, Berkeley, California 94720, USA. [5]Department of Diagnostic and Biological Sciences and Institute for Molecular Virology, [6]Department of Microbiology, University of Minnesota, Minneapolis, Minnesota 55455, USA. [7]Departments of Molecular and Cell Biology, Chemistry, and Howard Hughes Medical Institute, University of California, Berkeley, California 94720, USA. †Present addresses: Department of Cellular and Molecular Pharmacology, University of California, San Francisco, California 94158, USA (K.A.); Department of Physics and Center for Biophysics and Computational Biology, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801, USA (Y.R.C.).
*These authors contributed equally to this work.

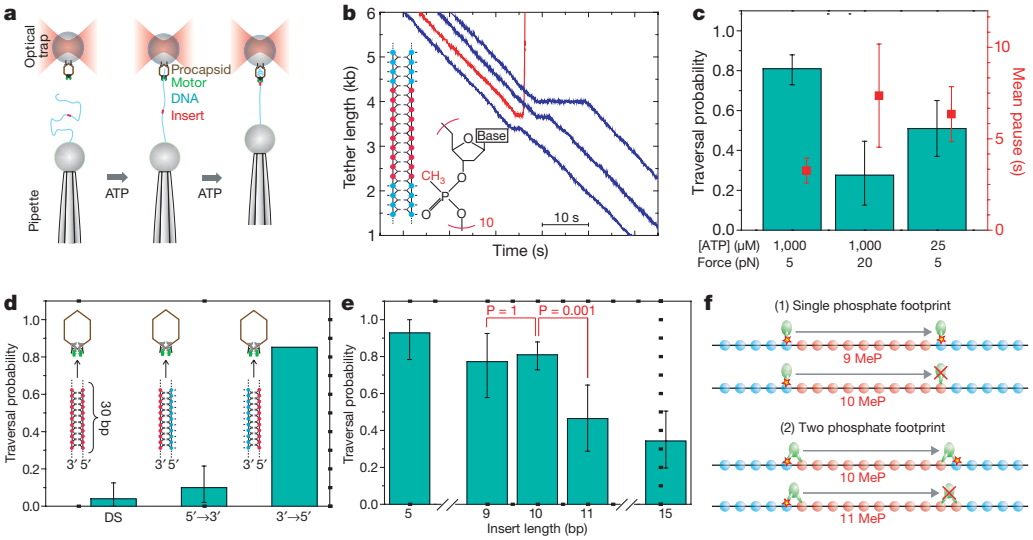


Figure 1 | Packaging of neutral DNA analogues. **a**, A prohead–motor complex bound to a microsphere is held in an optical trap while a micropipette holds a second microsphere bound to DNA containing a modified insert. A tether is formed and packaging is initiated when the beads are briefly brought into close proximity in the presence of ATP. **b**, Representative packaging traces of DNA containing 10-bp dsMeP at a constant load of 5 pN. Blue traces show traversal after a pause, and the red trace shows a terminal dissociation event after a pause. Inset are a schematic of the insert, with MeP nucleotides in red and unmodified nucleotides in blue, and the chemical structure of a MeP nucleotide. **c**, Force and ATP dependence of traversal probability and pause duration of 10-bp dsMeP inserts. **d**, Traversal probability of 30-bp dsMeP and DNA–MeP hybrid inserts at 5 pN. **e**, Traversal probability of dsMeP inserts at 5 pN force as a function of insert length. *P* values (two-tailed Fisher exact test) between 9 and 10 bp, and between 10 and 11 bp, are indicated. **f**, Translocation cycle length and footprint size limits from MeP length dependence. This scheme shows the position of a subunit that contacts the DNA before and after a full mechanochemical cycle, that is, 10 bp. Contact with a single phosphate would produce a drop in traversal probability between 9- and 10-bp dsMeP, whereas contact with two phosphates would produce the observed drop between 10- and 11-bp dsMeP. In **c**–**e** the traversal probability is plotted using the Laplace estimator[26], with 95% confidence intervals from the adjusted Wald method[27], and error bars of pause durations denote the s.e.m.

contact with adjacent phosphate charges[17], lending support to this interpretation.

We next investigated the specific role of these phosphates in the mechanochemical cycle by probing the base-pair-scale dynamics of the motor at an uncharged insert. Phosphate–motor interactions may have two possible roles in the mechanochemical cycle: they may provide the long-lived contacts that are required to keep the enzyme attached to the DNA, or they may have a sensory role, accelerating a chemical rate, such as ATP hydrolysis, upon detecting that the DNA is bound and properly oriented. These two roles of the phosphate charge, although not mutually exclusive, can be revealed by characteristic dynamics of the motor as it traverses the modified insert. If the phosphate provides load-bearing contacts, its absence will increase the dissociation rate of the motor, and the insert-induced pause will consist of a series of attempts to package followed by small slips. Alternatively, if the role of the phosphate is sensory, we expect the time between packaging steps to be lengthened, owing to the decreased rate of catalytic turnover.

To determine the dynamics of the motor as it crosses a neutral insert (10-bp dsMeP), we followed packaging using dual-trap optical tweezers with higher spatial and temporal resolution[6]. The pauses observed at low resolution are actually remarkably dynamic events, containing two types of sub-pauses that occur at distinct locations along the modified DNA insert (Fig. 2a). The first type of sub-pause, which we term 'upstream pauses' because it occurs at longer DNA tether lengths, is followed by either brief disengagement of the motor (slips) or packaging attempts. These attempts are themselves followed by a second class of sub-pause, which we term 'downstream pauses'. After slips from either the upstream or the downstream pauses, the motor typically recovers and repackages the DNA to the position of the upstream pauses. Occasionally the motor does not recover from a slip, resulting in a terminal slip. The branching probabilities of these events are shown in Fig. 2b.

The upstream sub-pauses occur in a uniform position on a given tether, ±1 bp (s.d.), and have longer average durations, 1.00 ± 0.08 s (s.e.m.; Fig. 2c, e), whereas the downstream sub-pauses occur at the end of attempts of different sizes, ±3 bp (s.d.), and have shorter durations, 80 ± 10 ms (s.e.m., Fig. 2d, f). The upstream- and downstream-pause time distributions are both well-described by single exponential

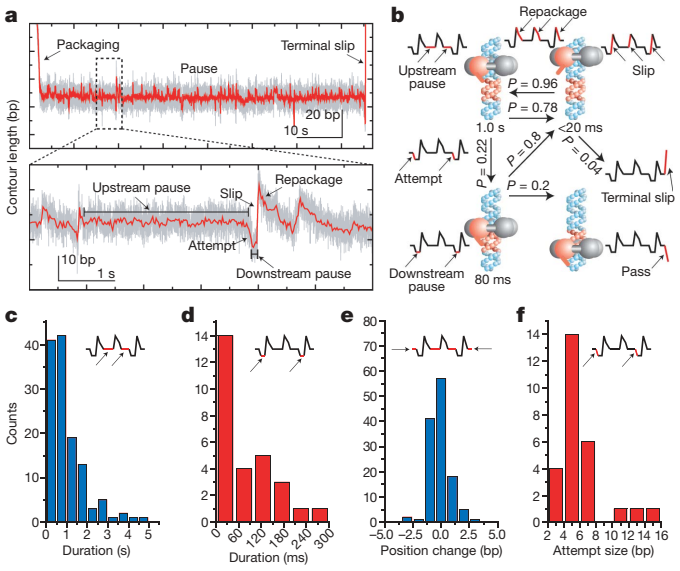


Figure 2 | High-resolution dynamics at a neutral-DNA insert. **a**, Base-pair scale dynamics at 10 bp of dsMeP consist of two classes of sub-pauses, upstream and downstream pauses, separated by attempts and punctuated by slips and repackaging events. **b**, A cartoon model of the dynamics of these pauses with average lifetimes and inter-conversion probabilities. The full statistics of these states are in Supplementary Table 3. **c**, **d**, Histograms of upstream (**c**) and downstream (**d**) pause durations. The distributions have n_{min} values—the ratio of the mean squared to the variance—of 1.1 ± 0.1 (s.d.) and 1.3 ± 0.4 (s.d.), respectively, and are thus well-described by single exponential decays. **e**, Histogram of position changes in the upstream pause. **f**, Histogram of the distance between upstream and downstream pauses.

decays (Fig. 2c, d), suggesting that each class of sub-pause corresponds to an individual kinetic process, as opposed to a weighted sum of different kinetic processes, and that each of them is governed by a single rate-limiting kinetic event. In combination with their distinct mean durations, these observations suggest that each class of sub-pause corresponds to one of the mechanochemical phases of the motor: ATP-loading dwell or stepping burst.

To identify the kinetic state of the motor associated with each class of sub-pause, we measured the dynamics of motors traversing 5 bp of neutral DNA, an insert shorter than the 10-bp span of the full mechanochemical cycle[6]. Whereas contacts during the burst phase occur every 2.5 bp, contacts during the dwell phase should only occur every 10 bp. Thus, all motors must make burst phase contacts with the neutral DNA, while a fraction of the motors will never make dwell phase contacts with the neutral DNA. If the upstream pause occurs during the dwell phase, then only a fraction of the motors will show such a pause. Only ~50% of the motors packaging 5 bp of dsMeP display a long pause (Supplementary Fig. 1), strongly suggesting that the upstream pause corresponds to the dwell phase, and hence the downstream pause to the burst phase. This conclusion is further supported by the average duration of the downstream pauses and the exponential nature of their distribution, which are both similar to those of the micro-dwells observed during the burst phase on charged DNA under similar conditions[6]. Finally, the highly uniform positions of the upstream pauses in a given tether indicate that this single-base-pair position is unique; suggesting that after slipping and repositioning, the motor resides at the boundary of the charged and neutral DNA during its dwell phase.

The multiple slip/attempt phenotype observed shows that removing the phosphate charge reduces the processivity of the motor (Fig. 2a). Because the motor slips from both upstream and downstream pauses, load-bearing phosphate contacts are made during both the dwell and the burst phases of the cycle. Notably, however, upstream pauses are on average tenfold longer than dwells on normal DNA under the same experimental conditions (~100 ms)[6]; thus, the absence of the phosphate charge slows the dwell phase, indicating a sensory role for this contact. The fact that a single kinetic event dominates the upstream pause duration, in sharp contrast to the ~four kinetic events that are known to be rate-limiting for the dwell phase on charged DNA[6], indicates that the mechanochemical cycle of the packaging motor contains a single kinetic checkpoint—a process that halts the chemical cycle until the DNA is correctly positioned. In contrast, the average duration of the contacts during the burst phase is not significantly modified on neutral DNA, and hence these contacts probably do not involve a significant sensory role.

These dynamics suggest a mechanism for the motor to cross neutral inserts. Successful traversal is a kinetic competition between the increased off-rate of the motor on neutral DNA and the time it takes to successfully complete the mechanochemical events necessary to traverse this modified DNA. There is a small probability of loading the necessary ATP molecules and starting the burst, out-competing the increased slipping rate on the neutral DNA; however, as illustrated by the probabilities in Fig. 2b, the large probability of recovering from a slip allows the motor to attempt to package the insert many times, amplifying this small probability to the large traversal probabilities observed in Fig. 1. This kinetic competition also explains the marked decrease in the ability of the motor to package 30 bp of neutral DNA because successful traversal would require several consecutive complete mechanochemical cycles to out-compete slipping. Finally, the branching probabilities observed (Fig. 2b) predict traversal probabilities consistent with the values in Fig. 1, providing further support for this model (Supplementary Discussion).

The fact that the motor resides for a finite time in the neutral DNA (Fig. 2) provides further evidence for additional, non-charge contacts. To test the role of sugars and bases in these interactions, we created an insert with these elements removed (Fig. 3a). These chemical modifications also disrupt the helical geometry of the DNA; thus, to isolate

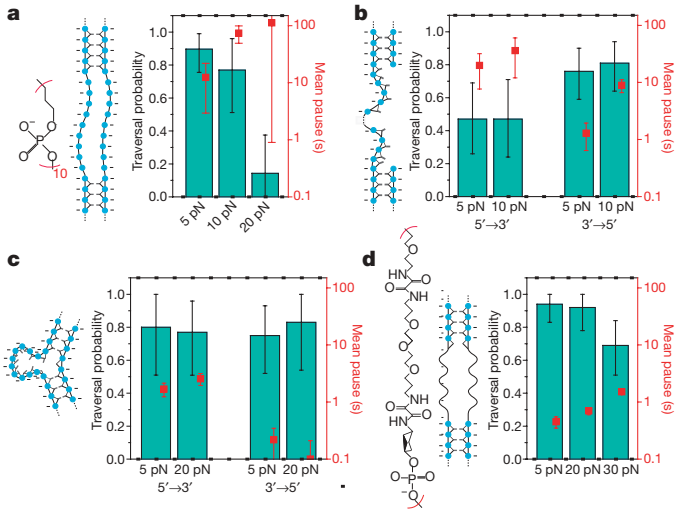


Figure 3 | Substrate promiscuity. **a–d**, The force dependence of traversal probabilities and pause durations of different modified DNA substrates. **a**, 10-nucleotide double-stranded abasic phosphate backbone. **b**, 20-nucleotide ssDNA (poly-AC) on 5′–3′ and 3′–5′ strands. **c**, 10-bp bulge (poly-AC) on 5′–3′ and 3′–5′ strands. **d**, Double-stranded linker. The traversal probabilities are plotted using the Laplace estimator[26], with 95% confidence intervals from the adjusted Wald method[27].The error bars of the pause durations are the s.e.m. (The results for other sizes of bulges and gaps are listed in Supplementary Table 2.)

the role of DNA geometry, we created inserts with no chemical modification yet with large disruptions to the helix; namely, single-stranded gaps (Fig. 3b) and unpaired bulges (Fig. 3c). Remarkably, the φ29 motor is capable of traversing all of these modifications (Fig. 3a–c), and the force dependence of the mean pause durations again rules out a purely diffusive model (Supplementary Table 4). Notably, the motor shows lower traversal probability and higher pause durations for gaps on the 5′–3′ strand than those on the 3′–5′ strand, with a similar trend seen in pause durations of the bulges, consistent with the important 5′–3′ contacts discussed above. Our findings are consistent with the successful packaging of short unpaired bulges in phage lambda[18], although they differ from reported results in φ29 (ref. 19) and T4 (ref. 20), in which single-stranded gaps were not packaged. However, disparate experimental methods probably account for these apparent differences (Supplementary Discussion).

Many of the modifications we have probed (Figs 1 and 3) change many features of the DNA simultaneously; thus, to extract the relative importance of the different chemical moieties while controlling for other changes, we performed a set of multivariate logistic regressions[21] on different subsets of our data. This analysis provides a quantitative ranking of the importance of the different contacts and their force dependence (Fig. 4 and Supplementary Fig. 2). These regressions confirm the importance of the phosphates every 10 bp on the 5′–3′ strand and show that these contacts are only important under the application of force. Regression analysis also shows less important contacts with phosphates on both strands on a smaller distance scale, which, remarkably, remain important when extrapolated to zero applied load. This residual importance may arise from the force the motor generates on the DNA, as opposed to the external load applied optically, supporting our conclusion that these contacts occur during the burst phase.

The regression analysis also provides information on the nature of the non-charge contacts, revealing that important, but minor, contacts are made with bases or sugars (Supplementary Fig. 2a). However, the analysis also predicts that removing all of the characteristic features of nucleic acid—phosphates, bases, sugars and native double-helix structure—will result in a reduced but finite traversal probability (Supplementary Discussion). Thus, a component of these additional, non-charge interactions is not specific to DNA. We tested this

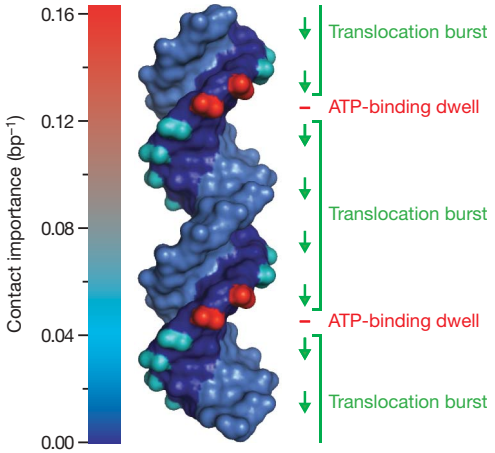


Figure 4 | The motor–DNA contacts. The relative importance of the motor–DNA contacts are marked using the program PyMOL with a quantitative colour scale with magnitudes inferred from the traversal probabilities at 5 pN for the measured modifications (Figs 1 and 3, Supplementary Table 2, Supplementary Discussion and Supplementary Fig. 2). The units of 'contact importance' correspond to the inverse of the distance over which the removal of the specific moiety would reduce the traversal probability to 50% (Supplementary Discussion). The mechanochemical phase of the motor as it moves along the DNA is indicated to the right.

surprising result by challenging the motor with a polymer lacking any resemblance to nucleic acid. Notably, the motor packages this insert as well (Fig. 3d), revealing a nonspecific component to the motor–DNA interaction. Such interactions may correspond to chemical contacts, that is, hydrogen bonds or hydrophobic interactions, which do not require a specific nucleic acid moiety. Alternatively, it is possible that these interactions are steric in nature. A prediction of such a steric drive mechanism is that amino acids essential for DNA translocation can be replaced by any residue with a bulky side chain, a behaviour suggested by mutational analysis of the DNA binding loop of the related hexameric helicase SV40 (ref. 22).

Taken together, our data indicate that the motor–DNA interaction involves a wide variety of contacts with the full complement of nucleic acid moieties, as well as contacts not specific to DNA. Furthermore, our data suggest that the type of motor–DNA interaction changes during the course of the mechanochemical cycle. During the dwell phase, when ATP is loaded, the motor maintains strong, load-bearing contact with the DNA via interactions with adjacent charges on the 5′–3′ strand every 10 bp. These contacts have a sensory role, coupling the mechanical and chemical cycles. In the burst phase, when the DNA is translocated, the motor makes contact with a variety of chemical moieties, including charges on both strands, bases or sugars, and additional non-DNA-specific contacts. Given the relatively long duration of the contacts during the dwell phase, specific, strong, ionic interactions may be preferable. In contrast, contacts during active translocation in the burst phase must be made and broken more quickly; thus, it would be advantageous to make weaker, more promiscuous contacts. The important motor–DNA interactions and the specific point in the mechanochemical cycle at which these contacts are made are summarized in Fig. 4.

Our findings have broad implications for the mechanism of the packaging motor. In particular, nonspecific contacts can have an important role in generating a step size that is a non-integer repeat of the chemical periodicity of DNA[6]. Moreover, the observation that important contacts are made below distances of ~10 bp provides evidence against an open-ring model of translocation in which only two subunits contact the DNA, thereby limiting the mechanism by which a non-integer step size can be generated[6]. In parallel, our findings have broader implications for the family of ASCE ring-ATPases. The use of different contacts during distinct portions of the mechanochemical cycle, as well as the use of sensory contacts, could act as a general mechanism by which multimeric motors synchronize and coordinate the hydrolysis cycles of their individual subunits with the position of their substrate. The use of nonspecific, perhaps steric, contacts by nucleic acid and polypeptide translocases[23] may reflect the existence of a conserved translocation mechanism shared by members of the ASCE superfamily; such nonspecific contacts may have facilitated the evolution of peptide translocases from nucleic acid translocases[24].

METHODS SUMMARY

Substrates were prepared by ligating custom oligonucleotide inserts with unmodified DNA. Single-molecule packaging assays were performed as described[6,7,25]. Insert-induced pauses were identified according to their location in the packaging trace.

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

Received 3 June; accepted 20 August 2009.

1. Iyer, L. M., Makarova, K. S., Koonin, E. V. & Aravind, L. Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. *Nucleic Acids Res.* **32,** 5260–5279 (2004).
2. Burroughs, A. M., Iyer, L. M. & Aravind, L. *Comparative Genomics and Evolutionary Trajectories of Viral ATP Dependent DNA-Packaging Systems* 48–65 (Basel, 2007).
3. Hopfner, K. P. & Michaelis, J. Mechanisms of nucleic acid translocases: lessons from structural biology and single-molecule biophysics. *Curr. Opin. Struct. Biol.* **17,** 87–95 (2007).
4. Grimes, S., Jardine, P. J. & Anderson, D. Bacteriophage ϕ29 DNA packaging. *Adv. Virus Res.* **58,** 255–294 (2002).
5. Morais, M. C. *et al.* Defining molecular and domain boundaries in the bacteriophage ϕ29 DNA packaging motor. *Structure* **16,** 1267–1274 (2008).
6. Moffitt, J. R. *et al.* Intersubunit coordination in a homomeric ring ATPase. *Nature* **457,** 446–450 (2009).
7. Chemla, Y. R. *et al.* Mechanism of force generation of a viral DNA packaging motor. *Cell* **122,** 683–692 (2005).
8. Thiviyanathan, V. *et al.* Structure of hybrid backbone methylphosphonate DNA heteroduplexes: effect of R and S stereochemistry. *Biochemistry* **41,** 827–838 (2002).
9. Strauss, J. K. & Maher, L. J. III. DNA bending by asymmetric phosphate neutralization. *Science* **266,** 1829–1834 (1994).
10. Eoff, R. L., Spurling, T. L. & Raney, K. D. Chemically modified DNA substrates implicate the importance of electrostatic interactions for DNA unwinding by Dda helicase. *Biochemistry* **44,** 666–674 (2005).
11. Kawaoka, J., Jankowsky, E. & Pyle, A. M. Backbone tracking by the SF2 helicase NPH-II. *Nature Struct. Biol.* **11,** 526–530 (2004).
12. SenGupta, D. J. & Borowiec, J. A. Strand-specific recognition of a synthetic DNA replication fork by the SV40 large tumor antigen. *Science* **256,** 1656–1661 (1992).
13. Mancini, E. J. *et al.* Atomic snapshots of an RNA packaging motor reveal conformational changes linking ATP hydrolysis to RNA translocation. *Cell* **118,** 743–755 (2004).
14. Dillingham, M. S., Soultanas, P. & Wigley, D. B. Site-directed mutagenesis of motif III in PcrA helicase reveals a role in coupling ATP hydrolysis to strand separation. *Nucleic Acids Res.* **27,** 3310–3317 (1999).
15. Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton* 1st edn 62 (Sinauer Associates, 2001).
16. Stanley, L. K. *et al.* When a helicase is not a helicase: dsDNA tracking by the motor protein EcoR124I. *EMBO J.* **25,** 2230–2239 (2006).
17. Enemark, E. J. & Joshua-Tor, L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442,** 270–275 (2006).
18. Pearson, R. K. & Fox, M. S. Effects of DNA heterologies on bacteriophage lambda packaging. *Genetics* **118,** 5–12 (1988).
19. Moll, W. D. & Guo, P. Translocation of nicked but not gapped DNA by the packaging motor of bacteriophage ϕ29. *J. Mol. Biol.* **351,** 100–107 (2005).
20. Oram, M., Sabanayagam, C. & Black, L. W. Modulation of the packaging reaction of bacteriophage t4 terminase by DNA structure. *J. Mol. Biol.* **381,** 61–72 (2008).
21. Agresti, A. *Categorical Data Analysis* Chs 5–6 (Wiley-Interscience, 2003).
22. Shen, J., Gai, D., Patrick, A., Greenleaf, W. B. & Chen, X. S. The roles of the residues on the channel β-hairpin and loop structures of simian virus 40 hexameric helicase. *Proc. Natl Acad. Sci. USA* **102,** 11248–11253 (2005).
23. Barkow, S. R., Levchenko, I., Baker, T. A. & Sauer, R. T. Polypeptide translocation by the AAA+ ClpXP protease machine. *Chem. Biol.* **16,** 605–612 (2009).
24. Mulkidjanian, A. Y., Makarova, K. S., Galperin, M. Y. & Koonin, E. V. Inventing the dynamo machine: the evolution of the F-type and V-type ATPases. *Nature Rev. Microbiol.* **5,** 892–899 (2007).

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

Acknowledgements We thank C. L. Hetherington, M. Kopaczynska, A. Spakowitz and J. M. Berger for critical discussions, and D. Reid, M. T. Couvillon and N. L. S. Chavez for preliminary work leading to this publication. K.A. acknowledges the PMMB fellowship through the Burroughs Wellcome Fund, A.T.P. the NIH Molecular Biophysics Training Grant, A.K. the Human Frontier Science Program Cross-Disciplinary Fellowship, J.R.M. the NSF Graduate Research Fellowship, and Y.R.C. the Burroughs Wellcome Fund Career Award at the Scientific Interface for funding. This research was supported in part by the National Institutes of Health (NIH) grants GM-071552, DE-003606 and GM-059604. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author Contributions K.A., A.T.P., A.K. and J.R.M. conducted the experiments; K.A., A.T.P., A.K., J.R.M. and Y.R.C. performed the analysis; S.G., P.J.J. and D.L.A. prepared and provided experimental materials; and K.A., A.T.P., A.K., J.R.M., Y.R.C., S.G., P.J.J. and C.B. wrote the paper. K.A., A.T.P., A.K. and J.R.M. contributed equally to this work.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.B. (carlos@alice.berkeley.edu.).

METHODS

DNA substrates. For each construct, two (or in the case of the gaps, three) modified or unmodified DNA oligonucleotides (listed in Supplementary Table 1) were purchased from Fidelity Systems or IDT. The oligonucleotides hybridize to form a DNA segment with 3-nucleotide overhangs. Using PCR, restriction digest and agarose gel purification, 4,187- and 4,008-bp DNA fragments with distinct non-palindromic 3-nucleotide overhangs were prepared. The primer for the 4,008-bp PCR fragment contains a terminal biotin. These fragments were ligated to the insert oligonucleotides and the 8.2-kb product was gel purified.

Single-molecule packaging. Single-molecule packaging assays were performed as previously described[4,7,25,28] in packaging buffer containing 50 mM Tris-HCl, pH 7.8, 50 mM NaCl, 5 mM $MgCl_2$ and 1 mM ATP, unless mentioned otherwise. The biotinylated DNA is bound to a 2.1-μm streptavidin-coated polystyrene bead held by a micropipette and the prohead–gp16–ATPase complex is attached to a second bead held in an optical trap. Packaging is initiated *in situ* by bringing the beads together[25,28]. The instrument[29] is run in force-feedback mode, in which the separation between the beads is adjusted to maintain constant tension in the DNA molecule. High-resolution packaging measurements were conducted on a dual trap instrument described previously[30] with materials prepared in the same fashion as described above, except 860-nm beads were used.

Analysis. The modified DNA insert is expected to reach the motor after packaging 4.2 kb. However, variability in attachment geometry of the DNA to the beads and other systematic errors introduce uncertainty in the measurement of tether length, so the insert was considered to be between the 2.5- and 4.5-kb positions. The longest pause in this interval was scored as the insert-induced pause. In traces in which a pause could not be identified, a pause of 0.1 s, the temporal resolution, was assigned. The rates of pausing ($0.017\ kb^{-1}$) and slipping ($0.02\ kb^{-1}$) during normal packaging[7] are small enough that these events do not significantly bias our measurements. All probabilities were calculated with the Laplace estimator[26], which is given by $P_{Laplace} = (x+1)/(n+2)$, in which x is the number of successes and n is the number of trials. This is a better estimator than the maximum-likelihood x/n, especially when n is small or the probabilities being estimated are near zero or one[26]. n_{min} values are a measure of the degree to which a distribution is exponential and provide a strict lower limit on the number of rate-limiting kinetic events[6]. These values were calculated by taking the ratio of the mean dwell time squared over the variance in the dwell times. n_{min} error bars are standard deviations and were calculated using a bootstrap method.

25. Smith, D. E. *et al.* The bacteriophage ϕ29 portal motor can package DNA against a large internal force. *Nature* **413,** 748–752 (2001).
26. Lewis, J. & Sauro, J. When 100% really isn't 100%: improving the accuracy of small-sample estimates of completion rates. *J. Usability Stud.* **1,** 136–150 (2006).
27. Agresti, A. & Coull, B. Approximate is better than ''exact'' for interval estimation of binomial proportions. *Am. Stat.* **52,** 119–126 (1998).
28. Rickgauer, J. P. *et al.* Portal motor velocity and internal force resisting viral DNA packaging in bacteriophage φ29. *Biophys. J.* **94,** 159–167 (2008).
29. Smith, S. B., Cui, Y. & Bustamante, C. Optical-trap force transducer that operates by direct measurement of light momentum. *Methods Enzymol.* **361,** 134–162 (2003).
30. Moffitt, J. R., Chemla, Y. R., Izhaky, D. & Bustamante, C. Differential detection of dual traps improves the spatial resolution of optical tweezers. *Proc. Natl Acad. Sci. USA* **103,** 9006–9011 (2006).

LETTERS

A human 5′-tyrosyl DNA phosphodiesterase that repairs topoisomerase-mediated DNA damage

Felipe Cortes Ledesma[1]*, Sherif F. El Khamisy[1,2]*, Maria C. Zuma[1], Kay Osborn[1] & Keith W. Caldecott[1]

Topoisomerases regulate DNA topology and are fundamental to many aspects of chromosome metabolism[1,2]. Their activity involves the transient cleavage of DNA, which, if it occurs near sites of endogenous DNA damage or in the presence of topoisomerase poisons, can result in abortive topoisomerase-induced DNA strand breaks[3–5]. These breaks feature covalent linkage of the enzyme to the DNA termini by a 3′- or 5′-phosphotyrosyl bond and are implicated in hereditary human disease[6–8], chromosomal instability and cancer[4,9], and underlie the clinical efficacy of an important class of anti-tumour poisons[3,9,10]. The importance of liberating DNA termini from trapped topoisomerase is illustrated by the progressive neurodegenerative disease observed in individuals containing a mutation in tyrosyl-DNA phosphodiesterase 1 (TDP1), an enzyme that cleaves 3′-phosphotyrosyl bonds[6–8]. However, a complementary human enzyme that cleaves 5′-phosphotyrosyl bonds has not been reported, despite the effect of DNA double-strand breaks containing such termini on chromosome instability and cancer[6–8]. Here we identify such an enzyme in human cells and show that this activity efficiently restores 5′-phosphate termini at DNA double-strand breaks in preparation for DNA ligation. This enzyme, TTRAP, is a member of the Mg^{2+}/Mn^{2+}-dependent family of phosphodiesterases. Cellular depletion of TTRAP results in increased susceptibility and sensitivity to topoisomerase-II-induced DNA double-strand breaks. TTRAP is, to our knowledge, the first human 5′-tyrosyl DNA phosphodiesterase to be identified, and we suggest that this enzyme is denoted tyrosyl DNA phosphodiesterase-2 (TDP2).

To identify new human tyrosyl DNA phosphodiesterase activities, we exploited the hypersensitivity of *Saccharomyces cerevisiae tdp1Δ rad1Δ* double-mutant cells to camptothecin (CPT), a topoisomerase I (Top1) poison that induces single-strand breaks (SSBs) with Top1 covalently linked to the 3′-terminus[3,10]. This strain lacks not only Tdp1 but also Rad1-Rad10 nuclease, which in yeast provides an alternative (endonucleolytic) pathway for removing Top1 from 3′-termini[11,12]. We transformed this strain with a human complementary DNA library and screened the resulting population of transformants for cellular resistance to CPT. Of six *tdp1Δ rad1Δ* transformants showing wild-type levels of CPT resistance, three contained cDNA clones encoding TDP1 and three contained cDNA clones encoding TTRAP (TRAF and TNF receptor-associated protein), a protein of unknown function and a putative member of the Mg^{2+}/Mn^{2+}-dependent phosphodiesterase superfamily, with the DNA repair protein apurinic/apyrimidinic (AP) endonuclease-1 (APE-1, also known as APEX1) being its closest relative[13,14] (Fig. 1a and Supplementary Fig. 1).

The *TDP1* and *TTRAP* cDNA clones recovered in the genetic screen suppressed the CPT sensitivity of *tdp1Δ rad1Δ* cells to a similar extent (Fig. 1b and data not shown). Although the pACT-TTRAP clones encoded TTRAP protein that lacked eight (pACT-TTRAP-2, pACT-TTRAP-3) or twenty-two (pACT-TTRAP-1) residues from the amino terminus (data not shown), full-length TTRAP similarly suppressed the CPT sensitivity of *tdp1Δ rad1Δ* (Fig. 1c, left). In contrast, human APE-1 protein failed to suppress this sensitivity, suggesting that the ability to complement CPT sensitivity in *tdp1Δ rad1Δ* cells is not a generic feature of metal-dependent phosphodiesterases (Fig. 1c, left). Conversely, whereas human APE-1 suppressed the sensitivity of AP endonuclease-defective *apn1Δ apn2Δ tpp1Δ* yeast cells to methyl methanesulphonate (MMS)-induced DNA base damage, human TTRAP did not, suggesting that the effect of TTRAP in these experiments was restricted to topoisomerase-mediated DNA damage (Fig. 1c, right). TTRAP contains four highly conserved motifs that putatively assign this protein to the metal-dependent phosphodiesterase superfamily (see Fig. 1a and Supplementary Fig. 1). We thus examined whether mutation of two predicted catalytic residues (Fig. 1a; Glu 152 and Asp 262) within two of these motifs affected the complementation of CPT sensitivity by TTRAP. Indeed, in contrast to wild-type TTRAP protein, neither TTRAP(E152A) nor TTRAP(D262A) restored CPT resistance in *tdp1Δ rad1Δ* cells (Fig. 1d).

We next purified recombinant human TTRAP, TDP1 and APE-1 from *Escherichia coli* (Fig. 2a, lanes 1–3), and incubated the purified proteins with a radiolabelled oligonucleotide duplex containing a nick with a single tyrosine covalently linked to the 3′-terminus by a phosphotyrosyl bond (Fig. 2a, inset). This is an established substrate for TDP1 that mimics the Top1-linked SSBs induced by CPT[6,8]. As expected, TDP1 cleaved the 3′-phosphotyrosyl bond and thereby converted the 3′-tyrosine terminus to a 3′-phosphate (Fig. 2a, lane 5). In contrast, recombinant human APE-1 failed to do so (Fig. 2a, lane 7). Notably, TTRAP cleaved the 3′-phosphotyrosyl bond, liberating DNA with a 3′-phosphate in a manner similar to TDP1, albeit at much higher enzyme concentrations and less efficiently (Fig. 2a, lane 6). TTRAP was also active on a double-strand break (DSB) substrate containing a 3′-tyrosine (Fig. 2b), albeit ~50-fold less efficiently and much slower than TDP1 (Fig. 2d, left panel and Supplementary Fig. 2). This activity was greatly reduced or lacking in the absence of Mg^{2+}, consistent with a dependency on metal cofactor for TTRAP activity (Supplementary Fig. 2). Moreover, the TTRAP activity detected here did not reflect a contaminant, because recombinant preparations of TTRAP(E152A) and TTRAP(D262A), purified in parallel with wild-type TTRAP, were largely or entirely inactive (Fig. 2b, lanes 6 and 7). We thus conclude that TTRAP possesses bona fide 3′-tyrosyl DNA phosphodiesterase activity, but that this activity is much weaker than that of TDP1, under the current experimental conditions at least.

In contrast to Top1, topoisomerase II (Top2) induces DNA DSBs in which the topoisomerase is covalently linked by a phosphotyrosyl bond to the 5′-terminus of the break[9]. Surprisingly, a human enzyme

[1]Genome Damage and Stability Centre, University of Sussex, Science Park Road, Falmer, Brighton, Sussex BN1 9RQ, UK. [2]Biochemistry Department, Faculty of Pharmacy, Ain Shams University, Abbassia 11566, Cairo, Egypt.
*These authors contributed equally to this work.

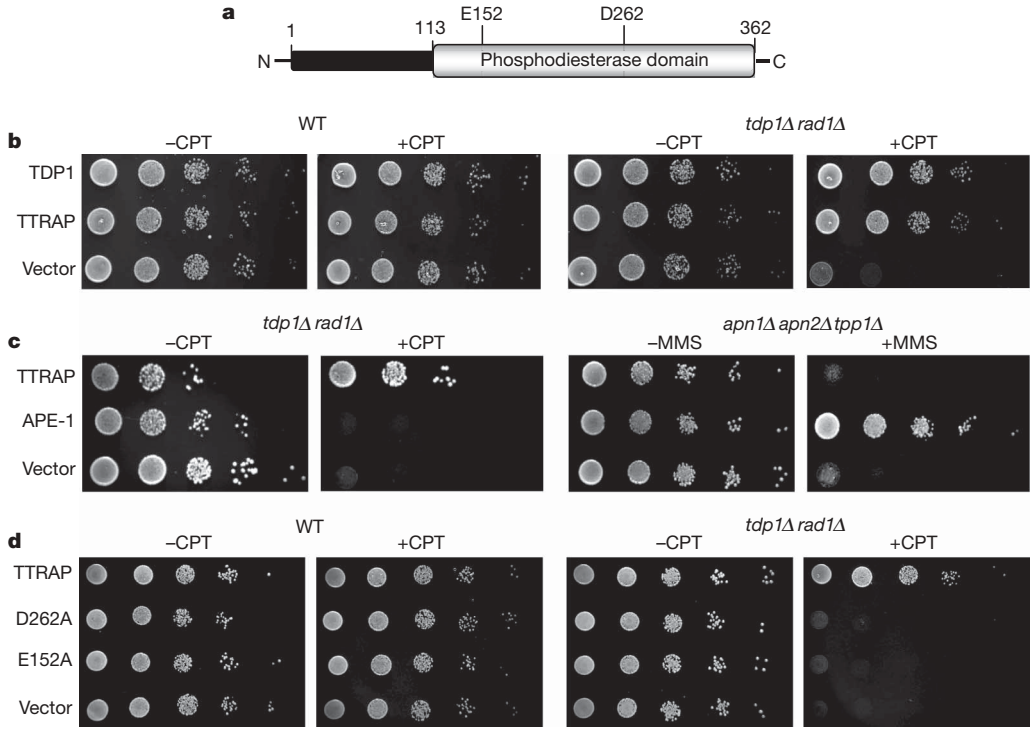


Figure 1 | Human TTRAP suppresses the sensitivity of yeast to Top1-induced DNA damage. a, Cartoon of TTRAP. **b**, TTRAP suppresses CPT sensitivity of *tdp1Δ rad1Δ* yeast. Tenfold serial dilutions of wild-type (WT, left) and *tdp1Δ rad1Δ* (right) cells containing pACT-TDP1-1, pACT-TTRAP-2 or pACT vector were plated onto media lacking (−) or containing (+) 20 μM CPT. **c**, Human APE-1 or full length TTRAP suppress the sensitivity of *tdp1Δ rad1Δ* or *apn1Δ apn2Δ tpp1Δ* cells to CPT or MMS, respectively. Tenfold serial dilutions of *tdp1Δ rad1Δ* (left) and *apn1 Δapn2Δ tpp1Δ* (right) transformed with pGBKT7-TTRAP, pAS-APE-1 or pGBKT7 vector were plated onto media lacking or containing 20 μM CPT or 100 μM MMS. **d**, Mutant TTRAP fails to suppress CPT sensitivity. Wild-type (left) or *tdp1Δ rad1Δ* (right) cells transformed with pVTU260 vector or pVTU260 encoding full-length untagged wild-type or mutant (E152A or D262A) TTRAP were analysed for CPT sensitivity as above.

that can preferentially or efficiently cleave DNA 5′-phosphotyrosyl bonds has not been reported, despite the established affect of this type of break on genetic stability and cancer[4,9]. Given that the 3′-tyrosyl DNA phosphodiesterase activity of TTRAP was relatively weak, we wondered whether this enzyme might be a 5′-tyrosyl DNA phosphodiesterase. To address this possibility we used a radiolabelled oligonucleotide duplex containing a DSB with a single tyrosine covalently linked to the 5′-terminus by a phosphotyrosyl bond (Fig. 2c, inset). Notably, whereas neither human APE-1 nor TDP1 cleaved the 5′-phosphotyrosyl bond, TTRAP led to the complete conversion of the DSB terminus to a 5′-phosphate (Fig. 2c, lanes 2–4). That the 5′-terminus of the TTRAP reaction product was a phosphate was confirmed using calf intestinal phosphatase, which converted the 5′-terminus of the reaction product to hydroxyl (Supplementary Fig. 3a). Notably, TTRAP activity on the 5′-phosphotyrosyl substrate was largely or entirely absent in reactions that lacked Mg^{2+} and contained EDTA (Supplementary Fig. 3b), or in reactions containing the phosphodiesterase mutant proteins TTRAP(E152A) or TTRAP(D262A) (Fig. 2c, lanes 5 and 6). Importantly, similar to TDP1 activity on 3′-phosphotyrosyl termini, the activity of TTRAP on 5′-phosphotyrosyl termini was robust and rapid (Fig. 2d and Supplementary Fig. 3c, d). In contrast, TTRAP was not active on a substrate containing an internal abasic site, a preferred substrate of APE-1, or on 5′-AMP DNA termini, a preferred substrate of aprataxin (Supplementary Fig. 4). We conclude from these experiments that TTRAP is a bona fide 5′-tyrosyl DNA phosphodiesterase.

We next examined whether we could detect 5′-tyrosyl DNA phosphodiesterase activity in human whole-cell extracts. Indeed, such activity was readily detected in A549 cell extract (Fig. 3, lanes 1–5). Moreover, this activity was reduced by ~50% in A549 cells in which TTRAP was depleted by ~80% by RNA interference (RNAi), and was increased by the addition of recombinant human TTRAP (Fig. 3). However, 5′-tyrosyl DNA phosphodiesterase activity in TTRAP-depleted extracts was not increased by the addition of recombinant TTRAP(D262A) or TTRAP(E152A) (Supplementary Fig. 5). Notably, wild-type and TTRAP-depleted A549 extracts showed similar levels of metal-independent 3′-tyrosyl DNA phosphodiesterase activity, and were equally able to repair a gapped DNA duplex that lacked damaged termini, suggesting that TTRAP-depleted extracts are competent for other types of DNA repair activity (Supplementary Fig. 5). These data indicate that human whole-cell extract possesses robust 5′-tyrosyl DNA phosphodiesterase activity, and suggest that TTRAP is a major source of this activity.

Finally, we examined whether the 5′-tyrosyl DNA phosphodiesterase activity of TTRAP fulfils a role in protecting against topoisomerase-induced DNA damage, *in vivo*. We first noted that the overexpression of TTRAP, but not of either TDP1 or the mutant proteins TTRAP(D262A) or TTRAP(E152A), measurably increased the resistance of *ISE2* cells—a strain of *S. cerevisiae* with increased drug permeability[15]—to the Top2 poison etoposide (Fig. 4a and Supplementary Fig. 6). We also noted that enhanced green fluorescent protein (EGFP)-tagged TTRAP accumulated in nuclear promyelocytic leukaemia (PML) bodies, as reported previously[16], and more importantly at sites of laser-induced UVA damage (Fig. 4b), consistent with a role for TTRAP in DNA repair. That this accumulation was relatively weak may reflect the fact that topoisomerase-induced DNA breaks are probably only a minor component of UVA laser-induced DNA damage[17–19]. We next examined the effect of TTRAP depletion on the sensitivity of human A549 cells to different types of DNA damage. Whereas TTRAP-depleted cells (see Fig. 3 for a typical level of depletion) had little or no sensitivity to CPT or MMS, these cells were hypersensitive to etoposide (Supplementary Fig. 7). These data suggest that TTRAP promotes the repair of Top2-induced DSBs in human cells. To explore this possibility further, we examined whether TTRAP depletion resulted in the accumulation of DNA strand breaks during exposure to topoisomerase poisons. Notably, TTRAP depletion

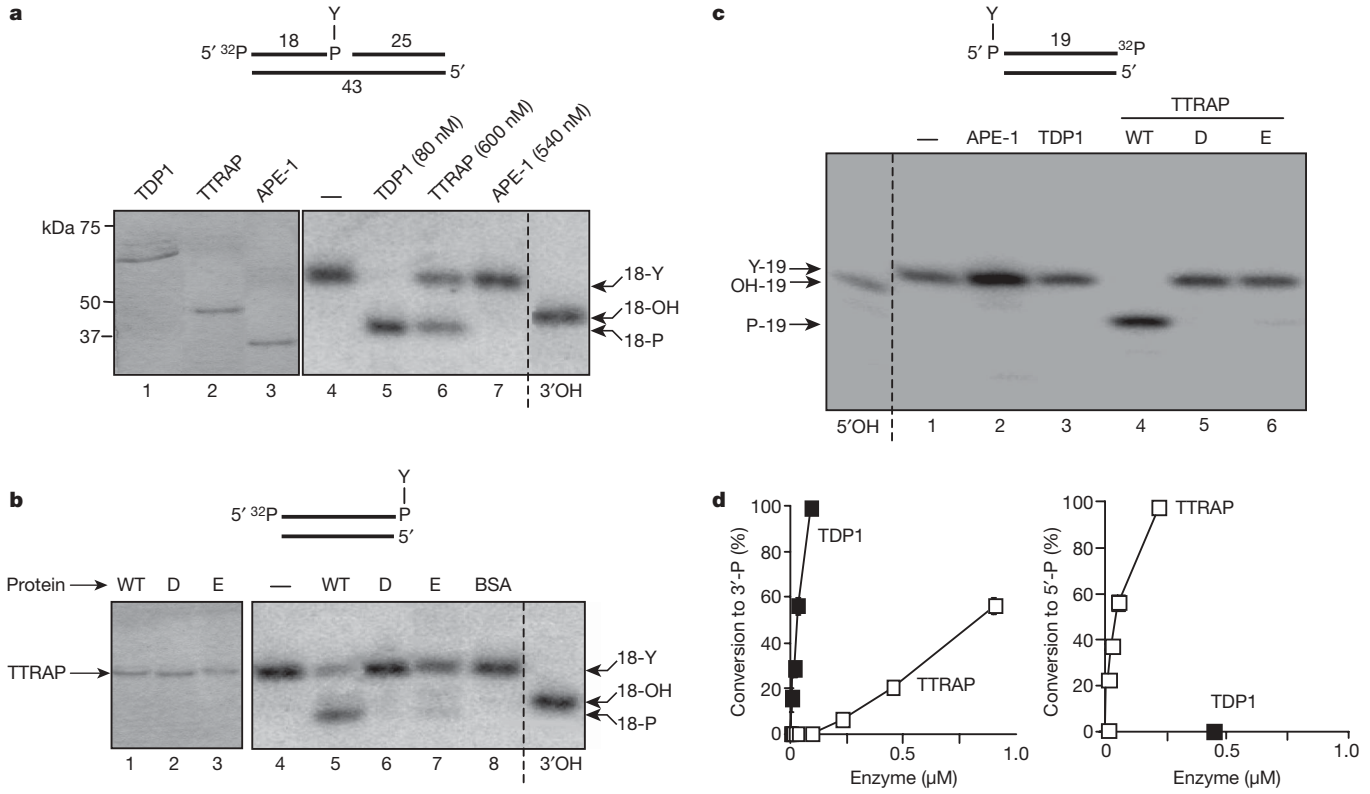


Figure 2 | TTRAP possesses 3′- and 5′-tyrosyl DNA phosphodiesterase (TDP) activity. **a**, 3′-TDP activity of TTRAP at SSBs. Lanes 1–3, SDS–PAGE of recombinant proteins. Lanes 4–7, ^{32}P-radiolabelled 43-base-pair (bp) oligonucleotide duplex (50 nM) containing a nick with a tyrosine (Y) linked to the 3′-terminus of the labelled 18-bp oligonucleotide (inset, top) was incubated for 1 h at 37 °C in the absence or presence of the indicated human proteins. Positions of ^{32}P-radiolabelled substrate (18-Y) and product (18-P) are indicated. An 18-bp oligonucleotide with a 3′-OH terminus (18-OH) was included as a marker (3′OH). **b**, 3′-TDP activity of TTRAP at DSBs. Lanes 1–3, analysis of human TTRAP (wild type), TTRAP(D262A) (D), or TTRAP(E152A) (E) by SDS–PAGE. Lanes 4–8, ^{32}P-radiolabelled 18-bp oligonucleotide duplex (50 nM) containing a DSB with a tyrosine linked covalently to the 3′-terminus of the labelled 18-bp oligonucleotide (inset, top) was incubated as above in the absence or presence of the indicated proteins (1 μM). **c**, 5′-TDP activity of TTRAP at DSBs. ^{32}P-radiolabelled 19-bp oligonucleotide duplex (50 nM) containing a DSB with a tyrosine linked covalently to the 5′-terminus of the labelled 19-bp oligonucleotide (inset, top) was incubated with of the indicated human proteins (150 nM) as above. Positions of ^{32}P-radiolabelled substrate (Y-19) and repair product (P-19) are indicated. A 19-bp oligonucleotide with a 5′-OH terminus (OH-19) was included as a marker (5′OH). **d**, Left, the 3′-TDP substrate (50 nM) in **b** was incubated with human TDP1 (0, 4.5, 9, 36 or 90 nM) or TTRAP (0, 36, 90, 225, 450 or 900 nM) for 1 h at 37 °C. Right, the 5′-TDP substrate (50 nM) in **c** was incubated with TTRAP (0, 15, 30, 60 or 240 nM) or 400 nM TDP1 for 1 h at 37 °C. Reaction products were quantified using ImageQuant and data are the mean ± s.e.m. of three experiments.

failed to increase the accumulation of DNA strand breaks during incubation with CPT (a Top1 poison), as measured either by alkaline-comet assays (Supplementary Fig. 7) or by the appearance of γ-H2AX (also known as γ-H2FAX) foci (Fig. 4c, left). This contrasted with the >10-fold increase in DNA strand breaks observed in mouse embryonic fibroblasts (MEFs) lacking Tdp1, which is the primary cellular 3′-tyrosyl DNA phosphodiesterase activity (Supplementary Fig. 7). However, treatment with etoposide resulted in the accumulation of ~twofold

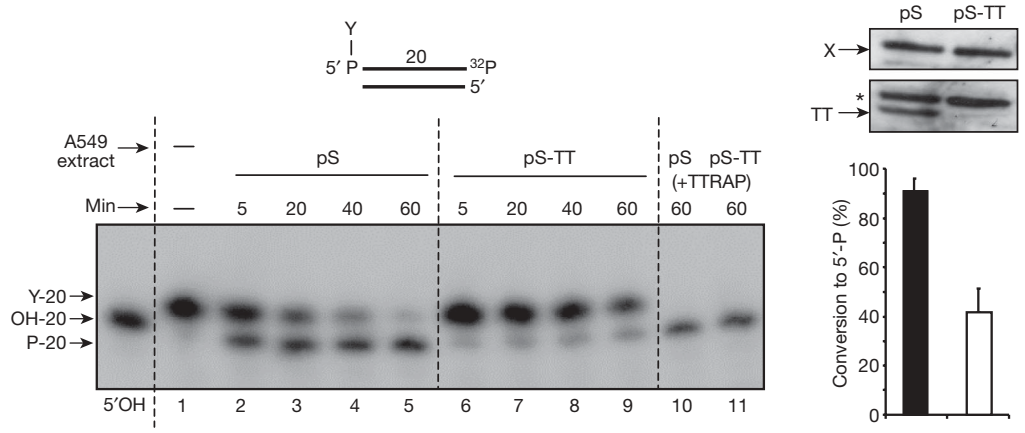


Figure 3 | Reduced 5′-TDP activity in TTRAP-deficient human cell extracts. Extract from A549 cells transfected with pSuper (pS) or pSuper-TTRAP (pS-TT) was incubated with a ^{32}P-radiolabelled 20-bp oligonucleotide duplex containing a DSB with a tyrosine linked covalently to the 5′-terminus of the labelled 20-bp oligonucleotide (inset, top). Reaction products were analysed by denaturing PAGE/phosphorimaging. The position of ^{32}P-radiolabelled substrate (Y-20), repair product (P-20), and 20-bp oligonucleotide 5′-OH marker (OH-20) are indicated. The percentage of substrate converted to reaction product was quantified (mean + s.e.m., $n = 4$) using ImageQuant (bottom right). Top right, western blot of XRCC1 (X) and TTRAP (TT) in the cell extracts used in these experiments. Asterisk denotes a nonspecific band.

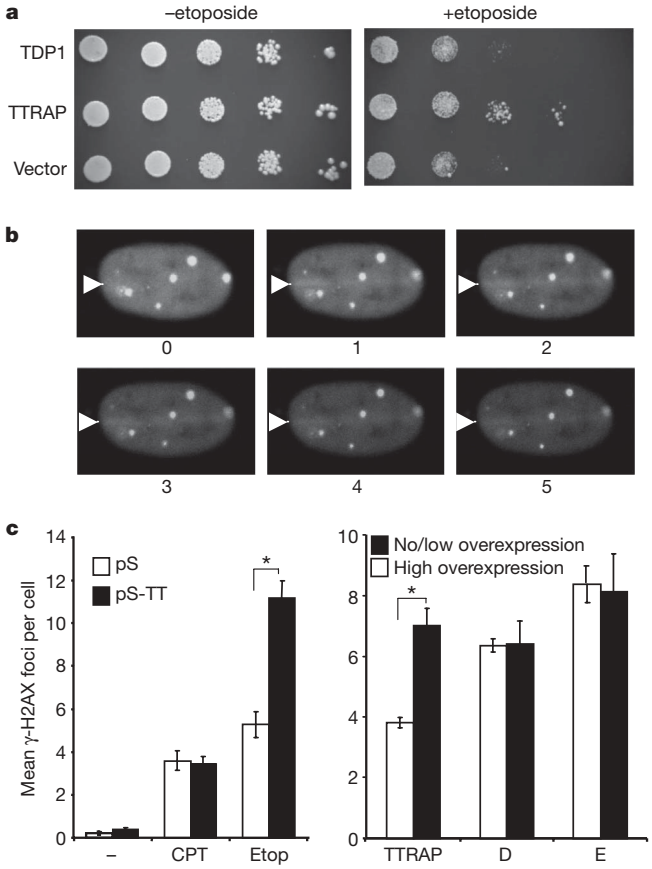


Figure 4 | TTRAP affects etoposide sensitivity and DNA damage accumulation. a, TTRAP increases etoposide resistance in yeast. Tenfold serial dilutions of *ISE2 S. cerevisiae* transformed with pACT-TDP1-1, pACT-TTRAP-2 or pACT vector were plated onto medium containing 500 μM etoposide. **b**, GFP–TTRAP accumulates at sites of UVA damage. Human A549 cells transiently transfected with pEGFP–TTRAP were damaged with a UVA laser and EGFP was imaged at the indicated times (min) (arrowheads denote position of laser track). **c**, TTRAP-depletion increases accumulation of etoposide-induced DSBs. A549 cells transfected with pSuper (pS) or pSuper-TTRAP (pS-TT) were treated with CPT (10 μM) or etoposide (etop, 10 μM) for 16 h and the average number of γ-H2AX foci per cell was quantified. Right, TTRAP overexpression decreases the accumulation of etoposide-induced DSBs. A549 cells transiently transfected with pcDNA3.1His-TTRAP, pcDNA3.1His-TTRAP(D262A) (D), or pcDNA3.1His-TTRAP(E152A) (E) were treated with etoposide as above and the average number of γ-H2AX foci quantified in cells either overexpressing recombinant TTRAP (open bars) or lacking visible recombinant TTRAP (filled bars). Data are the mean ± s.e.m.; $n = 3$). Asterisks denote statistically significant (paired *t*-test $P \leq 0.05$) differences.

more γ-H2AX foci in TTRAP-depleted cells than in mock-depleted control cells (Fig. 4c, left). Furthermore, the overexpression of wild-type TTRAP, but notably not of TTRAP(D262A) or TTRAP(E152A), reduced levels of etoposide-induced γ-H2AX foci in transiently transfected A549 cells (Fig. 4c, right and Supplementary Fig. 7). Together, these data support the notion that the 5′-tyrosyl DNA phosphodiesterase activity of TTRAP is required for the efficient repair of Top2-induced chromosomal DSBs.

It is noteworthy that the 5′-tyrosyl DNA phosphodiesterase activity of TTRAP can enable the repair of Top2-induced DSBs without the need for nuclease activity, creating a 'clean' DSB with 5′-phosphate termini that are ready for ligation. This contrasts with currently established mechanisms for DSB repair, which involve structure-specific nucleases such as CtIP (also known as RBBP8) and the MRE11–RAD50–NBS1 (MRN; NBS1 is also known as NBN) protein complex to 'trim' DSB termini[20,21]. TTRAP may thus provide an 'error-free' mechanism for direct end-joining of Top2-induced DSBs, a process that might have particular utility for maintaining genetic stability in long-lived non-cycling cells.

In summary, we have shown that human TTRAP is a 5′-tyrosyl DNA phosphodiesterase that is required for efficient repair of Top2-induced DNA double-strand breaks. Because TDP1 and TTRAP are complementary activities, together providing cells with an ability to remove trapped topoisomerase from both 3′- and 5′-DNA termini, we suggest that TTRAP is also designated tyrosyl DNA phosphodiesterase-2 (TDP2). It will now be of interest to determine how the tyrosyl DNA phosphodiesterase activity of TTRAP relates to the reported involvement of this protein in transcriptional regulation, apoptosis and embryonic development[13,22–24]. Given the central role fulfilled by topoisomerases in chromosome metabolism, our findings suggest that TTRAP may have other important functions in many aspects of cell biology, including the suppression of chromosome instability and cancer.

METHODS SUMMARY

Human TTRAP was recovered in a genetic screen for restoration of CPT resistance in *tdp1Δ rad1Δ* budding yeast cells by transformation with 16 μg of a pACT human cDNA library and plated on appropriate medium supplemented with 20 μM CPT (Sigma). For sensitivity assays, transformants were resuspended in 200 μl sterile H_2O and 10 μl drops of tenfold serial dilutions were plated onto medium lacking or containing the appropriate drug. The generation of oligonucleotide duplexes used for *in vitro* assays, DNA constructs used for expression of TTRAP in yeast, *E. coli* and mammalian cells, and the pSuper constructs used for short hairpin RNA (shRNA)-mediated depletion of TTRAP, are described in detail in the Methods. The recombinant human proteins APE-1, TDP1 and TTRAP used for *in vitro* assays were expressed in *E. coli* and purified by immobilised metal-chelate chromatagography (IMAC) followed by ion-exchange chromatography. The products of *in vitro* reactions were fractionated by denaturing PAGE and detected by phosphorimaging. The detection of EGFP–TTRAP accumulation at sites of UVA laser damage was achieved by seeding transfected A549 cells onto glass-bottom coverslips, followed by incubation with Hoechst 33258 dye and irradiation with a 351-nm UVA laser (approximately 0.35 J m^{-2}) using a Zeiss Axiovert. For clonogenic survival assays, A549 cells were plated into 10-cm dishes in duplicate and incubated at 37 °C for 4 h. Cells were then mock-treated or treated in medium at 37 °C with genotoxin (Sigma) for 1–3 h, and then incubated at 37 °C in drug-free medium for 14 days to allow the formation of macroscopic colonies. DNA strand breaks were quantified using the alkaline comet assay and by the quantification of γ-H2AX foci.

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

Received 29 April; accepted 20 August 2009.

1. Champoux, J. J. DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* **70,** 369–413 (2001).
2. Wang, J. C. Cellular roles of DNA topoisomerases: a molecular perspective. *Nature Rev. Mol. Cell Biol.* **3,** 430–440 (2002).
3. Li, T. K. & Liu, L. F. Tumor cell death induced by topoisomerase-targeting drugs. *Annu. Rev. Pharmacol. Toxicol.* **41,** 53–77 (2001).
4. Deweese, J. E. & Osheroff, N. The DNA cleavage reaction of topoisomerase II: wolf in sheep's clothing. *Nucleic Acids Res.* **37,** 738–748 (2009).
5. Pourquier, P. & Pommier, Y. Topoisomerase I-mediated DNA damage. *Adv. Cancer Res.* **80,** 189–216 (2001).
6. El-Khamisy, S. F. *et al.* Defective DNA single-strand break repair in spinocerebellar ataxia with axonal neuropathy-1. *Nature* **434,** 108–113 (2005).
7. Takashima, H. *et al.* Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. *Nature Genet.* **32,** 267–272 (2002).
8. Yang, S. W. *et al.* A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proc. Natl Acad. Sci. USA* **93,** 11534–11539 (1996).
9. Nitiss, J. L. Targeting DNA topoisomerase II in cancer chemotherapy. *Nature Rev. Cancer* **9,** 338–350 (2009).
10. Pommier, Y. Topoisomerase I inhibitors: camptothecins and beyond. *Nature Rev. Cancer* **6,** 789–802 (2006).
11. Vance, J. R. & Wilson, T. E. Yeast Tdp1 and Rad1-Rad10 function as redundant pathways for repairing Top1 replicative damage. *Proc. Natl Acad. Sci. USA* **99,** 13669–13674 (2002).
12. Liu, C., Pouliot, J. J. & Nash, H. A. Repair of topoisomerase I covalent complexes in the absence of the tyrosyl-DNA phosphodiesterase Tdp1. *Proc. Natl Acad. Sci. USA* **99,** 14970–14975 (2002).

13. Pype, S. *et al.* TTRAP, a novel protein that associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor-associated factors (TRAFs), and that inhibits nuclear factor-κB activation. *J. Biol. Chem.* **275,** 18586–18593 (2000).
14. Rodrigues-Lima, F., Josephs, M., Katan, M. & Cassinat, B. Sequence analysis identifies TTRAP, a protein that associates with CD40 and TNF receptor-associated factors, as a member of a superfamily of divalent cation-dependent phosphodiesterases. *Biochem. Biophys. Res. Commun.* **285,** 1274–1279 (2001).
15. Nitiss, J. L. *et al.* Amsacrine and etoposide hypersensitivity of yeast cells overexpressing DNA topoisomerase II. *Cancer Res.* **52,** 4467–4472 (1992).
16. Xu, G. L. *et al.* TTRAP is a novel PML nuclear bodies-associated protein. *Biochem. Biophys. Res. Commun.* **375,** 395–398 (2008).
17. Mielke, C., Christensen, M. O., Barthelmes, H. U. & Boege, F. Enhanced processing of UVA-irradiated DNA by human topoisomerase II in living cells. *J. Biol. Chem.* **279,** 20559–20562 (2004).
18. Mielke, C., Kalfalah, F. M., Christensen, M. O. & Boege, F. Rapid and prolonged stalling of human DNA topoisomerase I in UVA-irradiated genomic areas. *DNA Repair (Amst.)* **6,** 1757–1763 (2007).
19. Kingma, P. S. & Osheroff, N. The response of eukaryotic topoisomerases to DNA damage. *Biochim. Biophys. Acta* **1400,** 223–232 (1998).
20. Connelly, J. C. & Leach, D. R. Repair of DNA covalently linked to protein. *Mol. Cell* **13,** 307–316 (2004).
21. Hartsuiker, E., Neale, M. J. & Carr, A. M. Distinct requirements for the Rad32(Mre11) nuclease and Ctp1(CtIP) in the removal of covalently bound topoisomerase I and II from DNA. *Mol. Cell* **33,** 117–123 (2009).
22. Pei, H. *et al.* EAPII interacts with ETS1 and modulates its transcriptional function. *Oncogene* **22,** 2699–2709 (2003).
23. Esguerra, C. V. *et al.* Ttrap is an essential modulator of Smad3-dependent Nodal signaling during zebrafish gastrulation and left-right axis determination. *Development* **134,** 4381–4393 (2007).
24. Zucchelli, S. *et al.* Aggresome-forming TTRAP mediates pro-apoptotic properties of Parkinson's disease-associated DJ-1 missense mutations. *Cell Death Differ.* **16,** 428–438 (2009).

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

Acknowledgements This work was funded by the MRC (G0600776), the BBSRC (BB/C516595/1), and CR-UK (C6563/A10192). S.F.E.K. and F.C.L. were also funded by Fellowships from the Wellcome Trust (S.F.E.K.; 085284), Marie Curie (FCL; 2007-2-1-IEF-221222) and EMBO (FCL; ALTF 956-2006). We thank T. Wilson, H. Nash, M. Neale, J. Nitiss and E. Hoffmann for materials.

Author Contributions F.C.L. developed the genetic screen and conducted the mammalian cell experiments. M.C.Z. and F.C.L. conducted the yeast experiments. K.O. and S.F.E.K. prepared the recombinant proteins, and S.F.E.K. conducted the biochemical experiments. K.W.C., F.C.L. and S.F.E.K. designed and interpreted the experiments. K.W.C. coordinated the project and 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 K.W.C. (k.w.caldecott@sussex.ac.uk) and F.C.L. (fc55@sussex.ac.uk).

METHODS

Yeast strains, screening and sensitivity assays. All yeast strains used in this study have been previously described. Wild type (YW465), *tdp1Δ rad1Δ* (YW812) and *apn1Δ apn2Δ tpp1Δ* (YW778) were provided by T. E. Wilson[11,25]. The *ISE2* (JN362a) strain is a gift from M. Neale, originally obtained from J. Nitiss[26]. Screening of a human cDNA library was performed as follows. Yeast *tdp1Δ rad1Δ* cells were transformed with 16 μg of pACT human cDNA library[27] and plated on Synthetic Complete medium lacking leucine (SC-Leu) supplemented with 20 μM CPT (Sigma). A total of 7.5×10^5 clones were screened, as estimated from the transformation efficiency on SC-Leu lacking CPT. Healthy-looking colonies were isolated after 7 days, rescreened for growth in 20 μM CPT, and their pACT construct was isolated from *E. coli*. The phenotype was confirmed after retransforming these constructs into *tdpΔ rad1Δ* cells. Three full-length TDP1 clones were identified (denoted pACT-TDP1-1, -2 and -3), comprising nucleotides 234–2069, 97–2090 and 1–2101, respectively, of the published transcript variant 1 cDNA (NM_018319). Three incomplete clones of TTRAP were identified, one comprising nucleotides 90–1884 (pACT-TTRAP-1), and two comprising nucleotides 49–1889 (pACT-TTRAP-2 and -3) of the published cDNA sequence (NM_016614.2). For sensitivity assays, equal-sized colonies of the corresponding transformants were resuspended in 200 μl sterile H_2O, and 10 μl drops of tenfold serial dilutions were plated onto SC-Leu (for pACT transformants), SC-Tryp (for pAS/pGBKT7 transformants), or SC-Ura (for pVTU260 transformants) media lacking or containing the appropriate drug.

DNA constructs. Full-length *TTRAP* open-reading frame (ORF), flanked by NdeI and BamHI restriction sites, was cloned into pCRII-TOPO (Invitrogen) after PCR amplification from pACT-TTRAP-2 using the following primers (Operon) (the region of TTRAP missing from pACT-TTRAP-2 and included in the primer is underlined): 5′-AGGAAGCATATGGAGTTGGGGAGTTGCCTGGAGGGCGGGAGGGAGGCGGCGG-3′ (forward) and 5′-TGCAACGGATCCAATCAGGGCAAAACCCACAC-3′ (reverse). Mutations were introduced into pCRII-TTRAP with the Quickchange XL site-directed mutagenesis kit (Stratagene) using primers 5′-CCAGATGTGATATTTCTACAGGCAGTTATTCCCCCATATTATAGC-3′ (forward) and 5′-GCTATAATATGGGGGAATAACTGCCTGTAGAAATATCACATCTGG-3′ (reverse) for TTRAP(E152A), and 5′-GCTACAGTTATATTTGCAGGAGCTACAAATCTAAGGGATCGAGAG-3′ (forward) and 5′-CTCTCGATCCCTTAGATTTGTAGCGCCTGCAAATATAACTGTAGC-3′ (reverse) for TTRAP(D262A). pGBKT7 (Clontech) and pET16b (Novagen) constructs expressing TDP1 were previously described[6]. pGBKT7 (Clontech) and pET16b (Novagen) constructs expressing wild-type TTRAP, TTRAP(E152A) and TTRAP(E262A) were created by sub-cloning the *TTRAP* ORF from the appropriate pCRII-TOPO construct using NdeI and EcoRI. Untagged full-length human TTRAP yeast expression vectors were created by cloning the appropriate *TTRAP* ORF from pCRII-TOPO (NdeI-BamHI) into NheI-BamHI of pVTU260 (Euroscarf, provided by E. Hoffmann). pET16b-APE-1 has been described previously[28], and the *APE-1* ORF was cloned into BamHI-EcoRI sites of pAS2-1 (Clontech). pEGFP-TTRAP was created by cloning a BamHI-EcoRI fragment of pACT-TTRAP-2 into BglII-EcoRI of pEGFP-C3 (Clontech). Mammalian TTRAP expression constructs were created by cloning the appropriate *TTRAP* ORF from pCRII-TOPO into pcDNA3.1-HisC (Invitrogen) using EcoRI. For pSuper-TTRAP, the following oligonucleotides (Operon) containing an appropriate 19-base-pair (bp) region of homology to TTRAP (underlined) were annealed and sub-cloned into BglII-HindIII sites of pSuper (OligoEngine): 5′-GATCCCCGTACAGCCCAGATGTGATATTCAAGAGATATCACATCTGGGCTGTACTTTTTA-3′ (forward) and 5′-AGCTTAAAAAGTACAGCCCAGATGTGATATCTCTTGAATATCACATCTGGGCTGTACGGG-3′ (reverse).

Purification of recombinant human proteins and human cell extracts. Amino-terminal histidine-tagged recombinant human TTRAP, TTRAP(D262A), TTRAP(E152A), TDP1 and APE-1 were expressed from pET16b in BL21 (DE3) cells and purified by immobilised metal-chelate chromatography (IMAC) and ion-exchange chromatography. Proteins were then dialysed in buffer comprised of 25 mM Tris-HCl, pH 7.5, 1 mM dithiothreitol (DTT), 10% glycerol, and either 100 mM NaCl or 130 mM KCl. Purified human proteins were quantified on Coomassie-blue-stained polyacrylamide gels by comparison to bovine serum albumin (BSA) standards and verified by Bio-Rad Protein Assay Kit. Human whole-cell extracts were prepared from 3×10^6 A549 cells by lysis in 0.1 ml of 20 mM Tris-HCl, pH 7.5, 10 mM EDTA, 1 mM EGTA, 100 mM NaCl, 1% Triton X-100 and protease inhibitor cocktail (Sigma). The extract was clarified by centrifugation and quantified using a Bio-Rad Protein Assay Kit with BSA used as a standard.

Preparation of DNA substrates and *in vitro* repair reactions. Gel-purified oligonucleotides were 5′-labelled with ^{32}P using [γ-^{32}P]-ATP and T4 PNK or 3′-labelled using [α-^{32}P]-dCTP and Klenow DNA polymerase as described later. For the 19-bp oligonucleotide DSB 5′-phosphotyrosyl substrate, a 5′-Y-18-bp oligonucleotide (5′-Y-TCCGTTGAAGCCTGCTTT-3′) (Midland Certified Reagent Company) was annealed with a 19-bp oligonucleotide (5′-GAAAGCAGGCTTCAACGGA-3′) and the resulting 1-bp 5′ overhang filled-in with [α-^{32}P]-dCTP and Klenow DNA polymerase. For the 20-bp oligonucleotide 5′-phosphotyrosyl DSB substrate (used in reactions with whole-cell extract) the 5′-Y-18-bp oligonucleotide (above) was annealed to the 20-bp oligonucleotide (5′-AGAAAGCAGGCTTCAACGGA-3′), and the resulting 2-bp 5′ overhang filled-in as described earlier but using [α-^{32}P]-dCTP and ddTTP to inhibit degradation of the radiolabelled oligonucleotide by nonspecific nucleases. For the 43-bp oligonucleotide 3′-phosphotyrosyl SSB (nick) substrate, a radiolabelled 3′-Y-18-bp oligonucleotide (^{32}P-5′-TCCGTTGAAGCCTGCTTT-Y-3′) (provided by H. Nash) was annealed with a 25-bp oligonucleotide (5′-GACATACTAACTTGAGCGAAACGGT-3′) and a 43-bp oligonucleotide (5′-CCGTTTCGCTCAAGTTAGTATGTCAAAGCAGGCTTCAACGGAT-3′). For the 18-bp oligonucleotide 3′-phosphotyrosyl DSB substrate, the radiolabelled 3′-Y-18-bp oligonucleotide (above) was annealed with the 19-bp oligonucleotide (5′-AAAGCAGGCTTCAACGGAT-3′). For the AP endonuclease substrate, a 39-bp oligonucleotide (5′-GCGCAGCTAGCGGCGGATGGCXCCGTTGAAGCCTGCTTT-3′) containing an abasic site (X) obtained from Eurofins was 5′-labelled with ^{32}P and then annealed to a 40-bp oligonucleotide (5′-GAAAGCAGGCTTCAACGGAGCCATCCGCCGCTAGCTGCGC-3′). 5′-AMP substrates were prepared as described previously[29]. Reactions were initiated by mixing 4 μl of an appropriate dilution (in protein dialysis buffer) of the indicated recombinant human protein, or 4 μl of whole-cell extract (16 μg total protein), with 4 μl of 2× Master Mix (50 mM HEPES, pH 8.0, 260 mM KCl, 2 mM DTT and, unless otherwise indicated, 20 mM $MgCl_2$) containing 100 nM (for recombinant protein experiments) or 300 nM (for cell extract experiments) of the appropriate ^{32}P-labelled oligonucleotide duplex. For experiments using cell extract, 150 μM competitor single-stranded oligonucleotide (5′-CTAACTTGAGCGAAACGGT-3′) was present to reduce nonspecific nucleolytic degradation of the duplex substrate. All reactions were terminated by the addition of formamide loading buffer and fractionated by denaturing PAGE; images were analysed by phosphorimaging.

Cells, cell culture and transfection. A549 cells were cultured in DMEM (Gibco, Invitrogen) supplemented with 15% FCS and 1% glutamine. Cells were transfected using Genejuice (Novagene) according to the manufacturer's instructions. For the generation of wild-type and $Tdp1^{-/-}$ MEFs, $Tdp1^{-/+}$ animals were mated and 14-day embryos were dissociated, trypsinized and plated in DMEM with 10% FCS. Genotyping for the mutant *Tdp1* allele was performed on cell pellets, essentially as described[30]. For shRNA-mediated depletion of *TTRAP*, A549 lung carcinoma cells were co-transfected with 2 μg of pCD2E vector encoding G418 resistance and 1 μg of either pSuper or pSuper-TTRAP. After selection in 1.5 mg ml^{-1} G418 (Gibco, Invitrogen) for 6 days, cells were collected for analysis.

Immunoblotting, immunoflourescence and antibodies. For western blots, cells (10,000 per μl) were lysed in SDS–PAGE loading buffer and incubated at 90 °C for 5 min. Whole-cell extracts were fractionated by SDS–PAGE and transferred to nitrocellulose. Membranes were blocked for 1 h in PBS-Tween 20 (PBST) containing 5% non-fat dried milk (NFDM), and then incubated for 2 h with anti-XRCC1 monoclonal antibody (33-2-5) or affinity-purified anti-TTRAP polyclonal antibody (SY1340) at a 1/100 or 1/200 dilution, respectively, in PBST plus 5% NFDM. Membranes were rinsed in PBST and incubated in PBST plus 5% NFDM containing horseradish-peroxidase-conjugated anti-rabbit IgG or anti-mouse IgG (DAKO), as appropriate, at a 1/3,000 dilution for 1 h at room temperature. Membranes were then rinsed with PBST and antibody complexes detected by enhanced chemiluminescence (Amersham). Note that the anti-TTRAP polyclonal antibody was raised in rabbit by Eurogentec against full-length recombinant human TTRAP expressed in *E. coli*. For immunofluorescence analysis, cells were grown on coverslips, rinsed in PBS, fixed with 4% paraformaldehyde for 5 min at room temperature, permeabilized with 0.2% Triton X-100, and after rinsing in PBS twice, blocked in PBS plus 5% BSA. Cells were then incubated with appropriate primary antibodies in PBS plus 1% BSA for 1 h at room temperature. Anti-γ-H2AX (Ser 139) monoclonal antibody (Upstate) and anti-TTRAP rabbit polyclonal antibody (Abcam, ab33246) were used at 1/800 and 1/200 dilutions, respectively. After rinsing in PBST three times, cells were incubated for 30 min at room temperature in PBS plus 1% BSA containing appropriate secondary antibody (Alexa Fluor 555 goat anti-mouse IgG and Alexa Fluor 488 goat anti-rabbit IgG, Invitrogen) at a 1/200 dilution, and then rinsed again three times in PBST. Cells were counterstained with 4′,6′-diamidino-2-phenylindole (DAPI) and mounted with Vectashield (Vecta Laboratories).

Laser microirradiation. A549 cells were seeded onto glass-bottom coverslips (MatTek), transfected with 1 μg pEGFP-TTRAP, and 24 h later pre-incubated for 30 min with 10 μg ml^{-1} Hoechst dye 33258 at 37 °C. Selected cells were then

irradiated with a 351-nm UVA laser (approximately 0.35 J m^{-2}) focused through a ×40/1.2-W objective using a Zeiss Axiovert equipped with LSM 520 Meta, and images were taken every minute after irradiation.

Clonogenic survival assays. A549 cells transfected with empty pSuper or pSuper-TTRAP as described earlier were plated into 10-cm dishes (1,000 per plate) in duplicate and incubated at 37 °C for 4 h. Cells were mock-treated or treated in medium at 37 °C with the indicated concentration of CPT or etoposide for 3 h, or with MMS for 1 h. All drugs were from Sigma. Cells were then rinsed in PBS and incubated at 37 °C in drug-free medium for 14 days. Colonies were fixed with 90% ethanol and stained with 1% methylene blue. Survival was calculated as a percentage, using the equation $N_t/N_u \times 100$, in which N_t and N_u are the number of colonies on treated and untreated plates, respectively. Data are shown as the mean and s.e.m. of at least three independent experiments.

Alkaline single-cell agarose gel electrophoresis (comet) assays. DNA strand breaks were quantified using the alkaline comet assay as described previously[6]. Average tail moments from 100 cells per sample were obtained using Comet Assay III software (Perceptive Instruments), and data are shown as the mean and s.e.m. of at least three independent experiments.

γ-H2AX assays. Cells were mock-treated or treated with the indicated drug at the indicated concentration. After γ-H2AX immunofluorescence staining, the average number of γ-H2AX foci per cell was determined from 50 cells per sample using a Nikon Eclipse 50i microscope at ×100 magnification. Overexpression experiments were carried out as described above except that cells were transfected with 250 ng of the appropriate pcDNA3.1-His construct 1 day before incubation with the drug. Fifteen-to-forty cells per sample were scored, depending on the number of cells highly overexpressing TTRAP (see Supplementary Fig. 7 for representative examples). Data are the mean and s.e.m. of three independent experiments.

25. Vance, J. R. & Wilson, T. E. Repair of DNA strand breaks by the overlapping functions of lesion-specific and non-lesion-specific DNA 3′ phosphatases. *Mol. Cell. Biol.* **21,** 7191–7198 (2001).
26. Nitiss, K. C., Malik, M., He, X., White, S. W. & Nitiss, J. L. Tyrosyl-DNA phosphodiesterase (Tdp1) participates in the repair of Top2-mediated DNA damage. *Proc. Natl Acad. Sci. USA* **103,** 8953–8958 (2006).
27. Whitehouse, C. J. *et al.* XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair. *Cell* **104,** 107–117 (2001).
28. Parsons, J. L., Dianova, I. I. & Dianov, G. L. APE1 is the major 3′-phosphoglycolate activity in human cell extracts. *Nucleic Acids Res.* **32,** 3531–3536 (2004).
29. Reynolds, J. J. *et al.* Defective DNA ligation during short-patch single-strand break repair in ataxia oculomotor apraxia-1. *Mol. Cell. Biol.* **29,** 1354–1362 (2008).
30. Katyal, S. *et al.* TDP1 facilitates chromosomal single-strand break repair in neurons and is neuroprotective *in vivo*. *EMBO J.* **26,** 4720–4731 (2007).

NEWS

Older but not wiser

Almost all peer reviewers get worse, not better, over time, suggests a study presented on 10 September at the Sixth International Congress on Peer Review and Biomedical Publication in Vancouver, Canada.

Michael Callaham, editor-in-chief of the *Annals of Emergency Medicine* in San Francisco, California, analysed the scores that editors at the journal had given more than 1,400 reviewers between 1994 and 2008. The journal routinely has its editors rate reviews on a scale of one to five, with one being unsatisfactory and five being exceptional. Ratings are based on whether the review contains constructive, professional comments on study design, writing and interpretation of results, providing useful context for the editor in deciding whether to accept the paper.

V. WEHRMAN/CORBIS

The average score stayed at roughly 3.6 throughout the entire period. The most surprising result, however, was how individual reviewers' scores changed over time: 93% of them went down, which was balanced by fresh young reviewers coming on board and keeping the average score up. The average decline was 0.04 points per year.

"I was hoping some would get better, and I could home in on them. But there weren't enough to study," says Callaham. Less than 1% improved at any significant rate, and even then it would take 25 years for the improvement to become valuable to the journal, he says.

Others are not so convinced that older reviewers aren't wiser. "This is a quantitative review, which is fine, but maybe a qualitative study would show something different," says Paul Hébert, editor of the *Canadian Medical Association Journal* in Ottawa. A thorough review might score highly on the *Annals* scale, whereas a less thorough but more insightful review might not, he says. "When you're young you spend more time on it and write better reports. But I don't want a young person on a panel when making a multi-million-dollar decision."

Callaham agrees that a select few senior advisers are always very useful. But from his own observation, older reviewers do tend to cut corners. He notes that psychological research shows that experts in complex tasks typically reach a plateau and then stay there or slowly deteriorate. Perhaps by the time researchers are asked to review a paper at his journal, they are already experts. He suspects the same would hold true for journals across all fields.

Callaham also found that a mentoring programme at the *Annals*, in which new reviewers are paired up with senior ones, has only a temporary effect. Young reviewers assigned a mentor typically scored half a point better than non-mentored colleagues, but when the mentor's watchful eye disappeared after a year or so, this advantage evaporated. And the *Annals* dropped a separate peer-review training course some time ago, he notes, because although people loved it, it wasn't helping their scores. "It's kind of depressing," he says. ■

Nicola Jones

POSTDOC JOURNAL

A natural haven

Having worked as both a theorist and an experimentalist, I find theoretical work to be the most challenging emotionally because it's the most isolating. In my last postdoc, as an experimentalist, I relied on and interacted with other people for such daily tasks as preparing reagents. Work-based socializing was common. Now, as a theorist, I write computer code and solve equations by myself. Sometimes I go for days without any meaningful human interaction.

Isolation is not new to me. I grew up in a rural area, and when I felt sad or lonely I would head into the woods to look for birds, snakes and salamanders. Knowing the creatures around me made the world seem a little bit friendlier and less alien. But now, surrounded by the concrete of Baltimore, I no longer have an easy escape into the woods. So when a friend invited me to explore the Potomac River, about an hour away from Baltimore, I eagerly agreed.

Floating downriver on rubber inner-tyre tubes, we watched an osprey plunge into the water, saw a young bald eagle perched on a bank, heard the distinctive call of a barred owl and glimpsed a nighthawk hunting over the river at dusk. Thanks to this river retreat, my life of equations and computer code back in the city seems better, even days later. To paraphrase the author and naturalist Henry David Thoreau: not yet subdued by man, nature's presence refreshes him. ■

Sam Walcott is a postdoc in theoretical biophysics at Johns Hopkins University in Baltimore, Maryland.

IN BRIEF

CBI rallies industry

UK companies should play a bigger role in higher education, says a new report from the Confederation of British Industry (CBI). In *Stronger Together: Business and Universities in Turbulent Times*, the CBI, a lobbying organization, suggests that science, technology, engineering and mathematics (STEM) education would improve if industry provided, for example, increased input on course content and more work-experience opportunities. Students need more information about the "value business places on particular skills", says the report (see http://highereducation.cbi.org.uk). A 2009 CBI survey found that two-thirds of science, high-tech and information-technology UK firms believed the content of STEM degrees was not relevant to their needs.

Indians fast in protest

More than 1,500 science and engineering faculty members at the Indian Institutes of Technology (IITs) held a one-day hunger strike on 24 September.

The All-India IIT Faculty Federation is seeking withdrawal of two government rules: one barring IITs from recruiting professors with fewer than four years of teaching experience and assistant professors with fewer than three, and one capping the proportion of professors who can be promoted to a senior grade at 40%.

Federation president Muthuveerappan Thenmozhi says the rules undermine the autonomy of IITs and hamper recruitment. The federation also wants entry-level faculty salaries higher than the proposed 30,000 renminbi (US$625) per month.

Europe's research lagging

The quality, rate and results of scientific research in Europe lag behind those in the United States, according to a 21 September report from the Organisation for Economic Co-operation and Development (OECD). The *Economic Survey of the European Union 2009* report says government regulations stymie scientific collaboration and innovation because of a general lack of mobility of researchers across Europe. Many are forced to stay in a particular institution or country, it says, to retain research grants, for example. Nigel Pain, OECD senior economist and a report author, says easing mobility would facilitate research and innovation, and potentially create more research posts.

No place like home

Young eastern European scientists are returning to their home countries to set up labs — with mixed success. **Claire Ainsworth** tracks their progress.

D. TILLINGHAST/CORBIS

For Czech researcher Martin Anger, spring 2008 was a turning point. The 38-year-old cell biologist was just finishing a postdoc at the University of Oxford, UK, and was mulling over where he might set up his own lab. With two prestigious postdocs under his belt, he had to choose whether to apply to well-funded, well-known institutions in the United States and western Europe, or head home to the Czech Republic, and contend with low salaries, a developing research infrastructure and bitter political wrangling over science funding. He chose home.

"I believe people should go back and benefit their country," he says. "Also, the conditions for returning were good." Despite the drawbacks, Anger found decent funding and openings for young researchers.

Anger is not unlike many young researchers from the former Eastern bloc nations. Born in the late 1960s and 1970s, they witnessed the fall of the Berlin Wall in 1989 (see page 586) and headed overseas as PhDs or postdocs to expand their horizons and learn new research skills. Offered the option of remaining abroad, many are nevertheless returning home. Thanks to better funding in some cases from their home countries, and money from foreign grant funders, returning is now easier than a decade ago. But it's still not easy. Some have trouble garnering sufficient funding, or are battling old academic hierarchies that are bitterly opposed to change.

"I believe people should go back and benefit their country."
— Martin Anger

So what's the attraction? Many, like Anger, want to be part of their country's scientific development. Others also cite personal, family or cultural reasons for going back. "I never wanted to be an immigrant in another country," says Agnieszka Dobrzyn, a molecular biologist at the Nencki Institute of Experimental Biology in Warsaw. "Going abroad was always going to be a temporary measure."

In some cases, the past five to ten years have brought funding improvements. Individual governments have recognized the importance of enticing young scientists back to their native countries, and are tempting them with start-up money. In many cases, the funds come from deep-pocketed international sources such as the European Union, the Wellcome Trust, and the Howard Hughes Medical Institute, and play a significant part in making these homecomings a viable option.

The European Molecular Biology Organization (EMBO), for example, offers three-year Installation Grants; the Wellcome Trust has International Senior Research

Fellowships, which last five years and can be renewed. Anger received an installation grant together with a fellowship from the Academy of Sciences of the Czech Republic and a Marie Curie Reintegration Grant from the European Commission. Additional money from the academy has allowed him to buy the state-of-the-art live-cell imaging equipment needed for his lab at the Institute of Animal Physiology and Genetics in Liběchov.

His colleague Petr Svoboda, another EMBO and academy grant recipient, established his lab at the Institute of Molecular Genetics in Prague in 2007. Svoboda says the facilities and equipment there are world-class. "The infrastructure here is the same you would get in Germany or the United States," he says. "The physical environment we work in is exactly the same." The same is true of other former Eastern bloc nations, some researchers say. The past decade has brought significant improvement. "We have much better facilities and better access to equipment," says Arnold Kristjuhan, a lab head studying chromatin modifications at the University of Tartu in Estonia. "We have better international connections."

Juggling act

The situation isn't perfect, however. Researchers often have to manage several grants at once to stay afloat and to employ personnel. In the Czech Republic, for example, standard postdoc salaries are not very attractive. "If you want to have a postdoc, you have to assemble a salary from several different sources," says Svoboda.

Postdoc salaries amount to as little as a quarter or a sixth of those in the West, which makes it hard for new lab heads building their teams to attract foreign candidates. And although foreign grants can supply more generous salaries, this can put a postdoc in the uncomfortable position of earning substantially more than an institutional professor, says László Acsády, a Wellcome Trust fellow at the Institute of Experimental Medicine in Budapest. Although this is acceptable for a foreigner, says Acsády, realpolitik would preclude the presentation of such an offer to a Hungarian researcher. "Most of the time, if you can't get a Western European, you hire two eastern Europeans," he quips.

But money isn't the only problem. Many eastern European countries are still slowly restructuring their old-fashioned academic set-ups, in which senior professors would preside over huge labs of junior scientists and monopolize research funds. "The problem is the hierarchical academic system," says Gerlind Wallon, deputy director of EMBO and a manager for the Young Investigator Programme. "It's something they have to get rid of." The idea of a thirty-something researcher setting up shop independently is still alien to many of the old guard. "It's unbelievable sometimes — these professors are treated like gods," says Maciej Żylicz, president of the Foundation for Polish Science in Warsaw and head of the department of molecular biology at the International Institute of Molecular and Cell Biology (IIMCB) in Warsaw. Young, dynamic group leaders returning from abroad with large grants can infuriate incumbents, especially when the younger scientists start turning out more papers than their superiors, who are not used to being judged on their output.

Such mindsets, and some institutes' funding schemes, mean that young researchers sometimes struggle to establish an independent lab. The Western-style career path — starting with the PhD, followed by a postdoc or two, followed by a position as head of an independent lab — has been slow to catch on in some countries. "This scheme is absolutely lacking in Hungary," says Acsády. "It needs to be established."

V. STEPANETS

"The infrastructure here is the same as you would get in the United States."
— Petr Svoboda

Like many, Dobrzyn, had trouble finding an independent position when she returned home. Luckily for her, two forward-thinking institutions, the Nencki and the IIMCB, decided to run annual competitions for young group-leader positions. Dobrzyn won hers at the Nencki in 2007. Funding bodies such as EMBO and the Wellcome Trust are also pushing for change, insisting that institute heads guarantee the independent status of researchers before grants are awarded.

Once a researcher's lab is set up, problems accessing funding mean it can still be difficult to remain competitive at an international level. And that's not just because grants are scarce. Many eastern European institutions are still fumbling to grasp how international grants work and how to write effective applications. The Nencki has a team that hunts for funding opportunities and helps with applications. "It's a great help for us," says Dobrzyn.

Returning scientists typically face problems and political situations unique to their countries. The Czech government, for example, has dismayed basic researchers by announcing plans to cut the funding for the Czech science academy in order to fund more applied research. It is also attempting to impose a metric for measuring research that is linked to total output, rather than to the quality of original papers — a move that Czech scientists, including new lab heads, are trying to fight (see *Nature* **460,** 157; 2009). "There is a lack of a coherent research policy," complains Svoboda. "There is no strategy at the governmental level."

In Estonia, the problem isn't support for basic science, says Kristjuhan. It's the small size of the scientific community, which makes it harder to secure collegial advice on an experiment or draft manuscript. "Everyone is studying their own topic; it's hard to find [fellow] experts in the same institute or even within the country," Kristjuhan says.

Political drive

The desire for change has prompted young lab heads to mobilize and lobby politicians and government officials. Acsády, for example, belongs to an e-mail list of Hungarian group leaders who write letters and send petitions to government officials asking for change. And they have achieved results: he and his colleagues recently persuaded the government to earmark funding for Hungarian researchers who were unsuccessful in their bids for European Young Investigators Awards. Those who were awards finalists can now compete for the Hungarian government money.

National academies of sciences, international funders and politicians are also pushing for change. The Polish Foundation for Science, for example, has created three grant schemes to attract researchers at different stages of their career. It is also asking the Polish government to shift science funding more towards competitive grants, rather than lump-sum allocations to institutions. Stung by Poland's failure to win funding from the European Research Council last year (see *Nature* **453,** 558–559; 2008), politicians are listening. A new research council for basic science, which will manage grant money, is in the works, as is a scheme to encourage older professors to take early retirement. "We hope this will crash the hierarchy," says Żylicz.

Countering some of these predicaments is the apparent quality of these labs' science. The Wellcome Trust, which started offering fellowships to eastern European countries in 2002, reviewed the scheme in 2007. It was sufficiently impressed with the research results to change its policy and allow existing fellows to apply for grant renewal. "Many of these labs are at the cutting edge internationally," says Candace Hassall, head of basic careers at the Wellcome Trust.

Despite those ongoing thorny problems, many researchers remain sanguine. "This separation for 40 years cannot be undone in 20 years," says Anger. "It is still not easy, but it does change," says Dobrzyn. "It's worth it to come back." ■

Claire Ainsworth is a freelance writer based in Southhampton, UK.

For more on eastern Europe, see http://tinyurl.com/eeurope.