
TEAM tAO

Volume 436 Number 7052 pp753-888

In this issue (11 August 2005)

Editorials

Research Highlights

News

News Features

nature

www.nature.com/nature

Vol 436 | Issue no. 7052 | 11 August 2005

Keeping religion out of science class

President Bush's endorsement of 'intelligent design' has sparked a national debate in which scientists are well positioned to prevail.

Comments made last week by President George W. Bush have encouraged advocates of 'intelligent design', the idea that a hidden hand must lie somewhere behind the evolution of life. But the scientific community has responded energetically and effectively to Bush's statements. If researchers persevere, they can win this argument and keep religion out of the biology classroom.

In an informal 90-minute interview with newspaper reporters from his home state of Texas on 1 August, Bush talked at length about everything from abortion to illegal immigration. When asked about his view of intelligent design, the president was hesitant at first, re-iterating that, as governor of Texas, he had supported the right of local school boards to choose what is taught in the classroom. Prodded further to give an opinion, Bush said: "I'm not suggesting — you're asking me whether or not people ought to be exposed to different ideas, and the answer is yes."

A great deal has been made of the president's off-the-cuff endorsement of intelligent design. But it is worth noting the hesitancy of his comments: he didn't even mention intelligent design by name. Bush is a self-proclaimed evangelical Christian, and religious conservatives are often credited with helping him win last year's election. But he knows perfectly well that the introduction of religious ideas into science class is a potentially explosive issue in US politics — and not necessarily one that works in his favour.

The teaching of creationism in schools has been energetically advocated in parts of the United States for decades, but it has rarely help politicians to win elections, even in the heartland. Take Kansas, whose school board sought in 1999 to restrict the teaching of evolution in public schools. A vocal outcry by scientists, business leaders and others soon led to the electoral defeat of the board members in question and a repeal of the restrictions (see *Nature* **406,** 552; 2000).

Those Republicans who anticipate an encounter with the electorate are therefore cautious about the open embrace of intelligent design. Senator Rick Santorum (Republican, Pennsylvania), for example, is a Catholic and one of the most conservative members of the Senate, who faces a tough re-election battle next year. Just after Bush's remarks, he forthrightly declared: "I'm not comfortable with intelligent design being taught in the science classroom." Perhaps Santorum judges, wisely, that otherwise-conservative voters in the suburbs of Philadelphia and Pittsburgh would draw the line at backing a senator who wants to teach creationism to little Brad and Britney in biology class.

Even hard-line creationists are aware of the political price of appearing to be against science in a nation that puts so much stock in it. Part of the idea of intelligent design, of course, is subterfuge, an attempt to introduce religion under the guise of science. In Kansas, after the last row died down, a more conservative school board was elected once again, and now may approve the inclusion of intelligent design in its school curriculum. The progress of this effort — which scientists have once again made a concerted effort to oppose — will be a bellwether for the rest of the country.

"Part of the idea of intelligent design is subterfuge, an attempt to introduce religion under the guise of science."

As we report on page 761, scientific leaders have responded effectively to Bush's ill-advised comments. Researchers from a range of disciplines have spoken out vigorously, making it clear that the president's apparent willingness to allow intelligent design into the classroom is at odds with America's pressing need to improve science education. They should continue to speak out in all available forums.

This argument has reverberated in US politics since at least 1925, when John Scopes, a high-school teacher in Dayton, Tennessee, was convicted and fined for teaching evolution. The fight will go on — but science and reason can ultimately win. ■

Life in the old doc yet

The need to bring new blood into science shouldn't force out talented older researchers prematurely.

There is a common misperception that genius and exceptional scientific achievement are the preserve of the young. A News Feature on page 772 of this issue shows the limitations of this view: some scientists are performing creative and groundbreaking work into their 70s and beyond. Yet many of them feel frustrated at what they see as unjustified obstacles in their path — including, in some countries, mandatory retirement for university professors.

Mandatory retirement policies were generally introduced, along with decent pension provisions, to ensure dignity and leisure for the old, while opening up job opportunities for the young. But there is a big difference between being forced to work to avoid penury and doing so because you love what you do. Most scientists' work is far more than just a job — it is a vocation, pursued with a passion that cannot be switched off overnight.

And nor should it be. Carl Friedrich Gauss may have surprised his school-teacher at the age of seven when he worked out the factorial of 100 in his head. But although the German mathematician made many of his important discoveries before he was 20, he continued to make progress in a wide variety of fields into his seventies. Charles Darwin wrote two of his best books, *The Descent of Man* and *The Expression of the Emotions in Man and Animals*, in his sixties.

Science needs the energy and freshness of vision that youth can

bring, but it can also benefit from the perspective and experience of age. The qualitative research that has been done on the matter reinforces the common-sense view that groups of people from a wide range of backgrounds are better at problem solving than are homogeneous ones. Science needs to take steps to keep the door open to exceptional older minds.

This must not, however, mean closing the door on younger researchers. In some countries — Germany and Japan spring to mind — the security and immovability of a ruling caste of established senior professors sometimes does just that. Even in countries with more flexible university systems, such as the United States, it has never been tougher for young scientists to win faculty positions (see *Nature* **422,** 354–355; 2003).

At the same time, most countries face demographic and financial pressures that are forcing them to reconsider their general approach to mandatory retirement. Falling birth rates and higher life expectancy mean that there will soon be far fewer working people to support pensioners' income, from public or private sources. Poor stockmarket performance and inadequate provision for state pensions compound the problem. The subsequent discussion of higher retirement ages doesn't thrill the public. But in science, it should allow for a fuller consideration of measures that will tap the energy of older researchers who want to keep working.

It is time for policy-makers and university administrators to look more imaginatively at ways of structuring academic careers to introduce greater flexibility and alternative options. Just as science stands to benefit from arrangements that will better accommodate women with children, or attract more interest from ethnic minorities, it can benefit from active measures to help older researchers who still have what it takes.

Greater use of the 'emeritus professor' system, where researchers keep an association with an institution while drawing their salary from their pension, can tap the brains of older specialists who are no longer active in research. Research institutions and governments should also be searching for ways to accommodate those who want to remain active in research and have a contribution to make. ■

Tale of two tigers

Aspirant research hubs in southeast Asia have enjoyed contrasting fortunes.

When Malaysia and Singapore became politically distinct entities in 1965, a yawning chasm soon developed between the two nations' scientific performance — and it continues to grow.

Under the skilful tutelage of Philip Yeo, a businessman and former engineer, Singapore is creating one of the world's most dynamic research environments. The system built up by Yeo since he took control of the nation's science policy has given the country's biotechnology the promise of a bright future (see page 767). Despite some questions about the commitment of Singapore's government to free expression, the island nation has succeeded in attracting both leading scientists and substantial business investment from abroad, offering grants with few strings and a cosmopolitan research environment that is conducive to international collaboration.

Some concerns have been expressed about the reliance of Singapore's success on Yeo's own energy and influence. But there is ample evidence that he has put in place a meritocratic system that will comfortably outlive the oversight of one individual. In just two decades, Singapore has established itself as an important regional hub for biological research, with particular strengths in genomics and cancer research.

Malaysia's attempts to establish itself scientifically have been far less successful (see *Nature* **436,** 620–621; 2005). The country has sought in vain to establish a biotechnology industry, the most conspicuous failure being the BioValley project near Kuala Lumpur. New independent universities, emphasizing their dedication to cutting-edge research, have been built after considerable political negotiation, but they are struggling to establish themselves as internationally competitive.

Part of Malaysia's problem lies in educational and hiring policies that favour native ethnic Malays, at the expensive of native ethnic Chinese and Indians, as well as foreigners. It is understandable that the Malaysian government wants to take measures to preserve the opportunities for a group that makes up the majority of its population, but which, by some measures, remains economically and socially disadvantaged.

But accusations abound that the existing system is not sufficiently meritocratic, and that personal connections often predominate in decision-making, with regard to both hiring opportunities and the distribution of grants. Science also enjoys insufficient autonomy within the Malaysian government. And in instances such as the BioValley project, investment has been showered on infrastructure, rather than on people.

Malaysia, of course, is not the only country in the region where research suffers from political mismanagement. China, Japan and others are also prone to over-investment in facilities, as opposed to personnel; difficulties in the fair evaluation of grants; and problems in fully engaging with outside scientists from neighbouring countries and further afield.

"Singapore's success could serve as a focal point for more scientific collaboration in the region."

Singapore's success provides a useful model in this regard. It could also serve as a focal point for more scientific collaboration in the region. East Asian countries have never found it easy to work together. The Asia-Pacific International Molecular Biology Network (www.a-imbn.org), for example, has sought for years to build a firmer foundation in the manner of the European Molecular Biology Organization, which helped to build the European Molecular Biology Laboratory in Heidelberg, Germany.

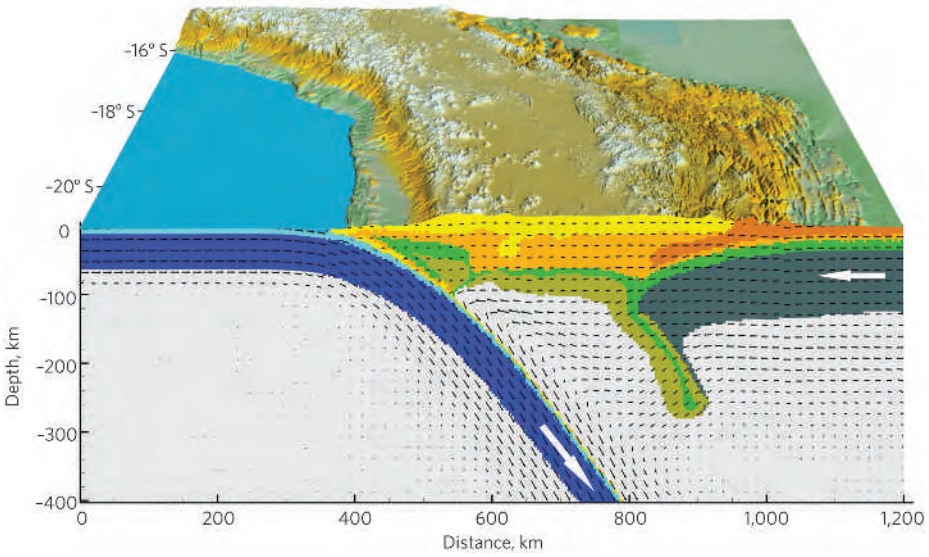
Singapore, with its recent track record of attracting top-notch researchers, would be the ideal location for any such regional collaboration. It could also act as a bridge between researchers in China, Japan and South Korea, where bitter rivalries have hamstrung such projects in the past. And it might enable Singapore to export some of the approaches which Yeo has so successfully championed in his own country. ■

RESEARCH HIGHLIGHTS

GEOPHYSICS

Up the Andes

Geology **33,** 617–620 and 621–624 (2005)
For more than 200 million years, the South American plate slid westwards over the Nazca plate without buckling. Then, in the past 30 million years, the spectacular plateau of the central Andes was pushed up. What changed? Using a computer model, Stephan Sobolev and Andrey Babeyko, both at Germany's National Research Centre for Geosciences in Potsdam, have shown that the most important factor was accelerated drift of the South American plate. The model reproduces features of the plates' interaction, including the rolling back of the subducting Nazca plate, and the forcing down of portions of the continental plate under the plateau.



S. SOBOLEV

NEUROSCIENCE

A life of regret

Nature Neurosci. doi: 10.1038/nn1514 (2005)
Regret springs from a region of the brain called the medial orbitofrontal cortex, according to a study that used functional magnetic resonance imaging to analyse peoples' response to a gamble.

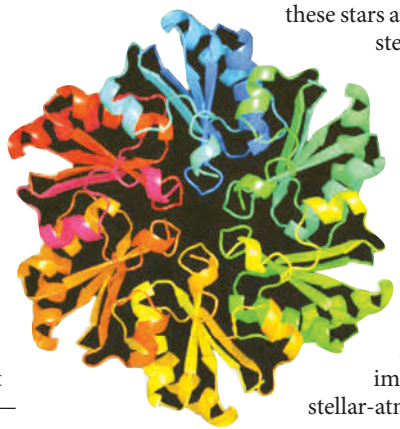
A team led by Raymond Dolan of the Wellcome Department of Imaging Neuroscience in London and Angela Sirigu of the Institute for Cognitive Science in Lyon offered 15 subjects the choice of two gambles, one riskier — but for higher stakes — than the other. When those who took the safe option were told they would have won the riskier bet, their medial orbitofrontal cortex became active. This area then fired before the regretful gambler's later choices, suggesting it is also involved in anticipating regret.

CELL BIOLOGY

Tiny organs

Science **309,** 936–938 (2005)
Scientists in California have revealed the molecular shape of a bacterial organelle. Its structure is uncannily similar to that of certain viruses.

Organelles are usually found in more complex eukaryotic cells — the plant chloroplast is one example — where they provide controlled environments for key biochemical reactions. But some bacteria contain protein shells that do a similar job. A team led by Todd Yeates at the University of California, Los Angeles, determined the structure of one such organelle, a carboxysome, which is involved in processing carbon. Its building blocks form hexagonal plates (pictured below) that are assembled into a polyhedral shell, similar to the protein coats of some viruses. The shell also has an electrically charged hole that may control what enters and leaves the structure.



E. YEATES

ASTRONOMY

Not so cool

Astrophys. J. **628,** 973–985 (2005)
Red supergiants begin life as stars at least 15 times more massive than our Sun, and then swell to hundreds of times the Sun's radius.

But observations have suggested that these stars are cooler than theories of stellar evolution can explain. A team of astronomers has now gone some way to raising these stars' temperatures and reconciling the two lines of evidence. By fitting new spectral measurements of the light from 74 supergiant stars to improved versions of stellar-atmosphere models called MARCS, they find that the stars are up to 400 K hotter than previously calculated, with temperatures typically around 3,000–4,000 K.

CELL BIOLOGY

Gas binding

Cell **122,** 195–207 (2005)
Nuclear receptors are a family of proteins that can switch some genes on and off, but the ligands they normally bind are known in only a few cases.

Henry Krause from the Charles H. Best Institute in Toronto and his colleagues have now identified the simple gases nitric oxide and carbon monoxide — short-lived signalling molecules important in many physiological functions — as natural ligands of the *Drosophila* nuclear receptor E75.

They show that the receptor includes an iron-containing haem molecule within its ligand-binding pocket. The state of oxidation of the haem affects the ability of the E75 receptor to bind the gases.

BIOGEOGRAPHY

Gluttony and sloths

Proc Natl Acad. Sci. USA doi: 10.1073/pnas.0502777102 (2005)
The Americas were once home to at least 19 genera of sloth species, some weighing as much as an Asian elephant. It has been unclear whether climate change or human hunters caused their demise, as the two hit much of the Western Hemisphere at roughly the same time. To resolve the issue, David

Steadman at the Florida Museum of Natural History in Gainsville and his colleagues focused on the Caribbean, where the climate changed 7,000 years before humans arrived. Carbon dating of bones showed that the sloths disappeared only after humans reached the islands. The finding adds weight to the idea that hungry humans wiped out the New World's large mammals.

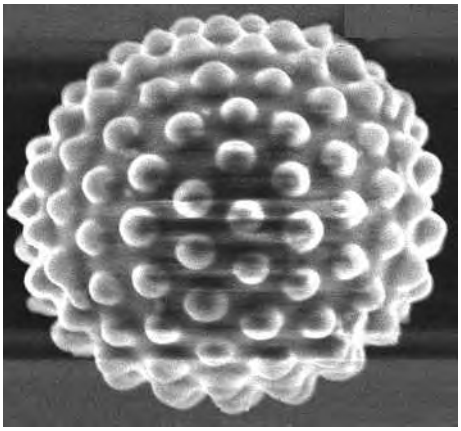
MATERIALS

Natural spirals

Science **309,** 909–911 (2005)

Spiral patterns similar to those seen in flower heads decorate the nanoparticles made by Zexian Cao's team at the Chinese Academy of Sciences in Beijing (pictured top right). Like the floral spirals, the microscopic swirls are described by Fibonacci number patterns.

The particles, measuring about 10 micrometres across, have a silver core with a silica coating. They are formed at 1270 K, slightly above the melting point of silver but well below that of silica. As the particles cool and shrink, stresses in the outer layer raise dimples on the shell. The patterns depend on the size of the particle, the thickness of its shell and the rate of cooling.



SCIENCE

BIOTECHNOLOGY

A clean break

Nature Methods doi:10.1038/nmeth787 (2005)

Separating a genetically engineered protein from the bacterial cell that made it has got easier, thanks to a trick developed by David Wood and his colleagues at Princeton University.

They fuse the gene encoding the target protein with genes encoding a sticky tail. The tail consists of a molecule called intein, which connects to the target protein, and a stretchy polypeptide similar to elastin. When the cell extract is heated, the tails get entangled, and the protein precipitates out of the mix. The intein links disintegrate in 4–10 hours, severing the protein's unwanted tail.

GENETICS

Noisy expression

Nature Genet. doi:10.1038/ng1616 (2005)

Fluctuating protein levels inside cells can arise from seemingly random variation in the rate that genes are expressed. To study the source of such noise, Alexander van Oudenaarden of the Massachusetts Institute of Technology and his colleagues looked at weakly active genes in the yeast *Saccharomyces cerevisiae*. Fluctuations in the activity of these genes became synchronized when they were engineered to be close together on the chromosome. This positional effect suggests that extrinsic factors, such as spatial variation within the cell's nucleus in the concentration of molecules that control gene activity, may cause the noise.

CLIMATE CHANGE

Greenland gets warmer

Geophys. Res. Lett. **32,** L14705 (2005)

Researchers have teased apart two factors affecting Greenland's climate, collecting data in agreement with climate-model predictions that global warming is having a disproportionately large effect on the island.

Temperature trends in Greenland are dominated by local climate patterns such as the North Atlantic Oscillation (NAO), which is thought to have cooled Greenland since the 1930s. To detect the change underlying the oscillation's effect, Petr Chylek of Los Alamos National Laboratory in New Mexico and Ulrike Lohmann of the Swiss Federal Institute of Technology in Zurich found a weather station on Greenland's northeast coast where temperatures for the past 30 years seem unaffected by the NAO. Temperatures at this station are rising 2.2 times faster than the global average.

JOURNAL CLUB

Peter Seeberger
Swiss Federal Institute of Technology Zurich, Switzerland

A new twist to an old reaction should break down barriers between chemists, argues one who has worked in two fields.

Every day, people in my laboratory are performing glycosylation reactions. This is the basic reaction that connects two sugars, and it is key to the study and exploitation of carbohydrate molecules. We need carefully synthesized structures to probe the biological role of carbohydrates and explore their therapeutic potential.

The problem is that the glycosylation reaction, despite being described in 1901, is one of the least predictable transformations in organic synthesis. The sugars link through a carbon and an oxygen atom. These can bond in two mirror-image configurations, producing two stereoisomers with different biological activity. Controlling the stereochemistry of some linkages is easy, but for some it is very hard. This was a problem when my group synthesized a carbohydrate vaccine for malaria, for example.

Recent work by Geert-Jan Boons and his colleagues at the University of Georgia (*Angew. Chem.* **44,** 947-949; 2005) describes a nice trick that will offer control over some reactions. A group needed to shield a reactive site in the first sugar is chosen so that it is also a chiral auxiliary — a chemical add-on that allows only one stereoisomer to form. The auxiliary forces the molecule into a ring-shaped intermediate that the incoming sugar can attack from only one side.

This shows that even in a seemingly mature field there is room for innovation. Organic chemistry routinely uses chiral auxiliaries, but carbohydrate chemistry has been seen as a separate discipline. Having worked in both fields, I am trying to break down the boundaries.

NEWS

Medics braced for fresh superbug

J. GUEZ/AFP/GETTY IMAGES

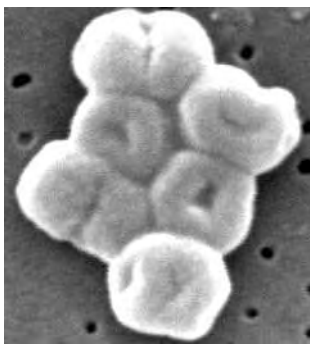
If you haven't heard of *Acinetobacter*, chances are you soon will. Reports that the bacterium is infecting US army personnel, mostly soldiers wounded in Iraq, are drawing attention to the little-known organism. Medical experts are concerned that if antibiotic overuse in hospitals is not curbed, drug-resistant strains of *Acinetobacter baumannii* could become a serious killer in intensive-care wards worldwide.

The latest data come from two US army medical centres, the Landstuhl Regional Medical Center in Germany and the Walter Reed Army Medical Center in Washington DC. Drug-resistant *A. baumannii* has infected at least 240 US soldiers since 2003, and five seriously ill patients who shared wards with them became infected and died.

"It's a big problem," says Colonel Bruno Petrucelli, the army's director of epidemiology and disease surveillance, pointing out that both civilian and military hospitals are being hit. "Hospitals are dangerous places to be."

Hard case: *Acinetobacter* (below) rapidly evolves drug resistance, and survives hygiene measures.

J. CARR/CDC



Resistance fighting

Detailed epidemiological data are sparse because surveillance of *Acinetobacter* is voluntary, but the number of outbreaks seems to be rising alarmingly. The UK Health Protection Agency (HPA) says that, in the past four years, outbreaks have been reported in 24 London hospitals alone. And Arjun Srinivasan, an epidemiologist from the Centers for Disease Control and Prevention in Atlanta, Georgia, who coordinates a programme on hospital-acquired infections in 300 US medical centres, says that the number of cases is rising there too, although they are still relatively uncommon.

"Twenty years ago, only around 1.5% of all cases of pneumonia acquired in the US hospitals we survey were caused by *Acinetobacter* infection," he says. "Now that number is 7% and most are drug-resistant." Similar increases have been seen for other types of infection, such as in blood or at surgical sites.

"40% of our patients who become infected with *Acinetobacter* die because of it."

Unlike methicillin-resistant *Staphylococcus aureus* (MRSA), another serious hospital-acquired, drug-resistant infection, *A. baumannii* is unlikely to infect many outside hospitals. "This bug is not the same as MRSA," says Tyrone Pitt, a bacteriologist at the HPA. "It is not a normal human pathogen." The bacterium normally lives in soil or water and is not very pathogenic in healthy people. But it can kill people who are already sick or whose immune systems are compromised. And it has several characteristics that could make it a serious threat in intensive-care wards. "The situation needs to be monitored," says Pitt.

Most worrying is the bacterium's ability to evolve resistance to a broad range of antibiotics unusually quickly, leaving few treatments for those who succumb. Scientists do not understand how this happens, but they know that the microbe has many ways to acquire resistance genes. And some strains naturally contain enzymes that break down some antibiotics.

Researchers are also studying how the bacterium can survive on dry surfaces, particularly plastics, for many weeks. This enables bacteria to hang around on wards even when strict hygiene is maintained.

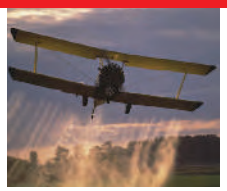
Those handling the outbreaks stress the importance of using antibiotics sparingly, to reduce the selective pressure on the bacterium to develop resistance. *Acinetobacter* acquires resistance to even more drugs than MRSA; doctors have been forced to turn to colimycin, an antibiotic abandoned in the 1960s because of its relative toxicity. On occasion, even this remedy has failed.

Petrucelli says that military field hospitals are particularly guilty of overusing antibiotics. Sterility is difficult to maintain, so doctors tend to throw broad-spectrum antibiotics at the wounded. "But this removes options later on," he warns.

Yehuda Carmeli, an infectious-disease physician at the Tel Aviv Sourasky Medical Center in Israel, describes the situation as "a real danger". The bacterium is endemic in many wards in his hospital, surviving hygiene measures that have seen off MRSA.

Carmeli also speculates that the bacterium has become more virulent as it has acquired antibiotic resistance. "We have calculated that 40% of our patients who become infected with *A. baumannii* die because of it," he says. ■

Alison Abbott



PESTICIDE-PROOF FLIES HAVE UNEXPECTED EDGE
Gene for DDT tolerance offers extra survival advantages.
www.nature.com/news

Preparations get under way for tsunami warning system

Tsunami experts from around the world met last week to hammer out a framework that will bring countries at risk of another Indian Ocean tsunami up to minimum safety standards. The meeting set the stage for upgrading tidal gauges throughout the region and for establishing an early-warning network by June 2006.

When a powerful tsunami devastated coastlines around the Indian Ocean on 26 December 2004, killing almost 300,000 people, there was no effective warning, even though the waves took several hours to reach some affected countries. Calls for a dedicated warning system in the region began almost as soon as the extent of the disaster became apparent.

As a result, the Intergovernmental Coordination Group of the Tsunami Warning and Mitigation System was set up by the United Nations in June. Its first meeting was held in Perth, Australia, last week. Representatives from the group's 27 member nations, and observer countries including Germany, Japan and the United States, selected a chair, P. S. Goel, who is secretary of the Indian government's Department of Ocean Development. Working groups were also set up to address the creation of networks for seismic and ocean monitoring, and to assess the respective needs of the countries involved.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

A. WEERAWONG/AP

An official checks for hints of a tsunami at the National Disaster Warning Centre in Thailand.

Since the disaster, 25 of the group's members have set up communication centres linked up to seismic-monitoring stations in Tokyo and Hawaii. Tidal gauges are also being deployed to provide further data on any developing tsunami. Six gauges sending real-time data have been established since December 2004, and the working group will decide when and where to place another 17 by the end of the year. A network of such gauges would allow information from one area to reach at-risk countries on the other side of the ocean.

The meeting also created a consortium to deploy deep-sea pressure sensors, which have been developed in the United States and Germany. There has previously been rivalry between the two countries over the technology. But when Germany presented plans to put two of its as yet untested buoys in Indonesian waters in October, the United States asked to place one of its buoys nearby, so that the systems can be calibrated. Member countries such as Australia and India will also now be able to use and refine the technology.

The possibility of setting up a regional tsunami-warning centre was not discussed. And international agencies have still to decide how to divide up resources to help countries, especially the smaller ones, prepare for future tsunamis. Such issues should be ironed out at the group's next meeting in Hyderabad, India, in December. ■

David Cyranoski

Scans provide picture of brain activity

Neuroscientists who use a popular scanning technique to look at brain activity can breathe a sigh of relief, researchers say.

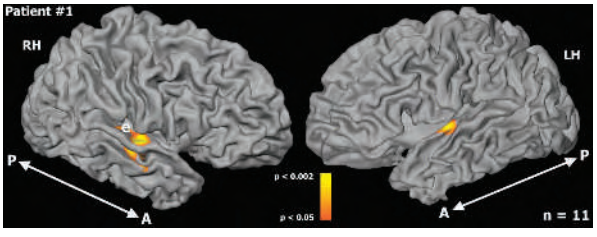
A paper in last week's *Science* suggests that, as well as identifying which parts of the brain are activated during certain tasks, the scans also provide a measure of actual brain-cell activity (R. Mukamel *et al. Science* 309, 951–954; 2005). Although researchers have been working on this assumption, it is the first study to show this in humans.

Functional magnetic resonance imaging, or fMRI, has been widely used to investigate which parts of the brain are active during anything from being in love to telling lies. But some researchers have questioned whether the technique reveals anything significant about the brain's workings. This is because fMRI does not measure brain-cell activity directly, but changes in blood flow within the brain in response to various stimuli.

The latest results could put those doubts to rest says the study's lead author, Rafael Malach of the Weizmann Institute of Science in Rehovot, Israel. "It's a great relief to know that these studies in which dozens of labs are now involved are not a waste of time," he says.

In 2001, researchers in Germany showed that fMRI scans could be correlated with brain-cell activity in anaesthetized monkeys (N. K. Logothetis *et al. Nature* 412, 150–157; 2001). But until now, no one had done the same thing for conscious human brains.

Malach joined forces with a group at the University of California, Los Angeles, led by neurosurgeon Itzhak Fried, that works with patients suffering from severe epilepsy. Before they can have surgery, the patients spend several days with electrodes wired into their brains so that the doctors can pinpoint the source of their seizures. Fried's team took advantage of this situation by recording the activity of the brain cells in two epilepsy patients who were shown nine minutes of the Clint Eastwood film *The Good, the Bad and the Ugly*.



AAAS

Models using brain-cell data show good agreement with real scans.

The researchers built a computer model to convert the measured activity into fMRI images. They then compared these predicted images with actual fMRI scans taken of 11 healthy volunteers who watched the same movie clip. In general, the model was accurate half to three-quarters of the time, but for certain points in the movie, agreement reached 90%. The correlation is encouraging, the team says, given the dramatic differences in the patients and their settings.

The results apply to only one area of the brain, and describe the reaction to just one kind of stimulus. But some neuroscientists are pleased that the study is looking at real human behaviour.

"They're actually getting a little bit closer to what the brain is doing, and I think that's great," says neurophysiologist David Leopold of the US National Institute of Mental Health in Bethesda, Maryland. ■

Erika Check

Four years on, no transgenes found in Mexican maize

Four years ago, the discovery of transgenes from genetically modified (GM) crops in traditional maize varieties in Oaxaca, Mexico, triggered an almighty row. A new survey suggests that measures taken since then to purge the crops of transgenes have been effective.

In the original paper, David Quist and Ignacio Chapela of the University of California, Berkeley, used the polymerase chain reaction (PCR) to detect two genetic sequences from GM maize in harvests from 2000 (D. Quist and I. H. Chapela *Nature* **414,** 541–543; 2001). Using a variant of the technique called inverse PCR, they also argued that the transgenes had integrated throughout the genomes of Mexico's maize varieties.

This was a shocking result, as it suggested that the 'contaminated' plants were not sporadic hybrids. Instead, it seemed that the transgenes were entrenched in the traditional varieties at the centre of natural genetic diversity for maize. "It was as if someone had gone to the United Kingdom and started replacing the stained-glass windows in the cathedrals with plastic," says Jorge Soberón, a former Mexican government scientist now at the University of Kansas in Lawrence and a co-author of the new study.

The inverse-PCR methodology used by Quist and Chapela soon came under fire, however, and *Nature* stated that it would not have published the paper if the criticisms had cropped up while the paper was under review. But even so, few experts questioned the basic finding that some transgenes had flowed into Mexican maize.

"The authors speculate that the fields were cleared of transgenes through education of local farmers and a reduction in GM imports."

L. ADDARIO/CORBIS

Maize unmodified: Mexican farmer Lorenzo Rebollo holds some of a recently harvested crop.

Despite an official moratorium on GM planting, this could have resulted from local farmers planting GM maize intended for food use that was imported from the United States. Unpublished work by Mexican government scientists also found transgenes, but a thorough and systematic confirmation was lacking.

That survey has now been done — and to the surprise of the authors, they found no transgenes at all. The sample of more than 150,000 seeds from 2003 and 2004 was negative for the same two transgenic sequences (S. Ortiz-García *et al. Proc. Natl Acad. Sci. USA* doi10.1073/pnas.0503356102; 2005). "I was convinced we were going to verify Quist and Chapela's results," says co-author Exequiel Ezcurra, head of the Biodiversity Research Center of the Californias at the San Diego Natural History Museum and former president of the National Institute of Ecology in Mexico City.

The authors speculate that transgenes were present in the fields in 2000, but dropped out of local maize varieties thanks to a programme of education for farmers and a reduction in GM maize imports. The researchers, led by Allison Snow of Ohio State University in Columbus, did not directly replicate Quist and Chapela's inverse-PCR methods. But Snow says that the apparent failure of the transgenes to persist down the generations contradicts the idea that they were entrenched in the genomes of the traditional maize varieties.

Brian Johnson, who follows developments in agricultural biotechnology for the government conservation agency English Nature, is unsurprised by this finding: "If there are transgenes in Mexico, or anywhere else, I would expect that they would be difficult to find — they would be rather sporadic." Johnson says he never believed they were permanently incorporated into the genes of traditional maize varieties.

Chapela stands by his findings, saying it is "naive" to believe an education programme could have such a dramatic effect. He claims that the commercial labs used by the research team to do the screening used conservative thresholds for declaring a match with the transgene sequences. But Bernd Schoel, director of research at one of those labs — Genetic ID of Fairfield, Iowa — says the screen was as sensitive as possible, given the sample size. ■

Emma Marris

Scientists attack Bush over intelligent design

C. DHARAPAK/AP

US scientists are again on the offensive against intelligent design, an idea that many see as thinly veiled creationism. The latest round stems from remarks by President George W. Bush, who on 1 August told a small group of reporters that he thought both evolution and intelligent design "ought to be properly taught" in US schools.

Scientists and science educators cried foul almost immediately, saying that such remarks could further the notion that intelligent design is a valid scientific alternative to evolution. Many researchers felt compelled to respond, even though Bush made his comments off the cuff in an informal setting. Because of the president's status, they say, his words could be used to introduce religious ideas into science classes.

"What the president has done is give impetus to people who would like to push their side of this agenda, and that's a real problem," says Fred Spilhaus, executive director of the American Geophysical Union (AGU), one of the first groups to respond to Bush's remarks. The AGU statement carried the headline "President confuses science and belief, puts schoolchildren at risk". The American Institute of Biological Sciences, the American Physical Society and the American Astronomical Society also released statements saying that intelligent design has no place in the science classroom.

Intelligent design — the notion that certain features of living organisms are so complex that they must have been shaped by an external intelligence — has enjoyed increasing prominence among the US public, although not among scientists (see *Nature* 434, 1062–1065; 2005). This is despite Bush's science adviser John Marburger stating on the record that intelligent design is not a scientific theory.

Lawrence Krauss, a theoretical physicist at Case Western Reserve University in Ohio and a frequent speaker on evolution issues, says that all scientists should be concerned. "Make no mistake — this is not an attack on evolution, but on science," he says.

Researchers are looking for new ways to fight the public-relations battle between science and intelligent design. Douglas Futuyma, an evolutionary biologist at the State University of New York at Stony Brook, says he would like to see a public-relations push mounted by an independent scientific group, such as the National Academy of Sciences.

Many experts say that scientists should get more involved in local politics — especially on school boards, where the conflicting views of scientists and advocates of intelligent design often play out. "Scientists have to be evangelical about explaining what science is, as well as its limitations," says Krauss.

> "The president has given impetus to people who would like to push their side of the agenda."

Kenneth Miller, a biologist at Brown University in Providence, Rhode Island, adds that scientists should highlight that there is no dissent over evolution within the scientific community and that if intelligent design had scientific merit, it would have been addressed by the vigorous and open scientific process. ■

Virginia Gewin

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

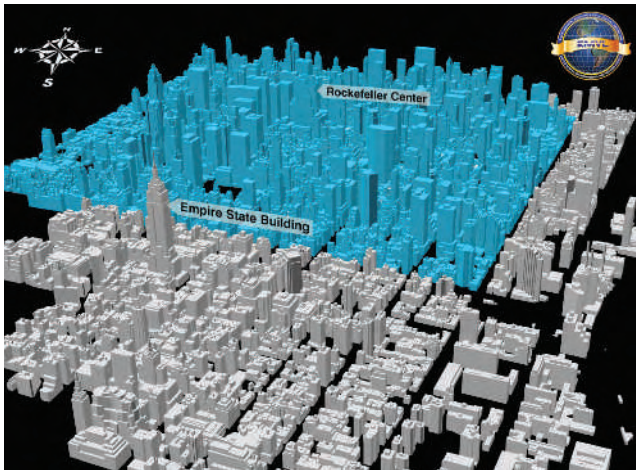
Off the cuff: President Bush said both evolution and intelligent design should be taught in schools.

SUNLIGHT USED TO SMELT ZINC
Solar technique could lead to cleaner, cheaper hydrogen.
www.nature.com/news

Anti-terror study tracks gas dispersal

Plumes of inert gas are rising above Manhattan this week as part of a research programme to improve the US response to terrorist attacks. Researchers will track the fate of seven different gases to test and refine models of how chemical, biological or radioactive material might spread through the city.

Emergency services responding to attacks in built-up areas already have to balance helping the injured, evacuating survivors and sealing the site for forensic examination. But responding to an explosive release of harmful airborne material would be even more difficult. To decide how far to move people from the scene and in which direction, authorities need some idea of how material will spread, and how that is affected by factors such as the weather.



EPA

Researchers will monitor the movement of test gases across New York.

Data from the New York experiments, principally funded by the Department of Homeland Security (DHS), will be fed into computer models that can be run in an emergency, to predict how a plume of material will disperse over the following hours and to inform emergency services about how to respond.

It is impossible to test every material, in every location and in all conditions. But Tom Keiss, acting programme manager for radiological countermeasures in the DHS Office of Research and Development, says the results from such experiments could lead to models accurate enough for use in an emergency.

"These models could tell us what sort of area first-responders should evacuate, and even the information you would give the public," he says, such as safe routes out of a city. "We consider this to be a sound investment that can save lives."

The gases are being released from an area of two square kilometres south of Central Park on six days (see graphic), determined by the weather, between 6 and 26 August. Using nearly 180 samplers and 35 weather stations, researchers from government and university labs will track the gases' movement through the streets and into buildings and the subway system.

The research is part of the four-year Urban Dispersion Programme, which began in 2003. Earlier tracer experiments took place in Oklahoma City and Salt Lake City, and a small-scale preliminary to the Manhattan experiments was carried out around Madison Square Garden in March.

Across the Atlantic, the British government is interested in tracer releases in London. It tested its models with data from a tracer experiment to examine air pollution in the city, and is deciding whether to fund a large-scale tracer-release programme in London.

"Experiments in New York are not much use to London or Paris. Our cities are very different physically, and have different climates," says Alan Robins, a specialist in environmental fluid mechanics at the University of Surrey, Guildford, and chair of the London project. ■

Tom Simonite

Japan accelerates global supercomputer war

TOKYO

Japanese groups are competing to reclaim for their country the honour of owning the world's fastest supercomputer.

Japan lost the title last autumn, when IBM's BlueGene/L computer — which has a maximum speed of 140 million million calculations per second (140 teraflops) — overtook the Yokohama-based Earth Simulator (see *Nature* 431, 618; 2004).

Last month, a group that includes Toshiba, the University of Tokyo and the Institute of Physical and Chemical Research announced plans for a computer that would be 73 times faster — capable of 10,000 million million calculations per second (10 petaflops). It said that the project could cost between ¥80 billion (US$700 million)and ¥100 billion, and would be ready by March 2011.

But experts are already expressing doubts. Tetsuya Sato, director of the Earth Simulator, says the project will be held back by the limits on semiconductor technology, which is reaching the narrowest dimensions of circuits that can be etched. "I am very sceptical," he says.

Erich Strohmaier, a computer scientist at Lawrence Berkeley National Laboratory, says difficulties with increased heat output and power consumption would also hamper the development of faster processors. But he predicts that technological advances could produce a machine of at least 3 petaflops by 2011, and doesn't rule out the Japanese team reaching their goal. "It's an ambitious project," he says. "With an exceptional system design, 10 petaflops seems within reach."

Meanwhile, Sato has his own plan — hooking the 40-teraflop Earth Simulator to a new 400-teraflop computer. The smaller machine would act as a 'macro-structure', performing large-scale calculations and directing the more intricate operations of the larger one, he told *Nature*.

Sato thinks that changing the networks' wiring in this way would allow the equivalent of 16 petaflops. "We are close to the technological limit," he says. "But this solution is based on physics, not technology." Strohmaier warns, however, that such hierarchical structures are a risk because they sacrifice a great asset of the Earth Simulator — "a tried and tested technology, which we understand and know how to program," he explains.

This computing power, if achieved, could be used to model climate change, drug metabolism and galaxy formation, for example. But it would be very expensive. It is not clear which of the projects, if either, will make it into the Japanese science ministry's budget requests, due this month. ■

David Cyranoski

NATURE|Vol 436|11 August 2005

ON THE RECORD

"This regime is intent on getting a bomb."

Paul Leventhal, founder of the Nuclear Control Institute, echoes the view of most analysts about Iran restarting chemical processing of uranium.

"You guys now have anthrax spores once again, so do be careful."

A dispute over property tax saw an NIH employee leave this voicemail for her Florida tax office, which promptly got her arrested.

SCORECARD

▲ **Science at the movies**
The silver screen is the latest weapon in the Pentagon's bid to bolster national defence. It is training scientists to write screenplays in the hope that films featuring glamorous researchers will draw more US students into science.

▼ **Fight against drugs**
Official figures for illegal drug use may be way off the truth. In Italy's Po valley, 15,000 users admit to taking cocaine at least once a month. But analysing river and sewage water for a byproduct of the drug suggests that the real number is closer to 40,000 doses — a day.

▼ **Ugly fish**
Forget size and taste, the latest goal in genetically engineered food is prettier fish. A team at the US Department of Agriculture is trying to create trout with lighter skin and smaller noses in a bid to make them more appealing to consumers.

OVERHYPED

Planet spotting
When astronomers at the California Institute of Technology reported a Solar-System object larger than Pluto two weeks ago, the news was followed by rumours that the announcement was rushed out when a hacker threatened to scoop the discovery. On his website, lead researcher Mike Brown reveals that someone used publicly available abstracts and telescope logs to piece together where the new 'planet' was located, although there was no threat to steal the credit. Brown says that checking the logs was "chilling" and "unethical". Maybe. Or you could just call it curiosity.

SIDELINES

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

G. HALL/CORBIS

Europe set for tough debate on curbing aircraft emissions

LONDON

Flight UA923 is a problem. When the Boeing 767 travels from London to Washington DC every week, it emits around a tonne of carbon dioxide for every passenger. The aircraft departs from a country that is intent on tackling such emissions, but lands in one that stands almost alone in resisting such measures. To further complicate matters, most of the emissions do not actually occur in the airspace of either nation. So how can UA923's emissions, and those from other flights, ever be regulated?

A first stab at an answer, at least in Europe, is likely to come soon. The European Commission (EC) is due to release a proposal on the issue in September, and details are starting to emerge. Emissions trading, already used to limit emissions from other European industries, will play a central role. New taxes are also likely. But although environmental groups and the airlines can agree on these points, a battle looms over a critical issue: the quantity of greenhouse gases that the industry should be allowed to emit.

International aviation is a pressing environmental concern. The industry emits around 3% of global greenhouse gases and is the fastest-growing source of emissions. Yet it is omitted from the Kyoto Protocol, which regulates emissions from most industrial nations. Total emissions from the European Union (EU), for example, dropped by around 5% between 1990 and 2003 — but contributions from the booming aviation industry rose by 75%.

The EC plans would cover all flights taking off in Europe. They would either see emissions included in Europe's existing carbon trading scheme, or in a stand-alone version for airlines. Under the existing scheme, which began in January, around 13,000 European firms monitor greenhouse-gas emissions. If companies produce more than their government-allotted quota, they must buy emissions credits from others that have emitted less than allowed.

PUNCHSTOCK

HURRICANES WHIP UP HUGE WAVES
40-metre monsters may account for mysteriously vanished ships.
www.nature.com/news

Some aspects of the plan to incorporate aviation are straightforward. As air travel is international, the European Union, not individual member states, will probably set emission targets. And the targets are likely to be based on the fuel burnt during each flight, not the distance flown, so that airlines will be rewarded for using more efficient engines.

Quantifying the impact of airline emissions may prove more difficult. As well as emitting carbon dioxide, aircraft exhaust gases promote the formation of ozone, another greenhouse gas. Aircraft contrails also create cirrus clouds, which have a warming effect. Together, these are believed to have a global-warming impact around two to four times greater than that of the carbon dioxide alone.

The EC's environment directorate, which is drafting the September document, told *Nature* that it nonetheless favours including carbon dioxide emissions only, as this is the basis for the existing trading scheme. Other emissions, such as the nitrogen oxides that promote ozone formation, would be tackled by extending existing taxes such as landing charges.

The final and thorniest part of the problem is the process of setting the industry an emissions target. An EC-commissioned report, published last month and led by a team from CE Delft, a Dutch environmental policy institute, explored the effect of setting allocations at 2008 levels. This would add only a few euros to the cost of an air ticket in the following years and is in line with industry thinking. "It's in the right ball park," says Robert Preston, executive officer of the British Air Transport Association.

"An analysis published this June predicts that current aviation growth will more than wipe out any emissions reductions from other industries."

But environmental groups take a different view. An analysis by the Tyndall Centre for Climate Change Research in Norwich, UK, commissioned by Friends of the Earth (FoE) and published this June, predicts that current aviation growth will more than wipe out any emissions reductions from other industries in coming decades. Even sticking to 2008 levels would leave airlines with much more than their current share of emissions, and possibly skew any carbon credit scheme, says Richard Dyer of FoE. The group wants airlines to cut emissions to 30% below 1990 levels by 2020, a massive challenge for an industry that is growing by 4% annually.

Better engines and improved routeing are only likely to cut emissions by 1–2% per year and it is unlikely that other industries could supply airlines with enough credits to make up the difference. So FoE argues that the cuts can only be achieved by introducing extra taxes, such as passenger duty, which reduce demand by forcing up ticket prices.

With such issues to be resolved, it is no surprise that the date for implementing trading is already slipping. The UK government, which is leading the debate as part of its current EU presidency, wants trading to begin in "2008 or as soon as possible afterwards". The EC environment directorate had also backed a 2008 target, but says now that 2012 may be a more realistic goal. ■

Jim Giles

DNA history collection kept whole by Venter

Genomics pioneer Craig Venter has bought an archive of historic molecular-biology documents, which he plans to put on public display at one of his research facilities in Maryland.

Venter, founder of The Institute for Genomic Research in Rockville, Maryland, bought the Jeremy Norman Molecular Biology Archive earlier this month for an undisclosed sum. Norman, a rare-book dealer in California, had assembled a collection of papers relating to the history of DNA (see *Nature* 411, 732–733; 2001). The archive includes a marked-up proof of the 1953 *Nature* paper by Francis Crick and James Watson reporting the double-helix structure of DNA.

In 2003, the auctioneers Christie's valued the collection at between US$2.2 million and $3.3 million, then cancelled a planned sale once it became known that the papers would be split up.

"We look forward to sharing this tremendous compilation of molecular-biology history with others," Venter said in a statement. "We hope to complement the collection with additional key works."

Fault-drillers dig down to earthquake active zone

US earthquake researchers drilling into the volatile San Andreas fault in California have reached the seismically active zone, opening a fresh vista on what they call "the earthquake machine".

After drilling intermittently for more than a year, the San Andreas Fault Observatory at Depth (SAFOD) team reached the seismically active zone, at a depth of nearly 4 kilometres, on 2 August. The borehole is in Parkfield, about halfway between Los Angeles and San Francisco on the 1,300-kilometre-long fault.

After logging the geochemical composition in the hole, the SAFOD team will cement steel casings in place and then insert seismometers. In 2007, after analysing seismic data from within the fault, the researchers will side-drill into newly identified active zones to place additional instruments.

M. ZOBACK/STANFORD

Cutting through: the project to drill into the San Andreas fault has hit pay dirt.

Avian flu vaccine 'feasible' in humans, trials show

The first human tests of a vaccine against the H5N1 avian flu virus have met with mixed results. In trials run by the US National Institute of Allergy and Infectious Diseases in Bethesda, Maryland, the vaccine generated a strong immune response. But the amount of modified virus required to elicit that response is so high that it effectively rules out large-scale production of the vaccine.

Annual flu vaccines require 15 micrograms of modified virus per dose; the H5N1 vaccine needs 90 micrograms. Even if the entire US flu-vaccine production capacity were switched to making an H5N1 vaccine, it could make enough vaccine for only 15 million people.

The agency is working to cut the dose needed, for example by adding adjuvants — chemicals that can boost the response to a vaccine.

First data trickle in from Mars Express water diviner

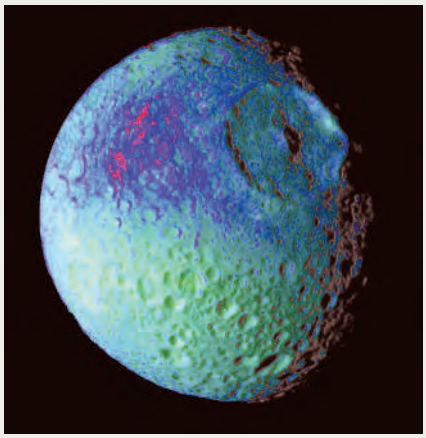
The MARSIS radar instrument on Mars Express, the European Space Agency's orbiting craft, has collected its first data on the surface and atmosphere of the red planet.

The findings are eagerly anticipated because MARSIS is the first orbiting experiment capable of detecting buried reservoirs of water ice up to 5 kilometres deep. Project scientists say that it will take a few more weeks of analysis to weed out the subsurface signals and obtain useful results.

MARSIS was supposed to have begun collecting data in April 2004, a few months after Mars Express arrived in orbit around the planet. But fears that the radar's three long booms could damage the craft as they unfolded delayed their deployment until June.

The radar will continue its first round of observations until mid-August.

NASA/JPL/SPACE SCIENCE INST



Cassini takes a closer look at Saturn's cratered moon

The Cassini spacecraft took this image of Saturn's heavily cratered moon Mimas on 2 August from a distance of just 228,000 kilometres. The 140-kilometre-wide Herschel crater appears in the upper right. Blue and violet colours represent areas where the surface of Mimas is weaker in infrared brightness than the rest of the surface, shown in green. The swaths of blue to the left of Herschel may represent debris kicked up from the interior of Mimas by the impact that created the crater.

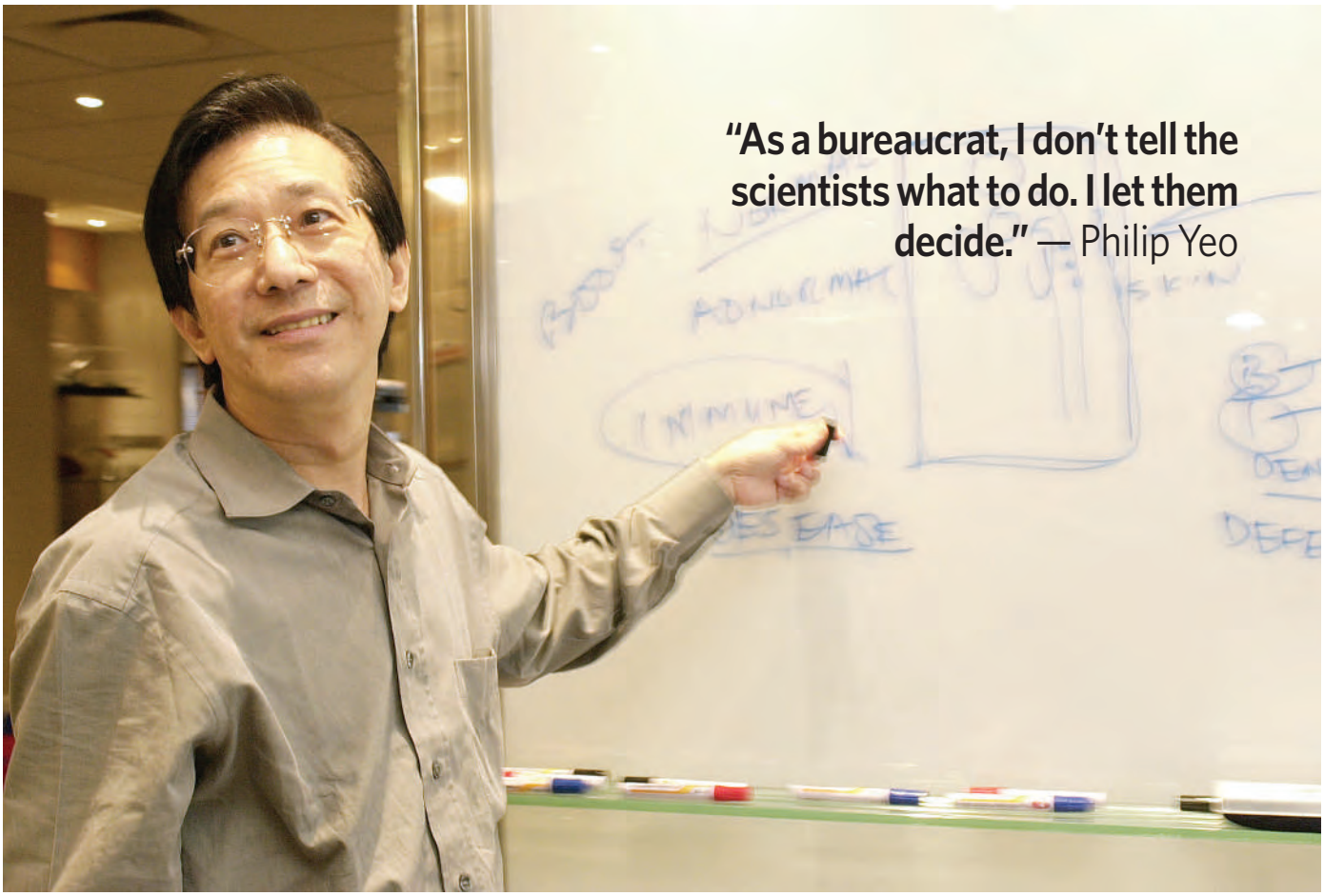
DESY tunes up electron laser for experiments

The free-electron laser at Germany's high-energy particle-physics laboratory, DESY, in Hamburg, opened for business on 3 August. The €117-million (US$145-million) facility produces intense pulses of tunable, short-wave ultraviolet radiation that last for 10 to 50 femtoseconds (one femtosecond is 10^{-15} seconds).

The pulses are generated when accelerated electrons are forced to wiggle along a winding path, radiating flashes of light at every turn. At least 200 scientists from around the world plan to use the pulses to directly observe, among other things, the formation of chemical bonds and the processes involved in the storage of magnetic data.

Correction

A mistake at our UK printers meant that the News story 'Drugs could head off a flu pandemic — but only if we respond fast enough' (*Nature* **436,** 614–615; 2005) contained an error in European copies of *Nature*. This said that even if a flu outbreak was controlled, up to half of Thailand's population would be infected. In fact, there would be only 200 cases in this scenario. In addition, a comment about US spending on anthrax surveillance was attributed to Ben Schwartz when it was actually made by Mark Lipsitch of Harvard University.



A*STAR

An irresistible force

Singapore's impressive advances in biomedicine are driven by the energetic personality of Philip Yeo. **David Cyranoski** meets a man who just can't stand still.

Modesty is not one of Philip Yeo's faults. As we tour Biopolis, the futuristic complex that is the centrepiece of Yeo's efforts to turn Singapore into a powerhouse for biomedical research, he announces that he has never failed in an attempt to headhunt a leading scientist to join the campus. "I have patience," he says with a wry smile.

Yeo's friends and admirers will be amused by this comment. Ask them for a description of him and the adjectives flow readily: persistent, passionate, provocative, even playful. But patience is rarely identified as a trait of this government official — a man whose charismatic influence on his nation's science policy has no obvious parallel anywhere in the world.

Since 1999, Yeo has chaired Singapore's main research organization, now known as the Agency for Science, Technology and Research, or A*STAR. Over the past couple of years, the agency has spent some US$300 million building Biopolis, which houses research institutes specializing in genomics, bioinformatics, bioprocessing and bioengineering. It is part of a multibillion-dollar investment in biomedicine, designed to nurture industries in biotechnology and drugs. Yeo has led from the front, enforcing an aggressive schedule and attracting some of the biggest names in biology to head the complex's lavishly equipped labs. "He has so much energy. If you tapped it, you could light a small city," says cell biologist Axel Ullrich, who was recruited from the Max Planck Institute of Biochemistry in Martinsried, Germany.

Biomedical research is just the latest endeavour to be swept up by the Yeo whirlwind. An engineer by training, his career took off in the 1970s when, as an official in the Ministry of Defence, he helped to turn a struggling manufacturer of munitions for the Singaporean military into an international supplier of armaments. Then, as chairman of the Economic Development Board from 1986, he oversaw the expansion of the semiconductor and electronics industries that has underpinned Singapore's economic miracle. And in a country where the boundaries between government and business are somewhat blurry, he has also become a wealthy electronics entrepreneur.

Island builder

By the late 1990s, Yeo's attention had shifted to the chemical industry. In his office, he proudly shows me satellite images documenting the year-by-year construction of an artificial island named Jurong. Today, it houses a huge petrochemical complex and has attracted some US$23 billion in investments from multinational companies. "It takes about five years to build each industry. Then I move on," he says.

Given this track record, Yeo has the trust of Singapore's leaders, who have given him almost total control over the nation's science budget. In a society known for its conservatism, he's a one-off — an impish character who darts out his tongue to punctuate each

provocative statement, and who clearly relishes every minute of his working day.

Three walls in his office are covered by white writing boards. As he explains his plans for Biopolis, Yeo jumps up and scribbles notes. Again, he shows me satellite images, before we drive to the complex for a whistle-stop tour. In addition to the gleaming labs, he makes a point of showing me the basement car park — a colour-coded expanse that wouldn't look out of place at an international airport — and a new French restaurant. "It is a self-contained city," says Yeo. "The only thing I can't control is the weather."

Yet Yeo has found biomedicine a tougher nut to crack than the subjects of his previous grand schemes. "With biology you have to read and read before you can absorb and understand," he says. When Biopolis was on the drawing board, he took Saturday tutorials on immunology from local researchers. "Immunology is related to sexy fields such as stem cells and cancer, but it has been neglected," he says. Next to his office is a room full of primary literature on all of these fields. When he finds something interesting, he sends it to scientists at his institutes. "I do my homework," he says.

Magnetic personality

For some of the big names who have been attracted to Biopolis, Yeo's personal enthusiasm for their work was a powerful draw. "He always sends PDFs of papers he's reading. I'm just trying to keep up with him," says David Lane, who in January was recruited from the University of Dundee in Scotland to head the Institute of Molecular and Cell Biology at Biopolis.

Even more popular is Yeo's policy of giving his principal investigators large five-year grants with no strings attached. "As a bureaucrat, I don't tell the scientists what to do. I let them decide," he says. The lucky recipients say they can get things done in months that elsewhere would take years just to clear the red tape. "It's fun," says Lane. "I've never before been anywhere they say you can do whatever you want."

In addition to headhunting big-hitters such as Lane, renowned for his work on the *p53* tumour-suppressor gene, Yeo has made a concerted effort to bring young people into science. Recruitment posters present biology as a hip activity. One, in a scene reminiscent of *The Simpsons*, pictures a student at a blackboard repeatedly scrawling: "I will not genetically manipulate the class hamster". Another features an attractive young woman — the 'DNA girl' — with a double-helix tattoo encircling her arm. The caption reads: "Got a burning passion for science? We'll take it to the highest level."

Those fast-tracked into A*STAR university scholarships get generous grants, but are given only three years to finish courses that take most students four. And Yeo thinks the brightest could be pushed still harder. "They could do it in two years," he says.

JTC CORP.

Powerhouse: Singapore's Biopolis is attracting top names in biological research.

The promotion of youth is a recurring theme. In Yeo's office, there are few greying civil servants; instead, he is surrounded by a coterie of bright, sharply dressed young staff. During our interview, they rush around providing background material to reinforce the points their boss is making.

But Yeo's aggressive youth policy has gained critics, as well as admirers. In early May, one Singaporean student working in the United States, writing a weblog under the pseudonym 'AcidFlask', criticized A*STAR's scholarship scheme and the expectations it places on students. A*STAR immediately threatened legal action, and the student — not himself an A*STAR scholar — apologized for a posting that was "defamatory of A*STAR, its Chairman, Mr. Philip Yeo, and its executive officers". With characteristic wit, Yeo issued a press release inviting the blogger to tea.

Yeo seems to relish such confrontations. In a newspaper interview given shortly after the incident, he laid into male students who break their contractual obligations to work for A*STAR for six years after their scholarships, saying he was tired of "whining Singaporean boys". He hands me a copy of the article. "I had to scold them," he says, adding that most A*STAR scholars are now foreigners and Singaporean women.

"Biopolis is a self-contained city. The only thing I can't control is the weather." — Philip Yeo

And during our interview, he can't help taking another shot at *Nature*'s coverage of his controversial 2001 move to close the Institute of Molecular Agrobiology. At the time, Yeo suggested that the journal should have paid the institute's running costs. Even now Yeo scoffs at the idea that the institute might have been able to offer agriculture-related intellectual property to Singapore's economy. "You've got to be kidding!" he exclaims.

Breaking the mould

It's not what I've come to expect from high-level government officials, but Yeo's supporters say that his combative and autocratic traits are part of the package that makes him so effective. It's difficult to imagine a similar character rising to the top in the more consensual environment of most national research agencies, says Ullrich. "Elsewhere, there is no space for people like Yeo," he notes. "He would become frustrated."

If Yeo can deliver on his promise to turn Singapore into a biotech and pharmaceutical powerhouse, its leaders will forgive him the occasional spat. Drug manufacturers lured to Singapore by generous tax breaks are already bringing in annual revenues of about US$9.5 billion a year. Yeo predicts that by 2015 the commercial spin-offs from home-grown research will have boosted this figure to US$15 billion.

The next phase of Yeo's master plan is the Centre for Molecular Medicine, where 150 researchers will bring advances in regenerative medicine, oncogenomics, immunology and epithelial biology to the clinic. Currently housed in Biopolis, it will move next year to the nearby National University of Singapore, to take advantage of its teaching hospital.

The centre will include a facility for the scaled-up production of stem cells for clinical use. "Nobody's doing that," says Yeo. "Industrialization is Singapore's strength." And by 2010, an expanded Biopolis will be joined by an information-technology hub called Fusionopolis, plus a complex designed to incubate spin-off businesses.

Lane sees the early signs of commercial success in the form of collaborations among institutes at Biopolis and industry. And with drug giants such Novartis and Eli Lilly establishing research bases at Biopolis, the future looks promising. "It's an amazing experiment here," says Lane. "I was overwhelmed by the opportunity."

Indeed, the biggest worry for many of Singapore's scientific leaders is what will happen should Yeo eventually retire. A youthful-looking 58, Yeo shows no signs of slowing down. And he is typically provocative when asked how Singaporean science would fare if he were no longer around. "I'm not indispensable," he says, "but I am irreplaceable." ■

David Cyranoski is *Nature*'s Asian-Pacific correspondent.

THE SECRET LIFE OF SPERM

Far from being mere DNA delivery boys, it's now becoming clear that sperm also ship a complex cargo of RNA and proteins that may be crucial for an embryo's early development. **Claire Ainsworth** reports.

D. WILSON/CORBIS

"What's in sperm?" demands Tim Karr. It's an unusual start to a lab visit, and I confess that I know little more than the sketchy textbook picture. I'm not alone, says Karr, and that is the whole problem. "How can we not know?" he asks. "It probably explains why we don't understand sex."

Karr, who works at the University of Bath, UK, is convinced that a more detailed knowledge of the molecular biology of sperm will help answer some of biology's most fundamental questions. This is what prompted him to spend the past three years dissecting fruitfly testes and developing methods to study the protein content —or proteome — of their sperm.

Sperm are amenable to detailed proteomic analysis because they contain no more than a few hundred proteins. But this apparent simplicity is deceptive. Karr's team is one of several to have found that there is far more to sperm than we thought. In addition to the DNA instructions that spell out a male's contribution to a new life, these sleek, whip-powered cells have in the past few years been shown to carry other pieces of cellular machinery, such as RNA and proteins. This discovery is changing our understanding of fertility, development and the evolution of sex. "It really challenges some basic ideas," says Stephen Krawetz, a biologist at Wayne State University School of Medicine in Detroit, Michigan.

Because sperm have to swim far and fast, biologists have come to view them like racing cars: streamlined and stripped down of all unnecessary bits and pieces. Generally speaking, the DNA in animal sperm is tightly packed inside a sleek head structure that contains little of the cytoplasm that fills most other cells. Behind the head is the midpiece, containing more than 50 power units called mitochondria that drive the lashing motion of the attached tail.

Egged on

Given the shortage of cytoplasm, and the lack of any detectable protein synthesis in mature sperm heads, biologists had long assumed that sperm contribute little to an embryo bar the father's genes. In contrast, the egg is replete with molecules such as proteins and RNAs that nourish and direct the development of the embryo. "The idea was that the egg was supplying everything and Dad was just tagging along with his DNA," says Krawetz.

Recent discoveries have revealed the error of this view. Studies now suggest that defects in sperm can disrupt embryo development even if the genes carried by the cells are perfectly normal[1]. And there are hints that faulty sperm could be the cause of a significant number of miscarriages, says David Miller, a reproductive biologist at the University of Leeds, UK. "So we know sperm is important."

But what does a sperm deliver? One popular misconception is that only the head enters the egg, while the tail is discarded. But in most

species, the entire cell enters the egg — midpiece, tail and all[2]. And in many mammals, midpiece and tail structures persist in the embryo for several cell divisions[3]. This results in a large number of proteins and other molecules being delivered to the egg. In mammals other than rodents, these include a piece of cellular machinery called the centrosome, which coordinates the molecular ropes cells use to haul chromosomes around during cell division[4].

Until 2002, this was thought to be an isolated example. But then a team headed by Anthony Lai at the University of Wales in Cardiff discovered that sperm also deliver a molecule called PLCζ that triggers the waves of calcium ions that activate a fertilized egg[5]. And a bigger surprise came when Krawetz and Miller studied sperm from 10 fertile men and found that they contained some 3,000 different kinds of messenger RNA[6]. Some of them coded for proteins needed for early embryo development; others were previously unknown, and had no equivalents in the egg.

Male delivery

This suggested that sperm could deliver RNAs that help direct an embryo's early development. Some biologists were sceptical, arguing that the RNAs were simply non-functional leftovers from the process of sperm development. But Krawetz, Miller and others have since gathered more evidence, and last year showed that a specific package of RNAs are indeed transferred from sperm to egg[7]. Earlier this year, Krawetz and his colleagues found that these include micro-RNAs, which don't code for proteins but are known to play a role in controlling gene activity[8].

It remains unclear what the transferred RNA does. The fact that cloning works, and the creation last year of a mouse by combining the nuclei of two eggs[9], both suggest it is not absolutely essential for embryo development. But these processes are grossly inefficient and often result in birth defects or abnormal gene activity, which hints that the paternal RNA may be important.

Miller suggests that messenger RNAs help protect paternal genes that are needed soon after fertilization from being shut down as sperm mature. Normally, most of a sperm's DNA is tightly wrapped up and gagged by proteins called protamines. The RNAs could stick to the genes that code for them and stop this, he argues. Another possibility is that paternal RNAs, particularly micro-RNAs, might be involved in controlling imprinting — the differential activation of genes according to whether they are inherited from the mother or the father. Certainly, the idea that the RNAs have no function is becoming a minority view. "Sperm are so sleek and have such powerful methods for eliminating everything that causes drag that I don't believe they are vestigial," says Gerald Schatten, a reproductive biologist at the University of Pittsburgh, Pennsylvania.

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Y. NIKAS/WELLCOME PHOTO LIB.

As well as DNA, sperm deliver at least one molecule that helps to activate a fertilized egg.

The new view of sperm as carriers of molecules crucial for early embryo development has thought-provoking implications for reproductive medicine. Comparing the RNA profiles of fertile and infertile men might reveal causes of unexplained infertility, says Miller.

Such studies may also raise questions about the wisdom of an *in vitro* fertilization technique called intracytoplasmic sperm injection, or ICSI, used to help men whose sperm do not fertilize their partner's eggs. ICSI involves injecting faulty or immature sperm — which might lack the normal complement of RNAs — directly into eggs. So far, there are no clear signs of problems among children conceived by ICSI, although long-term follow-up is needed to confirm the safety of the technique.

Now the action in sperm biology is moving from RNA to proteins. Last year, Christopher Barratt, a reproductive biologist at the University of Birmingham, UK, published the first proteomic study of male infertility. His aim was not to produce a complete proteome for sperm; instead, his team looked for differences in the protein profiles in the sperm of an infertile and a fertile man. The researchers found at least 20 proteins present in significantly different quantities[10], giving them a starting point to study cases of unexplained infertility and suggesting targets for new contraceptives.

Origins of life

The coming months could see the first publications from groups, including Karr's, that are conducting more comprehensive proteomic studies. Being able to compare the structure and content of the proteomes of sperm from different species should help researchers understand the evolution and origin of sperm. In particular, having a comprehensive catalogue of proteins to compare between different species may reveal how natural selection is operating on them, says Steve Dorus, a postdoc in Karr's lab. "It should give us some pretty powerful information about what our ancestors' core sperm attributes were."

"The view of sperm as carriers of molecules crucial for early embryo development has thought-provoking implications for reproductive medicine."

While Karr's group works on fruitfly sperm, Victor Vacquier, a reproductive biologist at the Scripps Institution of Oceanography in La Jolla, California, is producing a catalogue of the proteins found in the outer membrane of sperm of the purple sea urchin (*Strongylocentrotus purpuratus*). His particular interest lies in understanding how sperm and egg interact and recognize each other. Surprisingly, scientists know comparatively little about the molecules that interact when a sperm comes into contact with an egg's surface. But they do know that, in some species, such proteins can evolve extraordinarily rapidly[11]. Vacquier believes that these fast-evolving proteins may help explain why some animal populations become reproductively isolated, leading to the formation of new species.

The origin of species, the evolution of sex, the mysteries of infertility — these are some of the biggest and most intriguing questions in biology. Long dismissed as mere delivery boys, it seems that sperm are about to be put on the promotion fast track. ■

Claire Ainsworth is a senior news & features editor for *Nature*.

1. Loppin, B., Lepetit, D., Dorus, S., Couble, P. & Karr, T. L. *Curr. Biol.* **15,** 87–93 (2005).
2. Ankel-Simons, F. & Cummins, J. M. *Proc. Natl Acad. Sci. USA* **93,** 13859–13863 (1996).
3. Sutovsky, P. & Schatten, G. *Int. Rev. Cytol.* **195,** 1–65 (2000).
4. Simerly, C. *et al. Nature Med.* **1,** 47–52 (1995).
5. Saunders, C. M. *et al. Development* **129,** 3533–3544 (2002).
6. Ostermeier, G. C., Dix, D. J., Miller, D., Khatri, P. & Krawetz, S. A. *Lancet* **360,** 772–777 (2002).
7. Ostermeier, G. C., Miller, D., Huntriss, J. D., Diamond, M. P. & Krawetz, S. A. *Nature* **429,** 154 (2004).
8. Ostermeier, G. C., Goodrich, R. J., Moldenhauer, J. S., Diamond, M. P. & Krawetz, S. A. *J. Andrology* **26,** 70–74 (2005).
9. Kono, T. *et al. Nature* **428,** 860–864 (2004).
10. Pixton, K. L. *et al. Hum. Reprod.* **19,** 1438–1447 (2004).
11. Metz, E. C., Robles-Sikisaka, R. & Vacquier, V. D. *Proc. Natl Acad. Sci. USA* **95,** 10676–10681 (1998).

NATURE|Vol 436|11 August 2005

Breaking the age barrier

Many scientists continue to run productive and innovative research programmes well beyond typical retirement age. But in many countries, tough retirement laws make staying in the lab a challenge. **Laura Bonetta** reports.

When Jack Strominger and his wife moved to an apartment just a 15-minute walk from the lab, Strominger thought he would be able to do his reading and writing from home. Yet he rarely makes it out of the office before six in the evening. The head of a 14-member group at Harvard University, Strominger has an ambitious research agenda by any standards, but even more so because he turned 80 this month.

The immunologist is best known for his work with Don Wiley on the proteins that the body uses to distinguish its own cells from foreign ones. But these days, he enjoys applying his years of experience to new problems, such as autoimmune diseases. "My goal has always been to stay original," he says. "When I can't do that any more, maybe I will spend more time riding horses at our farm."

Eighty-year-old Janet Rowley, a cancer geneticist at the University of Chicago, has also toyed with the idea of slowing down. "I was thinking of going part-time but wanted to see what would happen to my latest grant," she says. "We got funding for another three years, so now I can postpone that decision." Rowley made her mark by discovering that pieces of chromosomes in cancer patients can break off and join on to other chromosomes. She had something of a late start in research, having worked as a clinician for several years and then as a part-time researcher while raising four sons. But she has more than made up for lost time. "I have had a front-row seat in a never-ending suspense story," she says.

Such tales are not unique. Many researchers continue to lead active and innovative research programmes in their 70s and 80s, quashing the notion that science is a young person's game. "Age matters, but not that much. And it matters less in the life sciences than in physics," says Paula Stephan. An economist at Georgia State University, Atlanta, Stephan co-wrote a book in 1992 called *Striking the Mother Lode in Science: The Importance of Age, Place and Time*, in which she examined the relationship between age and productivity based on publication record.

There may even be some advantage to age. "It can maybe help you focus on larger problems," says Joseph Gall of the Carnegie Institution in Baltimore, Maryland, who at 77 still works at the bench all day. "If you feel you have made your mark, you can sit back and look at the bigger picture. In a way that is what happened with our work on the Cajal body." First described more than 100 years ago, this cell structure was brought out of obscurity in the past decade by Gall's work showing that it is the assembly site in cells for key proteins that modify RNA molecules to make them function.

Generation gap

But there are drawbacks too. "It is definitely harder to recruit students and postdocs," says Strominger. "But it is understandable. My grandchildren are almost the age of my students. It is harder to relate to your grandfather."

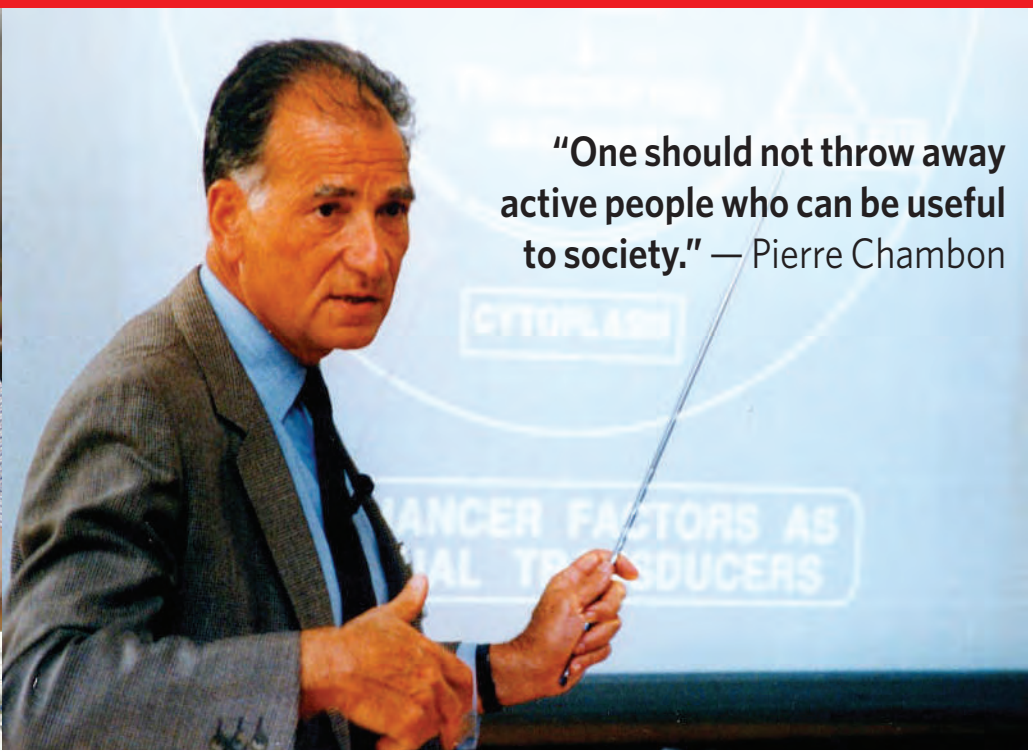
Writing grants can also begin to get tedious after so many years . At 86, Herman Eisen, an immunologist at the Massachusetts Institute of Technology (MIT) in Cambridge, decided to let his last federal grant run out. "It was unrealistic to keep writing grants," he says. "For one thing, I had stopped taking postdocs because I could not commit to them for several years. And also on some level it felt embarrassing; I would be competing with former postdocs and students."

But Eisen is not out of work. MIT let him keep his lab space, where he continues to plug away at the bench aided by a part-time assistant and a succession of undergraduate students. "It is still very fascinating. The problems I choose today are precisely those that appeal to me a lot," Eisen says.

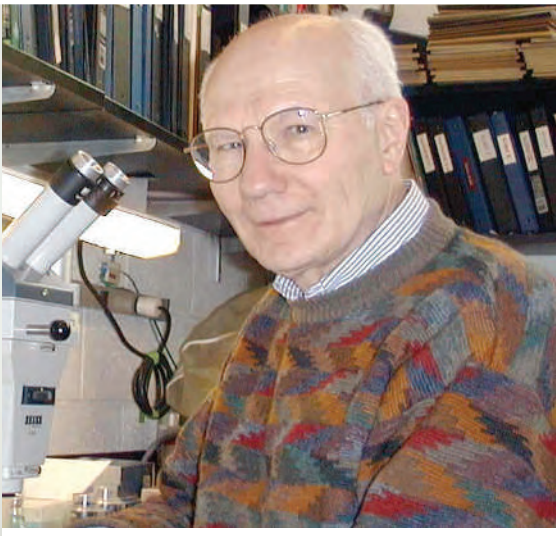
The lure of finding one more piece of the puzzle is what keeps thoughts of leaving research at bay for many ageing scientists. "I can't think of a day that I did not want to come to the lab. Maybe when my papers start getting rejected, I will start thinking about doing something else," says Gall.

J. CHASE/HARVARD UNIV. NEWS OFFICE (TOP); UNIV. CHICAGO

Scientists in the United States can continue working as long as they can get grants because the country did away with mandatory retirement in 1994. Australia, New Zealand and some provinces of Canada have followed suit. But in many European countries and in Japan, scientists working at government-funded universities have little choice but to retire sometime between the ages of 60 and 70, regardless of their level of productivity. As a result, the United States has benefited from the influx of several superstar foreign scientists trying to escape this fate.



B. BOULAY, IGBMC STRASBOURG



J. GALL

Willing workforce: (clockwise from top left) Jack Strominger, Pierre Chambon, Joseph Gall and Janet Rowley, who have all continued their research projects well beyond retirement age.

Swiss Nobel laureate Kurt Wüthrich's move to the Scripps Research Institute in La Jolla, California, in 2001 caused a scandal in the European press. Five years before reaching retirement age, which is 65 in Switzerland, Wüthrich tried to negotiate a five-year extension to his appointment with the Swiss Federal Institute of Technology in Zurich. "But on very friendly terms, my request was denied," says Wüthrich, who will be 67 in October. As a result, he accepted one of the many job offers he had obtained abroad.

The situation quickly changed in 2002 when Wüthrich received the Nobel Prize in Chemistry. The Swiss parliament passed an extraordinary law to let him keep his post in Switzerland, although he had already set up shop in the United States. Wüthrich, who now splits his time between the two positions, says that the new law has not yet helped other Swiss academics.

Perhaps one reason for the resistance to change, suggests Wüthrich, is that senior scientists in Europe acquire many benefits that are difficult for universities to maintain if a scientist is not productive. "My chair in Zurich carried a non-competitive endowment for the salaries and consumables for a team of eight scientists and technicians. If I stop being productive at 55 but last my time until 65, this huge investment is not well used," he says.

Cold shoulder

Japanese scientists face a similar plight. "The system in Japan is considered to be based on achievement, but the reality is not so," says Yoshiaki Ito, who at 63 traded in his retirement from Kyoto University for a position at the Institute of Molecular and Cell Biology in Singapore. According to Ito, Japanese colleagues who want to continue working past retirement age have to scramble to find positions at private institutes or companies, or head abroad. "I never thought of stopping work. My work is going so well right now," says Ito, who believes he is hot on the heels of a genetic pathway fundamental to cancer. "To set a mandatory retirement age at 63 is cruel. It is like cutting your head off in the middle of your career."

In Germany, the situation is a little better, as retired professors can apply for grants, provided their home institution lets them have some lab space to continue working. But that too is far from ideal, according to immunologist Klaus Rajewsky, who left the University of Cologne in 2001 to take up a post at Harvard University. "The last years working in Cologne were often psychologically depressing. My colleagues, who were same age as I, were wondering how to deal with retirement. Some would talk about trying to get a place in someone else's lab to keep working," recalls Rajewsky. "I came to the United States and had a contract without a time limit and all those discussions came to an end. It was a relief."

"To set a mandatory retirement age at 63 is cruel. It is like cutting your head off in the middle of your career." — Yoshiaki Ito

Forced out

Scientists in Europe and Japan are sceptical about whether their universities will do away with mandatory retirement anytime soon. At the University of Tokyo, the decision to extend the retirement age from 60 to 65 by 2013 drew fire because many felt it would slow opportunities for new faculty appointments (see *Nature* **407,** 550; 2000).

The end of mandatory retirement in the United States had raised similar concerns, but Ronald Ehrenberg, professor of industrial and labour relations at Cornell University in Ithaca, New York, says they are misplaced. "The consensus is that people who would have retired before age 70 in the absence of mandatory retirement are continuing to retire, and of the people who were constrained by mandatory retirement, some are staying longer," he says. "People who stay past age 70 tend to be those individuals most driven by work and most productive, especially in sciences where you have a system in which you have to generate research grants to do your work."

Pierre Chambon, founder of the Institute of Genetics and Molecular and Cellular Biology in Strasbourg, thinks that judging a person by his or her age makes little sense. "Some people get old faster than others," says the biologist, who had to retire from the institute's directorship in 2002. "One should not throw away active people who can be useful to society."

This attitude is not unique to scientists. In May 2005, the UK-based HSBC Bank published a survey of some 12,000 people in ten different countries to gauge attitudes to ageing and retirement. The report, *The Future of Retirement*, claims that 80% of people worldwide want to scrap mandatory retirement, a trend partly attributed to the increase in life expectancy.

Chambon managed to find a way to keep working, albeit on a voluntary basis, by founding an institute that is affiliated with the University of Strasbourg but that operates in a semi-private way. "It's like telling someone who has always climbed mountains: 'Now, you cannot climb mountains any more, not even walk in the mountains'," says 74-year-old Chambon. "I am not pretending I can climb the peaks as fast as I did 20 years ago, but I still want to be in the mountains." ■

Laura Bonetta is a freelance writer based in Bethesda, Maryland.

BUSINESS

Ancient foil maker wraps up mobile-phone market

Few high-tech companies celebrate links with the world as it was before about 1980 — never mind with traditions crafted 305 years ago. But a family firm that has been going strong since the end of the seventeenth century last week won a Japanese award for its work with metal films, most of which, these days, end up in mobile phones.

Fukuda Metal Foil & Powder started up in Kyoto in 1700 in a business not dissimilar to the one it runs today — putting intricate gold and silver patterns on luxurious folding screens and pottery. By keeping a close eye on technological progress and market changes, it has become a world leader in supplying rolled copper-foil components to the burgeoning mobile-phone industry.

"Fukuda has built up good expertise over its long history," explains Osamu Nagashima of the business school at Ritsumeikan University near Kyoto, who knows the company well. "But more importantly and unusually, it has developed the ability to explore new technology."

Privately owned Fukuda amassed sales of ¥36 billion (US$323 million) last year and now employs around 1,000 people, mostly in Kyoto and Suzhou, China. It produces a huge array of metal products — 1,000 types of metal powder and 200 kinds of foil. They go into everything from car parts to food packaging. But the global market for copper-foil parts used in mobile phones — of which Fukuda claims a 40% share — is perhaps its most striking success.

The company received its award on 4 August from the Japanese prime minister, Junichiro Koizumi, in recognition of its outstanding craftsmanship, and its success in adapting traditional skills to the information age.

"Products we thought would sell well haven't always been a success," says Fukuda's president Yasuhiko Hayashi. "A lot of our lines have grown unexpectedly with the advent of new demands. Our solid research base has enabled us to tailor our products to them."

Before he joined the company, Hayashi, a chemical engineer, had been interested in metal powders but spent most of his career working on copper foil, acquiring the ability to judge a foil's quality merely by shaking it in his hand. Now, he says, he is trying to realize his dream of inventing different types of powder.

Traditional crafts have won Yasuhiko Hayashi's company a niche supplying mobile-phone parts.

YOMIURI SHIMBUN OSAKA

"What matters is the size and shape" of the powder's particles, Hayashi explains. The powders made by Fukuda can have particles ranging in diameter from 1 to 250 micrometres. Small differences in size and shape can greatly affect a powder's properties, such as whether the particles will stick together when compressed and heated.

Fukuda's scientists and engineers have an annual research and development budget of US$7.3 million, and work with universities and research institutes throughout Japan. For example, together with Mitsuo Kawasaki, a surface photochemist at Kyoto University, Fukuda has recently produced nanoscale particles — less than 0.1 micrometres in diameter — in copper alloys. This involves a process in which the particles are suspended in an organic solvent and then broken up with a laser.

One cloud on the horizon, Fukuda's managers say, is the competition with big Japanese corporations for technical staff. Growing competition is also coming from China, South Korea and Taiwan. Hayashi sees even faster product development as the key to the future, as well as energetic pursuit of new markets, such as in fuel cells. "We were getting into a rut," Hayashi says. "We need to come up with new stuff." ■

Ichiko Fuyuno

IN BRIEF

PORK BARREL Two major food companies announced that they will start winding down the amounts of antibiotics fed to pigs and chickens used in their products. The North American arm of Compass Group, an international catering company with 400,000 employees, and Smithfield Foods, the world's largest pork producer, based in Richmond, Virginia, said they would stop using medically important antibiotics to promote growth in pigs, and set up mechanisms to report and steadily reduce all antibiotic use. The agreement was developed with Environmental Defense, the New York-based green group. Antibiotic use in farm animals is thought to contribute to the growing human resistance to important antibiotics.

SPECIAL RELATIONSHIP Qinetiq, a British research company that runs former Ministry of Defence laboratories, has made two major acquisitions in the United States. The company announced on 3 August that it is to buy two information-systems contractors based in Virginia — Apogen Technologies for $288 million, and Planning Systems for $42 million. Both companies sell computer and security systems to the Pentagon, the Department of Homeland Security and other US government departments. The expansion of its US business comes as Qinetiq gears up for an expected public offering in which the British government will relinquish its share of the company.

HYBRID OPTION Japanese car giant Toyota says that it expects sales of hybrid vehicles — which boost fuel economy by combining a battery and a conventional engine — to soar over the next five years. The company's North American president Jim Press says that sales of the vehicles could reach 600,000 a year in the United States alone by early in the next decade. The company, which plunged into hybrid technology more enthusiastically than its competitors (see *Nature* **435**, 1026–1027; 2005), dominates the hybrid market and is on track to sell at least 100,000 of its popular Prius models in the United States this year.

Harry Potter and the recessive allele

SIR — We are bombarded with news of genetic discoveries on an almost daily basis, but people without a formal knowledge of heredity and genetics can have difficulty in deciphering and applying this information. Education and debate across all ages would undoubtedly help, but how can we teach children these concepts?

We believe that successful lessons for younger children can be achieved using analogies of direct interest and relevance. Most children are familiar with J. K. Rowling's stories about the young wizard Harry Potter (whose latest exploit, *Harry Potter and the Half-Blood Prince*, was published by Bloomsbury in July). They are set in a world like our own, but populated by a minority of people with supernatural powers (wizards and witches) and a majority of people without (muggles).

"With the use of these examples, the concepts of mendelian genetics can be introduced to children as young as five."

— J. M Craig, R. Dow, M. A. Aitken

Wizards or witches can be of any race, and may be the offspring of a wizard and a witch, the offspring of two muggles ('muggle-born'), or of mixed ancestry ('half-blood').

This suggests that wizarding ability is inherited in a mendelian fashion, with the wizard allele (W) being recessive to the muggle allele (M). According to this hypothesis, all wizards and witches therefore have two copies of the wizard allele (WW). Harry's friends Ron Weasley and Neville Longbottom and his arch-enemy Draco Malfoy are 'pure-blood' wizards: WW with WW ancestors for generations back. Harry's friend Hermione is a powerful muggle-born witch (WW with WM parents). Their classmate Seamus is a half-blood wizard, the son of a witch and a muggle (WW with one WW and one WM parent). Harry (WW with WW parents) is not considered a pure-blood, as his mother was muggle-born.

There may even be examples of incomplete penetrance (Neville has poor wizarding skills) and possible mutations or questionable paternity: Filch, the caretaker, is a 'squib', someone born into a wizarding family but with no wizarding powers of their own.

We believe that, with the use of these examples, the concepts of mendelian genetics can be introduced to children as young as five, and then built on by gradually introducing specific terms such as 'gene' and 'allele', and relating these to chromosomes and DNA. At every stage, the children's familiarity with the Harry Potter characters can be used as a hook to engage them in discussing concepts of heredity and genetics.

Jeffrey M. Craig*‡, Renee Dow†, MaryAnne Aitken†‡
*Chromosome Research, Murdoch Childrens Research Institute, Royal Childrens Hospital, †Genetics Education, Murdoch Childrens Research Institute, Royal Childrens Hospital, ‡Department of Paediatrics, University of Melbourne, Royal Childrens Hospital, Flemington Road, Parkville, Victoria 3052, Australia

Tunnel vision all in the mind at Hothouse High

SIR — Your News Feature "Hothouse High" (*Nature* **435,** 874–875; 2005) quotes current students of Thomas Jefferson High School for Science and Technology as wishing they had more opportunities to take music and other non-science courses.

As an alumna from the class of 1993, and a musician, I am dismayed, because this was not the case at all while I attended Jefferson. If anything, there was a strong multidisciplinary approach to teaching, which I believe actually helped those of us who went on to pursue a science career. I disagree with Chris Colin's statement that "tunnel vision at Jefferson hindered thinking about other interests and career choices" and I suggest that the only tunnel vision at Jefferson in danger of affecting students might be their own.

I was saddened by the declaration that two-thirds of the seniors at Jefferson now say they wouldn't choose it again. Perhaps the wrong students are attending Jefferson for the wrong reasons? There is pressure, especially from parents worried about college admissions, for their children to attend Jefferson regardless of their interest in science. Not to mention that Jefferson is a public school, free to attend and free from the violence endemic to most public schools in the greater Washington DC area.

Samantha G. Zeitlin
Department of Medicine, Moores-UCSD Cancer Center, University of California, San Diego, La Jolla, California 92093, USA

Misconduct: acceptable practices differ by field

SIR — The US Office of Science and Technology Policy wisely states, in its definition of misconduct, that there must be a significant departure from the accepted practices of the relevant research community.

In their Commentary article on the integrity of science, "Scientists behaving badly" (*Nature* **435,** 737–738; 2005), Brian C. Martinson and colleagues use a survey that blends responses from investigators doing basic discovery science with those doing clinical research, as if the accepted and acceptable practices in these two areas did not differ markedly. However, that is not the case.

For instance, in basic science, there are no specific regulations other than keeping research records (misbehaviour no. 16) for a particular time period. In clinical research, on the other hand, good laboratory practice requires certain types of record-keeping, the omission of which is not just careless but potentially sanctionable, as it would place an investigator out of compliance with Federal regulations.

Also, in basic research, intuition (misbehaviour no. 15) is an important, and perhaps in the end a researcher's best, guide to distinguishing between data and noise. In clinical research, by contrast, intuition should never be used for deciding what data can or cannot be included in an experiment. Indeed, every subject counts, and excluding subjects without appropriate documentation would be a serious and sanctionable matter.

Finally, Martinson and his colleagues consider "changing the design, methodology or results of a study in response to pressure from a funding source" as a single misbehaviour (misbehaviour no. 10). They ignore the fact that changing research design and methodology — although never the results — is precisely what investigators are supposed to do, to satisfy the criticisms of scientific review groups and the funding agencies that they represent.

Frederick Grinnell
Ethics in Science and Medicine Program, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9039, USA

Cannabis on a downer

SIR — Why did you use a down arrow to score the recent US Supreme Court decision banning users from growing cannabis at home ("Scorecard" *Nature* **435,** 724; 2005)?

The issue of medical marijuana in the United States is controversial. Although there is still no clear definitive scientific evidence for either side of the argument, you seemed to pass judgement on this issue in a one-line statement.

Martin J. Neumann
Department of Nuclear, Plasma and Radiological Engineering, College of Medicine, University of Illinois, Urbana, Illinois 61801, USA

The arrows used in Sidelines are intended to record changes to the current prospects of the topic under discussion, not to pass an editorial judgement, although in this case their meaning was ambiguous — News Editor, *Nature*.

BOOKS & ARTS

The rovers' tale

How NASA scientists overcame the odds to find signs of water on Mars.

Roving Mars: Spirit, Opportunity, and the Exploration of the Red Planet
by Steve Squyres
Hyperion: 2005. 432 pp. $25.95

Gregory Benford

Roving Mars is a deftly and dramatically written history of the Mars rovers, Spirit and Opportunity. It is also a primer on how to do exotic geology at a distance of 100 million miles using robots. Steve Squyres knows how to render scenes and intricate technical detail to build tension, without losing sight of the thrill and grind of the groundbreaking work.

"Eleven years had passed since I had started trying to send hardware to Mars, and in all that time I hadn't seen a single plan for Mars exploration survive for more than about eighteen months before there was some sort of cataclysm," writes Squyres. Chief among these was the loss of the 1999 Mars Climate Orbiter: "The Mars program had become so screwed up that nobody had caught a high-school mistake like mixing up English and metric units."

NASA doesn't escape criticism over the rover mission either. According to Squyres, NASA's rules meant that "cutting corners and taking chances" were the Jet Propulsion Laboratory's only management tools. After the losses of the Climate Orbiter and the Polar Lander, Squyres' team had 34 months to try for Mars again — six months too little, by the usual rules.

They proposed twin rover missions. A team member advised him: "Our top three problems are schedule, schedule and schedule. There's nothing in fourth place." The Polar Lander failed because NASA broke an old rule: "Test as you fly, fly as you test." So the 2003 team set out to test everything and put every design aspect through fresh trials. But NASA gave them more problems, wanting to send only one rover. As Squyres shows — with a novelist's dialogue skills, quoting from meetings and arguments — this idea "doubled our risk and cut our science in half" because it removed the advantage of flying to two very different locations.

Squyres' account also captures the emotions involved: "I do my best to play the steely-eyed space explorer dude," but the first time he saw Spirit move, "crawling slowly forward over blue plastic mats on the high bay floor, it brought tears to my eyes."

Squyres' missions had enormous bad luck.

JPL/NASA

On a roll: during testing for manoeuvrability in the lab, the rovers overcame a series of obstacles.

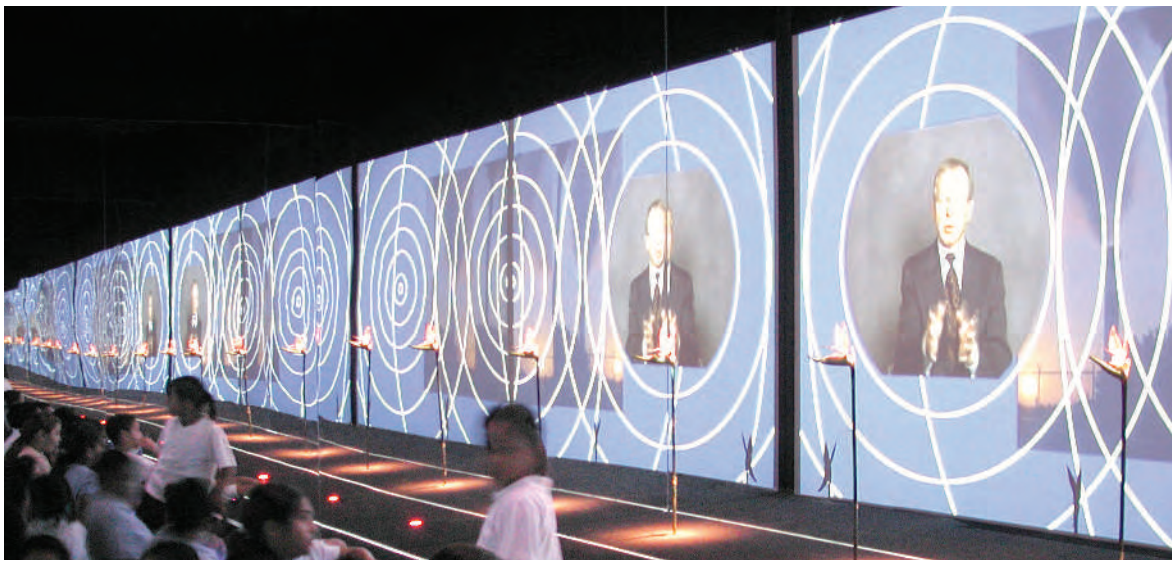
The team had to trim experiments and patch problems right up to the launch date. The largest solar flare ever recorded erupted during their flight. The team got spooked, even after Spirit's successful landing at Gusev Crater. For the second landing, many people, including Squyres, wore the same outfit they wore for the first one. "We may be feeling looser than we did last time, but we still need all the good mojo we can get."

Yet both missions succeeded, and the rovers, engineered for a 90-day mission, are still on the go now after more than 500 days in action. Opportunity ("the glamour rover") even landed in a small depression, Eagle Crater, its first pictures showing true bedrock — what Squyres describes as "a three-hundred-million-mile interplanetary hole in one".

Then came the grind of moving the rovers on a martian day schedule and coaxing pictures from the software: "I love these rovers, but damn, they're slow." It takes them a full martian day to do a simple task that a human explorer on Mars could do in a minute. The drama came as a result of hard work, the long march of rovers across rugged terrain. Moving at a few tens of metres a day, the rovers sought signs of ancient surface water and found it, principally in the 'blueberries' that proved to be haematite, a mineral often formed in the presence of water. Then the rovers reached 140 metres a day and exceeded their engineering lifetimes. Cross-laminated ripples told the research team that Opportunity had found clear signs of layered terrain laid down (perhaps "in a ruby red sea under a pink Martian sky") by flowing water, and now sealed in the geological record. Spirit later discovered layered terrain with clear signs of layered deposition and erosion by water. These findings confirmed the growing view that a warm, wet early Mars was a plausible site for life.

Squyres shows both management skills and an eye for the far horizon, an unusual combination. More geology can be done, but he implies that biology is the next challenge. He concludes: "There are many things I could wish for our rovers, but in the end, there's only one thing that matters. What I really want, more than anything else, is boot prints in our wheel tracks at Eagle Crater." ■

Gregory Benford is a professor of physics at the University of California, Irvine, and is the author of *The Martian Race*.



B. ROGERS/NAT. HIST. MUS. OF LOS ANGELES COUNTY

Looking ahead: a series of experts outline some of the problems facing southern California.

EXHIBITION

Down and out

Collapse?
At the Natural History Museum of Los Angeles County until 15 January 2006
www.nhm.org/exhibitions/collapse/index.html

Philip Campbell

"Collapse? — Are we next?" ask the flags attached to the streetlights as you cruise down Sunset Boulevard. And as you attempt to drive along the freeways, moving ever more slowly across a city faced with increasing water shortages, random gunfire and neighbouring districts that are racially and culturally divided, the appropriate response seems to be:"You bet!"

The flags promote an exhibition at the Natural History Museum of Los Angeles County that provides a strong taste of the threats to our survival, and a sobering sense of past civilizations' failures to meet them. Stimulated by Jared Diamond's book *Collapse* (Viking/Allen Lane, 2005; reviewed in *Nature* **433,** 15–16; 2005), the exhibition is an excellent attempt to persuade citizens to face up to the problems that increasingly confront them. The reality is that Angelenos have little to fear, provided that they and their nation avoid the mistakes highlighted in the exhibition.

If there is one message that emerges clearly from Diamond's comprehensive survey of past and present societies in trouble, it is the importance of anticipating and responding appropriately to the portents of collapse. This means you need to capture a population's attention. Ruefully, Vanda Vitali, the exhibition's originator, acknowledges that Los Angeles is not a museum-going city, at least in comparison with Europe. But the exhibition has attracted strong media coverage and a healthy attendance, not least from schools. The museum is set to launch an interactive extension to accommodate feedback and spontaneous discussions, alongside monthly debates involving citizens and their leaders.

Diamond's unusual approach to history, as previously exhibited in the phenomenally successful *Guns, Germs, and Steel* (W. W. Norton, 1997), is to treat it in a scientific fashion akin to evolutionary biology and cosmology. Unable to tweak the world in order to test theories, he conducts copious surveys and analyses in order to discern key principles. Thus *Collapse* surveys many past societies, from Easter Island to Greenland, analysing in detail how each failed or succeeded in coping with some or all of five factors: climate change; decline in support from neighbours or trading partners; hostile neighbours; loss of environmental services; and, critically, how the society dealt with the problems facing it.

In contrast to Diamond's comprehensive coverage, Vitali and her team have achieved a triumph of economy. The failure of the Mayas is highlighted by a sizeable reconstruction of the pyramid and temple of Tikal, set starkly alongside blown-up photos of the ruins, taken before they were excavated. This is in mournful and salutary contrast to the richness of Mayan life as portrayed on its plates and vessels.

Comfortingly, the exhibition depicts a past success, too: the top-down management of forest resources by the Shogun of Japan's Tokugawa region in the nineteenth century. This is effectively represented by the simplicity of a Japanese domestic interior decorated with scrolls describing Japanese forest management and emblems of Samurai leadership.

The visitor is then transported to a contemporary society under threat: Australia. Here I feel that the exhibition failed to convey just how daunting is the litany of threats now facing that arid and hypersaline continent.

At this point the exhibition turns from the concrete to the conceptual, focusing one by one on the five factors identified by Diamond. Anyone familiar with museums knows of that sense of fraying patience and concentration when overly didactic exhibitors confront the visitor with text piled on text. This exhibition takes a radical step to avoid that risk: the concepts are illustrated with almost no text and with the most simple Bayeux tapestry-like cartoon narratives. This controversial strategy pays off. The concepts are explicit enough and are easily comprehended. And they are popular, according to Vitali. As one grandmother said to her: "It allows me to look intelligent to my grandchildren as I tell them the stories."

So far, so good. But these displays would leave little lasting impression were the exhibition not framed by today's people, captured on video for the exhibition. At the outset, Montana is depicted, the place from where Diamond embarks on his own narrative: spacious, inspiring nature undermined by the quietly devastating effects of logging, mining and holiday-home ownership by people with no stake in the area. And at the end of the exhibition, there is southern California: a 12-minute cycle of voices from experts on water, power, climate and planning, explaining the choices that need to be made if the region is to cope with the increasing pressures on support systems that its citizens could all too easily take for granted. Encouragingly, Vitali tells of visitors staying and listening several times over.

Diamond's book showed that the threat to southern California is as nothing compared with the grave dangers that ultimately extinguished the Mayans — and, for that matter, the Easter Islanders, the Chaco culture of what is now the southwest United States and Mexico, and the Greenlanders. It also shows, however, how relationships with neighbours, threats to water from climate change, and environmental

degradation are unavoidable challenges. It is hard to imagine a more successful illumination of those challenges than this exhibition.

Travelling exhibitions are money-losers. But this exhibition deserves to be seen beyond Los Angeles. Southern California is famous enough, thanks to Hollywood, for a presentation of its problems to serve a purpose even outside the United States. Better still would be for local museums to adapt the exhibition, soliciting video narratives of foreseeable threats from their local citizens and leaders. Next stop Australia? ■

Philip Campbell is editor-in-chief of *Nature*.

Chemistry to die for

The Elements of Murder: A History of Poison
by John Emsley
Oxford University Press: 2005. 436 pp.
£18.99, $30

Roger P. Smith

The title is a bit of a put-on. In *The Elements of Murder*, John Emsley, a chemist and science writer, slowly works his way through the periodic table, discussing those elements that are toxic enough to have caused human poisonings, whether accidental, or as emphasized here, deliberate. Rather more amusing is the (unintentional?) quip in the introduction: "Murder by poison may be a dying art…"

The major thrust of this effort centres on what professional toxicologists invariably, but unsatisfactorily, classify as 'heavy metals'. These are the electropositive natural constituents of the Earth's crust, with a density greater than 5. The classification has always been unsatisfactory because it lumps together some 40 elements of wildly different acute or chronic toxicity, including some that have a very low toxicity, or are toxic but still required in small amounts for normal biological function. No toxicologist, however, would dispute the importance of mercury, arsenic and lead in any work on poisons, and a great deal of attention is lavished on them here, together with antimony and thallium. Because Emsley excludes non-metallic organic compounds, the work is not a general history of poison.

The Elements of Murder is obviously a labour of love, but it is less certain whether a niche readership will emerge that will reciprocate the author's affection for the material. The book is both authoritative and meticulously researched, and I found remarkably little to quarrel with in the factual content. Emsley knows what he is talking about. The book's formidable girth includes an extensive bibliography, a brief appendix, a glossary and a useful index. But the lack of citations in the text will limit the book's use as a primary reference source. On the other hand, it is too technical and comprehensive to be a good beach read.

Emsley gives us much more than the collection of anecdotes about grisly murders that the title and dust-jacket seem to promise. Indeed, even the most ghoulish of readers will grow weary from the sheer number of cases reported. Because homicidal arsenic poisoning has been carried out with regularity since antiquity, there are many more examples for arsenic than for the other elements. One would think that a rather smaller group might have made the point. How many cases can one read about before imagining that one is personally experiencing the symptoms? I was surprised by the intimate details that Emsley has uncovered (and perhaps unfortunately chosen to include) about activities that were supposedly conducted in the greatest of secrecy so many years ago.

The book also includes detailed information on the amounts of the elements found in common foodstuffs, their industrial applications, common chemical reactions, distributions of the element in the body, methods of detection, common forms found in the environment, medical uses, and much more. It's not that this material is irrelevant, just that it is all a little too much. Tucked away into this mass of information are many fascinating tidbits. For example, can it be that thallium is one of the weapons of mass destruction supposedly concealed by Saddam Hussein? The book includes reports on the poisonings, alleged or authentic, occupational or sinister, of many famous people, such as Napoleon, Isaac Newton, Handel, Beethoven, King George III and others.

Toxicological classifications always include a 'miscellaneous' category, and the final chapter has such a group. It contains much shorter treatments of a few more heavy metals and some other assorted elements, many of which were never involved in homicides. But there is no mention of iron, which has caused many tragic accidental deaths.

This is a lovely book, but perhaps sitting down and reading it from cover to cover is not the best way to appreciate it. ■

Roger Smith is the Irene Heinz Given professor of pharmacology and toxicology emeritus, Dartmouth Medical School, Hanover, New Hampshire 03755, USA.

Creatures and craters

Whether you are planning an exotic vacation or looking for a location for your next field trip, two recently published books should provide some ideas.

Nature's Stronghold: The World's Great Wildlife Reserves **by Laura and William Riley (Princeton University Press, $49. 50, £32.50) is a compendium of more than 600 of the world's wildlife reserves.**

The book features reserves from 80 countries, giving details of such well known ones as America's Yellowstone National Park and Tanzania's Serengeti reserve, as well as lesser known specialist reserves such as Russia's Shulgan-Tash Zapovednik, which was set up to protect the last wild Burzyan honeybees in Russia. Vu Quang reserve in Vietnam is home to the Vu Quang ox — the first new mammal species reported for 50 years. The European bison was once nearly hunted to extinction but a remnant herd, originally protected by Hitler's deputy Hermann Göring, now inhabits Poland's Bialowieza National Park and a transborder reserve in Belarus. And the endangered jaguar (see picture) finds safe haven at the Cockscomb Basin Jaguar Sanctuary in Belize.

This illustrated book gives details of what to see and when to visit for each of the reserves featured, making it an interesting read whether you are planning a trip or are just an armchair traveller.

More intrepid travellers might find *The Volcano Adventure Guide* (Cambridge University Press, £30, $50) more exciting. Rosaly Lopes' book gives advice on planning a visit to an active volcano, and has detailed guides to 42 of the world's most spectacular examples.

With chapters on the types of eruption as well as safety and survival rules, this is another book that should interest those staying at home as well as amateur and professional volcanologists alike. M.P.

NEWS & VIEWS

NEUROSCIENCE

Neurons and navigation

György Buzsáki

Where is the geometry of the environment represented in the brain? The entorhinal cortex, where neurons fire repeatedly when an animal's position coincides with the vertices of a grid of triangles, looks like a good bet.

Navigators will be familiar with the principles underlying the question tackled by Hafting *et al.*[1] on page 801 of this issue. The authors' aim was to clarify where in the brain information about an animal's whereabouts is integrated. Their experimental subjects were rats; their approach involved recording from neurons in a part of the brain known as the entorhinal cortex as the rats explored specially designed enclosures.

The simplest navigational method is 'dead reckoning', which requires continuous monitoring of the distance and course taken from some known place. To calculate distance, mechanisms for monitoring both speed and time elapsed are needed. But unless position is recalibrated, errors accumulate over time, hence the need also for celestial or map-based navigation, in which geographical position is determined by 'fixes' of stars or landmarks. A detailed map of the environment can be created through tessellation, which exploits the principle from floor tiling that every part of a surface can be covered by certain polygons of the same size — squares or equilateral triangles — without any gaps or overlaps. Once these polygons are calibrated with the fixed landmarks, they can then provide local reference points (Fig. 1a).

All mammals seem to use both dead reckoning and map-based methods of navigation. To investigate how this information is represented in the brain, researchers have focused on so-called 'place cells' in the hippocampus — a structure that has indirect connections to several areas of the cerebral cortex[2,3]. But the entorhinal cortex, which occupies a crucial position between the hippocampus and the cortical structures that provide visuo-spatial information, has increasingly come in for attention. Hafting *et al.*[1] now report that information about position, distance and direction can be integrated in the dorsocaudal medial region of the entorhinal cortex (dMEC).

The most striking of their results is the spatial regularity of neuron discharges in the dMEC: activation occurs in a grid-like pattern, with each cell becoming repeatedly active when the rat walks through the vertices of equilateral triangles (Fig. 1b). This grid of virtual triangles covers the entire surface of the recording environment, with the resolution determined by the size of the tiles. Hafting *et al.* found not only that individual entorhinal neurons tessellate a rat's enclosure in a way that corresponds to a grid of triangles, but also that the triangular representation varies from cell to cell.

Like the multiple layers of Marrakech mosaics, representing the evolving styles of different centuries, different neurons can display rotated or displaced versions of the same tessellating structure. Furthermore, neuronal grid size increases from the dorsal to the ventral part of the dMEC from 30 to 70 cm, indicating that the spatial resolution of the environment varies systematically within the entorhinal cortex. By integrating the information from such linearly transformed grids, spatial resolution can, in principle, be increased (Fig. 1c), and the rat's position accurately predicted from the activity of just a handful of dMEC neurons[4].

Hafting *et al.* also found that the tessellating features of individual neurons persist in environments of different sizes or shapes, and even after the removal of visual cues or when a rat enters a new environment — a strong indication of a rigidly wired representation of spatial metrics. On the other hand, the orientation grid of all dMEC neurons could be altered simultaneously and instantaneously by the coherent rotation of visual landmarks (such orientation information may be supplied by another linear scheme, the previously described head-direction, or 'compass', system[5]).

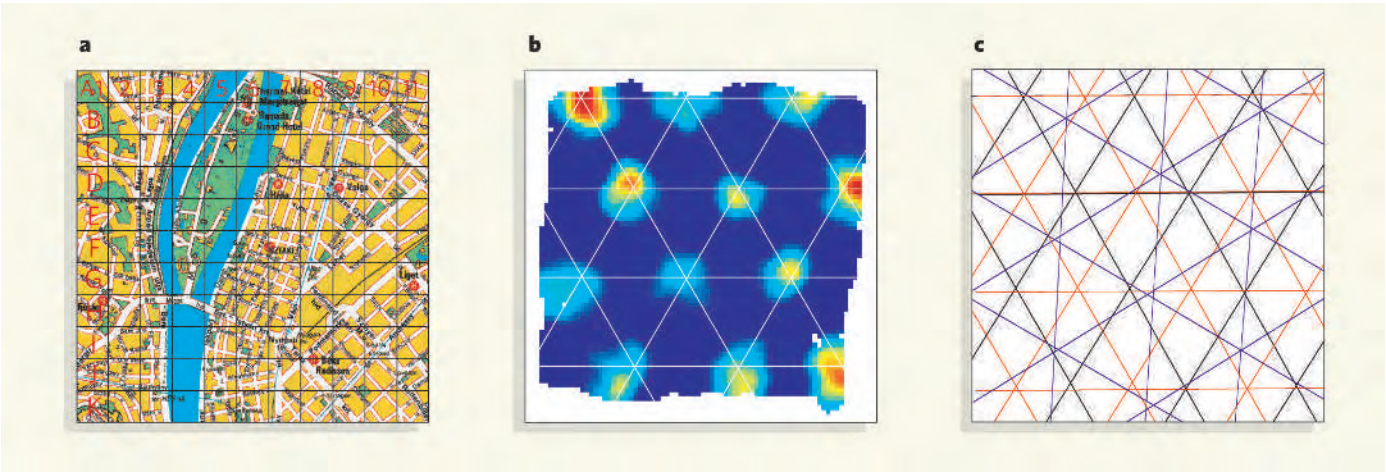


Figure 1 | Tessellation and navigation. **a**, Tessellation of a city map by squares provides information about position, distance and direction, allowing specific places to be easily located. **b**, Hafting *et al.*[1] find that as a rat explores an experimental enclosure, the discharge rate of a neuron in the dorsocaudal medial entorhinal cortex increases at regular intervals corresponding to the vertices of a triangular grid. **c**, Integration of information from several grid components (that is, from the outputs of several neurons) can increase the spatial resolution of the environment. Three triangular grids are represented here, with red displaced and blue rotated relative to a neuron grid shown in black.

All in all, although neurons in many other systems fire in relation to the animal's position and orientation, Hafting and colleagues' findings suggest that the representation of position, distance and direction may come together in the dMEC. Their observations are compatible with both dead-reckoning and map-based navigation: dMEC neurons showed the same grid-like pattern of activity even in complete darkness, but also followed rearrangements of visible landmarks in the test enclosures.

What about the issues raised by this discovery of tessellation patterns? The superficial layer of the dMEC monitored by Hafting *et al.* is only a small part of the entorhinal cortex — do neurons in other layers and parts of this brain region show similar grid patterns or operate in a qualitatively different manner? If they do possess tessellation features, what functions do these neurons serve? How are the grid patterns described by Hafting *et al.* formed in the first place? Is the 30–70-cm metric specific to the rat and the environment in which it was reared, or is it universal in mammals? Is the flexible representation of position that is evidently encoded in the hippocampus crucial in setting up and modifying the rigid grid system of the entorhinal cortex? Finally, how do these results relate to the non-spatial responses of neurons observed in the lateral entorhinal cortex[6]?

There is also the long-standing issue of the relationship between navigation and memory. Ample clinical evidence indicates that the hippocampus and entorhinal cortex are involved in encoding and recalling memories, in particular episodic memories[7–9] — events that are specific to an individual and that are embedded in both a spatial and a temporal context. Are such memories processed by different neuronal mechanisms from those in the dMEC, or do the tessellation features described by Hafting *et al.* play a part? One potential link is the 'distance' relations (the coupling strengths) between the adjacent elements of an episode, which allow for recollection of the exact sequence of what happened where and when.

However, the map-based navigation supported by the dMEC grid patterns does not require a time component, whereas the key feature of episodic memories is their temporal context[7]. So another organizational feature of the hippocampal–entorhinal system might be involved. This is the 'theta oscillation' of collective neuronal activity, which has a frequency of 5–10 Hz, and which may provide the temporal metrics needed for both dead-reckoning navigation and the temporal ordering of the elements of episodic memory. This mechanism could conceivably bring grid cells with different sizes and orientations into temporal register with each other. If it does, we may begin to think about how elementary mechanisms such as space tessellation and temporal coordination, which initially supported dead-reckoning navigation, evolved to represent memories[10]. ■

György Buzsáki is at the Center for Molecular and Behavioral Neuroscience, Rutgers, State University of New Jersey, 197 University Avenue, Newark, New Jersey 07102, USA.
e-mail: buzsaki@axon.rutgers.edu

1. Hafting, T., Fyhn, M., Molden, S., Moser, M.-B. & Moser, E. I. *Nature* **436,** 801–806 (2005).
2. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Clarendon, Oxford, 1978).
3. McNaughton, B. L. *et al. J. Exp. Biol.* **199,** 173–185 (1996).
4. Fyhn, M., Molden, S., Witter, M. P., Moser, E. I. & Moser, M. B. *Science* **305,** 1258–1264 (2004).
5. Muller, R. U., Ranck, J. B. Jr & Taube, J. S. *Curr. Opin. Neurobiol.* **6,** 196–206 (1996).
6. Hargreaves, E. L., Rao, G., Lee, I. & Knierim, J. J. *Science* **308,** 1792–1794 (2005).
7. Tulving, E. *Annu. Rev. Psychol.* **53,** 1–25 (2002).
8. Squire, L. R. *Psychol. Rev.* **99,** 195–231 (1992).
9. Eichenbaum, H. *Nature Rev. Neurosci.* **1,** 41–50 (2000).
10. Buzsáki, G. *Hippocampus* (in the press).

MAGNETOSPHERIC PHYSICS

Turbulence on a small scale

Melvyn L. Goldstein

The four-spacecraft Cluster mission has identified small-scale vortices in Earth's magnetosphere. The observation reveals processes that transfer energy and momentum from the solar wind to the magnetosphere.

Turbulence is ubiquitous. It can be seen by gently stirring cream into coffee, or by observing the white caps and surf at the beach, and it causes the drag on cars and aeroplanes. The phenomenon is also widespread in magnetized plasmas — ionized gases that contain a magnetic field — such as the interstellar and intergalactic media, the solar wind and Earth's magnetosphere.

Although magnetic fields complicate the description of turbulence, certain characteristics, among them vortices, can arise in fluid and magnetofluid turbulence. On page 825 of this issue, Sundkvist *et al.*[1] report the detection of small vortices in the 'dayside cusp' of Earth's magnetosphere by the four-spacecraft Cluster, a joint European Space Agency and NASA mission. The authors describe their observations as evidence for a phenomenon known as drift–kinetic Alfvénic turbulence.

The solar wind is the extension of the corona, the hot outer atmosphere of the Sun, into interplanetary space. Near Earth, the speed of the solar wind is typically around 400 kilometres per second, with variations of a couple of hundred kilometres per second that

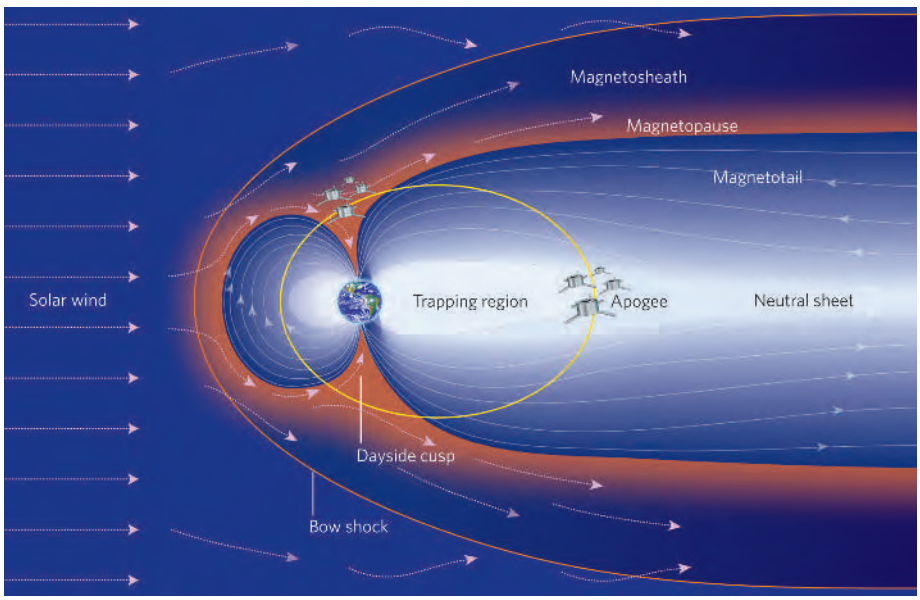


Figure 1 | Transverse cut of Earth's magnetosphere. A two-dimensional cut through the magnetosphere, with the Sun to the left. Solar-wind plasma is heated as it passes through the bow shock caused by its encounter with Earth's magnetic field. Some of the plasma is funnelled down the dayside cusps into the Northern and Southern Hemispheres. The Cluster orbits are shown as they appear in February/March of each year, the time period of the observations reported by Sundkvist *et al.*[1]. When the four spacecraft are on the dayside, they pass through the cusp regions; when near apogee on the nightside, they cross the magnetotail and neutral sheet. (The diagram is not to scale; in reality, the apogee of Cluster's orbits occurs at some 19 Earth radii.)

depend on solar activity. That speed exceeds by about a factor of ten the speed of sound and the speed of the most common magnetic waves, known as Alfvén waves. (Alfvén waves are highly correlated fluctuations in both the fluid-velocity and magnetic fields.)

As the supersonic, super-Alfvénic solar wind encounters Earth's magnetic field, a bow shockwave is produced at about 10–15 Earth radii in front of Earth (Fig. 1). Behind the bow shock, the hot solar-wind plasma can flow down towards the ionosphere through the dayside cusp. This cusp forms the boundary between magnetic field lines that are closed on the dayside (the side of Earth exposed to the Sun) and magnetic field lines that are open and have been swept back into the lobes of the nightside magnetosphere.

The flow of plasma down the funnel-like cusp has been conjectured either to excite turbulence locally or to amplify the turbulence carried by the shocked solar-wind plasma. Savin *et al.*[2] noted that the flow down the cusp should generate vortices. Nykyri *et al.*[3] found, using Cluster magnetometer data from March 2001, evidence that the cusp contained magnetic turbulence.

Cluster's four spacecraft orbit such that, at the point where they are farthest from Earth — at apogee — their positions form a regular tetrahedron (Fig. 1). This formation is ideal for distinguishing between unchanging spatial features and features that evolve with time. In the spring of 2002, at the time of Sundkvist and colleagues' measurements[1], the separation of the spacecraft at apogee was about 100 kilometres (compared with around 500 kilometres for earlier Cluster measurements[3]), allowing the resolution of smaller spatial features. When the spacecraft entered the cusp, they were still relatively close together, but no longer traced out the points of a regular tetrahedron (see Fig. 3e on page 827).

Turbulence is often described as a process in which large-scale eddies cascade down to smaller-scale eddies until a scale is reached at which dissipation sets in. In magnetized plasmas, because of the large variety of possible small-scale wave modes, it is not clear how that cascade progresses to the dissipation range. Understanding these processes would enable us to determine how energy flows in a turbulent magnetofluid from large scales to the smaller, kinetic scale — and thus heats Earth's ambient plasma. Furthermore, where vortices form, materials in initially separate regions of space become mixed, which transfers energy, momentum and material from one region of the magnetosphere to another.

In the solar wind itself, measurements from single spacecraft[4,5] indicate that as the scales of the turbulent fluctuations approach the dissipation scale, the kinetic Alfvén wave becomes the predominant wave mode. The distinguishing characteristic of such waves is that they have a small electric field that is parallel to the direction of the local magnetic field. Sundkvist and colleagues[1] have now analysed the temporal evolution of the magnetic field in the dayside cusp region. There they show that kinetic Alfvén waves interact nonlinearly with so-called drift waves caused by gradients in plasma density and magnetic fields, and provide evidence that this interaction produces distinctive small-scale turbulent features known as drift–kinetic Alfvén vortices. Measurements of the velocity shears across the plasma flow direction made by Cluster's thermal plasma instrument indicate amplitudes that exceed those required for vortex production. The authors' interpretation is further bolstered by a vortex model constructed by the authors that accurately reproduces the observed behaviour of the magnetic field.

The Cluster observations are the first measurements in space to indicate that small-scale vortices are formed as eddies reach the dissipation scale. At the time of the observations, Cluster 1 and 2 were aligned with the plasma flow. Data from those two spacecraft indicated that the observed structures were quasi-stationary. The other two spacecraft were not aligned with the flow and could be used to deduce that the transverse radial scales of the structures were a few proton gyroradii (the radius of the circle described by a proton moving across a background magnetic field — in this region of the cusp, about 25 kilometres).

The small spatial scales involved make measurements of a turbulent cascade's dissipation — and of the transfer of the energy contained in magnetic fields and particle motion into the heating of the ambient plasma — difficult in both terrestrial laboratories and in space. Cluster observations indicate the existence of a turbulent vortical boundary layer that enhances the transfer of momentum and energy from the solar wind to the magnetosphere. The four-spacecraft Cluster, with its ability to distinguish between spatial and temporal effects, has opened a new window on the study of turbulence, both in the magnetosphere and in the solar wind. In the near future, missions such as Magnetospheric Multiscale (MMS), with its even smaller spacecraft separation and higher time-resolution for plasma measurements, will further enhance our understanding of the generation and dissipation of magnetofluid turbulence. ■

Melvyn L. Goldstein is at the NASA Goddard Space Flight Center, Code 612.2, Greenbelt, Maryland 20771, USA.
e-mail: melvyn.l.goldstein@nasa.gov

1. Sundkvist, D. *et al. Nature* **436,** 825–828 (2005).
2. Savin, S. P. *et al. JETP Lett.* **74,** 547–551 (2001).
3. Nykyri, K. *et al. Ann. Geophys.* **22,** 2413–2429 (2004).
4. Leamon, R. J. *et al. Astrophys. J.* **537,** 1054–1062 (2000).
5. Bale, S. *et al. Phys. Rev. Lett.* **94,** 215002 (2005).

PLANT BIOLOGY

Engineered male sterility

Muhammad Sarwar Khan

The phenomenon of 'cytoplasmic male sterility' in plants has long been exploited to enhance the productivity of certain crops. An innovative genetic-engineering system promises to widen applicability of the approach.

Among the main items on the wish-list of plant breeders are these. First, the ability to artificially suppress pollination and so prevent a plant's self-fertilization, thereby encouraging cross-pollination and higher-yielding seed through an effect known as 'hybrid vigour'. Second, the ability to genetically engineer such suppression of male fertility into elements in the cytoplasm, rather than the nucleus — the result is transmission of desirable characteristics through genes in the female line, and those genes cannot 'escape' uncontrollably via pollen. Third, the ability to selectively restore male fertility. Although farmers want high-yielding hybrid seeds, for certain crops that seed has to produce fertile plants.

These aims can be achieved by exploiting a phenomenon known as 'cytoplasmic male sterility'. As they describe in *Plant Physiology*, Ruiz and Daniell[1] have demonstrated a promising new way of achieving all three goals. Their approach was tested in tobacco plants. It involves inserting a gene (*phaA*) from the bacterium *Acinetobacter* into plant chloroplasts, with — upstream of that gene — a 'promoter', *psbA*, and other regulatory elements.

The *phaA* gene encodes an enzyme, β-ketothiolase. The authors show that in their transgenic plants the enzyme accumulates in leaves and anthers, the pollen-producing structures, and alters the course of synthesis of fatty acids. A starting point in fatty-acid synthesis is acetyl-CoA, which under normal circumstances is converted to malonyl-CoA. However, β-ketothiolase overrides the usual enzyme involved, acetyl-CoA carboxylase, to produce acetoacetyl-CoA instead (Fig. 1a, b, overleaf). Correct lipid metabolism is essential to the normal development of pollen, not least the pollen wall[2,3]. Ruiz and Daniell[1] found that β-ketothiolase accelerates anther development and, among other consequences, causes the pollen grains to collapse — and thereby results in male

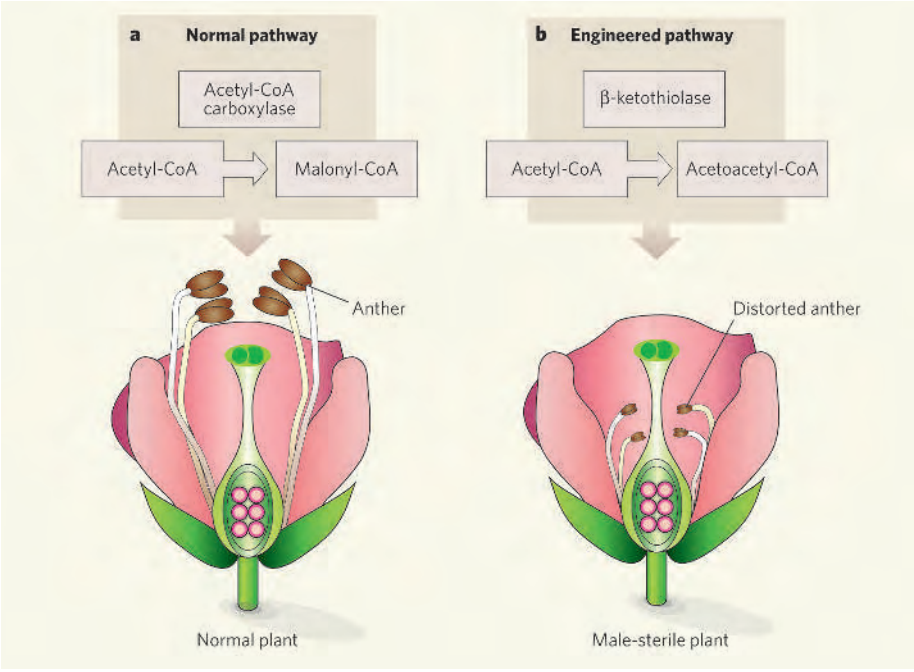


Figure 1 | Engineering cytoplasmic male sterility with β-ketothiolase[1]. **a**, In chloroplasts, acetyl-CoA, the substrate for the synthesis of fatty acids, is normally converted by acetyl-CoA carboxylase to yield malonyl-CoA. This pathway results in the correct development of anthers, pollen grains and seeds. **b**, In the transgenic chloroplasts, β-ketothiolase out-competes acetyl-CoA carboxylase for acetyl-CoA, with acetoacetyl-CoA being produced instead. The upshot is distorted anthers and failure of pollen development. Fertility in plants grown from the resulting hybrid seeds is restored under continuous illumination, with reversion to the normal pathway.

sterility. But the transgenic plants were otherwise unaffected.

So much for producing male sterility. How about restoring it? The hybrid seed itself is valuable for growing certain ornamental species, for example, or for producing vegetables. But in cases such as oilseed rape, sunflower or maize, where the crop germinates from the second-generation seed of the hybrid plants, fertility has to be restored.

In some plants, the nuclear genome overrides cytoplasmic male sterility to restore male function, but this process often works inefficiently and has deleterious effects on plant growth because it interferes with general metabolism and development[4,5]. In the case investigated by Ruiz and Daniell[1], no nuclear-encoded restoring factor is involved. This is where the *psbA* promoter and associated regulatory elements come in, because they confer light-sensitivity on the gene they control[6].

The authors hypothesized that, even though both acetyl-CoA carboxylase and β-ketothiolase are controlled by light-inducible promoters, under continuous illumination the carboxylase would gain the upper hand, so restoring normal fatty-acid synthesis and male fertility. That turned out to be the case, at least to some extent. When grown under continuous illumination for 10 days, a sample of transgenic plants produced four flowers with viable pollen, and in due course viable seed.

Genetic transformation of chloroplasts in the cytoplasm has several advantages over nuclear transgenic technologies[7]. Apart from transgene containment[8–11], those advantages include a comparatively high level of transgene expression, yielding proteins that are properly folded and fully functional; lack of side effects, such as stunting or other abnormalities[12]; and elimination of the laborious back-crossing that is needed with nuclear transformation to introduce cytoplasmic male sterility into élite plant lines. The new method is likely to be especially advantageous when applied to crop plants with longer generation times, such as cotton, maize and rice.

As to future research, it is not clear from Ruiz and Daniell's study how sterility was reversed by continuous illumination because the two competing enzymes are both light-regulated. That, then, is one aspect that calls for further investigation. ■

Muhammad Sarwar Khan is at the National Institute for Biotechnology and Genetic Engineering, PO Box 577, Postcode 38000 Jhang Road, Faisalabad, Pakistan.
e-mail: sarwarkhan_40@hotmail.com

1. Ruiz, O. N. & Daniell, H. *Plant Physiol.* **138,** 1–15 (2005).
2. Yui, R. *et al. Plant J.* **34,** 57–66 (2003).
3. Ariizumi, T. *et al. Plant J.* **39,** 170–181 (2004).
4. Hernould, M. *et al. Plant Mol. Biol.* **36,** 499–508 (1998).
5. Goetz, M. *et al. Proc. Natl Acad. Sci. USA* **98,** 6522–6527 (2001).
6. Staub, J. M. & Maliga, P. *EMBO J.* **12,** 601–606 (1993).
7. Daniell, H., Kumar, S. & Duformantel, N. *Trends Biotechnol.* **23,** 238–245 (2005).
8. Hagemann, R. in *Molecular Biology and Biotechnology of Plant Organelles* (eds Daniell, H. & Chase, C.) 87–108 (Springer, Dordrecht, 2004).
9. Daniell, H. *Nature Biotechnol.* **20,** 581–587 (2002).
10 Khan, M. S., Khalid, A. M. & Malik, K. A. *Trends Biotechnol.* **23,** 217–220 (2005).
11. Maliga, P. *Nature* **422,** 31–32 (2003).
12. Daniell, H. *et al. J. Mol. Biol.* **311,** 1001–1009 (2001).

50 YEARS AGO

"A cæsium resonator" — A frequency standard based on the natural resonant frequency of the cæsium atom has been constructed at the National Physical Laboratory and used for calibrating the quartz clock standards with an accuracy of ±1 part in 10^9 (0.0001 sec. per day)... The resonance observed is that due to transitions between the hyperfine structure energy-levels caused by the interaction between the nuclear and electron spins... all the allowed transitions have been observed, and from measurements made under a variety of conditions the frequency of the central line and zero field was found to be 9 192 631 830 ± 10 c./s... The results reported here appear to be the first examples of the use of an atomic frequency standard with an accuracy greater than that obtained by using a unit of astronomical time. They also represent the highest accuracy ever achieved in the measurement of any physical quantity in terms of a definitive standard. **L. Essen & J. V. L. Parry**

From *Nature* 13 August 1955.

100 YEARS AGO

"A Modern Utopia." By H. G. Wells. — It is instructive to watch the growth, both in power and in hopefulness, of Mr. Wells's criticism of life. In the "Time Machine" his forecast of the future of humanity was frankly appalling; in "When the Sleeper Wakes," more lurid (albeit far more probable) than the worst imaginings of "reforming" socialists... "Mankind in the Making" contained much vigorous criticism and many sensible and practical suggestions. In the present book Mr. Wells has become still more moderate and practicable and hopeful, without in the least derogating from his ingenuity and originality. We sincerely hope, therefore, he will not, as he threatens, stick henceforth to his "art or trade of imaginative writing," but will continue from time to time to regale and stimulate us with sociological speculations.

From *Nature* 10 August 1905.

50 & 100 YEARS AGO

OCEANOGRAPHY

A bigger nitrogen fix

Nicolas Gruber

Biologically useful nitrogen is delivered to the upper ocean from the depths by vertical transport processes — or such was the received wisdom. In fact, bacteria that convert atmospheric nitrogen may be just as crucial.

Next to light, nitrogen is the main factor that limits the biological productivity of primary producers in the sea. All photosynthetic organisms need nitrogen to grow, yet most cannot use it in its simplest form, molecular N_2. But biologically useful, 'fixed' forms of nitrogen are scarce over much of the illuminated upper ocean. Writing in *Global Biogeochemical Cycles*, Capone *et al.*[1] highlight a little-considered mechanism for supplying the upper ocean with nitrogen — bacteria that fix molecular nitrogen from the atmosphere.

The scarcity of fixed nitrogen in the illuminated ocean results from a continuous loss of nitrogen to the dark, abyssal ocean, resulting primarily from the sinking of organic nitrogen (Fig. 1). If this lost nitrogen were not recycled to the upper ocean, primary production in the sea would plummet within a few years. So identifying and quantifying the replenishment mechanism is essential for understanding marine productivity. Furthermore, as the biologically induced cycling of nitrogen is closely coupled to the cycling of carbon, these processes are also of prime relevance for understanding the oceanic carbon cycle, and hence atmospheric carbon dioxide levels.

Over much of the ocean, upward mixing and transport of nitrate (NO_3^-) constitutes the primary supply pathway (Fig. 1). Nitrate is the dominant form of fixed nitrogen and is generated in the dark ocean from the mineralization of sinking organic nitrogen. In most regions of the tropical and subtropical ocean, however, vertical transport is insignificant, and mixing tends to be inefficient. Surprisingly, biological productivity in these regions is often much higher than might be expected from the poor vertical supply of nitrate, implying the involvement of other, unidentified nitrogen sources[2]. One candidate source was the biological fixation of molecular N_2 from the atmosphere, but based on the measurements then available, this was generally considered unimportant.

Capone and colleagues[1] now demonstrate, in the most exhaustive and comprehensive study so far, that over large regions of the tropical and subtropical Atlantic, biological N_2 fixation is indeed substantial. In fact, it provides the ecosystem of the illuminated ocean with a source of nitrogen that rivals the vertical supply of nitrate. The contribution cements a paradigm shift that has been occurring over the past decade in models of marine nitrogen cycling (see ref. 3 for an example).

The authors focused[1] on the contribution of *Trichodesmium*, a cyanobacterium that is the most conspicuous and best-studied marine N_2 fixer. In a painstaking effort, they measured the N_2 fixation rates of gently collected colonies of *Trichodesmium* at more than 150 stations during six cruises to the tropical and subtropical Atlantic. Averaged over all stations and integrated over the depth of the upper ocean at each station, the annual N_2 fixation rate was 87 ± 14 millimol of N per square metre. This value is of the same order of magnitude as the vertical flux of nitrate into the upper ocean in the region studied.

What makes Capone and colleagues' study particularly compelling is that they estimated N_2 fixation rates using an array of independent methods, each with their own strengths and weaknesses. This results in an unprecedented level of confidence in the estimates obtained. In particular, the authors used measurements of the ratio of the two stable isotopes of nitrogen, ^{14}N and ^{15}N, to confirm that about half of the organic nitrogen in the surface ocean stems from atmospheric N_2 (Fig. 1).

Assuming that their estimate of the biological N_2 fixation rate is representative of most of the tropical and subtropical North Atlantic, Capone *et al.* estimate[1] that *Trichodesmium* annually adds between 1.6 and 2.4 teramol (a teramol is 10^{12} mol) of fixed nitrogen to this region, almost an order of magnitude larger than earlier estimates of N_2 fixation over the whole Earth[4]. This extrapolation is consistent with several indirect geochemical estimates[5–7] of the biological fixation rate that are based on anomalies in the relative abundance of nitrate and phosphate (PO_4^{3-}); these ranged, with one exception[7], from 2 to 6 teramol of nitrogen per year. As these estimates are integrated explicitly over time and space, they are less prone to the difficulties associated with extrapolating local rate measurements. They are, however, based on a number of assumptions — some of which, such as the exact amount of nitrogen and phosphorus required by primary producers that do not fix N_2 — are difficult to verify.

Capone *et al.* also show that their extrapolation is broadly consistent with the value inferred from a region-wide analysis of the ^{15}N-isotopic mass balance. To determine the rate of N_2 fixation, they combine their estimate derived from the nitrogen isotopic ratio — that some 50% of the organic nitrogen in the upper ocean comes from atmospheric N_2 — with an estimate for the turnover rate of organic nitrogen. This value is in fact somewhat larger than the value the authors find based on *Trichodesmium*. But the discrepancy might be explained by the fact that mass-balance analysis includes all sources of N_2 fixation, including recently discovered single-cell cyanobacteria[8] and a cyanobacterium that

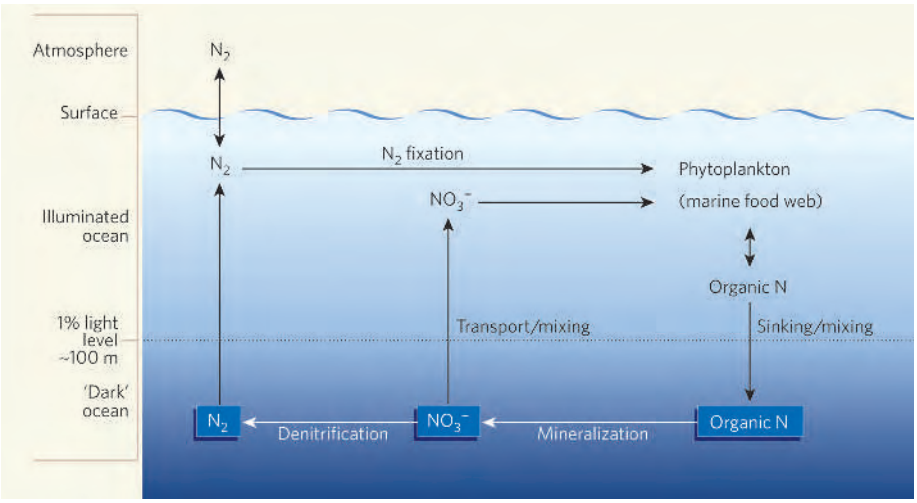


Figure 1 | The marine nitrogen cycle. Some of the organic nitrogen fixed by phytoplankton (primary producers), and then cycled through the food web of the upper ocean, is lost to the dark ocean by sinking and mixing. Over most of the ocean this loss is compensated by the upward transport and mixing of nitrate that is generated from the mineralization of the sinking organic nitrogen. Capone *et al.*[2] show that in certain regions, such as the tropical and subtropical North Atlantic, biological N_2 fixation can supply as much nitrogen to the upper ocean as vertical nitrate transport, and can compensate for the fixed nitrogen lost through denitrification.

lives symbiotically within a marine diatom[9].

Not only does N_2 fixation provide a pathway for adding new nitrogen to the illuminated ocean, but it is also the main source of fixed nitrogen to the ocean as a whole. On long timescales, this source could compensate for the effects of denitrification, a respiratory process that converts fixed nitrogen back to N_2 (Fig. 1). It has been suggested[10] that past estimates of the loss of fixed nitrogen from the ocean need to be revised substantially upwards. If the rate of N_2 fixation is as low as was estimated two decades ago, this would imply that the present-day marine nitrogen budget is seriously out of balance. The much higher estimates of N_2 fixation proposed on the basis of geochemical methods had brought the budget back to near-balance[5], but without direct measurements that conclusion remained tentative. The new convergence of estimates for the Atlantic provides good evidence that N_2 fixation in the ocean may occur on a large enough scale to balance losses of fixed nitrogen.

A fascinating corollary is the question of how marine N_2 fixation and denitrification are coupled. With both processes occurring at such high rates, the residence time of fixed nitrogen in the ocean can be only a few thousand years. This would require a well-established balance between the two processes in order to avoid large swings in the biological productivity of the oceans. Because all organisms need phosphorus as well as nitrogen to survive, the marine phosphorus content almost certainly plays an important role here. Room for surprises in the marine nitrogen cycle remains, but one conclusion is clear: that cycle is much more dynamic than was thought only a few years ago. ■

Nicolas Gruber is at the Institute of Geophysics and Planetary Physics, and the Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, California 90095, USA.
e-mail: ngruber@igpp.ucla.edu

1. Capone, D. G. *et al. Glob. Biogeochem. Cycles* **19,** GB2024, doi:10.1029/2004GB002331 (2005) .
2. Emerson, S. *et al. Nature* **389,** 951–954 (1997).
3. Capone, D. G. *Curr. Opin. Microbiol.* **4,** 341–348 (2001).
4. Capone, D. G. & Carpenter, E. J. *Science* **217,** 1140–1142 (1982).
5. Gruber, N. & Sarmiento, J. L. *Glob. Biogeochem. Cycles* **11,** 235–266 (1997).
6. Michaels, A. F. *et al. Biogeochemistry* **35,** 181–226 (1996).
7. Hansell, D., Bates, N. & Olson, D. *Mar. Chem.* **84,** 243–265 (2004).
8. Zehr, J. *et al. Nature* **412,** 635–638 (2001).
9. Carpenter, E. J. *Mar. Ecol. Prog. Ser.* **188,** 273–283 (1999).
10. Codispoti, L. A. *et al. Sci. Mar.* **65,** 85–105 (2001).

CANCER

Two in one

Anton Berns

As cancer develops, at least two cell processes are disrupted — cell growth is promoted, and cell death inhibited. It seems that mutated versions of the notorious cancer-promoting protein MYC can accomplish both at once.

The MYC gene is one of the classic cancer-promoting 'oncogenes'. It is overexpressed in many types of tumour, and the MYC protein it encodes causes a surge in the proliferation of cells. But it has another effect: it enhances programmed cell death, or 'apoptosis'. So under normal circumstances, the extra cell divisions MYC causes when overexpressed are cancelled out by a rise in cell fatalities. In MYC-associated tumours, however, there is usually a mutation in an ancillary protein that disrupts the apoptosis pathway, releasing the brakes on cell proliferation. In this issue, Hemann *et al.* (page 807)[1] report that MYC does not need to rely on a partner-in-crime to cause tumours — it can itself be mutated in a way that interferes with its apoptotic function.

The authors examined MYC function in Burkitt's lymphoma, a malignancy of immune cells called B cells, which normally produce immunoglobulins in response to an immune attack. The cancer is characterized by gross overexpression of MYC because the position of the MYC gene is swapped with that of the immunoglobulin gene. As this latter gene is often switched on in B cells, and as the 'translocated' MYC gene is slotted in next to the regulatory sequences for the immunoglobulin gene (the 'promoter'), it too is switched on. Translocated MYC genes often harbour specific point mutations that tend to result in altered amino acids in a specific part of the MYC protein (the amino-terminal domain). This clustering of mutations led to the suspicion that they augment the tumorigenic potential of MYC. However, *in vitro* assays with these mutants failed to show either enhanced cancer-promoting activity or reduced apoptosis[2].

To assess the cancer-promoting potential of MYC mutants, the authors used mice that had been irradiated to permit reconstitution of their immune system. They reconstituted these cells from progenitor cells, called haematopoietic stem cells (HSC), that had been manipulated to express either normal (wild-type) or mutated MYC, and then followed the formation of tumours. When HSC expressing wild-type MYC were used, lymphomas developed in a small fraction of the mice only after a long period. By contrast, most of the mice grafted with HSC carrying mutant MYC quickly developed tumours, indicating that the MYC mutants are far more tumorigenic than wild-type MYC.

To identify the mechanism underlying this difference, the authors checked whether the gene mutants had different effects on the apoptotic signalling pathways induced by MYC overexpression. They found that the ARF–p53 pathway, a known target of MYC, was induced equally by wild-type and mutant MYC. However, the apoptosis-promoting protein BIM, which is part of a different pathway, was

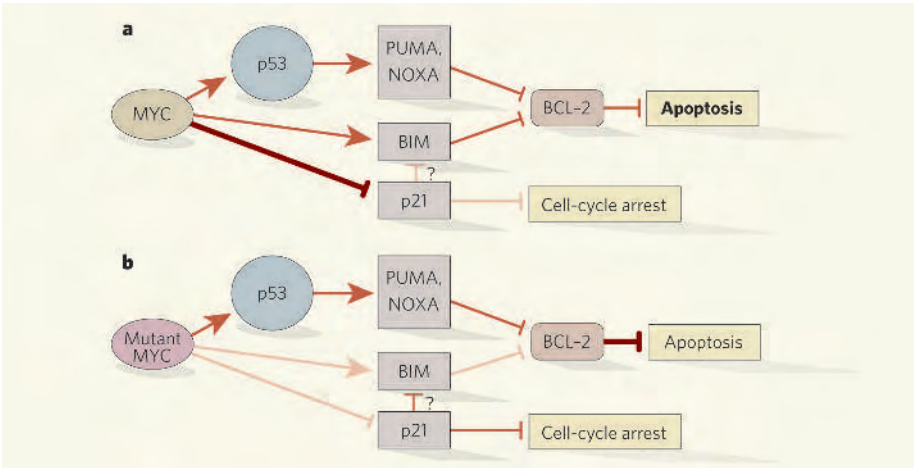


Figure 1 | The machinations of MYC. The pathways induced by wild-type (**a**) and mutant (**b**) MYC. Arrows indicate activation, and T-bars indicate inhibition, with the depth of colour showing signal intensity. Hemann *et al.*[1] find that mutant MYC fails to induce BIM. BIM acts through BCL-2, a key player in protecting cells from programmed cell death (apoptosis). If BIM is not induced, apoptosis is prevented. The same result is observed if wild-type MYC is overexpressed in cells in which either the p53 or the BIM pathway is disabled by targeted disruption of BIM or p53. The signalling through these pathways seems to be additive, and therefore partial inactivation of either one might lead to a similar outcome. Question marks indicate speculative connections.

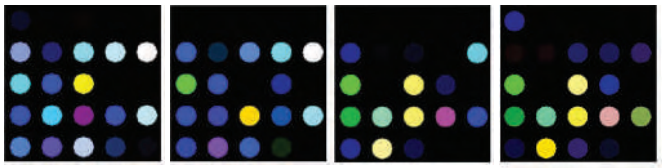
ORGANIC CHEMISTRY

Dotty solutions

What are you drinking? One answer to this frequently asked question might be found by a sensor that uses an array of chemically sensitive dyes to identify organic compounds dissolved in water.

The innovative array developed by Chen Zhang and Kenneth S. Suslick (see *J. Am. Chem. Soc.* doi:10.1021/ja052606z; 2005) uses 36 dots of dyes that change colour in response to pH, molecular polarity and Lewis basicity (how readily a molecule donates an electron pair). These properties are strongly influenced by water, making it tricky to identify trace molecules in solution. So Zhang and Suslick dampen the effects of water by using hydrophobic dyes on a hydrophobic membrane.

The combination of colour changes in the dye dots when they are dunked in solution forms a 'fingerprint' of the compounds present. The authors confirm this using a variety of common organic molecules at concentrations as low as one micromol per litre. And, although a breakdown of components is not possible with the array, complex mixtures of organic molecules do excite a unique response — as the authors show by testing a number of similar aqueous solutions found in their refrigerators (see images).



AM. CHEM. SOC.

Zhang and Suslick concede that the recognition of flavours is still some way off. That would require the incorporation into the array of hydrophobic dyes that are sensitive to salt or sugar, for instance. For now, answers to questions of taste, at least, will remain on the tip of the human tongue.

Richard Webb

activated only by wild-type MYC. The mutants' failure to activate BIM seems to contribute to their enhanced tumorigenicity, because both wild-type and mutant MYC were equally oncogenic when HSC lacking BIM were used. This implies that BIM normally constrains the carcinogenic potential of wild-type MYC, consistent with previous observations[3].

Activation of BIM by wild-type MYC does not require p53 signalling, as BIM levels were elevated in p53-deficient cells overexpressing wild-type MYC. Therefore, the oncogene activates BIM through an independent route. The authors' model predicts that BIM expression is not induced in Burkitt's lymphoma cells carrying MYC mutations. Indeed, high levels of BIM were found in all seven human Burkitt's lymphoma samples examined that carried a wild-type MYC gene, but in only one of seven Burkitt's lymphoma samples carrying a mutant gene.

Although the MYC mutants seemed to have little effect on p53 itself, they did activate one of its downstream effectors, an inhibitor of cell division called p21. How does this effect on p21 occur, and does it contribute to MYC-driven signals? Wild-type MYC inhibits p21 production by binding to the p21 gene promoter in a complex with the protein Miz-1 and recruiting further repressor proteins[4,5]. Perhaps the mutations disrupt MYC's interaction with the repressors. Or, as overexpression of mutant MYC also suppresses its own promoter, perhaps there is no wild-type MYC to form complexes with Miz-1, thereby relieving the suppression of p21. Consistent with this notion, overexpression of wild-type MYC reduced p21 levels.

BIM and p21 levels seem to be inversely regulated in this system, but it remains unclear whether wild-type MYC directly activates BIM or does so through p21. In the latter case, an interesting scenario emerges in which p21 acts upstream of BIM[6], serving as a switch to determine whether a cell will stop dividing or undergo apoptosis.

Both wild-type and mutant MYC have similar activating effects on three components of the p53-controlled apoptotic pathway — Bax, PUMA and NOXA. However, activation of these apoptosis-promoting factors does not trigger apoptosis if BIM activation is compromised (Fig. 1). Similarly, BIM activation by wild-type MYC does not induce apoptosis if the p53 pathway is disabled. The picture that emerges suggests that impairment of either the p53 or the BIM signalling route is enough to make wild-type MYC as oncogenic as the MYC mutants. The apoptotic signals conveyed by either pathway seem to be similar and additive. In mouse models, loss of one copy of the BIM gene confers strong resistance to apoptosis[3], and it will be interesting to learn whether such a loss is sufficient to abolish the difference between wild-type and mutant MYC in inducing lymphomas.

As chemotherapy usually activates the p53 pathway, these observations prompt a comparison of the response to chemotherapy between lymphoma patients carrying a translocated mutant MYC gene and normal p53 and patients carrying a translocated wild-type MYC and mutant p53. For patients with a mutant MYC, one might argue that, despite the reduced apoptosis resulting from suppression of BIM, forceful activation of the p53 pathway might potently induce apoptosis. However, it is possible that p21 would divert the signal towards cell-cycle arrest rather than death[6]. This comparison would also show whether the resistance to chemotherapy observed in mouse models of Burkitt's lymphoma[7] that overexpress MYC and are deficient in p53 faithfully mimics the response of human lymphomas with similar characteristics. If this is the case, it will offer the opportunity to refine the treatment of Burkitt's lymphoma. ■

Anton Berns is in the Division of Molecular Genetics and Centre for Biomedical Genetics, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands.
e-mail: a.berns@nki.nl

1. Hemann, M. T. *et al. Nature* **436,** 807–811 (2005).
2. Chang, D. W., Claassen, G. F., Hann, S. R. & Cole, M. D. *Mol. Cell. Biol.* **20,** 4309–4319 (2000).
3. Egle, A., Harris, A. W., Bouillet, P. & Cory, S. *Proc. Natl Acad. Sci. USA* **101,** 6164–6169 (2004).
4. Wu, S. *et al. Oncogene* **22,** 351–360 (2003).
5. Seoane, J., Le, H. V. & Massagué, J. *Nature* **419,** 729–734 (2002).
6. Collins, N. L. *et al. Mol. Cell. Biol.* **25,** 5282–5291 (2005).
7. Schmitt, C. A. *et al. Cell* **109,** 335–346 (2002).

MANTLE GEOCHEMISTRY

Big lessons from little droplets

Claude Herzberg

How does Hawaii look deep below the surface? Like viewing an object at a different magnification, studies of minuscule inclusions in volcanic rocks on the surface provide a fresh perspective on the question.

Mantle plumes are thought to be roughly elongate cylinders of rock that buoyantly rise up from deep within the Earth, manifesting themselves at the surface in features such as the Hawaiian islands and Iceland. Much attention has centred on Hawaii, because it is constructed from Earth's largest volcanoes distributed along two geochemically distinct alignments, and there is considerable debate about what these distinctions reveal about the underlying plume that feeds them. On page 837 of this issue[1], Ren *et al.* add to that

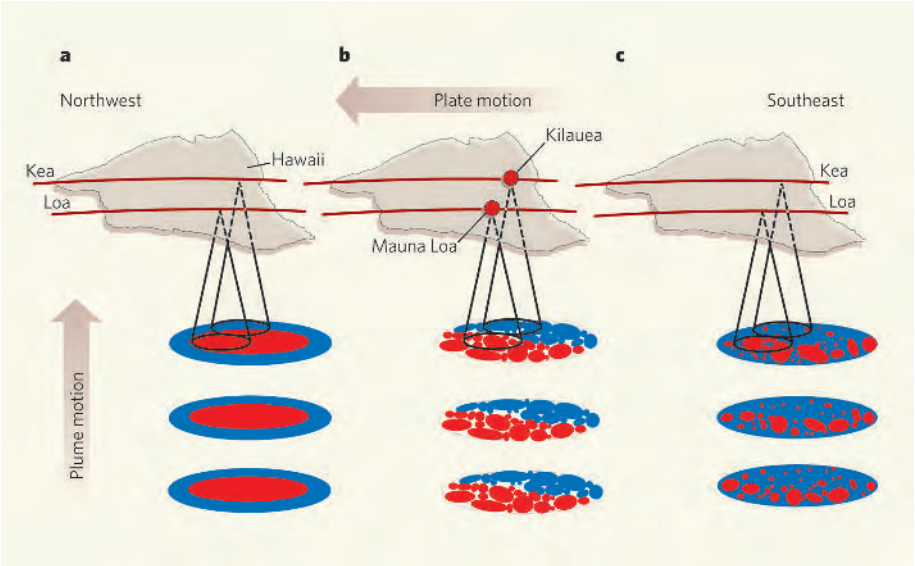


Figure 1 | Models of the geochemical structure of the Hawaiian mantle plume. The diagram depicts three possible plume structures below the two alignments (Kea and Loa) of Hawaiian volcanoes, which are surface expressions of the tectonic plate on which they sit. **a**, A concentrically zoned and vertically continuous plume[6]. **b**, A bilateral bundle of filaments vertically continuous on the 50–500-kilometre scale[7], with undefined spaces between filaments. Mauna Loa and Kilauea are two volcanoes on the Loa and Kea trends, respectively. **c**, A partly ordered structure, with streaks of red in a matrix of blue (modified from ref. 8). Red, average geochemical properties of volcanoes along the Loa track[3–8]; blue, average geochemical properties of volcanoes along the Kea track[3–8]. Ren *et al.*[1] favour the partly ordered structure (**c**), which they believe is defined by streaks of pyroxenite (red, pyroxene-rich rocks) in a heterogeneous peridotite matrix (blue, olivine-rich rocks).

debate: they show that interpretations depend on the size of the material that is subject to chemical analysis.

The Hawaiian plume is roughly 100 kilometres in diameter, and rises from the lower mantle below a depth of 660 kilometres[2]. It begins to partly melt in the final stages of its upward journey, and lava erupts at the surface to make the Hawaiian volcanoes. These are heterogeneous in that each volcano has its own geochemical identity, which can vary with time[3–5]. However, lavas from Kilauea, currently the most active volcano, have a geochemistry similar to that of older lavas that erupted along the 'Kea' alignment of volcanoes[4,5], which include Mauna Kea, Kohala and Haleakala to the northwest. The geochemistry of Kea volcanoes differs from that of the 'Loa' volcanoes[4,5], which include Loihi, Mauna Loa, Kahoolawe and Koolau, a parallel alignment displaced to the west (see Fig. 1 on page 837).

These differences, expressed in radiogenic isotopes, and trace and major elements, have been interpreted to reflect the three-dimensional spatial organization of the chemical constituents in the plume before melting took place. This is referred to as the geochemical structure of the Hawaiian plume[1,6–8]. But there is no consensus about the form of this structure; different models are summarized in Figure 1.

One model has geochemical heterogeneities ordered in a concentric and vertically continuous structure[6] (Fig. 1a); Loa volcanoes sample the centre, Kea volcanoes sample the periphery. Another model has geochemical heterogeneities ordered in filaments clustered together like spaghetti[7] (Fig. 1b), which is consistent with the expectation of stretching and shearing of heterogeneities in a plume[9]; Loa and Kea volcanoes sample a bilateral distribution of filaments[7]. A third model has a more random but partly ordered distribution of geochemical heterogeneities both vertically and laterally[8] (Fig. 1c). The origin of these heterogeneities is a separate issue. However, there is a consensus that it has something to do with near-surface magmatism, crust production and sediment accumulation in the early Earth, subduction of this outer unit into the deep mantle, and an upward return in a hot plume that melts to make Hawaii.

All three models are interpretations of the geochemistry of whole-rock samples — 'hand specimens' — of centimetre scale. The approach taken by Ren and co-workers[1] differs: using microanalytical techniques, they have acquired geochemical data on melt inclusions, about 10–100 micrometres in size, contained in millimetre-sized crystals of olivine, a common mineral in some volcanic rocks. They show that major-element and trace-element compositions of olivine-hosted inclusions from a single rock extracted from a single volcano can have the properties of both Loa and Kea volcanoes. They suggest that either the Loa or the Kea component is represented by whole rocks, but both are represented in olivine-hosted inclusions regardless of the specific geographical location of the volcano. This study supports the more random model[8], although some hybrid of the ordered models (Fig. 1a, b) is still required to yield the Loa and Kea trends (Fig. 1c).

How is it possible that whole-rock samples from an individual volcano support either a Loa or a Kea source, but the inclusions support both sources? Why does the scale of the geochemical observation yield different interpretations? The answers lie in an understanding of how rock melts. The Hawaiian plume melts by forming millimetre-sized liquid droplets that inherit the geochemical properties of their source rock. Each drop of melt usually mixes with other drops during transport to the surface. Although it is not clear exactly where mixing takes place, it is known to homogenize the geochemistry on the centimetre scale of the hand specimen and the kilometre scale of an individual volcano. A rock from Kilauea can be thought of as a blend of one melt drop from a Loa source and 100 melt drops from a Kea source. However, even a well-mixed rock can contain olivines that crystallized from melt droplets before mixing took place. In this way, crystal growth can entomb the drops as tiny inclusions before they become part of the blend.

The large geochemical heterogeneities reported by Ren and co-workers for Hawaii are similar to those of inclusion studies for lavas from Iceland[10]. However, the Iceland case has been interpreted somewhat differently, as being the result of the continuous removal of small melt fractions from a single source composition[10], a process called fractional melting. The authors of that study[10] acknowledge that some of their data might also be explained by variable source compositions, as do Ren *et al.* for Hawaii. And it is likely that some of the geochemical variability reported by Ren *et al.* can be explained by fractional melting.

Further studies are evidently called for. What is becoming clear is that complementary inclusion and whole-rock geochemical studies expand the scale of observation in a way that is comparable to viewing an object with variable magnification. ■

Claude Herzberg is in the Department of Geological Sciences, Rutgers, State University of New Jersey, 610 Taylor Road, Piscataway, New Jersey 08854, USA.
e-mail: herzberg@rci.rutgers.edu

1. Ren, Z.-Y., Ingle, S., Takahashi, E., Hirano, N. & Hirata, T. *Nature* **436,** 837–840 (2005).
2. Morgan, W. J. *Nature* **230,** 42–43 (1971).
3. Frey, F. A. & Rhodes, J. M. *Phil. Trans. R. Soc. Lond. A* **342,** 121–136 (1993).
4. Lassiter, J. C., DePaolo, D. J. & Tatsumoto, M. *J. Geophys. Res.* **101,** 11769–11780 (1996).
5. Hauri, E. H. *Nature* **382,** 415–419 (1996).
6. DePaolo, D. J., Bryce, J. G., Dodson, A., Shuster, D. L. & Kennedy, B. M. *Geochem. Geophys. Geosyst.* **2,** doi:10.1029/2000GC000139 (2001).
7. Abouchami, W. *et al. Nature* **434,** 851–856 (2005).
8. Blichert-Toft, J., Weis, D., Maerschalk, C., Agranier, A. & Albarède, F. *Geochem. Geophys. Geosyst.* **4,** doi:10.1029/2002GC000340 (2003).
9. Farnetani, C. & Samuel, H. *Geophys. Res. Lett.* **32,** doi:10.1029/2005GL022360 (2005).
10. Maclennan, J. *et al. Geochem. Geophys. Geosyst.* **4,** doi:10.1029/2003C000558 (2003).

BRIEF COMMUNICATIONS

Structure of the blue cornflower pigment

Packaging red-rose anthocyanin as part of a 'superpigment' in another flower turns it brilliant blue.

The same anthocyanin pigment makes roses red but cornflowers blue[1], a phenomenon that has so far not been entirely explained. Here we describe the X-ray crystal structure of the cornflower pigment, which reveals that its blue colour arises from a complex of six molecules each of anthocyanin and flavone, with one ferric iron, one magnesium and two calcium ions. We believe that this tetrametal complex may represent a previously undiscovered type of supermolecular pigment.

An anthocyanin pigment was first isolated in 1913 as a red oxonium salt from the blue cornflower *Centaurea cyanus*[2], and later the same pigment was discovered in a red rose[3]. The colour variation was variously ascribed to a difference in flower-petal pH (ref. 3), an association of the pigment with metal ions[4] or with another pigment[5]. However, the blue-cornflower pigment (later named protocyanin[6,7]) was found to contain iron and magnesium[8,9] in complex with anthocyanin and flavone[10,11]; calcium is also essential for the reconstruction of protocyanin[12].

We determined the crystal structure of reconstructed protocyanin at 1.05 Å resolution. (For methods, see supplementary information.) In the refined molecule, the four metal ions Fe^{3+}, Mg^{2+} and two Ca^{2+} align along the pseudo-three-fold axis (Fig. 1a) and are coordinated to six anthocyanin (cyanidin 3-*O*-(6-*O*-succinyl glucoside)-5-*O*-glucoside) and six flavone (apigenin 7-*O*-glucuronide-4′-*O*-(6-*O*-malonyl-glucoside)) molecules. The inner Fe^{3+} and Mg^{2+} ions are each coordinated to three anthocyanin molecules, respectively, and the two outer Ca^{2+} ions are each coordinated to three flavone molecules (see supplementary information).

The two sites of the central nuclei have the same electron density, which corresponds to the average for Fe^{3+} and Mg^{2+} (results not shown). This suggests that the sites are occupied by either Fe^{3+} or Mg^{2+} because of the random orientation of the molecule along the pseudo-three-fold axis. To verify this, we investigated the X-ray structures of metal-substituted FeMgBa-, FeMnBa- and FeCdBa-protocyanins (results not shown). (Barium derivatives are more stable than calcium derivatives for isolation as crystals.)

In the FeMgBa-protocyanin, the electron density of the inner nuclei is almost the same as in protocyanin, whereas in the FeMnBa- and FeCdBa-protocyanins it is the average of Fe^{3+} and of the substituted ions, respectively. The distances between the two metals and the coordinating oxygen atoms vary according to the radii of the substituted metals. These results indicate that the two metals are heterogeneous. We conclude that the inner nuclei of protocyanin consist of Fe^{3+} and Mg^{2+} ions.

In the protocyanin molecule, the anthocyanins and flavones self-associate in pairs. The Fe^{3+} and Mg^{2+} ions are coordinated to different anthocyanin fragments in an anthocyanin pair (Fig. 1b, left); the two Ca^{2+} ions are coordinated with separate flavone fragments in a flavone pair (Fig. 1b, centre). Stacking of anthocyanin and flavone (co-pigmentation) is also evident (Fig. 1b, right). The C–C and C–O bond lengths in the B-ring of anthocyanin (data not shown) indicate that the anthocyanin is in the 4′-keto-quinoidal form.

Blue flower colours are caused mainly by delphinidin-type anthocyanins — for example, commelinin[13]. In protocyanin, however, the blue colour is produced by a cyanidin-type anthocyanin. The chelation of Fe^{3+} and Mg^{2+} with the 4′-keto-quinoidal base of anthocyanin is important for the blueing in protocyanin. The two Ca^{2+} ions coordinate with the flavones to form components that bring about co-pigmentation as well as stabilization of the molecule. The blue colour in protocyanin is therefore developed by a tetranuclear metal complex, which may represent a new type of supermolecular pigment.

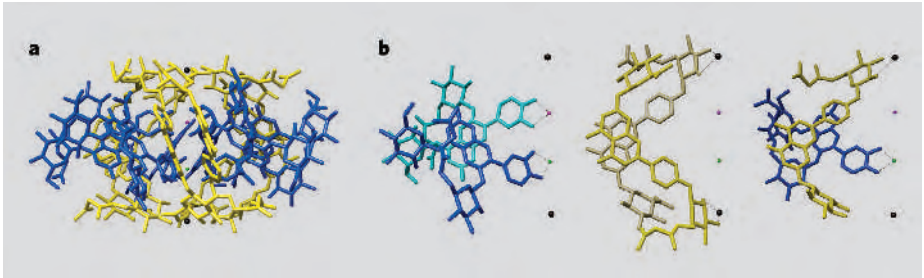


Figure 1 | Crystal structure of the protocyanin molecule. **a,** Side view along the pseudo-three-fold axis. Blue, anthocyanin; yellow, flavone glycoside; spheres: red, Fe^{3+}; green, Mg^{2+}; black, Ca^{2+}. For stereo view, see supplementary information. **b,** Side views of stacking pigment pairs. Left, one anthocyanin molecule binds Fe^{3+} (red), the other binds Mg^{2+} (green); centre, two flavones each bind to one Ca^{2+} (black); right, flavone and anthocyanin molecules bind to Ca^{2+} and Mg^{2+}, respectively.

Masaaki Shiono*, Naohiro Matsugaki†, Kosaku Takeda‡

*Department of Physics, Graduate School of Science, Kyushu University, Hakozaki 6-10-1, Fukuoka 812-8581, Japan
†Photon Factory, High Energy Accelerator Research Organization, Oho 1-1, Tsukuba 305-0801, Japan
‡Department of Biology, Tokyo Gakugei University, Koganei, Tokyo 184-8501, Japan
e-mail: takedks@ss.iij4u.or.jp

1. Brouillard, R. & Dangles, O. in *The Flavonoids, Advances in Research since 1986* (ed. Harborne, J. B.) 565–588 (Chapman and Hall, London, 1994).
2. Willstätter, R. & Everest, R. W. *Justus Liebigs Ann. Chem.* **401,** 189–232 (1913).
3. Willstätter, R. & Mallison, H. *Justus Liebigs Ann. Chem.* **408,** 147–162 (1915).
4. Shibata, K., Shibata, Y. & Kashiwagi, I. *J. Am. Chem. Soc.* **41,** 208–220 (1919).
5. Robinson, G. M. & Robinson, R. *Biochem. J.* **25,** 1687–1705 (1931).
6. Bayer, E. *Chem. Ber.* **91,** 1115–1122 (1958).
7. Bayer, E., Egeter, H., Fink, A., Nether, K. & Wegmann, K. *Angew. Chem.* **78,** 834–841 (1966).
8. Hayashi, K., Saito, N. & Mitsui, S. *Proc. Jpn Acad.* **37,** 393–397 (1961).
9. Saito, N., Mitsui, S. & Hayashi, K. *Proc. Jpn Acad.* **37,** 485–490 (1961).
10. Kondo, T. *et al. Angew. Chem. Int. Ed. Engl.* **33,** 978–979 (1994).
11. Kondo, T., Ueda, M., Isobe, M. & Goto, T. *Tetrahedron Lett.* **39,** 8307–8310 (1998).
12. Takeda, K. *et al. Phytochemistry* **66,** 1607–1613 (2005).
13. Kondo, T. *et al. Nature* **358,** 515–518 (1992).

Supplementary information accompanies this communication on *Nature*'s website.
Competing financial interests: declared none.
doi:10.1038/436791a

COLORECTAL CANCER

Mutations in a signalling pathway

Protein kinases are enzymes that are important for controlling cellular growth and invasion[1], and their malfunction is implicated in the development of some tumours. We analysed human colorectal cancers for genetic mutations in 340 serine/threonine kinases and found mutations in eight genes, including in three members of the phosphatidylinositol-3-OH kinase (PI(3)K) pathway. The discovery of this mutational activation of a key cell-signalling pathway may provide new targets for therapeutic intervention.

Although genetic alterations in tyrosine kinases have previously been firmly implicated in tumorigenesis, only a few serine/threonine kinases (STKs) are known to be mutated in human cancers[1–4]. We selected 340 genes encoding STKs from the human genome[5] and analysed them for mutations in tumours from colorectal cancer patients (for details, see supplementary information). As the catalytic domains of these genes are most likely to harbour mutations that activate the gene product[1], we focused on stretches (exons) containing the kinase domains.

These exons were amplified by using polymerase chain reaction (PCR) on template DNA derived from 24 colorectal cancers and were then sequenced directly. Any observed changes were evaluated against DNA from patient-matched normal tissue to identify somatic (tumour-specific) mutations. The entire coding regions of those genes found to contain mutations were then further evaluated in a larger panel of 180 colorectal tumours. A total of 23 changes, including 20 non-synonymous point mutations, one insertion and one splice-site alteration, were identified (see supplementary information).

The gene mutations affected eight different proteins: six were in mitogen-activated protein-kinase kinase-4 (MKK4/JNKK1), six in myosin light-chain kinase-2 (MYLK2), three in phosphoinositide-dependent protein kinase-1 (PDK1, of which two mutations affect the same residue in the kinase domain), two in p21-activated kinase 4 (PAK4), two in v-*akt* murine thymoma viral oncogene homologue-2 kinase (AKT2), and two in MAP/microtubule affinity-regulating kinase-3 (MARK3); there was one alteration in cell-division cycle-7 kinase (CDC7) and another in a hypothetical casein kinase (PDIK1L). Eighteen of the 23 somatic mutations occurred at evolutionarily conserved residues.

MKK4/JNKK1 is altered in a variety of tumour types[6], but no mutations in any of the other genes have previously been found in colorectal cancers. Three of the altered genes, *PDK1*, *AKT2* and *PAK4*, encode proteins involved in the PI(3)K signalling pathway[7,8] (Fig. 1), and two of these (*AKT2* and *PAK4*) are overexpressed in human cancers[1].

We tested whether any of the three kinases could have been altered by amplification, another mechanism for kinase activation. Quantitative PCR analysis of 146 colorectal tumours showed co-amplification of *AKT2* and *PAK4* on chromosome 19q13.2 in two samples, which we confirmed by digital karyotyping[9] and fluorescent *in situ* hybridization (see supplementary information).

We also evaluated other non-STK members of the PI(3)K pathway in the same 146 samples (Fig. 1, and see supplementary information)[7] and found one mutation in the insulin-related receptor INSRR, one in the v-Erb-B erythroblastic leukaemia viral oncogene homologue ERBB4, seven in the phosphatase-and-tensin homologue PTEN, and three cases of amplification of the insulin-receptor substrate IRS2 (see supplementary information). When these alterations are compared with those previously discovered in the phosphoinositide-3-kinase p110α catalytic subunit PIK3CA (ref. 10), their distribution is striking: all but two of the 58 alterations were in different tumours ($P = 0.02$, χ^2 test), indicating that the mutated genes probably have equivalent tumorigenic effects and are operating through the same pathway.

Overall, nearly 40% of colorectal tumours had alterations in one of eight PI(3)K-pathway genes: as most of these encode protein kinases, they could serve as sites for therapeutic intervention. Also, targeting the downstream genes *PDK1* or *AKT2* could be effective against the much larger fraction of tumours that contain mutations in PIK3CA or PTEN.

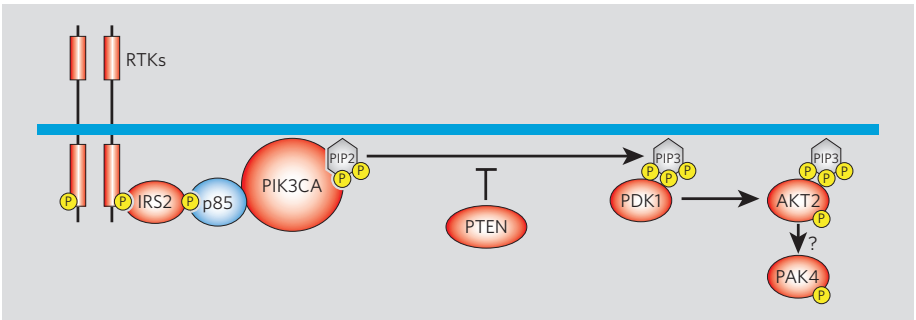


Figure 1 | The phosphatidylinositol-3-OH kinase (PI(3)K) signalling pathway and mutations in its components found in colorectal cancers. In the pathway (reviewed in ref. 7), receptor tyrosine kinases (RTKs) recruit IRS adaptor proteins that induce proper assembly of the p85/PIK3CA complex; the PIK3CA enzyme then phosphorylates phosphatidylinositol-4,5- bisphosphate (PIP2) to phosphatidylinositol-3,4,5-trisphosphate (PIP3). (The enzyme PTEN normally reverses this process under appropriate circumstances.) PDK1 is then recruited to the cell surface by PIP3 and phosphorylates and activates AKT2; the activation of PAK4, a downstream step in the pathway, is dependent on PI(3)K signalling, presumably through AKT2. Members of the pathway highlighted in red were found to be genetically altered in the colorectal cancers examined here. Phosphate groups are indicated in yellow. Intermediates: IRS2, insulin-receptor substrate-2; PIK3CA, the phosphoinositide-3-kinase p110α catalytic subunit; PTEN, phosphatase-and-tensin homologue; PDK1, phosphoinositide-dependent protein kinase-1; AKT2, v-*akt* murine thymoma viral oncogene homologue-2 kinase; PAK4, p21-activated kinase 4.

D. Williams Parsons*, Tian-Li Wang*, Yardena Samuels*, Alberto Bardelli*†, Jordan M. Cummins*, Laura DeLong*, Natalie Silliman*, Janine Ptak*, Steve Szabo*, James K. V. Willson‡, Sanford Markowitz§, Kenneth W. Kinzler*, Bert Vogelstein*, Christoph Lengauer*, Victor E. Velculescu*

*The Sidney Kimmel Comprehensive Cancer Center and The Howard Hughes Medical Institute, The Johns Hopkins University Medical Institutions, Baltimore, Maryland 21231, USA
e-mail: velculescu@jhmi.edu
‡Harold C. Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA
§Howard Hughes Medical Institute and Ireland Cancer Center, University Hospitals of Cleveland and Case Western University, Cleveland, Ohio 44106, USA
†Present address: Institute for Cancer Research and Treatment, University of Torino Medical School, 10060 Candiolo, and FIRC Institute of Molecular Oncology, 20139 Milan, Italy

1. Blume-Jensen, P. & Hunter, T. *Nature* **411,** 355–365 (2001).
2. Bardelli, A. *et al. Science* **300,** 949 (2003).
3. Davies, H. *et al. Nature* **417,** 949–954 (2002).
4. Futreal, P. A. *et al. Nature Rev. Cancer* **4,** 177–183 (2004).
5. Manning, G., Whyte, D. B., Martinez, R., Hunter, T. & Sudarsanam, S. *Science* **298,** 1912–1934 (2002).
6. Teng, D. H. *et al. Cancer Res.* **57,** 4177–4182 (1997).
7. Vivanco, I. & Sawyers, C. L. *Nature Rev. Cancer* **2,** 489–501 (2002).
8. Wells, C. M., Abo, A. & Ridley, A. J. *J. Cell Sci.* **115,** 3947–3956 (2002).
9. Wang, T. L. *et al. Proc. Natl Acad. Sci. USA* **99,** 16156–16161 (2002).
10. Samuels, Y. *et al. Science* **304,** 554 (2004).

Supplementary information accompanies this communication on *Nature*'s website.
Competing financial interests: declared online.
doi:10.1038/436792a

ARTICLES

The map-based sequence of the rice genome

International Rice Genome Sequencing Project*

Rice, one of the world's most important food plants, has important syntenic relationships with the other cereal species and is a model plant for the grasses. Here we present a map-based, finished quality sequence that covers 95% of the 389 Mb genome, including virtually all of the euchromatin and two complete centromeres. A total of 37,544 non-transposable-element-related protein-coding genes were identified, of which 71% had a putative homologue in *Arabidopsis*. In a reciprocal analysis, 90% of the *Arabidopsis* proteins had a putative homologue in the predicted rice proteome. Twenty-nine per cent of the 37,544 predicted genes appear in clustered gene families. The number and classes of transposable elements found in the rice genome are consistent with the expansion of syntenic regions in the maize and sorghum genomes. We find evidence for widespread and recurrent gene transfer from the organelles to the nuclear chromosomes. The map-based sequence has proven useful for the identification of genes underlying agronomic traits. The additional single-nucleotide polymorphisms and simple sequence repeats identified in our study should accelerate improvements in rice production.

Rice (*Oryza sativa* L.) is the most important food crop in the world and feeds over half of the global population. As the first step in a systematic and complete functional characterization of the rice genome, the International Rice Genome Sequencing Project (IRGSP) has generated and analysed a highly accurate finished sequence of the rice genome that is anchored to the genetic map. Our analysis has revealed several salient features of the rice genome:

- We provide evidence for a genome size of 389 Mb. This size estimation is ~260 Mb larger than the fully sequenced dicot plant model *Arabidopsis thaliana*. We generated 370 Mb of finished sequence, representing 95% coverage of the genome and virtually all of the euchromatic regions.
- A total of 37,544 non-transposable-element-related protein-coding sequences were detected, compared with ~28,000–29,000 in *Arabidopsis*, with a lower gene density of one gene per 9.9 kb in rice. A total of 2,859 genes seem to be unique to rice and the other cereals, some of which might differentiate monocot and dicot lineages.
- Gene knockouts are useful tools for determining gene function and relating genes to phenotypes. We identified 11,487 *Tos17* retrotransposon insertion sites, of which 3,243 are in genes.
- Between 0.38 and 0.43% of the nuclear genome contains organellar DNA fragments, representing repeated and ongoing transfer of organellar DNA to the nuclear genome.
- The transposon content of rice is at least 35% and is populated by representatives from all known transposon superfamilies.
- We have identified 80,127 polymorphic sites that distinguish between two cultivated rice subspecies, *japonica* and *indica*, resulting in a high-resolution genetic map for rice. Single-nucleotide polymorphism (SNP) frequency varies from 0.53 to 0.78%, which is 20 times the frequency observed between the Columbia and Landsberg *erecta* ecotypes of *Arabidopsis*.
- A comparison between the IRGSP genome sequence and the 6.3× *indica* and 6× *japonica* whole-genome shotgun sequence assemblies revealed that the draft sequences provided coverage of 69% by *indica* and 78% by *japonica* relative to the map-based sequence.

Rice has played a central role in human nutrition and culture for the past 10,000 years. It has been estimated that world rice production must increase by 30% over the next 20 years to meet projected demands from population increase and economic development[1]. Rice grown on the most productive irrigated land has achieved nearly maximum production with current strains[1]. Environmental degradation, including pollution, increase in night time temperature due to global warming[2], reductions in suitable arable land, water, labour and energy-dependent fertilizer provide additional constraints. These factors make steps to maximize rice productivity particularly important. Increasing yield potential and yield stability will come from a combination of biotechnology and improved conventional breeding. Both will be dependent on a high-quality rice genome sequence.

Rice benefits from having the smallest genome of the major cereals, dense genetic maps and relative ease of genetic transformation[3]. The discovery of extensive genome colinearity among the Poaceae[4] has established rice as the model organism for the cereal grasses. These properties, along with the finished sequence and other tools under development, set the stage for a complete functional characterization of the rice genome.

The International Rice Genome Sequencing Project

The IRGSP, formally established in 1998, pooled the resources of sequencing groups in ten nations to obtain a complete finished quality sequence of the rice genome (*Oryza sativa* L. ssp. *japonica* cv. Nipponbare). Finished quality sequence is defined as containing less than one error in 10,000 nucleotides, having resolved ambiguities, and having made all state-of-the-art attempts to close gaps. The IRGSP released a high-quality map-based draft sequence in

*Lists of participants and affiliations appear at the end of the paper

December 2002. Three completely sequenced chromosomes have been published[5–7], as well as two completely sequenced centromeres[8–10]. As the IRGSP subscribed to an immediate-release policy, high-quality map-based sequence has been public for some time. This has permitted rice geneticists to identify several genes underlying traits, and revealed very large and previously unknown segmental duplications that comprise 60% of the genome[11–13]. The public sequence has also revealed new details about the syntenic relationships and gene mobility between rice, maize and sorghum[13–15].

Physical maps, sequencing and coverage

The IRGSP sequenced the genome of a single inbred cultivar, *Oryza sativa* ssp. *japonica* cv. Nipponbare, and adopted a hierarchical clone-by-clone method using bacterial and P1 artificial chromosome clones (BACs and PACs, respectively). This strategy used a high-density genetic map[16], expressed-sequence tags (ESTs)[17], yeast artificial chromosome (YAC)- and BAC-based physical maps[18–20], BAC-end sequences[21] and two draft sequences[22,23]. A total of 3,401 BAC/PAC clones (Table 1) were sequenced to approximately tenfold sequence coverage, assembled, ordered and finished to a sequence quality of less than one error per 10,000 bases. A majority of physical gaps in the BAC/PAC tiling path were bridged using a variety of substrates, including PCR fragments, 10-kb plasmids and 40-kb fosmid clones. A total of 62 unsequenced physical gaps, including nine centromere and 17 telomere gaps, remain on the 12 chromosomes (Table 2). Chromosome arm and telomere gaps were measured, and the nine centromere gaps were estimated on the basis of CentO satellite DNA content. The remaining gaps are estimated to total 18.1 Mb.

Ninety-seven percent of the BAC/PACs and gap sequences (3,360) have been submitted as finished quality in the PLN division of GenBank/DDBJ/EMBL. These and the remaining draft-sequenced clones were used to construct pseudomolecules representing the 12 chromosomes of rice (Fig. 1). The total nucleotide sequence of the 12 pseudomolecules is 370,733,456 bp, with an N-average continuous sequence length of 6.9 Mb (see Table 1 for a definition of N-average length). Sequence quality was assessed by comparing 1.2 Mb of overlapping sequence produced by different laboratories. The overall accuracy was calculated as 99.99% (Supplementary Table 2). The statistics of sequenced PAC/BAC clones and pseudomolecules for each chromosome are shown in Table 1.

The genome size of rice (*O. sativa* ssp. *japonica* cv. Nipponbare) was reported to have a haploid nuclear DNA content of 394 Mb on the basis of flow cytometry[24], and 403 Mb on the basis of lengths of anchored BAC contigs and estimates of gap sizes[20]. Table 2 shows the calculated size for each chromosome and the estimated coverage. Adding the estimated length of the gaps to the sum of the non-overlapping sequence, the total length of the rice nuclear genome was calculated to be 388.8 Mb. Therefore, the pseudomolecules are expected to cover 95.3% of the entire genome and an estimated 98.9% of the euchromatin. An independent measure of genome coverage represented by the pseudomolecules was obtained by searching for unique EST markers[19]; of 8,440 ESTs, 8,391 (99.4%) were identified in the pseudomolecules.

Centromere location

Typical eukaryotic centromeres contain repetitive sequences, including satellite DNA at the centre and retrotransposons and transposons in the flanking regions. All rice centromeres contain the highly repetitive 155–165 bp CentO satellite DNA, together with centromere-specific retrotransposons[25,26]. The CentO satellites are located within the functional domain of the rice centromere[10,26]. Complete sequencing of the centromeres of rice chromosomes 4 and 8 revealed that they consist of 59 kb and 69 kb of clustered CentO repeats (respectively)[8–10], tandemly arrayed head-to-tail within the clusters. Numerous retrotransposons, including the centromere-specific RIRE7, are found between and around the CentO repeats. CentO clusters show differences in length and orientation for the two centromeres.

BLASTN analysis of the pseudomolecules indicated that about 0.9 Mb of CentO repeats (corresponding to more than 5,800 copies of the satellite) were sequenced and found to be associated with centromere-specific retroelements. Locations of all CentO sequences correspond to genetically identified centromere regions (Supplementary Table 3). Our pseudomolecules cover the centromere regions on chromosomes 4, 5 and 8, and portions of the centromeres on the remaining chromosomes (Fig. 1).

Gene content, expression and distribution

We masked the pseudomolecules for repetitive sequences and used the *ab initio* gene finder FGENESH to identify only non-transposable-element-related genes. A total of 37,544 non-transposable-element protein-coding sequences were predicted, resulting in a density of one gene per 9.9 kb (Supplementary Tables 4 and 5). As the ability to identify unannotated and transposable-element-related genes improves, the true protein-coding gene number in rice will doubtless be revised.

Full-length complementary DNA sequences are available for rice[27], and provide a powerful resource for improving gene model structure derived from *ab initio* gene finders[28]. Of the 37,544 non-transposable-element-related FGENESH models, 17,016 could be supported by a total of 25,636 full-length cDNAs (Supplementary Table 6).

A total of 22,840 (61%) genes had a high identity match with a rice EST or full-length cDNA. On average, about 10.7 EST sequences were present for each expressed rice gene. A total of 2,927 genes aligned well with ESTs from other cereal species, and 330 of these genes matched only with a non-rice cereal EST (Supplementary Fig. 1). Except for the short arms of chromosomes 4, 9 and 10, which are known to be highly heterochromatic, the density of expressed genes is greater on the distal portions of the chromosome arms compared with the regions around the centromeres (Supplementary Fig. 2).

A total of 19,675 proteins had matches with entries in the Swiss-Prot database; of these, 4,500 had no expression support. Domain searches revealed a minimum of one motif or domain present in 63% of the predicted proteins, with a total of 3,328 different domains present in the predicted rice proteome. The five most abundant domains were associated with protein kinases (Supplementary Table 7). Fifty-one per cent of the predicted proteins could be associated with a biological process (Supplementary Fig. 3a), with metabolism (29.1%) and cellular physiological processes (11.9%) representing the two most abundant classes.

Approximately 71% (26,837) of the predicted rice proteins have a homologue in the *Arabidopsis* proteome (Supplementary Fig. 4). In a reciprocal search, 89.8% (26,004) of the proteins from the *Arabidopsis* genome have a homologue in the rice proteome. Of the 23,170 rice genes with rice EST, cereal EST, or full-length cDNA support, 20,311 (88%) have a homologue in *Arabidopsis*. Fewer putative homologues were found in other model species: 38.1% in *Drosophila*, 40.8% in human, 36.5% in *Caenorhabditis elegans*, 30.2% in yeast, 17.6% in *Synechocystis* and 10.2% in *Escherichia coli*.

There are profound differences in plant architecture and biochemistry between monocotyledonous and dicotyledonous angiosperms. Only 2,859 rice genes with evidence of transcription lack homologues in the *Arabidopsis* genome. We investigated these to learn what functions they encoded. The vast majority had no matches, or most closely matched unknown or hypothetical proteins. The grasses have a class of seed storage proteins called prolamins that is not found in dicots. There are also families of hormone response proteins and defence proteins, such as proteinase inhibitors, chitinases, pathogenesis-related proteins and seed allergens, many of which are tandemly repeated (Supplementary Table 8). Nevertheless, with a large number of proteins of unknown function, the most interesting

differences between the genome content of these two groups of angiosperms remain to be discovered.

Tos17 is an endogenous *copia*-like retrotransposon in rice that is inactive under normal growth conditions. In tissue culture, it becomes activated, transposes and is stably inherited when the plant is regenerated[29]. There are only two copies of *Tos17* in the rice cultivar Nipponbare. These features, together with its preferential insertion into gene-rich regions, make *Tos17* uniquely suitable for the functional analysis of rice genes by gene disruption. About 50,000 *Tos17*-insertion lines carrying 500,000 insertions have been produced[30]. A total of 11,487 target loci were mapped on the 12 pseudomolecules (Supplementary Fig. 5), with at least one insertion detected in 3,243 genes. The density of *Tos17* insertions is higher in euchromatic regions of the genome[30], in contrast to the distribution of high-copy retrotransposons, which are more frequently found in pericentromeric regions. A similar target site preference has been reported for T-DNA insertions in *Arabidopsis*[31].

Tandem gene families

One surprising outcome of the *Arabidopsis* genome analysis was the large percentage (17%) of genes arranged in tandem repeats[32]. When performing a similar analysis with rice, the percentage was comparable (14%). However, manual curation on rice chromosome 10 showed one gene family encoding a glycine-rich protein with 27 copies and one encoding a TRAF/BTB domain protein with 48 copies[33]. These tandemly repeated families are interrupted with other genes and are not included in strictly defined tandem repeats. We therefore screened for all tandemly arranged genes in 5-Mb intervals. Using these criteria, 29% of the genes (10,837) are amplified at least once in tandem, and 153 rice gene arrays contained 10–134 members (Supplementary Fig. 6). Sixty five per cent of the tandem arrays with over 27 members, and 33% of all the arrays with over 10 members, contain protein kinase domains (Supplementary Table 9).

Non-coding RNA genes

The nucleolar organizer, consisting of 17S–5.8S–25S ribosomal DNA coding units, is found at the telomeric end of the short arm of chromosome 9 (ref. 34) in *O. sativa* ssp. *japonica*, and is estimated to comprise 7 Mb (ref. 35). A second 17S–5.8S–25S rDNA locus is found at the end of the short arm of chromosome 10 in *O. sativa* ssp. *indica*[34]. A single 5S cluster is present on the short arm of chromosome 11 in the vicinity of the centromere[36], and encompasses 0.25 Mb.

A total of 763 transfer RNA genes, including 14 tRNA pseudogenes were detected in the 12 pseudomolecules. In comparison, a total of 611 tRNA genes were detected in *Arabidopsis*[32]. Supplementary Fig. 7 shows the distribution of these tRNA genes in each chromosome. Chromosome 4 has a single tRNA cluster[6], and chromosome 10 has two large clusters derived from inserted chloroplast DNA[7]. Except for regions of intermediate density on chromosomes 1, 2, 8 and 12, there seem to be no other large clusters.

MicroRNAs (miRNAs), a class of eukaryotic non-coding RNAs, are believed to regulate gene expression by interacting with the target messenger RNA[37]. miRNAs have been predicted from *Arabidopsis*[38] and rice[39], and we mapped 158 miRNAs onto the rice pseudomolecules (Supplementary Table 10). Among other non-coding RNAs, we identified 215 small nucleolar RNA (snoRNA) and 93 spliceosomal RNA genes, both showing biased chromosomal distributions, in the rice genome (Supplementary Table 11).

Organellar insertions in the nuclear genome

Mitochondria and chloroplasts originated from alpha-proteobacteria and cyanobacteria endosymbionts. A continuous transfer of organellar DNA to the nucleus has resulted in the presence of chloroplast and mitochondrial DNA inserted in the nuclear chromosomes. Although the endosymbionts probably contained genomes of several Mb at the time they were internalized, the organellar genomes diminished so that the present size of the mitochondrial genome is less than 600 kb, and that of the chloroplast is only 150 kb. Homology searches detected 421–453 chloroplast insertions and 909–1,191 mitochondrial insertions, depending upon the stringency adopted (Supplementary Fig. 8 and Supplementary Table 12). Thus, chloroplast and mitochondrial insertions contribute 0.20–0.24% and 0.18–0.19% of the nuclear genome of rice, respectively, and correspond to 5.3 chloroplast and 1.3 mitochondrial genome equivalents. The distribution of chloroplast and mitochondrial insertions over the 12 chromosomes indicates that mitochondrial and chloroplast transfers occurred independently. Two chromosomes harbour more insertions than the others (Supplementary Fig. 8 and Supplementary Table 12), with chromosome 12 containing nearly 1% mitochondrial DNA and chromosome 10 containing approximately 0.8% chlor-

Table 1 | Classification and distribution of sequenced PAC and BAC clones* on the 12 rice chromosomes

Chr	Sequencing laboratory†	PAC	BAC	OSJNBa/b	OJ	OSJNO	Others‡	Total§	Pseudomolecule (bp)	N-average length‖ (bp)	Accession no.
1	RGP, KRGRP	251	77	42	23	4	0	397	43,260,640	9,688,259	AP008207
2	RGP, JIC	117	16	80	142	4	0	359	35,954,074	7,793,366	AP008208
3	ACWW, TIGR	1	8	263	47	1	10	330	36,189,985	5,196,992	AP008209
4	NCGR	2	7	275	7	0	0	291	35,489,479	1,427,419	AP008210
5	ASPGC	67	11	113	87	0	0	278	29,733,216	3,086,418	AP008211
6	RGP	169	20	78	14	0	0	281	30,731,386	8,669,608	AP008212
7	RGP	102	19	68	97	0	0	286	29,643,843	14,923,781	AP008213
8	RGP	113	23	56	83	2	0	277	28,434,680	14,872,702	AP008214
9	RGP, KRGRP, BIOTEC, BRIGI	72	24	72	50	5	0	223	22,692,709	5,219,517	AP008215
10	ACWW, TIGR, PGIR	1	5	172	6	0	21	205	22,683,701	2,124,647	AP008216
11	ACWW, TIGR, IIRGS, PGIR, Genoscope	10	6	236	3	2	1	258	28,357,783	1,087,274	AP008217
12	Genoscope	2	6	179	79	0	2	268	27,561,960	7,600,514	AP008218
	Total	907	222	1634	638	18	34	3453	370,733,456	6,928,182	

Chr, chromosome.

*PAC, Rice Genome Research Program PAC; BAC, Rice Genome Research Program BAC; OSJNBa/b, Clemson University Genomics Institute BAC; OJ, Monsanto BAC; OSJNO, Arizona Genomics Institute fosmid (http://www.genome.arizona.edu/orders/direct.html?library = OSJNOa); Others, artificial gap-filling clones designated as OSJNA and OJA.

†ACWW (Arizona Genomics Institute, Cold Spring Harbor Laboratory, Washington University Genome Sequencing Center, University of Wisconcin) Rice Genome Sequencing Consortium; ASPGC, Academia Sinica Plant Genome Center; BIOTEC, National Center for Genetic Engineering and Biotechnology; BRIGI, Brazilian Rice Genome Initiative; IIRGS, Indian Initiative for Rice Genome Sequencing; JIC, John Innes Centre; KRGRP, Korea Rice Genome Research Program; NCGR, National Center for Gene Research; PGIR, Plant Genome Initiative at Rutgers; RGP, Rice Genome Research Program; TIGR, The Institute for Genomic Research.

‡Constructs derived by joining (mostly from the clone gap regions) sequence from PCR fragments, Monsanto or Syngenta sequences and the neighbouring clone sequences.

§A total of 2,494 BAC and 907 PAC clones were used for draft and finished sequencing. Monsanto draft-sequenced BACs underlie 638 finished clones. The Syngenta draft sequence contributed to the assemblies of 140 IRGSP clone sequences. Thirty-four sequence submissions are artificial constructs derived by joining a regional sequence (mostly from the clone gap regions) from PCR fragments, Monsanto or Syngenta sequences with the neighbouring clone sequences. This also includes 93 clones submitted as phase 1 or phase 2 to the HTG section of GenBank.

‖N-average length: the average length of a contiguous segment (without sequence or physical gaps) containing a randomly chosen nucleotide.

oplast DNA. It is clear that several successive transfer events have occurred, as insertions of less than 10 kb have heterogeneous identities. The longest insertions, however, systematically show >98.5% identity to organellar DNA (Supplementary Table 13), indicating recent insertions for both chloroplast and mitochondrial genomes.

Transposable elements

The rice genome is populated by representatives from all known transposon superfamilies, including elements that cannot be easily classified into either class I or II (ref. 40). Previous estimates of the transposon content in the rice genome range from 10 to 25% (refs 21, 40). However, the increased availability of transposon query sequences and the use of profile hidden Markov models allow the identification of more divergent elements[41] and indicate that the transposon content of the *O. sativa* ssp. *japonica* genome is at least 35% (Table 3). Chromosomes 8 and 12 have the highest transposon content (38.0% and 38.3%, respectively), and chromosomes 1 (31.0%), 2 (29.8%) and 3 (29.0%) have the lowest proportion of transposons. Conversely, elements belonging to the IS*5*/*Tourist* and IS*630*/Tc1/*mariner* superfamilies, which are generally correlated with gene density, are prevalent on the first three chromosomes and least frequent on chromosomes 4 and 12.

Class II elements, characterized by terminal inverted-repeats and including the *h*AT, CACTA, IS*256*/*Mutator*, IS*5*/*Tourist*, and IS*630*/Tc1/*mariner* superfamilies, outnumber class I elements, which include long terminal-repeat (LTR) retrotransposons (Ty1/*copia*, Ty3/*gypsy* and *TRIM*) and non-LTR retrotransposons (LINEs and SINEs, or long- and short-interspersed nucleotide elements, respectively), by more than twofold (Table 3). However, the nucleotide contribution of class I is greater than that of class II, due mostly to the large size of LTR retrotransposons and the small size of IS*5*/*Tourist* and IS*630*/Tc1/*mariner* elements. The inverse is the case for maize, for which class I elements outnumber class II elements[42]. Given their larger sizes, differential amplification of LTR elements in maize compared with rice is consistent with the genomic expansion found between orthologous regions of rice and maize[15,33].

Most class I elements are concentrated in gene-poor, heterochromatic regions such as the centromeric and pericentromeric regions (Supplementary Table 14). In contrast, members of some transposon superfamilies, including IS5/*Tourist*, IS*630*/Tc1/*mariner* and LINEs, have a significant positive correlation with both recombination rate and gene density. There is an effect of average element length associated with these patterns: short elements generally show a positive correlation with recombination rate and gene density, and are under-represented in the centromere regions, whereas larger elements have higher centromeric and pericentromeric abundance.

Intraspecific sequence polymorphism

Map-based cloning to identify genes that are associated with agronomic traits is dependent on having a high frequency of polymorphic markers to order recombination events. In rice, most of the segregating populations are generated from crosses between the two major subspecies of cultivated rice, *Oryza sativa* ssp. *japonica* and *O. sativa* ssp. *indica*. Although several studies on the polymorphisms detected between *japonica* and *indica* subspecies have been reported[6,43,44], the analysis reported here uses an approach that ensures comparison of orthologous sequences. *O. sativa* ssp. *indica* cv. Kasalath and *O. sativa* ssp. *japonica* cv. Nipponbare are the parents of the most densely mapped rice population[16]. BAC-end sequences were obtained from a Kasalath BAC library of 47,194 clones. Only high quality, single-copy sequences were mapped to the Nipponbare pseudomolecules, and only paired inverted sequences that mapped within 200 kb were considered. A total of 26,632 paired Kasalath BAC-end sequences were mapped to the 12 rice pseudomolecules (Supplementary Table 15). Kasalath BAC clones spanned 308 Mb or 79% of the Nipponbare genome. Sequence alignments with a PHRED quality value of 30 covered 12,319,100 bp (3%) of the total rice genome. A total of 80,127 sites differed in the corresponding regions in Nipponbare and Kasalath. The frequency of SNPs varied between chromosomes (0.53–0.78%). Insertions and deletions were also detected. The ratio of small insertion/deletion site nucleotides (1–14 bases) against the alignment length (0.20–0.27%) was similar among the different chromosomes, and there was no preference for the direction of insertions or deletions. The main patterns of base substitutions observed between Nipponbare and Kasalath are shown in Supplementary Table 16. Transitions (70%) were the most prominent substitutions; this is a substantially higher fraction than found between *Arabidopsis* ecotypes Columbia and Landsberg *erecta*[32].

Class 1 simple sequence repeats in the rice genome

Class 1 simple sequence repeats (SSRs) are perfect repeats >20 nucleotides in length[45] that behave as hypervariable loci, providing a rich source of markers for use in genetics and breeding. A total of 18,828 Class 1 di, tri and tetra-nucleotide SSRs, representing 47 distinctive motif families, were identified and annotated on the rice genome (Supplementary Fig. 9). Supplementary Table 17 provides information about the physical positions of all Class 1 SSRs in relation to widely used restriction-fragment length polymorphisms (RFLPs)[16,46] and previously published SSRs[45]. There was an average of 51 hypervariable SSRs per Mb, with the highest density of markers occurring on chromosome 3 (55.8 SSR Mb^{-1}) and the lowest occurring on chromosome 4 (41.0 SSR Mb^{-1}). A summary of information about the Class 1 SSRs identified in the rice pseudomolecules appears

Table 2 | Size of each chromosome based on sequence data and estimated gaps

Chr	Sequenced bases (bp)	Gaps on arm regions No.	Gaps on arm regions Length (Mb)	Telomeric gaps* (Mb)	Centromeric gap† (Mb)	rDNA‡ (Mb)	Total (Mb)	Coverage§ (%)	Coverage‖ (%)
1	43,260,640	5	0.33	0.06	1.40		45.05	99.1	96.0
2	35,954,074	3	0.10	0.01	0.72		36.78	99.7	97.7
3	36,189,985	4	0.96	0.04	0.18		37.37	97.3	96.8
4	35,489,479	3	0.46	0.20			36.15	98.7	98.2
5	29,733,216	6	0.22	0.05			30.00	99.3	99.1
6	30,731,386	1	0.02	0.03	0.82		31.60	99.8	97.2
7	29,643,843	1	0.31	0.01	0.32		30.28	98.9	97.9
8	28,434,680	1	0.09	0.05			28.57	99.7	99.5
9	22,692,709	4	0.13	0.14	0.62	6.95	30.53	98.8	74.3
10	22,683,701	4	0.68	0.13	0.47		23.96	96.6	94.7
11	28,357,783	4	0.21	0.04	1.90	0.25	30.76	99.1	92.2
12	27,561,960	0	0.00	0.05	0.16		27.77	99.8	99.2
All	370,733,456	36	3.51	0.81	6.59	7.20	388.82	98.9	95.3

*Estimated length including the telomeres, calculated with the average value of 3.2 kb for each chromosome[24].
†Estimated length of centromere-specific CentO repeats on each chromosome[26].
‡Represents the estimated length of the17S–5.8S–25S rDNA cluster on Chr 9 (ref. 35) and the 5S cluster on Chr 11 (ref. 24).
§Coverage of the pseudomolecules for the euchromatic regions in each chromosome.
‖Coverage of the pseudomolecules over the full length of each chromosome.

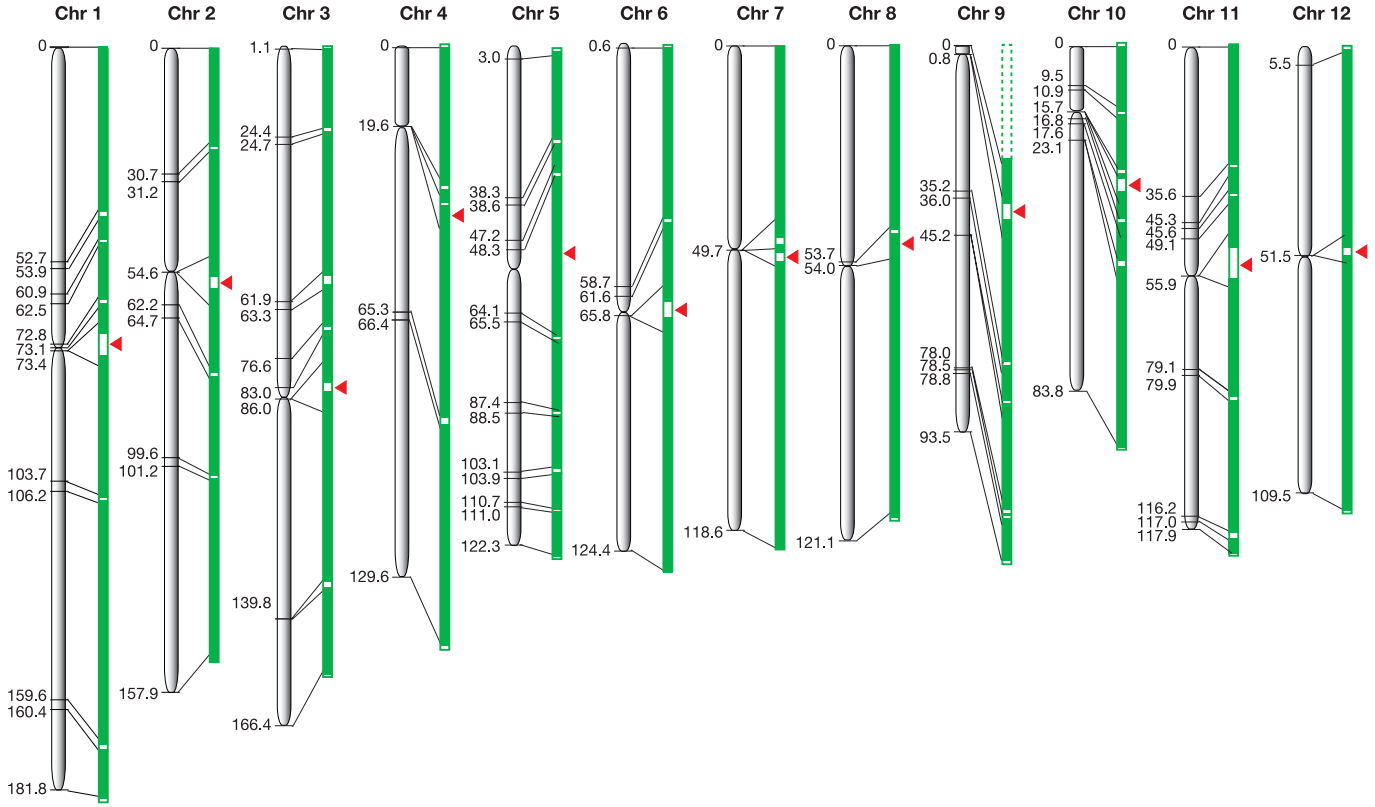


Figure 1 | Maps of the twelve rice chromosomes. For each chromosome (Chr 1–12), the genetic map is shown on the left and the PAC/BAC contigs on the right. The position of markers flanking the PAC/BAC contigs (green) is indicated on the genetic map. Physical gaps are shown in white and the nucleolar organizer on chromosome 9 is represented with a dotted green line. Constrictions in the genetic maps and arrowheads to the right of physical maps represent the chromosomal positions of centromeres for which rice CentO satellites are sequenced. The maps are scaled to genetic distances in centimorgans (cM) and the physical maps are depicted in relative physical lengths. Please refer to Table 2 for estimated lengths of the chromosomes.

in Supplementary Table 18. Several thousand of these SSRs have already been shown to amplify well and be polymorphic in a panel of diverse cultivars[45], and thus are of immediate use for genetic analysis.

Genome-wide comparison of draft versus finished sequences

Two whole-genome shotgun assemblies of draft-quality rice sequence have been published[23,47], and reassemblies of both have just appeared[48]. One of these is an assembly of 6.28 × coverage of *O. sativa* ssp. *indica* cv. 93-11. The second sequence is a ~6 × coverage of *O. sativa* ssp. *japonica* cv. Nipponbare[23,48]. These assemblies predict genome sizes of 433 Mb for *japonica* and 466 Mb for. *indica*, which differ from our estimation of a 389 Mb *japonica* genome. Contigs from the whole-genome shotgun assembly of 93-11 and Nipponbare[48] were aligned with the IRGSP pseudomolecules. Non-redundant coverage of the pseudomolecules by the *indica* assembly varied from 78% for chromosome 3 to 59% for chromosome 12, with an overall coverage of 69% (Supplementary Table 19). When genes supported by full-length cDNA coverage were aligned to the covered regions, we found that 68.3% were completely covered by the *indica* sequences. The average size of the *indica* contigs is 8.2 kb, so it is not surprising that many did not completely cover the gene models defined here. The coverage of the Nipponbare whole-genome shotgun assembly varied from 68–82%, with an overall coverage of 78% of the genome, and 75.3% of the full-length cDNAs supported gene models.

We undertook a detailed comparison of the first Mb of these assemblies on 1S (the short arm of chromosome 1) with the IRGSP chromosome 1 (Supplementary Fig. 10 and Supplementary Table 20). The numbers from this comparison agree with the whole-genome comparison described above. In addition, we observed that a substantial portion of the contigs from each assembly were non-homologous, misaligned or provided duplicate coverage. Indeed, the whole-genome shotgun assembly differed by 0.05% base-pair mismatches for the two aligned regions from the same Nipponbare cultivar. The two assemblies were further examined for the presence of the CentO sequence (Supplementary Table 21). Sixty-eight per cent of the copies observed in the 93-11 assembly and 32% of the CentO-containing contigs in the whole-genome shotgun Nipponbare assembly were found outside the centromeric regions. In contrast, the CentO repeats were restricted to the centromeric regions in the IRGSP pseudomolecules. It is unlikely that there are dispersed centromeres in *indica* rice; misassembly of the whole-genome shotgun sequences is a more likely explanation for dispersed CentO repeats. These observations indicate that the draft sequences, although providing a useful preliminary survey of the genome, might not be adequate for gene annotation, functional genomics or the identification of genes underlying agronomic traits.

Concluding remarks

The attainment of a complete and accurate map-based sequence for rice is compelling. We now have a blueprint for all of the rice chromosomes. We know, with a high level of confidence, the distribution and location of all the main components—the genes, repetitive sequences and centromeres. Substantial portions of the map-based sequence have been in public databases for some time, and the availability of provisional rice pseudomolecules based on this sequence has provided the scientific community with numerous opportunities to evaluate the genome, as indicated by the number of publications in rice biology and genetics over the past few years. Furthermore, the wealth of SNP and SSR information provided here

and elsewhere will accelerate marker-assisted breeding and positional cloning, facilitating advances in rice improvement.

The syntenic relationships between rice and the cereal grasses have long been recognized[4]. Comparing genome organization, genes and intergenic regions between cereal species will permit identification of regions that are highly conserved or rapidly evolving. Such regions are expected to yield crucial insights into genome evolution, speciation and domestication.

METHODS

Physical map and sequencing. Nine genomic libraries from *Oryza sativa* ssp. *japonica* cultivar Nipponbare were used to establish the physical map of rice chromosomes by polymerase chain reaction (PCR) screening[19], fingerprinting[20] and end-sequencing[21]. The PAC, BAC and fosmid clones on the physical map were subjected to random shearing and shotgun sequencing to tenfold redundancy, using both universal primers and the dye-terminator or dye-primer methods. The sequences were assembled using PHRED (http://www.genome.washington.edu/UWGC/analysistools/Phred.cfm) and PHRAP (http://www.genome.washington.edu/UWGC/analysistools/Phrap.cfm) software packages or using the TIGR Assembler (http://www.tigr.org/software/assembler/).

Sequence gaps were resolved by full sequencing of gap-bridge clones, PCR fragments or direct sequencing of BACs. Sequence ambiguities (indicated by PHRAP scores less than 30) were resolved by confirming the sequence data using alternative chemistries or different polymerases. We empirically determined that a PHRAP score of 30 or above exceeds the standard of less than one error in 10,000 bp. BAC and PAC assemblies were tested for accuracy by comparing computationally derived fingerprint patterns with experimentally determined patterns of restriction enzyme digests. Sequence quality was also evaluated by comparing independently obtained overlapping sequences.

Small physical gaps were filled by long-range PCR. Remaining physical gaps were measured using fluorescence *in situ* hybridization analysis. We used the length of CentO arrays[26] to estimate the size of each of the remaining centromere gaps.

Annotation and bioinformatics. Gene models were predicted using FGENESH (http://www.softberry.com/berry.phtml?topic = fgenesh) using the monocot trained matrix on the native and repeat-masked pseudomolecules. Gene models with incomplete open reading frames, those encoding proteins of less than 50 amino acids, or those corresponding to organellar DNA were omitted from the final set. The coordinates of transposable elements, excluding MITEs (miniature inverted-repeat transposable elements), were used to mask the pseudomolecules.

Conserved domain/motif searches and association with gene ontologies were performed using InterproScan (http://www.ebi.ac.uk/InterProScan/) in combination with the Interpro2Go program. For biological processes, the number of detected domains was re-calculated as number of non-redundant proteins.

The predicted rice proteome was searched using BLASTP against the proteomes of several model species for which a complete genome sequence and deduced protein set was available. Each rice chromosome was searched against the TIGR rice gene index (http://www.tigr.org/tdb/tgi/ogi/) and against gene index entries that aligned to gene models corresponding to expressed genes. In addition, five cereal gene indices (http://www.tigr.org/tdb/tgi/) were searched against the rice chromosomes, and gene index matches were recorded. We searched the *Oryza sativa* ssp. *japonica* cv. Nipponbare collection of full-length cDNAs (ftp://cdna01.dna.affrc.go.jp/pub/data/), after first removing the transposable-element-related sequences, against the FGENESH models.

Gene models with rice full-length cDNA, EST or cereal EST matches but without identifiable homologues in the *Arabidopsis* genome were searched for conserved domains/motifs using InterproScan, and for homologues in the Swiss-Prot database (http://us.expasy.org/sprot/) using BLASTP. All proteins with positive blast matches were further compared with the nr database (http://www.ncbi.nlm.nih.gov/blast/html/blastcgihelp.html#protein_databases), using BLASTP to eliminate truncated proteins and those with matches to other dicots.

Tandem gene families. The rice genome was subjected to a BLASTP search as previously described[32]. The search was also performed by permitting more than one unrelated gene within the arrays, and the limit of the search was set to 5-Mb intervals to exclude large chromosomal duplications.

Non-coding RNAs. Transfer-RNA genes were detected by the program tRNAscan SE (http://www.genetics.wustl.edu/eddy/tRNAscan-SE/). The miRNA registry in the Rfam database (http://www.sanger.ac.uk/Software/Rfam/) was used as a reference database for miRNAs. In addition, experimentally validated miRNAs of other species, excluding *Arabidopsis* miRNAs, were used for BLASTN queries against the pseudomolecules. Spliceosomal and snoRNAs were retrieved from the Rfam database and used for queries. BLASTN was used to find the location of snoRNAs and spliceosomal RNAs in the pseudomolecules.

Organellar insertions. *Oryza sativa* ssp. *japonica* Nipponbare chloroplast (GenBank NC_001320) and mitochondrial (GenBank BA000029) sequences were aligned with the pseudomolecules using BLASTN and MUMmer[49].

Transposable elements. The TIGR *Oryza* Repeat Database, together with other published and unpublished rice transposable element sequences, was used to create RTEdb (a rice transposable element database)[50] and determine transposable element coordinates on the rice pseudomolecules. In the case of *h*AT, IS*256*/*Mutator*, IS*5*/*Tourist* and IS*630*/Tc1/*mariner* elements, family-specific profile hidden Markov models were applied using HMMER[41] (http://hmmer.wustl.edu/). The remaining superfamilies were annotated using RepeatMasker (http://www.repeatmasker.org/).

***Tos17* insertions.** Flanking sequences of transposed copies of 6,278 *Tos17* insertion lines were isolated by modified thermal asymmetric interlaced (TAIL)-PCR and suppression PCR, and screened against the pseudomolecule sequences.

SNP discovery. BAC clones from an *O. sativa* ssp. *indica* var. Kasalath BAC library were end-sequenced. Sequence reads were omitted if they contained more than 50% nucleotides of low quality or high similarity to known repeats. The remaining sequences were subjected to BLASTN analysis against the pseudomolecules. Gaps within the alignments were classified as small insertions/deletions.

SSR loci. The Simple Sequence Repeat Identification Tool (http://www.gramene.org/) was used to identify simple sequence repeat motifs, and the physical position of all Class 1 SSRs was recorded. The copy number of SSR markers was estimated using electronic (e)-PCR to determine the number of independent hits of primer pairs on the pseudomolecules.

Whole-genome shotgun assembly analysis. Contigs from the BGI 6.28 × whole genome assembly of *O. sativa* ssp. *indica* 93-11 (GenBank/DDBJ/EMBL accession number AAAA02000001–AAAA02050231) and the Syngenta 6 × whole genome assembly of *O. sativa* ssp. *japonica* cv. Nipponbare (AACV01000001–AACV01035047; ref. 48) were aligned with the pseudomolecules using MUMmer[49]. The number of IRGSP Nipponbare full-length cDNA-supported gene models completely covered by the aligned contigs was tabulated. The 155-bp CentO consensus sequence was used for BLAST analysis against the 93-11 and Nipponbare whole-genome shotgun contigs, and the coordinates of the positive hits recorded. Locations of centromeres for each *indica* chromosome were obtained with the CentO sequence positions on the IRGSP pseudomolecule of the corresponding chromosome. A detailed comparison of the BGI-assembled and -mapped Syngenta contigs (AACV01000001–AACV01000070) and the 93-11 contigs (AAAA02000001–AAAA02000093) was obtained by BLAST analysis against the IRGSP chromosome 1 pseudomolecule.

Detailed procedures for the analyses described above can be found in the Supplementary Information.

Table 3 | Transposons in the rice genome

	Copy no. (× 10^3)	Coverage (kb)	Fraction of genome (%)
Class I			
LINEs	9.6	4161.3	1.12
SINEs	1.8	209.9	0.06
Ty1/*copia*	11.6	14266.7	3.85
Ty3/*gypsy*	23.5	40363.3	10.90
Other class I	15.4	12733.3	3.43
Total class I	**61.9**	**71734.4**	**19.35**
Class II			
*h*AT	1.1	1405.9	0.38
CACTA	10.8	9987.3	2.69
IS*630*/Tc1/*mariner*	67.0	8388.3	2.26
IS*256*/*Mutator*	8.8	13485.7	3.64
IS*5*/*Tourist*	57.9	12095.8	3.26
Other class II	18.2	2703.6	0.73
Total class II	**163.8**	**48066.6**	**12.96**
Other TEs	**23.6**	**6797.7**	**1.80**
Total TEs	**249.3**	**129019.3***	**34.79**

TE, transposable element.
*Total length; corrected for 2420.7 kb in overlaps of multiple, non-nested elements.

Received 29 December 2004; accepted 25 May 2005.

1. Peng, S., Cassman, K. G., Virmani, S. S., Sheehy, J. & Khush, G. S. Yield potential trends of tropical rice since the release of IR8 and the challenge of increasing rice yield potential. *Crop Sci.* **39,** 1552–1559 (1999).
2. Peng, S. *et al.* Rice yields decline with higher night temperature from global warming. *Proc. Natl Acad. Sci. USA* **101,** 9971–9975 (2004).

3. Sasaki, T. & Burr, B. International Rice Genome Sequencing Project: the effort to completely sequence the rice genome. *Curr. Opin. Plant Biol.* **3,** 138–141 (2000).
4. Moore, G., Devos, K. M., Wang, Z. & Gale, M. D. Cereal genome evolution: Grasses, line up and form a circle. *Curr. Biol.* **5,** 737–739 (1995).
5. Sasaki, T. *et al.* The genome sequence and structure of rice chromosome 1. *Nature* **420,** 312–316 (2002).
6. Feng, Q. *et al.* Sequence and analysis of rice chromosome 4. *Nature* **420,** 316–320 (2002).
7. Rice Chromosome 10 Sequencing Consortium, In-depth view of structure, activity, and evolution of rice chromosome 10. *Science* **300,** 1566–1569 (2003).
8. Wu, J. *et al.* Composition and structure of the centromeric region of rice chromosome 8. *Plant Cell* **16,** 967–976 (2004).
9. Zhang, Y. *et al.* Structural features of the rice chromosome 4 centromere. *Nucleic Acids Res.* **32,** 2023–2030 (2004).
10. Nagaki, K. *et al.* Sequencing of a rice centromere uncovers active genes. *Nature Genet.* **36,** 138–145 (2004).
11. Guyot, R. & Keller, B. Ancestral genome duplication in rice. *Genome* **47,** 610–614 (2004).
12. Simillion, C., Vandepoele, K., Saeys, Y. & Van de Peer, Y. Building genomic profiles for uncovering segmental homology in the twilight zone. *Genome Res.* **14,** 1095–1106 (2004).
13. Paterson, A. H., Bowers, J. E. & Chapman, B. A. Ancient polyploidization predating divergence of the cereals, and its consequences for comparative genomics. *Proc. Natl Acad. Sci. USA* **101,** 9903–9908 (2004).
14. Salse, J., Piegu, B., Cooke, R. & Delseny, M. New *in silico* insight into the synteny between rice (*Oryza sativa* L.) and maize (*Zea mays* L.) highlights reshuffling and identifies new duplications in the rice genome. *Plant J.* **38,** 396–409 (2004).
15. Lai, J. *et al.* Gene loss and movement in the maize genome. *Genome Res.* **14,** 1924–1931 (2004).
16. Harushima, Y. *et al.* A high-density rice genetic linkage map with 2275 markers using a single F_2 population. *Genetics* **148,** 479–494 (1998).
17. Yamamoto, K. & Sasaki, T. Large-scale EST sequencing in rice. *Plant Mol. Biol.* **35,** 135–144 (1997).
18. Saji, S. *et al.* A physical map with yeast artificial chromosome (YAC) clones covering 63% of the 12 rice chromosomes. *Genome* **44,** 32–37 (2001).
19. Wu, J. *et al.* A comprehensive rice transcript map containing 6591 expressed sequence tag sites. *Plant Cell* **14,** 525–535 (2002).
20. Chen, M. *et al.* An integrated physical and genetic map of the rice genome. *Plant Cell* **14,** 537–545 (2002).
21. Mao, L. *et al.* Rice transposable elements: a survey of 73,000 sequence-tagged-connectors. *Genome Res.* **10,** 982–990 (2000).
22. Barry, G. F. The use of the Monsanto draft rice genome sequence in research. *Plant Physiol.* **125,** 1164–1165 (2001).
23. Goff, S. A. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* **296,** 92–100 (2002).
24. Ohmido, N., Kijima, K., Akiyama, Y., de Jong, J. H. & Fukui, K. Quantification of total genomic DNA and selected repetitive sequences reveals concurrent changes in different DNA families in *indica* and *japonica* rice. *Mol. Gen. Genet.* **263,** 388–394 (2000).
25. Dong, F. *et al.* Rice (*Oryza sativa*) centromeric regions consist of complex DNA. *Proc. Natl Acad. Sci. USA* **95,** 8135–8140 (1998).
26. Cheng, Z. *et al.* Functional rice centromeres are marked by a satellite repeat and a centromere-specific retrotransposon. *Plant Cell* **14,** 1691–1704 (2002).
27. Kikuchi, S. *et al.* Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice. *Science* **301,** 376–379 (2003).
28. Castelli, V. *et al.* Whole genome sequence comparisons and "full-length" cDNA sequences: a combined approach to evaluate and improve *Arabidopsis* genome annotation. *Genome Res.* **14,** 406–413 (2004).
29. Hirochika, H., Sugimoto, K., Otsuki, Y., Tsugawa, H. & Kanda, M. Retrotransposons of rice involved in mutations induced by tissue culture. *Proc. Natl Acad. Sci. USA* **93,** 7783–7788 (1996).
30. Miyao, A. *et al.* Target site specificity of the *Tos17* retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. *Plant Cell* **15,** 1771–1780 (2003).
31. Alonso, J. M. *et al.* Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* **301,** 653–657 (2003).
32. Arabidopsis Genome Initiative, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408,** 796–815 (2000).
33. Song, R., Llaca, V. & Messing, J. Mosaic organization of orthologous sequences in grass genomes. *Genome Res.* **12,** 1549–1555 (2002).
34. Shishido, R., Sano, Y. & Fukui, K. Ribosomal DNAs: an exception to the conservation of gene order in rice genomes. *Mol. Gen. Genet.* **263,** 586–591 (2000).
35. Oono, K. & Sugiura, M. Heterogeneity of the ribosomal RNA gene clusters in rice. *Chromosoma* **76,** 85–89 (1980).
36. Kamisugi, Y. *et al.* Physical mapping of the 5S ribosomal RNA genes on rice chromosome 11. *Mol. Gen. Genet.* **245,** 133–138 (1994).
37. Bartel, D. P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–297 (2004).
38. Wang, X. J., Reyes, J. L., Chua, N. H. & Gaasterland, T. Prediction and identification of *Arabidopsis thaliana* microRNAs and their mRNA targets. *Genome Biol.* **5,** R65 (2004).
39. Wang, J. F., Zhou, H., Chen, Y. Q., Luo, Q. J. & Qu, L. H. Identification of 20 microRNAs from *Oryza sativa*. *Nucleic Acids Res.* **32,** 1688–1695 (2004).
40. Turcotte, K., Srinivasan, S. & Bureau, T. Survey of transposable elements from rice genomic sequences. *Plant J.* **25,** 169–179 (2001).
41. Eddy, S. R. Profile hidden Markov models. *Bioinformatics* **14,** 755–763 (1998).
42. Messing, J. *et al.* Sequence composition and genome organization of maize. *Proc. Natl Acad. Sci. USA* **101,** 14349–14354 (2004).
43. Shen, Y. J. *et al.* Development of genome-wide DNA polymorphism database for map-based cloning of rice genes. *Plant Physiol.* **135,** 1198–1205 (2004).
44. Feltus, F. A. *et al.* An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. *Genome Res.* **14,** 1812–1819 (2004).
45. McCouch, S. R. *et al.* Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* **9,** 257–279 (2002).
46. Causse, M. A. *et al.* Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics* **138,** 1251–1274 (1994).
47. Yu, J. *et al.* A draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* **296,** 79–92 (2002).
48. Yu, J. *et al.* The genomes of *Oryza sativa*: A history of duplications. *PLoS Biol.* **3,** e38 (2005).
49. Delcher, A. L. *et al.* Alignment of whole genomes. *Nucleic Acids Res.* **27,** 2369–2376 (1999).
50. Juretic, N., Bureau, T. E. & Bruskiewich, R. M. Transposable element annotation of the rice genome. *Bioinformatics* **20,** 155–160 (2004).

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

Acknowledgements Work at the RGP was supported by the Ministry of Agriculture, Forestry and Fisheries of Japan. Work at TIGR was supported by grants to C.R.B. from the USDA Cooperative State Research, Education and Extension Service–National Research Initiative, the National Science Foundation and the US Department of Energy. Work at the NCGR was supported by the Chinese Ministry of Science and Technology, the Chinese Academy of Sciences, the Shanghai Municipal Commission of Science and Technology, and the National Natural Science Foundation of China. Work at Genoscope was supported by le Ministère de la Recherche, France. Funding for the work at the AGI and AGCoL was provided by grants to R.A.W. and C.S. from the USDA Cooperative State Research, Education and Extension Service–National Research Initiative, the National Science Foundation, the US Department of Energy and the Rockefeller Foundation. Work at CSHL was supported by grants from the USDA Cooperative State Research, Education and Extension Service–National Research Initiative and from the National Science Foundation. Work at the ASPGC was supported by Academia Sinica, National Science Council, Council of Agriculture, and Institute of Botany, Academia Sinica. The IIRGS acknowledges the Department of Biotechnology, Government of India, for financial assistance and the Indian Council of Agricultural Research, New Delhi, for support. Work at Rice Gene Discovery was supported by BIOTECH and the Princess Sirindhorn's Plant Germplasm Conservation Initiative Program. Work at PGIR was supported by Rutgers University. The BRIGI was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos - Ministério de Ciência e Tecnologia (FINEP-MCT), Fundação de Amparo a Pesquisa do Rio Grande do Sul (FAPERGS) and Universidade Federal de Pelotas (UFPel). Work at McGill and York Universities was supported by the National Science and Engineering Research Council of Canada and the Canadian International Development Agency. Funding for H.H. at the National Institute of Agrobiological Sciences was from the Ministry of Agriculture, Forestry, and Fisheries of Japan, and the Program for Promotion of Basic Research Activities for Innovative Biosciences. Funding at Brookhaven National Laboratory was from The Rockefeller Foundation and the Office of Basic Energy Science of the United States Department of Energy. We would like to thank G. Barry and S. Goff for their help in negotiating agreements that permitted the sharing of materials and sequence with the IRGSP. We also acknowledge the work of G. Barry, S. Goff and their colleagues in facilitating the transfer of sequence information and supporting data.

Author Information The genomic sequence is available under accession numbers AP008207–AP008218 in international databases (DDBJ, GenBank and EMBL). Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Takuji Sasaki (tsasaki@nias.affrc.go.jp).

International Rice Genome Sequencing Project (Participants are arranged by area of contribution and then by institution.)
Physical Maps and Sequencing: Rice Genome Research Program (RGP) Takashi Matsumoto[1], Jianzhong Wu[1], Hiroyuki Kanamori[1], Yuichi Katayose[1], Masaki Fujisawa[1], Nobukazu Namiki[1], Hiroshi Mizuno[1], Kimiko Yamamoto[1], Baltazar A. Antonio[1], Tomoya Baba[1], Katsumi Sakata[1], Yoshiaki Nagamura[1], Hiroyoshi Aoki[1], Koji Arikawa[1], Kohei Arita[1], Takahito Bito[1], Yoshino Chiden[1], Nahoko Fujitsuka[1], Rie Fukunaka[1], Masao Hamada[1], Chizuko Harada[1], Akiko Hayashi[1], Saori Hijishita[1], Mikiko Honda[1], Satomi Hosokawa[1], Yoko Ichikawa[1], Atsuko Idonuma[1], Masumi Iijima[1], Michiko Ikeda[1], Maiko Ikeno[1], Kazue Ito[1], Sachie Ito[1], Tomoko Ito[1], Yuichi Ito[1], Yukiyo Ito[1], Aki Iwabuchi[1], Kozue Kamiya[1], Wataru Karasawa[1], Kanako Kurita[1], Satoshi Katagiri[1], Ari Kikuta[1], Harumi Kobayashi[1], Noriko Kobayashi[1], Kayo Machita[1], Tomoko Maehara[1], Masatoshi Masukawa[1], Tatsumi Mizubayashi[1], Yoshiyuki Mukai[1], Hideki Nagasaki[1], Yuko Nagata[1], Shinji Naito[1], Marina Nakashima[1], Yuko Nakama[1], Yumi Nakamichi[1], Mari Nakamura[1], Ayano Meguro[1], Manami Negishi[1], Isamu Ohta[1], Tomoya Ohta[1], Masako Okamoto[1], Nozomi Ono[1], Shoko Saji[1], Miyuki Sakaguchi[1], Kumiko Sakai[1], Michie Shibata[1], Takanori Shimokawa[1], Jianyu Song[1], Yuka Takazaki[1], Kimihiro Terasawa[1], Mika Tsugane[1], Kumiko Tsuji[1], Shigenori Ueda[1], Kazunori Waki[1], Harumi Yamagata[1], Mayu Yamamoto[1], Shinichi Yamamoto[1], Hiroko Yamane[1], Shoji Yoshiki[1], Rie Yoshihara[1], Kazuko Yukawa[1], Huisun Zhong[1], Masahiro Yano[1], Takuji Sasaki (Principal Investigator)[1]; **The Institute for Genomic Research (TIGR)** Qiaoping Yuan[2], Shu Ouyang[2], Jia Liu[2], Kristine M. Jones[2], Kristen Gansberger[2], Kelly Moffat[2], Jessica Hill[2], Jayati Bera[2], Douglas Fadrosh[2], Shaohua Jin[2], Shivani Johri[2], Mary Kim[2], Larry Overton[2], Matthew Reardon[2], Tamara Tsitrin[2], Hue Vuong[2], Bruce Weaver[2], Anne Ciecko[2], Luke Tallon[2], Jacqueline Jackson[2], Grace Pai[2], Susan Van Aken[2], Terry Utterback[2], Steve Reidmuller[2], Tamara Feldblyum[2], Joseph Hsiao[2], Victoria Zismann[2], Stacey Iobst[2], Aymeric R. de Vazeille[2], C. Robin Buell (Principal Investigator)[2]; **National Center for Gene Research Chinese Academy of Sciences (NCGR)** Kai Ying[3], Ying Li[3], Tingting Lu[3], Yuchen Huang[3], Qiang Zhao[3], Qi Feng[3], Lei Zhang[3], Jingjie Zhu[3], Qijun Weng[3], Jie Mu[3], Yiqi Lu[3], Danlin Fan[3], Yilei Liu[3], Jianping Guan[3], Yujun Zhang[3], Shuliang Yu[3], Xiaohui Liu[3], Yu Zhang[3], Guofan Hong[3], Bin Han (Principal Investigator)[3]; **Genoscope** Nathalie Choisne[4], Nadia Demange[4], Gisela Orjeda[4], Sylvie Samain[4], Laurence Cattolico[4], Eric Pelletier[4], Arnaud Couloux[4], Beatrice Segurens[4], Patrick Wincker[4], Angelique D'Hont[5], Claude Scarpelli[4], Jean Weissenbach[4], Marcel Salanoubat[4], Francis Quetier (Principal Investigator)[4]; **Arizona Genomics Institute (AGI) and Arizona Genomics Computational Laboratory (AGCoI)** Yeisoo Yu[6], Hye Ran Kim[6], Teri Rambo[6], Jennifer Currie[6], Kristi Collura[6], Meizhong Luo[6], Tae-Jin Yang[6], Jetty S. S. Ammiraju[6], Friedrich Engler[6], Carol Soderlund[6], Rod A. Wing (Principal Investigator)[6]; **Cold Spring Harbor Laboratory (CSHL)** Lance E. Palmer[7], Melissa de la Bastide[7], Lori Spiegel[7], Lidia Nascimento[7], Theresa Zutavern[7], Andrew O'Shaughnessy[7], Sujit Dike[7], Neilay Dedhia[7], Raymond Preston[7], Vivekanand Balija[7], W. Richard McCombie (Principal Investigator)[7]; **Academia Sinica Plant Genome Center (ASPGC)** Teh-Yuan Chow[8], Hong-Hwa Chen[9], Mei-Chu Chung[8], Ching-San Chen[8], Jei-Fu Shaw[8], Hong-Pang Wu[8], Kwang-Jen Hsiao[10], Ya-Ting Chao[8], Mu-kuei Chu[8], Chia-Hsiung Cheng[8], Ai-Ling Hour[8], Pei-Fang Lee[8], Shu-Jen Lin[8], Yao-Cheng Lin[8], John-Yu Liou[8], Shu-Mei Liu[8], Yue-Ie Hsing (Principal Investigator)[8]; **Indian Initiative for Rice Genome Sequencing (IIRGS), University of Delhi South Campus (UDSC)** S. Raghuvanshi[11], A. Mohanty[11], A. K. Bharti[11,13], A. Gaur[11], V. Gupta[11], D. Kumar[11], V. Ravi[11], S. Vij[11], A. Kapur[11], Parul Khurana[11], Paramjit Khurana[11], J. P. Khurana[11], A. K. Tyagi (Principal Investigator)[11]; **Indian Initiative for Rice Genome Sequencing (IIRGS), Indian Agricultural Research Institute (IARI)** K. Gaikwad[12], A. Singh[12], V. Dalal[12], S. Srivastava[12], A. Dixit[12], A. K. Pal[12], I. A. Ghazi[12], M. Yadav[12], A. Pandit[12], A. Bhargava[12], K. Sureshbabu[12], K. Batra[12], T. R. Sharma[12], T. Mohapatra[12], N. K. Singh (Principal Investigator)[12]; **Plant Genome Initiative at Rutgers (PGIR)** Joachim Messing (Principal Investigator)[13], Amy Bronzino Nelson[13], Galina Fuks[13], Steve Kavchok[13], Gladys Keizer[13], Eric Linton Victor Llaca[13], Rentao Song[13], Bahattin Tanyolac[13], Steve Young[13]; **Korea Rice Genome Research Program (KRGRP)** Kim Ho-Il[14], Jang Ho Hahn (Principal Investigator)[14]; **National Center for Genetic Engineering and Biotechnology (BIOTEC)** G. Sangsakoo[15], A. Vanavichit (Principal Investigator)[15]; **Brazilian Rice Genome Initiative (BRIGI)** Luiz Anderson Teixeira de Mattos[16], Paulo Dejalma Zimmer[16], Gaspar Malone[16], Odir Dellagostin[16], Antonio Costa de Oliveira (Principal Investigator)[16]; **John Innes Centre (JIC)** Michael Bevan[17], Ian Bancroft[17]; **Washington University School of Medicine Genome Sequencing Center** Pat Minx[18], Holly Cordum[18], Richard Wilson[18]; **University of Wisconsin–Madison** Zhukuan Cheng[19], Weiwei Jin[19], Jiming Jiang[19], Sally Ann Leong[20]

Annotation and Analysis: Hisakazu Iwama[21], Takashi Gojobori[21,22], Takeshi Itoh[22,23], Yoshihito Niimura[24], Yasuyuki Fujii[25], Takuya Habara[25], Hiroaki Sakai[23,25], Yoshiharu Sato[22], Greg Wilson[26], Kiran Kumar[27], Susan McCouch[26], Nikoleta Juretic[28], Douglas Hoen[28], Stephen Wright[29], Richard Bruskiewich[30], Thomas Bureau[28], Akio Miyao[23], Hirohiko Hirochika[23], Tomotaro Nishikawa[23], Koh-ichi Kadowaki[23] & Masahiro Sugiura[31]

Coordination: Benjamin Burr[32]

Affiliations for participants: [1]National Institute of Agrobiological Sciences/Institute of the Society for Techno-innovation of Agriculture, Forestry and Fisheries, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan. [2]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, Maryland 20850, USA. [3]Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences (CAS), 500 Caobao Road, Shanghai 200233, China. [4]Centre National de Séquençage, INRA-URGV, and CNRS UMR-8030, 2, rue Gaston Crémieux, CP 5706, 91057 EVRY Cedex, France. [5]UMR PIA, Cirad-Amis, TA40-03 avenue Agropolis, 34398 Montpellier Cedex 05, France. [6]Department of Plant Sciences, BIO5 Institute, The University of Arizona, Tucson, Arizona 85721, USA. [7]Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11723, USA. [8]Institute of Botany, Academia Sinica, 128, Sec. 2, Yen-Chiu-Yuan Rd, Nankang, Taipei 11529, Taiwan. [9]National Cheng Kung University, No. 1, Ta-Hsueh Road, Tainan 701, Taiwan. [10]National Yang-Ming University, 155, Sec. 2, Li-Nong St, Peitou, Taipei 112, Taiwan. [11]Department of Plant Molecular Biology, University of Delhi South Campus, New Delhi 110021, India. [12]National Research Centre on Plant Biotechnology, Indian Agricultural Research Institute, New Delhi 110012, India. [13]Waksman Institute, Rutgers University, Piscataway, New Jersey 08854, USA. [14]National Institute of Agricultural Science and Technology, RDA, Suwon, 441-707 Republic of Korea. [15]Rice Gene Discovery Unit, Kasetsart University, Nakron Pathom 73140, Thailand. [16]Centro de Genomica e Fitomelhoramento, UFPel, Pelotas, RS, I 96001-970, Brazil. [17]John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK. [18]Washington University Genome Sequencing Center, 3333 Forest Park Boulevard, St. Louis, Missouri 63108, USA. [19]University of Wisconsin, Department of Horticulture, Madison, Wisconsin 53706, USA. [20]University of Wisconsin, Department of Plant Pathology, Madison, Wisconsin 53706, USA. [21]Center for Information Biology and DNA Data Bank of Japan, National Institute of Genetics, Mishima 411-8540, Japan. [22]Biological Information Research Center, National Institute of Advanced Industrial Science and Technology, Koto-ku, Tokyo 135-0064, Japan. [23]National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8602, Japan. [24]Medical Research Institute, Tokyo Medical and Dental University, Bunkyo-ku, Tokyo 113-8510, Japan. [25]Japan Biological Information Research Center, Japan Biological Informatics Consortium, Koto-ku, Tokyo 135-0064, Japan. [26]Plant Breeding Dept, Cornell University, Ithaca, New York 14850-1901, USA. [27]Cold Spring Harbor Laboratory, PO Box 100, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. [28]Department of Biology, McGill University, 1205 Dr Penfield Avenue, Montreal, Quebec H3A 1B1, Canada. [29]Department of Biology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada. [30]Biometrics and Bioinformatics Unit, International Rice Research Institute, DAPO Box 7777, Metro Manila, Philippines. [31]Graduate School of Natural Sciences, Nagoya City University, Nagoya 467-8501, Japan. [32]Biology Department, Brookhaven National Laboratory, Upton, New York 11973, USA.

ARTICLES

Microstructure of a spatial map in the entorhinal cortex

Torkel Hafting[1]*, Marianne Fyhn[1]*, Sturla Molden[1]†, May-Britt Moser[1] & Edvard I. Moser[1]

The ability to find one's way depends on neural algorithms that integrate information about place, distance and direction, but the implementation of these operations in cortical microcircuits is poorly understood. Here we show that the dorsocaudal medial entorhinal cortex (dMEC) contains a directionally oriented, topographically organized neural map of the spatial environment. Its key unit is the 'grid cell', which is activated whenever the animal's position coincides with any vertex of a regular grid of equilateral triangles spanning the surface of the environment. Grids of neighbouring cells share a common orientation and spacing, but their vertex locations (their phases) differ. The spacing and size of individual fields increase from dorsal to ventral dMEC. The map is anchored to external landmarks, but persists in their absence, suggesting that grid cells may be part of a generalized, path-integration-based map of the spatial environment.

Navigation in mammals depends on a distributed, modularly organized brain network[1–7]. The network computes and represents positional and directional information, as indicated by the existence of 'place cells'[1] and 'head-direction cells'[3] in the hippocampal and parahippocampal cortices. The function of hippocampal place cells has received particular attention[1,2]. Place cells respond to a wide variety of spatial inputs, including extrinsic landmarks[8,9] and translational and directional movement signals[10–13]. Along with the strong involvement of the hippocampus in spatial learning[1,14], the convergent expression of metric positional and directional information in place cells pointed, early on, to a pivotal role for the hippocampus in spatial computation and representation[1–7,14,15]. However, whereas place cells were originally thought to have general navigational functions, it became clear that sensory information about space is differentiated by the hippocampus into a multitude of context-specific representations[16–21], which can be retrieved independently from degraded versions of the original input[22,23]. This diversification of stored information is consistent with a principal role for the hippocampus in event- or context-specific memory[24,25], and raises the possibility that context-independent position information is computed upstream of the hippocampus[5,6], for example, by algorithms that integrate self-motion information into a metric and directionally oriented representation that is valid in all contexts[4–6].

A map with such properties has not been identified, but recent work has shown that location is represented accurately before the hippocampus in the superficial layers of the dorsocaudal region of the medial entorhinal cortex (dMEC)[26]. Whereas place cells in the hippocampus usually have a single confined firing field[1], upstream dMEC neurons have multiple discrete fields of similar amplitude[26]. To determine whether the dMEC has a map-like structural organization, we recorded spike activity in this area while rats ran in enclosures that were large enough to capture possible regularities in the spatial organization of neural activity (14 rats, 211 cells).

Grid cells have tessellating firing fields

To visualize the spatial structure of firing fields in dMEC neurons, we first tested the rats in a circular enclosure with a diameter of 2 m (45 neurons from 6 rats). Individual neurons in layer II of the dMEC had multiple discrete firing fields with distinct inhibitory surrounds (Fig. 1)[26]. The expanded recording environment revealed a striking spatial organization of the subfields that was not apparent in the smaller, conventionally sized enclosures used previously. In every isolated principal neuron, the firing field formed a grid of regularly tessellating triangles spanning the whole recording surface (Fig. 1b, left and middle columns). All nodes of activity were sharply delineated from the background, although the individual peak firing rates varied. The regular nature of the activity distribution was verified by spatial autocorrelation analyses, which for all cells showed a tessellating pattern similar to that of the original rate maps (Fig. 1b, right column; Supplementary Fig. S1).

To examine the geometric structure of the grid, we measured the separation of the peaks in the autocorrelogram. If the unit is an equilateral triangle, the central peak of the autocorrelogram should be surrounded by six equidistant peaks forming the vertices of a regular hexagon. The analysis confirmed this prediction. First, within each firing grid, the distance from the central peak of the autocorrelogram to the nearest six peaks was nearly constant. Although the average of this distance varied from 39 to 73 cm across different cells of different rats, the standard deviation (s.d.) within a single grid was only 3.2 cm (mean s.d. across 45 cells). Additional hexagons of equidistant firing peaks were formed at multiples of the distance to the nearest hexagon, implying that the pattern was regular across the entire field. Second, the angular separation of the vertices of the inner hexagon was in multiples of 60 degrees (Fig. 1c; s.d. = 7.1 degrees), as expected if the unit was an equilateral triangle.

The location of the grid vertices (the phase) remained stable across recording trials (Supplementary Fig. S2)[26]. The rats were tested twice in the large cylinder at an interval of 15 min. The peak firing rates in each subfield on the first trial correlated weakly but significantly with the peak rates in the same areas on the second trial (mean correlation ± s.e.m., $r = 0.35 \pm 0.06$, $t(25) = 6.1$, where 25 degrees of freedom are indicated in parentheses, $P < 0.001$), implying that activity was distributed non-randomly across the nodes of the firing lattice.

The structure of the grid was not confined by the boundaries of the enclosure (Fig. 1d). When the environment was expanded, the

[1]Centre for the Biology of Memory, Norwegian University of Science and Technology, 7489 Trondheim, Norway. †Present address: Department of Physiology, University of Oslo, PO Box 1103 Blindern, 0317 Oslo, Norway.
*These authors contributed equally to this work.

number of activity nodes increased, but their density remained constant ($t(28) = 1.2$, $P > 0.20$; 29 cells from 3 rats), suggesting that grids may potentially have infinite size (Supplementary Fig. S3).

Grid cells are topographically organized

Grid cells in the dMEC showed a striking topographic organization. Grids recorded at the same electrode location shared a number of metric properties, including spacing, orientation (direction) and field size (Fig. 2). Spacing was expressed for each grid as the distance from the central peak to the vertices of the inner hexagon in the autocorrelogram (the median of the six distances). While spacing varied by more than 30 cm in the entire cell population, the s.d. among cells recorded on the same tetrode was only 2.1 cm (mean across 12 sets of simultaneously recorded cells, $n = 42$). The orientation of the grid was expressed in the autocorrelogram as the angle ϕ between a camera-defined reference line (0 degrees) and a vector to the nearest vertex of the inner hexagon in the counterclockwise direction (Fig. 2e inset). The entire range of orientations was represented in the population as a whole (from 1 to 59 degrees), but among cells recorded on the same tetrode, orientation varied minimally (s.d. = 1.8 degrees). The size of the individual fields was estimated as the area covered by the central peak of the autocorrelogram, using a threshold of $r = 0.2$. Field sizes ranged from 326 to 709 cm^2 in the population as a whole. The s.d. among cells recorded at the same electrode location was 42 cm^2.

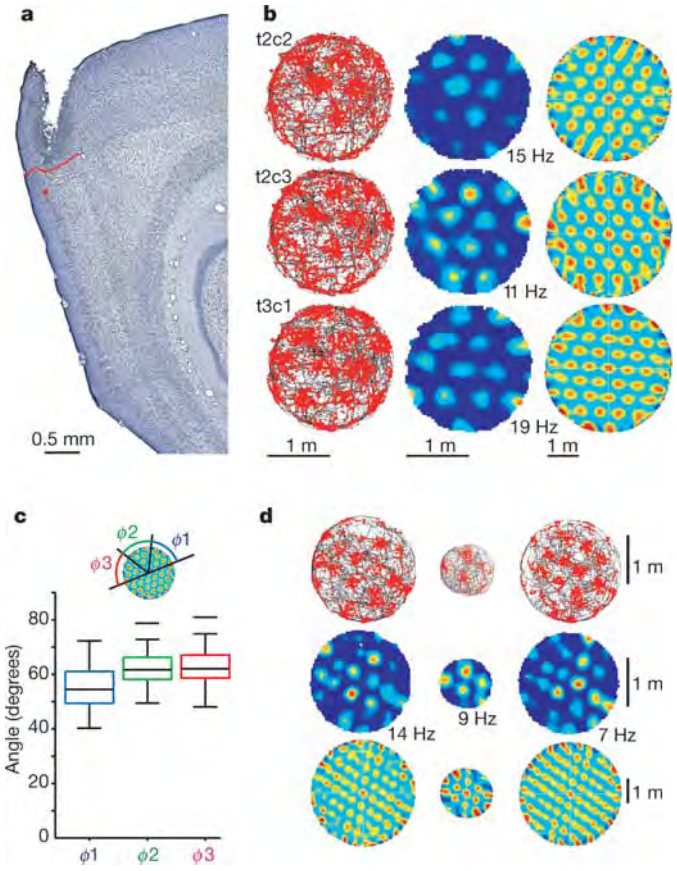


Figure 1 | Firing fields of grid cells have a repetitive triangular structure. **a**, Sagittal Nissl-stained section indicating the recording location (red dot) in layer II of the dMEC. Red line indicates border to postrhinal cortex. **b**, Firing fields of three simultaneously recorded cells at the dot in **a** during 30 min of running in a large circular enclosure. Cell names refer to tetrode (t) and cell (c). Left, trajectory of the rat (black) with superimposed spike locations (red). Middle, colour-coded rate map with the peak rate indicated. Red is maximum, dark blue is zero. Right, spatial autocorrelation for each rate map (see Supplementary Methods). The colour scale is from blue ($r = -1$) through green ($r = 0$) to red ($r = 1$). **c**, Box plot showing distribution of angles $\phi 1$, $\phi 2$ and $\phi 3$ between the central peak of the autocorrelogram and the vertices of a hexagon defined by the nearest six peaks. The diagram shows median angles (horizontal lines inside boxes), interquartile distances (boxes), upper and lower limits, and outliers (horizontal lines). **d**, Discharge maps (as in **b**) showing similar triangular structure in enclosures of different size (left, large; middle, small; right, large).

Although spacing, orientation and field size were almost invariant at individual recording locations, spacing and field size increased with distance from the postrhinal border, resulting in more dispersed fields at more ventral electrode positions. This pattern was observed both within animals (Fig. 2a–e) and between animals (Fig. 2f–h; Supplementary Fig. S4). Figure 2 shows the difference between grid cells at two electrode locations, one near the postrhinal border and one 560 μm deeper, in a rat with a double dMEC implant tested in the large circular enclosure. The spacing of the grid was consistently denser at the dorsal position (Fig. 2e). In the dorsal cells, the spacing ranged from 39.1 to 43.0 cm (14 cells recorded over 3 days). In the more ventrally located cells, the spacing ranged from 48.1 to 52.2 cm (5 cells recorded over 2 days). The increase in spacing at the ventral position was accompanied by an increase in the size of the individual fields (dorsal 353–583 cm^2; ventral 511–637 cm^2; $t(17) = 2.4$, $P < 0.05$). A similar topographic arrangement was revealed when one set of tetrodes was lowered tangentially along the border of layers II and III (yellow circles in Fig. 2b). Over a distance of 100 μm, the spacing increased from between 41.9 and 45.9 cm ($n = 3$) to between 44.9 and 49.7 cm ($n = 4$). These results were replicated in two other rats with double implants (17 cells) as well as in a larger sample of rats running in a smaller square box (1 m^2; 11 rats, 57 cells; Fig. 2f–h and Supplementary Fig. S4). In the latter group, the distance from the postrhinal border correlated significantly with both the spacing of the grid ($r = 0.82$, degrees of freedom, d.f. = 55, $P < 0.001$) and the size of the individual fields ($r = 0.79$, $P < 0.001$). The field size was quadratically proportional to the spacing ($r = 0.75$, $P < 0.001$).

Like spacing and field size, orientation of the grid varied between electrode locations (Fig. 2 and Supplementary Fig. S4). In Fig. 2, the angle with the camera-defined horizontal line was between 26.7 and 32.7 degrees at the recording location near the postrhinal border ($n = 14$) and between 19.7 and 22.1 degrees at the more ventral location on the contralateral side ($n = 5$) (Fig. 2e). However, although the entire range of orientations was represented across animals, we were unable to detect any systematic change from dorsal to ventral in dMEC ($r = -0.13$, $P = 0.30$).

Although grids of neighbouring cells had similar spacing, field size and orientation, their phases (the vertex locations) were apparently not related. Collectively, grids from a small number of units recorded simultaneously at the same electrode position filled up the entire space of the recording arena (Fig. 2). Because neighbouring cells had similar grid spacing and grid orientation, a slight phase shift in each of the grids was sufficient to superimpose the vertices of the grids almost completely (Fig. 3a). Cross-correlation of the rate maps of cell pairs recorded simultaneously at the same electrode location yielded a regular multi-peaked surface with a structure very similar to that of each cell's autocorrelogram, except that the peaks were offset from the origin in most cases (Fig. 3b). Cross-correlation of cells recorded at different locations (with different spacing and orientation) gave cross-correlograms with more dispersed peaks and lower peak amplitudes (Supplementary Fig. S5). Among the co-localized cells, the average phase shift, expressed as the distance from the origin to the nearest peak in the cross-correlogram, was evenly distributed, extending from 0 to a maximum of 0.5 of the spacing of the corresponding autocorrelograms, both in the group as a whole and in individual recordings (Fig. 3c). The distribution of phase shifts did not deviate significantly from uniformity ($\chi(4) = 0.84$ with bins corresponding to 20% of the maximally possible grid spacing) and was not related to the distance from the postrhinal border ($r = 0.13$, d.f. = 43, n.s.). These results suggest that the complete surface of the

environment was represented at each dorsoventral level of the dMEC, for every spacing and every orientation of the map.

Grids are anchored to external cues

How does the spatial map in dMEC contribute to navigation? A key question is whether locations of discharge in grid cells are determined by external landmarks (allothetic cues) or by information generated by the rat's own movement (idiothetic cues). The stability of the grid vertices across successive trials in the same enclosure (Supplementary Fig. S2) suggests that allothetic cues exert a significant influence. To test their influence more directly, we rotated a cue card on the wall of a circular test box while recording from grid cells in dMEC (3 rats, 24 cells). Distal cues were masked by curtains. When the cue card was rotated 90 degrees, the grid rotated similarly (24/24 cells; Fig. 4; Supplementary Fig. S6). The correlation between the rate maps of the initial baseline trial and the rotation trial was substantially lower than between trials with the cue card in the same position (Fig. 4b; $t(23) = 4.6$, $P < 0.001$). There was no change in grid spacing or field size (Fig. 4c and d; $t < 0.5$). Counter-rotating the rate map from the rotation trial reversed the drop in spatial correlation (Fig. 4a and b; $t(23) = 1.6$, $P = 0.10$), supporting the notion that phase and orientation are set by external landmarks.

Grid structure persists after cue removal

To determine whether allothetic cues were also necessary for maintaining grid-like activity during exploration, we asked whether grids were expressed after all visual cues were removed. Spike activity was recorded in total darkness for 30 min after an initial period of 10 min with the lights on (4 rats, 33 cells; Fig. 5 and Supplementary Fig. S7). The grid was maintained in darkness. Darkness had no significant effect on the spacing of the grid, and there was no change in average firing rate or spatial information per spike ($t < 1.5$, $P > 0.15$), suggesting that the existence of a grid-like firing structure was independent of allothetic information. Yet, in the majority of the cells, the onset of total darkness caused a weak dispersal or displacement of the vertices, expressed as a moderate decrease in the spatial correlation of the rate maps (Fig. 5b; $t(29) = 6.5$, $P < 0.001$). The decrease in spatial correlation is consistent with a role for allothetic cues in determining the phase of the grid, that is, in aligning the grid to the external reference frame. Stationary non-visual cues contributed minimally to the firing positions in darkness (Supplementary Fig. S8).

Grid development in a novel environment

The persistence of firing structure after perturbations of external sensory input suggests that grids may to a large extent be based on hardwired network mechanisms. If this is true, they may stabilize rapidly in a novel environment. We tested this prediction by recording grid cells for 30 min while rats explored a square enclosure in a novel room (room N; 7 rats, 24 cells). Before and after this trial, the rats ran in a similar enclosure in a familiar room (F and F′). Visual inspection suggested that the grid pattern was expressed from the outset (Fig. 6a). Firing locations were mostly stable from the first time the rat passed through the area, both in light (Fig. 6; Supplementary

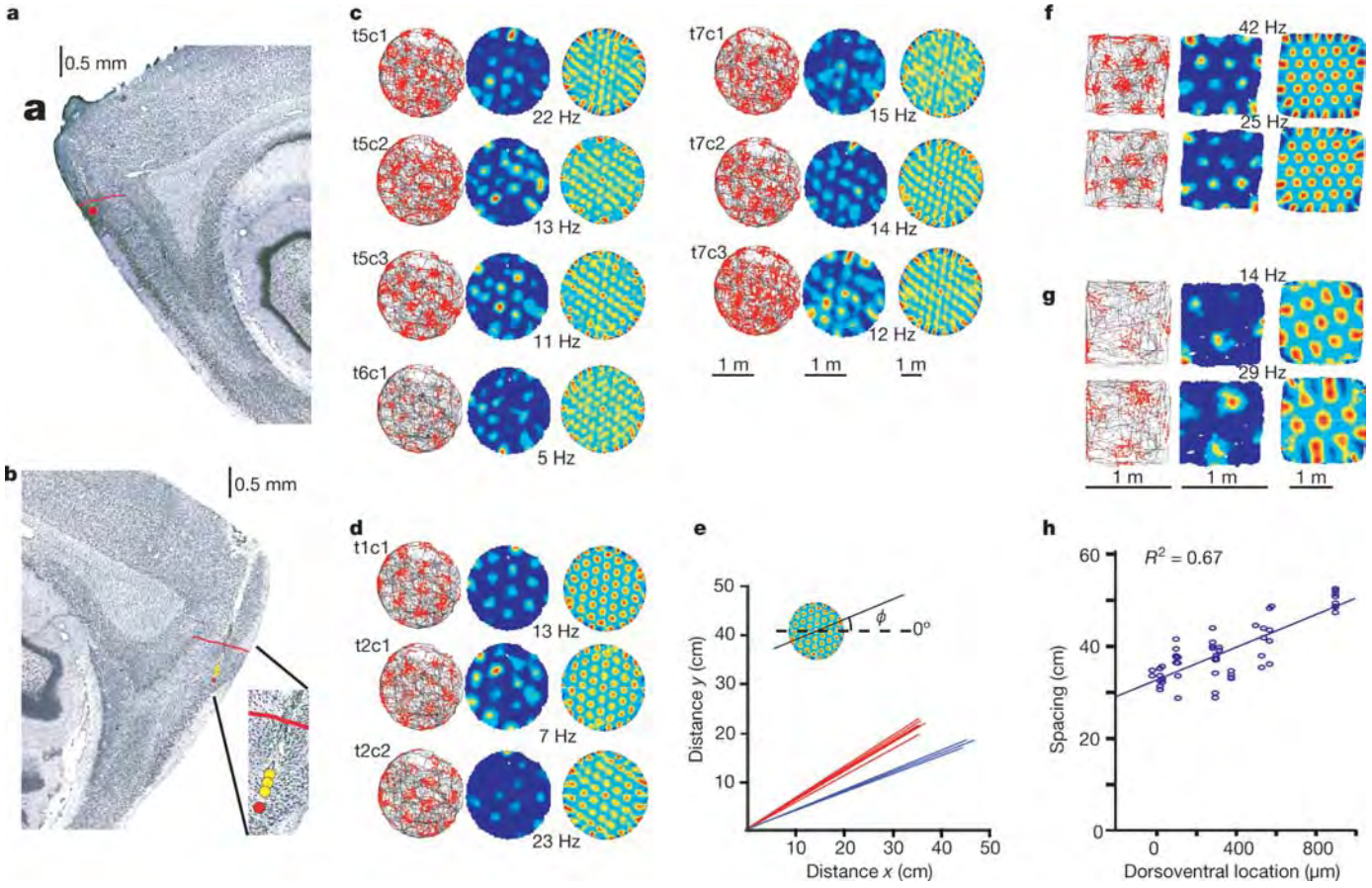


Figure 2 | Grid cells recorded simultaneously at two electrode locations in the same rat. **a**, **b**, Sagittal sections showing recording locations (dots) in layer II at the dorsal extreme of the left dMEC (**a**) and 560 μm more ventrally in the right dMEC (**b**). Red line indicates border to postrhinal cortex. **c**, **d**, Trajectory maps (left), rate maps (middle) and spatial autocorrelograms (right) for cells recorded at positions indicated by red dots in **a** and **b**, respectively. **e**, Vector representation of grid spacing (length) and grid orientation (direction) in **c** (red) and **d** (blue). Each vector refers to one cell. Direction is relative to a fixed horizontal reference (inset). Note invariance of spacing and orientation at each location. **f**, **g**, Discharge maps (as in **c** and **d**) for representative pairs of cells at the dorsalmost (**f**) and ventralmost (**g**) recording locations tested. These rats ran in a smaller square enclosure. **h**, Grid spacing as a function of distance from the postrhinal cortex for all animals tested in the square enclosure (Supplementary Fig. S4).

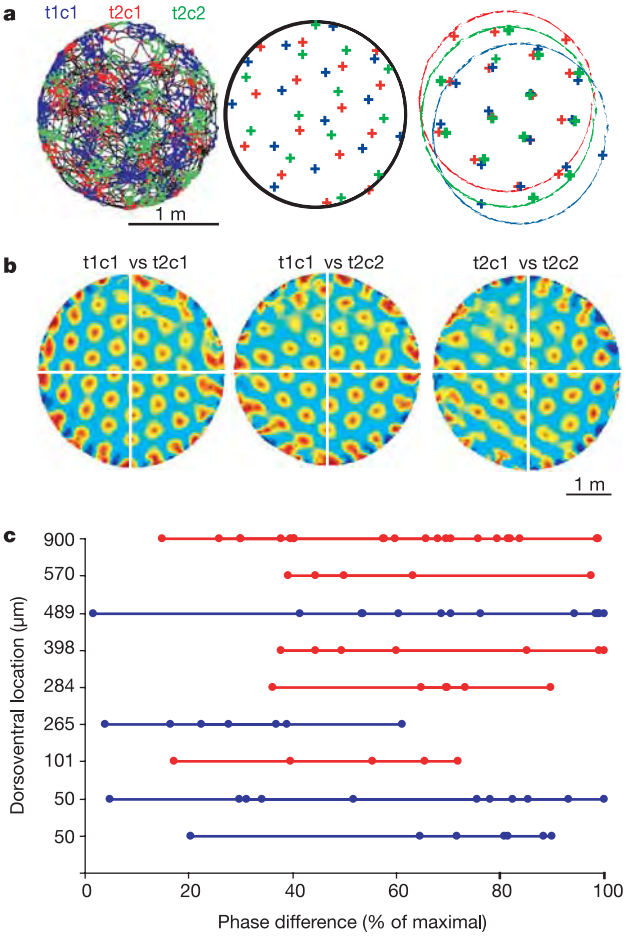


Figure 3 | Distributed spatial phase of co-localized grid cells. **a**, Grids of the three cells in Fig. 2d, each with a separate colour. Left, trajectory maps. Middle, peak locations. Right, peaks are offset to visualize similarity in spacing and orientation. **b**, Spatial cross-correlations for the same three cells. **c**, Distribution of phase differences between co-localized neuron pairs, expressed as distance from the origin to the nearest peak in their cross-correlogram. Each circle indicates one cell pair. Simultaneously recorded cell pairs are connected (only recordings with ≥5 cell pairs). Blue, large cylinder; red, small square enclosure.

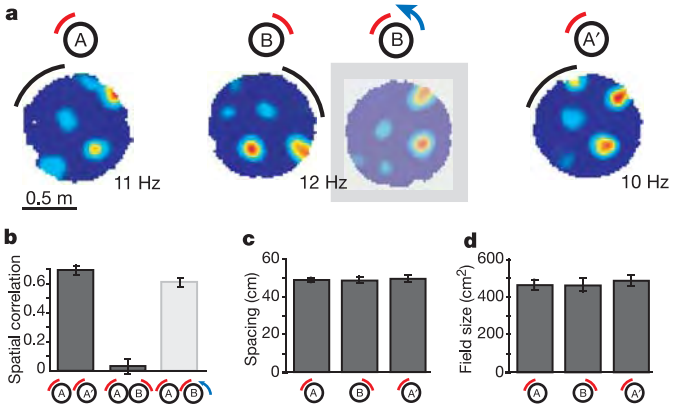


Figure 4 | Grids are aligned to environment-specific landmarks. **a**, Rate maps for a representative cell after rotation of the cue card (arc) on the small cylinder. Left and right, cue card in original position (A and A′). Middle pair, cue card rotated 90° (B). Shaded map, rate map counter-rotated 90°. **b–d**, Firing properties of grid cells on trials with the cue card in A and B positions (all rats; means ± s.e.m.).

Fig. S9) and in total darkness (Supplementary Fig. S7). The development of firing was estimated by correlating for each grid cell the final firing rate at each position in the environment (estimated during the last 10 min of the trial) with the firing rate at the same locations (if visited) for blocks of 2 min throughout the preceding 20 min. Firing locations during the first 2 min were significantly correlated with firing locations at the end of the trial (mean correlation ± s.e.m. = 0.53 ± 0.04; Fig. 6b). However, the correlation was weaker during the first blocks than later (Fig. 6b; Block effect: $F_{9,207} = 8.5$, $P < 0.001$), suggesting that stabilization of the grid, like stabilization of hippocampal place fields[18,27], requires some minimum exposure. The delay may reflect the time needed to set phase and orientation in relation to context-specific landmarks. The orientation of the grid in the new room was different from the orientation in the familiar room (mean change = 20.8 ± 3.2 degrees). The orientation drifted only minimally between the two tests in the familiar room (2.3 ± 0.5 degrees), as expected if grids are aligned with external landmarks.

Discussion

This study points to the dMEC as part of a neural map of the spatial environment. The basic unit of the map is the grid cell, whose multiple discrete firing fields invariantly form a stable, regularly tessellating structure of equilateral triangles. Grid spacing, grid orientation and field size were topographically organized, with minimal variation locally, but significant variation across the surface of the dMEC. Spacing and field size increased from the dorsal to the ventral part of dMEC, raising the possibility that this gradient is continuous with the broad, single-peaked firing fields in the intermediate band of the MEC[26]. In contrast, the phase of the grid appeared to vary randomly among cells recorded at the same brain location, implying that the entire surface of the environment was represented within a local cell ensemble with a common grid spacing and orientation. These observations point to a modular local organization of the entorhinal spatial map, similar to the columnar organization of other areas of neocortex[28,29]. The mosaic organization of the superficial layers of dMEC[30] represents a possible substrate for the modularity of the spatial map.

A fundamental property of the entorhinal spatial map is the apparent universality of several key parameters. The triangular firing structure and the spacing of the grid cells were impervious to displacement or removal of external sensory input, in much the same way that units in more ventromedial parts of MEC preserve firing fields after geometric transformation of the local environment[31]. The persistent firing structure of the population is reminiscent of the coherence across environments expressed by head-direction cells[3,12,32,33] and suggests that in both systems the same neurons may perform the same algorithms in different environments. The maintenance of grid structure during visual deprivation points to path integration[7,34–36] as a probable motor of the spatial periodicity. Whether the position vector is computed in the dMEC itself or outside remains to be determined, but the convergence in the dMEC of directional input from the dorsal presubiculum[37] and visuospatial input from the postrhinal cortex[38,39] puts the dMEC in a unique position to perform this computation. A direct involvement of the dMEC in path integration is consistent with the disruption of return paths[40] and spatial search patterns[41] during navigation in rats with entorhinal lesions that include the dorsocaudal pole. However, although the spatial structure of firing may be determined by self-motion cues, phase and orientation were controlled by the specific landmarks of each environment, in the same way that specific cues are needed to initialize the directional preferences of the head-direction cells[3]. These alignment processes presumably involve associative learning.

The representation of place, distance and direction in the same network of dMEC neurons would permit the computation of a continuously updated metric representation of the animal's location. The underlying algorithms may have similarities with those

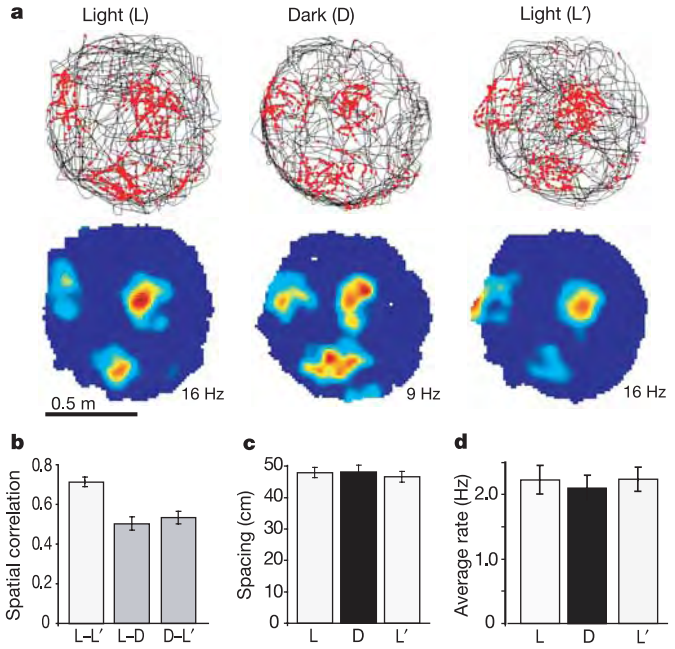


Figure 5 | Grids persist in darkness. a, Trajectory and rate maps for a representative dMEC cell after onset of darkness. Room lights were on for 10 min (L), off for 10 min (D), and on for another 10 min (L′). **b–d**, Firing properties of grid cells in L, D and L′ (all rats; means ± s.e.m.).

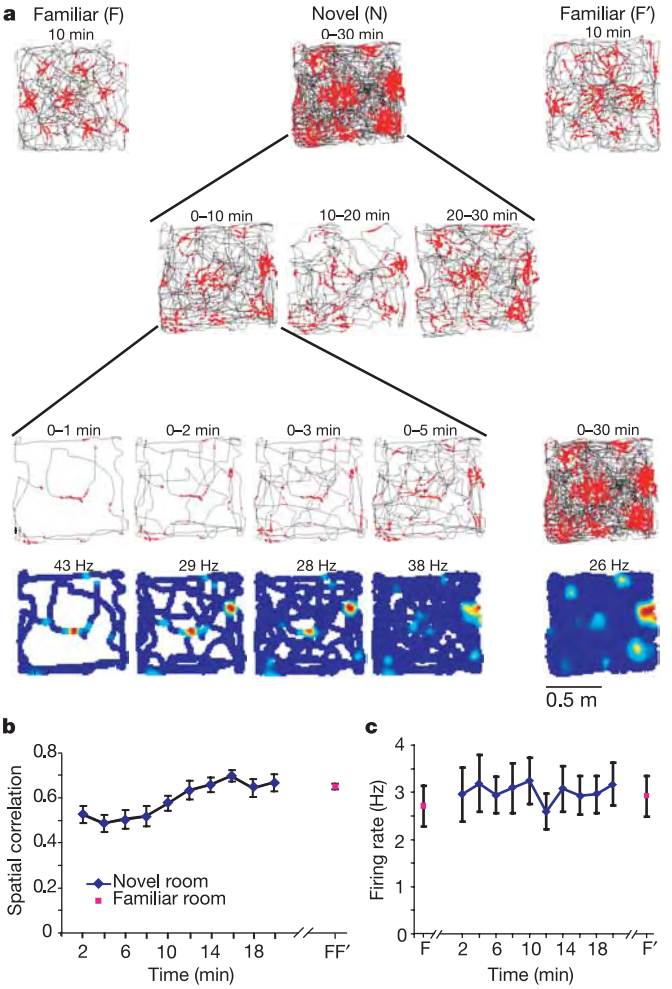


Figure 6 | Grid structure of dMEC cells is expressed instantly in a novel environment. a, Trajectory and rate maps for a grid cell in a familiar room, a novel room, and the familiar room (top row). The middle trial is broken into blocks to illustrate development. **b**, Development of grid structure as indicated by spatial correlation with the final firing field (last 10 min) across preceding 2-min blocks of the trial (all rats; means ± s.e.m.). **c**, Stability of firing rate across 2-min blocks, normalized to the average value for the last 10 min.

proposed previously for recurrent networks of head-direction cells[42] and hippocampal place cells[15]. By way of their intrinsic connections[43–45], entorhinal neurons may be able to support a two-dimensional continuous attractor-based representation of the environment, in which activity is moved between successive mutually connected cell ensembles by a path-integration-based mechanism[2,15]. The constant spacing, field size and orientation of grids in neighbouring cells suggest that conjunctions of active neurons are repeated regularly as the animal moves over a surface. If the start location is retained, a cyclically organized map might be sufficient to signal location in environments of any size, and work equally well at the edges and in the centre of the map[5].

One potential problem with the grid structure is that the population code at each vertex of the grid may be relatively similar. This potential ambiguity could be resolved by using the small, but reliable rate differences between the vertices, or perhaps more efficiently, by integrating over grids with different spacing and orientation. The mechanism by which information about distance and direction is extracted is not known, but successful read-out would probably depend significantly on temporal integration of neural activity. Sustained firing in single cells[46] and reverberatory circuits[47] may contribute to such integration. Although the wide distribution of grid phase allows the animal's position to be predicted from the collective firing of small local cell ensembles[26], it remains to be determined whether read-out occurs within the entorhinal cortex, or in one or several of its hippocampal or parahippocampal target structures[4–6].

The results suggest that the entorhinal cortex may support some spatial computations that were previously attributed to the hippocampus[1,2,14,15]. Although hippocampal place cells respond to translational and directional information[10–13], similar influences were observed upstream in layer II of the dMEC. The contextual specificity of hippocampal representations[16–21] suggests that during encoding, the hippocampus associates output from a generalized, path-integration-based coordinate system with landmarks or other features specific to the individual environment or context. Through backprojections to the superficial layers of the entorhinal cortex[39], associations stored in the hippocampus may reset the path integrator as errors accumulate during exploration[7,36,48,49]. Anchoring the output of the path integrator to external reference points stored in the hippocampus or other cortical areas may enable alignment of entorhinal maps from one trial to the next, even when the points of departure are different.

METHODS

Subjects. Fourteen male Long Evans rats (350–450 g) kept at ~90% of free-feeding body weight were implanted above the dMEA with a microdrive containing four tetrodes[26]. Six of these rats received an additional microdrive in the contralateral hemisphere either above the dMEA ($n = 5$) or above the dorsal CA3 of the hippocampus[18] ($n = 1$). Tetrodes were made of 17-μm-diameter polyimide-coated platinum–iridium (90%–10%) wire.

Enclosures and rooms. Neuronal activity was recorded during pellet chasing in enclosures of different size and shape: a large circular box (200 cm in diameter, 50 cm high), a small circular box (90 cm in diameter, 50 cm high), a square box (100 cm × 100 cm × 50 cm high), and a linear track (235 cm long, 10 cm wide and 50 cm above the floor). A cue card (45 cm × 50 cm or 20 cm × 30 cm) was displayed at a constant location in each box.

Training and testing. Tetrodes were moved in daily steps of ~50 μm or less towards layer II of the dMEA until multiple distinguishable large-amplitude units were obtained. Six rats were tested on alternating trials in two circular enclosures of different size at the same location. In three rats, the cue card was

rotated 90 degrees and back on separate trials in the small circle. Four animals were tested in a black square box with room lights on and off on alternating trials. The rat was not taken out of the recording enclosure between these trials. Seven animals were tested on alternating trials in two square boxes, one in a familiar room and one in a novel room. An eighth animal was tested in the novel room with lights off during the first 30 min. Three rats were trained to run back and fourth on a linear track with food rewards at the ends. These rats were tested on alternating trials in light and darkness.

Spike sorting and analysis of firing structure. Position and unit data were recorded with an Axona recording station. Spike sorting was performed off-line for all data using graphical cluster-cutting software[26]. Firing-rate maps, spatial correlations, spatial autocorrelations and spatial cross-correlations were computed as described in Supplementary Methods.

Histology and topography. The majority of the electrodes were positioned in layer II at the caudal end of the dorsolateral band of the MEC[26]. A smaller subset was located at the border between layers II and III. Data from these recordings were not separated out because cells in layers II and III appear to have similar grid-like firing properties[50]. Most data were recorded at the bottom of the electrode trace. When this was not the case, the recording position was extrapolated using the read-out of the tetrode turning protocol corrected for shrinkage[26]. Distance between the electrode position and the postrhinal border was measured tangentially along layer II.

Received 16 February; accepted 5 May 2005.
Published online 19 June 2005.

1. O'Keefe, J. & Nadel, L. *The Hippocampus as a Cognitive Map* (Clarendon, Oxford, 1978).
2. McNaughton, B. L. *et al.* Deciphering the hippocampal polyglot: the hippocampus as a path integration system. *J. Exp. Biol.* **199,** 173–185 (1996).
3. Taube, J. S. Head direction cells and the neurophysiological basis for a sense of direction. *Prog. Neurobiol.* **55,** 225–256 (1998).
4. Redish, A. D. & Touretzky, D. S. Cognitive maps beyond the hippocampus. *Hippocampus* **7,** 15–35 (1997).
5. Redish, A. D. *Beyond the Cognitive Map: From Place Cells to Episodic Memory* (MIT Press, Cambridge, 1999).
6. Sharp, P. E. Complimentary roles for hippocampal versus subicular/entorhinal place cells in coding place, context, and events. *Hippocampus* **9,** 432–443 (1999).
7. Etienne, A. S. & Jeffery, K. J. Path integration in mammals. *Hippocampus* **14,** 180–192 (2004).
8. O'Keefe, J. & Conway, D. H. Hippocampal place units in the freely moving rat: why they fire where they fire. *Exp. Brain Res.* **31,** 573–590 (1978).
9. O'Keefe, J. & Burgess, N. Geometric determinants of the place fields of hippocampal neurons. *Nature* **381,** 425–428 (1996).
10. Gothard, K. M., Skaggs, W. E. & McNaughton, B. L. Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *J. Neurosci.* **16,** 8027–8040 (1996).
11. Sharp, P. E., Blair, H. T., Etkin, D. & Tzanetos, D. B. Influences of vestibular and visual motion information on the spatial firing patterns of hippocampal place cells. *J. Neurosci.* **15,** 173–189 (1995).
12. Knierim, J. J., Kudrimoti, H. S. & McNaughton, B. L. Place cells, head direction cells, and the learning of landmark stability. *J. Neurosci.* **15,** 1648–1659 (1995).
13. Jeffery, K. J., Donnett, J. G., Burgess, N. & O'Keefe, J. M. Directional control of hippocampal place fields. *Exp. Brain Res.* **117,** 131–142 (1997).
14. Nadel, L. The hippocampus and space revisited. *Hippocampus* **1,** 221–229 (1991).
15. Samsonovich, A. & McNaughton, B. L. Path integration and cognitive mapping in a continuous attractor neural network model. *J. Neurosci.* **17,** 5900–5920 (1997).
16. Muller, R. U. & Kubie, J. L. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. *J. Neurosci.* **7,** 1951–1968 (1987).
17. Bostock, E., Muller, R. U. & Kubie, J. L. Experience-dependent modifications of hippocampal place cell firing. *Hippocampus* **1,** 193–205 (1991).
18. Leutgeb, S., Leutgeb, J. K., Treves, A., Moser, M.-B. & Moser, E. I. Distinct ensemble codes in hippocampal areas CA3 and CA1. *Science* **305,** 1295–1298 (2004).
19. Markus, E. J. *et al.* Interactions between location and task affect the spatial and directional firing of hippocampal neurons. *J. Neurosci.* **15,** 7079–7094 (1995).
20. Frank, L. M., Brown, E. N. & Wilson, M. Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* **27,** 169–178 (2000).
21. Wood, E. R., Dudchenko, P. A., Robitsek, R. J. & Eichenbaum, H. Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* **27,** 623–633 (2000).
22. Nakazawa, K. *et al.* Requirement for hippocampal CA3 NMDA receptors in associative memory recall. *Science* **297,** 211–218 (2002).
23. Lee, I., Yoganarasimha, D., Rao, G. & Knierim, J. J. Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature* **430,** 456–459 (2004).
24. Rolls, E. T. & Treves, A. *Neural Networks and Brain Function* (Oxford Univ. Press, Oxford, 1998).
25. Squire, L. R., Stark, C. E. & Clark, R. E. The medial temporal lobe. *Annu. Rev. Neurosci.* **27,** 279–306 (2004).
26. Fyhn, M., Molden, S., Witter, M. P., Moser, E. I. & Moser, M. B. Spatial representation in the entorhinal cortex. *Science* **305,** 1258–1264 (2004).
27. Wilson, M. A. & McNaughton, B. L. Dynamics of the hippocampal ensemble code for space. *Science* **261,** 1055–1058 (1993).
28. Mountcastle, V. B. The columnar organization of the neocortex. *Brain* **120,** 701–722 (1997).
29. Rockland, K. S. & Ichinohe, N. Some thoughts on cortical minicolumns. *Exp. Brain Res.* **158,** 265–277 (2004).
30. Ikeda, J., Mori, K., Oka, S. & Watanabe, Y. A columnar arrangement of dendritic processes of entorhinal cortex neurons revealed by a monoclonal antibody. *Brain Res.* **505,** 176–179 (1989).
31. Quirk, G. J., Muller, R. U., Kubie, J. L. & Ranck, J. B. Jr The positional firing properties of medial entorhinal neurons: description and comparison with hippocampal place cells. *J. Neurosci* **12,** 1945–1963 (1992).
32. Taube, J. S., Muller, R. U. & Ranck, J. B. Jr Head-direction cells recorded from the postsubiculum in freely moving rats. II. Effects of environmental manipulations. *J. Neurosci.* **10,** 436–447 (1990).
33. Goodridge, J. P. & Taube, J. S. Preferential use of the landmark navigational system by head direction cells in rats. *Behav. Neurosci.* **109,** 49–61 (1995).
34. Mittelstaedt, M. L. & Mittelstaedt, H. Homing by path integration in a mammal. *Naturwissenschaften* **67,** 566–567 (1980).
35. Gallistel, C. R. *The Organization of Learning* (MIT Press, Cambridge Massachusetts, 1990).
36. Biegler, R. Possible uses of path integration in animal navigation. *Anim. Learn. Behav.* **28,** 257–277 (2000).
37. van Haeften, T., Wouterlood, F. G., Jorritsma-Byham, B. & Witter, M. P. GABAergic presubicular projections to the medial entorhinal cortex of the rat. *J. Neurosci.* **17,** 862–874 (1997).
38. Witter, M. P., Groenewegen, H. J., Lopes da Silva, F. H. & Lohman, A. H. Functional organization of the extrinsic and intrinsic circuitry of the parahippocampal region. *Prog. Neurobiol.* **33,** 161–253 (1989).
39. Witter, M. P. & Amaral, D. G. in *The Rat Nervous System* 3rd edn (ed. Paxinos, G.) 637–703 (Academic, San Diego, 2004).
40. Parron, C. & Save, E. Evidence for entorhinal and parietal cortices involvement in path integration in the rat. *Exp. Brain Res.* **159,** 349–359 (2004).
41. Steffenach, H.-A., Witter, M. P., Moser, M.-B. & Moser, E. I. Spatial memory in the rat requires the dorsolateral band of the entorhinal cortex. *Neuron* **45,** 301–313 (2005).
42. Skaggs, W. E., Knierim, J. J., Kudrimoto, H. & McNaughton, B. L. in *Advances in Neural Information Processing Systems* (eds Tesauro, G., Touretzky, D. S. & Leen, T. K.) Vol. 7, 173–180 (MIT Press, Cambridge, Massachusetts, 1995).
43. Lingenhohl, K. & Finch, D. M. Morphological characterization of rat entorhinal neurons *in vivo*: soma-dendritic structure and axonal domains. *Exp. Brain Res.* **84,** 57–74 (1991).
44. Germroth, P., Schwerdtfeger, W. K. & Buhl, E. H. Ultrastructure and aspects of functional organization of pyramidal and nonpyramidal entorhinal projection neurons contributing to the perforant path. *J. Comp. Neurol.* **305,** 215–231 (1991).
45. Dhillon, A. & Jones, R. S. Laminar differences in recurrent excitatory transmission in the rat entorhinal cortex *in vitro*. *Neuroscience* **99,** 413–422 (2000).
46. Egorov, A. V., Hamam, B. N., Fransen, E., Hasselmo, M. E. & Alonso, A. A. Graded persistent activity in entorhinal cortex neurons. *Nature* **420,** 173–178 (2002).
47. Iijima, T. *et al.* Entorhinal-hippocampal interactions revealed by real-time imaging. *Science* **272,** 1176–1179 (1996).
48. Lorincz, A. & Buzsaki, G. Two-phase computational model training long-term memories in the entorhinal-hippocampal region. *Ann. NY Acad. Sci.* **911,** 83–111 (2000).
49. Fyhn, M., Molden, S., Hollup, S., Moser, M.-B. & Moser, E. I. Hippocampal neurons responding to first-time dislocation of a target object. *Neuron* **35,** 555–566 (2002).
50. Sargolini, F., Molden, S., Witter, M. P., Moser, E. I. & Moser, M.-B. Place representation in the deep layers of entorhinal cortex. *Soc. Neurosci. Abstr.* **30,** 330.9 (2004).

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

Acknowledgements We are grateful to the members of the Centre for the Biology of Memory as well as W. E. Skaggs, J. Lisman and G. Einevoll for discussions. We also thank the technical team of the Centre for their assistance. This work is supported by the Norwegian Research Council's Centre of Excellence scheme.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.I.M. (edvard.moser@ntnu.no).

ARTICLES

Evasion of the p53 tumour surveillance network by tumour-derived *MYC* mutants

Michael T. Hemann[1]*, Anka Bric[1]*, Julie Teruya-Feldstein[2], Andreas Herbst[1], Jonas A. Nilsson[3], Carlos Cordon-Cardo[2], John L. Cleveland[3], William P. Tansey[1] & Scott W. Lowe[1,4]

The c-Myc oncoprotein promotes proliferation and apoptosis, such that mutations that disable apoptotic programmes often cooperate with *MYC* during tumorigenesis. Here we report that two common mutant *MYC* alleles derived from human Burkitt's lymphoma uncouple proliferation from apoptosis and, as a result, are more effective than wild-type *MYC* at promoting B cell lymphomagenesis in mice. Mutant MYC proteins retain their ability to stimulate proliferation and activate p53, but are defective at promoting apoptosis due to a failure to induce the BH3-only protein Bim (a member of the B cell lymphoma 2 (Bcl2) family) and effectively inhibit Bcl2. Disruption of apoptosis through enforced expression of Bcl2, or loss of either *Bim* or *p53* function, enables wild-type *MYC* to produce lymphomas as efficiently as mutant *MYC*. These data show how parallel apoptotic pathways act together to suppress MYC-induced transformation, and how mutant MYC proteins, by selectively disabling a p53-independent pathway, enable tumour cells to evade p53 action during lymphomagenesis.

Human tumours frequently show deregulated expression of the *c-Myc* proto-oncogene[1–3]. In Burkitt's lymphoma, this deregulation occurs through reciprocal translocations that juxtapose *c-Myc* with an immunoglobulin (Ig) promoter, leading to gross overexpression of *c-Myc* messenger RNA in the B cell lineage[4,5]. In addition, point mutations are often found in the translocated *MYC* alleles, clustering in a conserved region known as MYC box I (refs 6–8). Although some mutations can increase MYC stability and transforming activity *in vitro*, their impact on the pathogenesis of Burkitt's lymphoma is unclear[9–16]. In fact, translocated *c-Myc* genes are subject to hypermutation *in vivo* that can also alter non-coding sequences, raising the possibility that these mutations are a consequence and not a cause of tumour development[17,18].

To examine the effects of *MYC* mutation on lymphoma development *in vivo*, we used a system for rapidly generating tissue-specific transgenic mice (see Supplementary Fig. 1). Two mutant *MYC* alleles commonly observed in Burkitt's lymphoma (P57S and T58A)[13] were cloned into a murine stem cell virus (MSCV)-based vector that co-expresses green fluorescent protein (GFP). Haematopoietic stem cells (HSCs) derived from normal fetal livers were transduced with retroviruses expressing either wild-type or mutant *MYC*, and the genetically modified stem cells were then used to reconstitute the haematopoietic system of lethally irradiated recipient animals. This adoptive transfer strategy has several advantages over traditional transgenic approaches, such as: (1) transgene expression is limited to the haematopoietic system; (2) various combinations of transgenes or stem cell genotypes can be rapidly analysed; and (3) only a fraction of the injected stem cells are infected, allowing tumorigenesis to occur in the context of an essentially normal haematopoietic compartment. Also, the strong MSCV promoter produces high levels of MYC expression that recapitulate that observed in Burkitt's lymphoma (data not shown), and the co-expressed GFP enables assessment of infection efficiency and visualization of tumorigenesis using fluorescence imaging.

MYC mutants show enhanced oncogenicity *in vivo*

Although the control MSCV vector was not oncogenic (data not shown), recipients of stem cells infected with wild-type *MYC* developed tumours at low penetrance after a long latency (Fig. 1a; 2 out of 13 recipient mice at >100 days). The tumours that arose were aggressive pre-B cell lymphomas (Fig. 1b, c; see also Supplementary Fig. 2a), similar to those produced by the tissue-specific overexpression of *MYC* in transgenic mice[19]. Recipients of stem cells infected with the tumour-derived *MYC* point mutants also developed pre-B cell lymphomas, but at a significantly higher penetrance (Fig. 1a; 9 out of 12 for P57S and 8 out of 11 for T58A; $P < 0.005$ for each mutant versus wild-type *MYC*) and a significantly reduced latency (56 ± 9 days for P57S and 66 ± 9 days for T58A). Flow cytometry for GFP fluorescence showed that the increased tumorigenicity of these mutants was not due to differences in stem cell infection efficiency or transgene expression (see Supplementary Fig. 2b).

As deregulated *MYC* expression coordinately induces proliferation and apoptosis[20], the increased oncogenicity of the mutant MYC proteins might reflect the altered activation of one or both of these processes. To investigate the proliferation of cells expressing wild-type and mutant *MYC*, we assessed BrdU incorporation in the bone marrow of pre-malignant mice after stem cell reconstitution. Although all *MYC* constructs induced an increase in proliferation relative to uninfected controls (data not shown), wild-type and mutant *MYC* showed indistinguishable BrdU incorporation profiles (Fig. 2a; $25 \pm 3\%$ for wild-type MYC, $24 \pm 6\%$ for P57S and $27 \pm 5\%$ for T58A BrdU-positive cells). To examine the impact of *MYC* expression on apoptosis, we compared the ability of wild-type and mutant *MYC* to induce tumours in the presence of a strong apoptotic block. Recipient mice were reconstituted with stem cells co-infected with a retrovirus expressing *MYC* constructs and a retrovirus expressing the anti-apoptotic protein Bcl2. Notably, when co-expressed with *Bcl2*, wild-type and mutant *MYC* produced

[1]Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, New York 11724, USA. [2]Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, New York 10021, USA. [3]Department of Biochemistry, St Jude Children's Research Hospital, Memphis, Tennessee 38105, USA. [4]Howard Hughes Medical Institute, Cold Spring Harbor, New York 11724, USA.
*These authors contributed equally to this work.

aggressive B cell lymphomas with indistinguishable latency (34 ± 2 days for wild-type *MYC*, 35 ± 2 days for T58A and 33 ± 2 days for P57S) and penetrance (Fig. 2b, c). In addition, propidium iodide staining for DNA content revealed no differences in cell cycle profiles between lymphomas expressing wild-type and mutant *MYC* (Fig. 2d). Thus, in the absence of apoptosis, wild-type and mutant *MYC* are equally oncogenic.

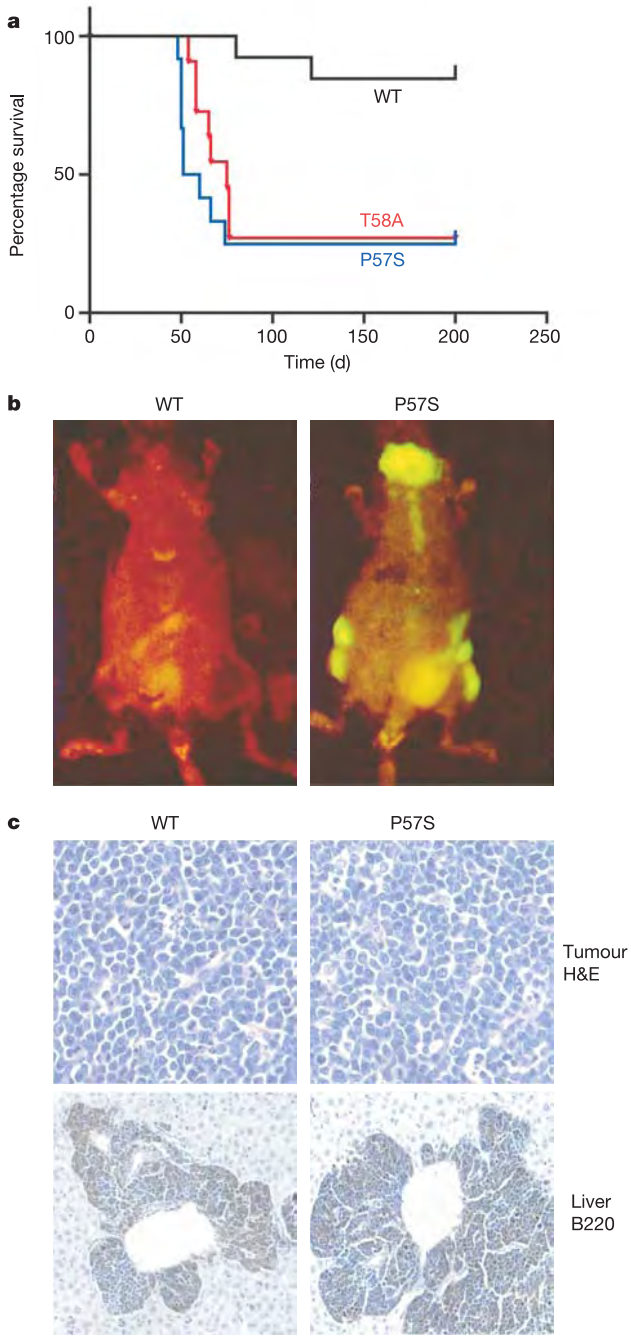


Figure 1 | Tumour-derived *MYC* mutants show enhanced oncogenicity *in vivo*. **a**, Kaplan–Meier curve showing survival at various times after adoptive transfer. **b**, *In vivo* GFP imaging shows disseminated lymphomas in a mouse 60 days after reconstitution with HSCs transduced with P57S, but not wild-type *MYC*. **c**, Haematoxylin/eosin staining and immunohistochemical staining for B220 (a B cell lineage marker) of lymph node and liver sections of animals harbouring wild-type *MYC* and P57S lymphomas, showing an aggressive B cell disease with perivascular infiltration of B220-positive tumour cells into the liver.

Mutant MYC induces p19^{ARF} and p53, but not Bim

Oncogene-induced apoptosis is often mediated by the induction of p19^{ARF} and subsequent stabilization of p53, which is thought to sense hyperproliferative signals and prevent aberrant proliferation[21]. To investigate the basis for the apoptotic defect of mutant MYC, we examined p19^{ARF} and p53 expression in murine embryonic fibro-

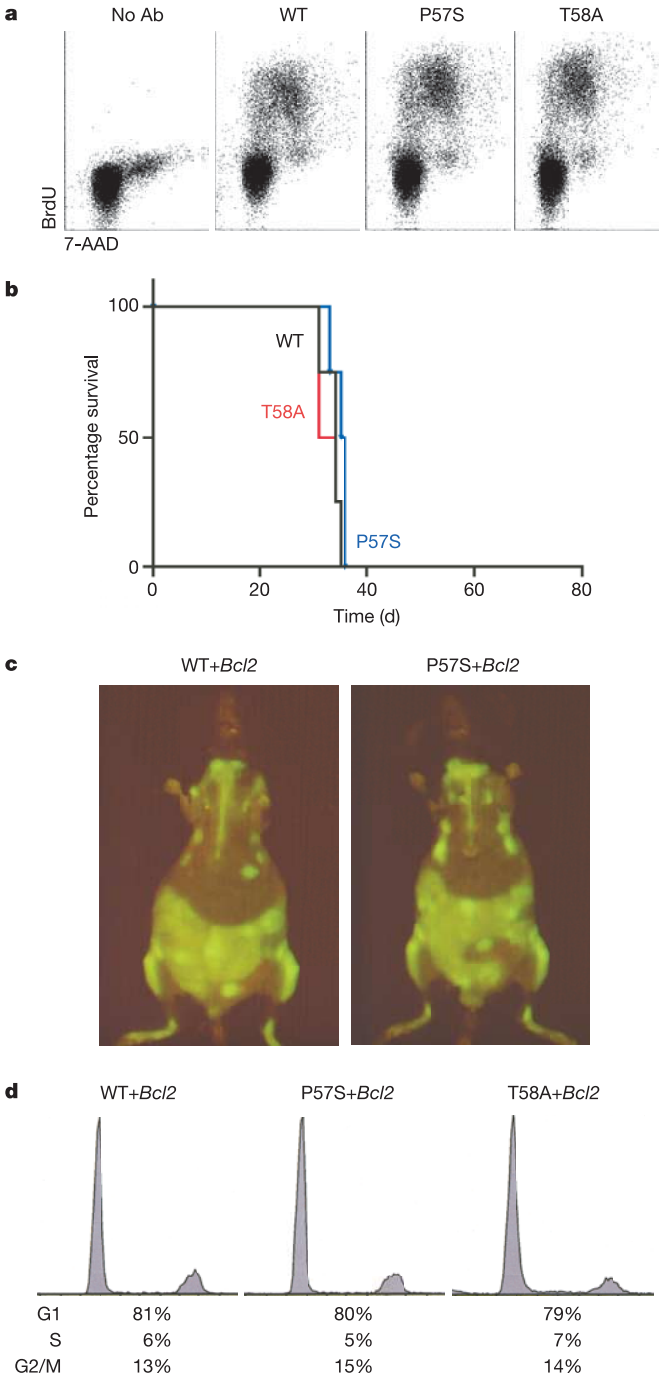


Figure 2 | Wild-type and mutant *MYC* show apoptotic, but not proliferative, differences *in vivo*. **a**, Flow cytometry showing BrdU incorporation in GFP-positive bone marrow cells 24 days after reconstitution. Data are representative of three independent experiments. **b**, Kaplan–Meier curve showing survival following adoptive transfer with HSCs co-infected with the indicated *MYC*-expressing retroviruses and a *Bcl2*-expressing retrovirus. **c**, *In vivo* GFP imaging showing disseminated lymphomas in mice 35 days after adoptive transfer with HSCs co-transduced with wild-type or mutant *MYC* and *Bcl2*. **d**, Representative histograms of DNA content of wild-type *MYC*/*Bcl2* and mutant *MYC*/*Bcl2* lymphomas.

blasts (MEFs) and HSC populations infected with wild-type and mutant *MYC* retroviruses. Interestingly, both P57S and T58A induced $p19^{ARF}$ and p53 as well as (or better than) wild-type *MYC* (Fig. 3a, c). This p53 was transcriptionally active, as cells expressing either wild-type or mutant MYC showed a similar increase in p53 phosphorylation and expression of the p53 transcriptional targets *Bax*, *Puma* and *Noxa* (Fig. 3c and data not shown). Furthermore, the apoptotic defect was unrelated to the established ability of *MYC* to repress Bcl2 and Bcl-XL[22], because no consistent effect of *MYC* on Bcl2 expression was observed and both wild-type and mutant *MYC* were capable of repressing Bcl-XL. Although $p21^{Cip1}$ levels were reduced in HSCs expressing wild-type MYC, they were substantially higher in cells expressing mutant MYC. However, this effect was not consistently seen in MEFs (Fig. 3c and data not shown), suggesting that the impact of $p21^{Cip1}$ repression on MYC-induced apoptosis[23] might contribute to, but is not essential for, the mutant MYC phenotype.

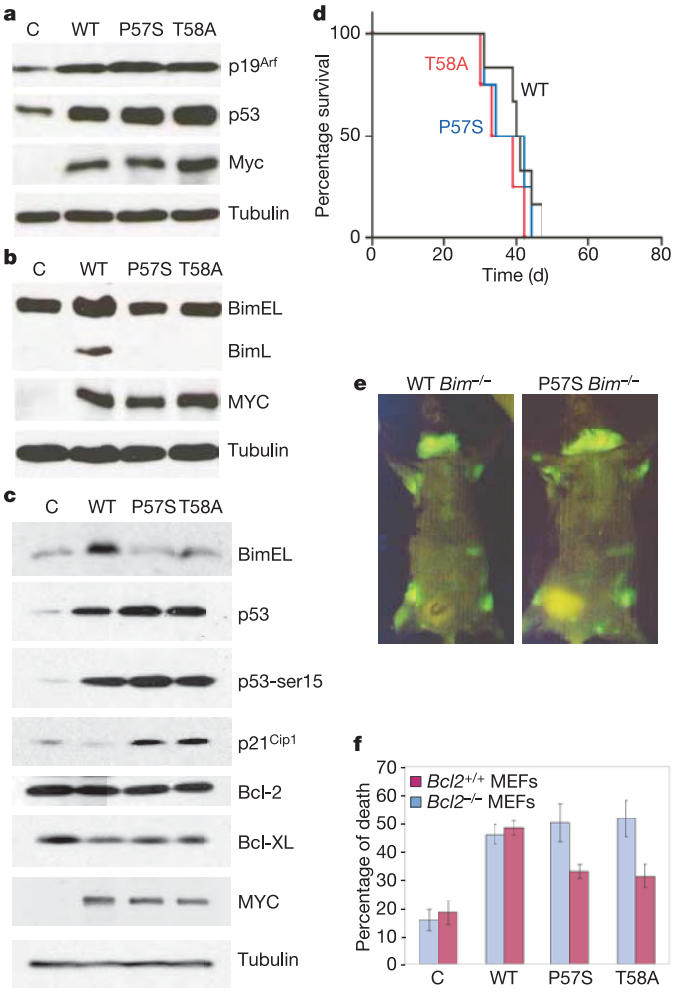


Figure 3 | Impaired Bim induction contributes to the increased oncogenicity of mutant *MYC* alleles. **a**, **b**, Western blot analysis showing ARF and p53 levels (**a**) or Bim levels (**b**) in MEFs stably infected with control (labelled C), wild-type *MYC* or mutant *MYC* vectors. The long (BimL) and extra long (BimEL) isoforms of Bim are shown. **c**, Western blot analysis of HSCs using antibodies against the indicated proteins. **d**, Kaplan–Meier curve showing free–free survival following adoptive transfer with $Bim^{-/-}$ HSCs infected with the indicated *MYC*-expressing retroviruses. **e**, *In vivo* GFP imaging showing disseminated lymphomas in mice 34 days after adoptive transfer with $Bim^{-/-}$ HSCs transduced with mutant or wild-type *MYC*. **f**, Wild type ($Bcl2^{+/+}$) and $Bcl2^{-/-}$ MEFs transduced with the indicated retroviruses were incubated in low serum and viability was assessed by trypan blue exclusion. Data represent the mean ± standard deviation of three independent experiments.

The pro-apoptotic, BH3-only protein Bim has recently been shown to inhibit *MYC*-induced lymphomagenesis, and Bim expression is elevated in B cells overexpressing MYC (ref. 24). Accordingly, Bim was acutely induced in response to wild-type *MYC* in MEFs, FL5.12 cells and HSCs (Fig. 3b, c; see also Supplementary Fig. 3a). This effect was not dependent on *MYC* induction of p53 because Bim was induced in $p53^{-/-}$ MEFs (data not shown). However, both tumour-derived *MYC* mutants were unable to induce Bim efficiently. Furthermore, mice reconstituted with $Bim^{-/-}$ stem cells expressing wild-type *MYC* produced lymphomas at a latency and penetrance indistinguishable from those reconstituted with $Bim^{-/-}$ stem cells expressing mutant *MYC* (Fig. 3d, e; 40 ± 5 days for wild-type *MYC*, 36 ± 6 days for T58A and 38 ± 6 days for P57S), implying that the failure of these mutant *MYC* alleles to induce Bim contributes to their increased oncogenicity. Thus, *MYC* mutations do not affect the 'sensing' of MYC-induced hyperproliferative signals by the $p19^{ARF}$–p53 pathway, but promote lymphomagenesis by abrogating a parallel apoptotic signal transmitted from MYC to Bim.

Bim was initially identified by virtue of its ability to bind Bcl2, and it acts as a potent inhibitor of the pro-survival members of the Bcl2 family[25]. Thus, failure of mutant *MYC* to induce Bim should result in a defective ability to suppress Bcl2 activity. Such a model predicts that suppression of Bcl2 activity might rescue the apoptotic defect of mutant *MYC*. To test this, we assessed the ability of wild-type and mutant *MYC* to promote apoptosis in cells derived from *Bcl2* knockout mice. Wild-type ($Bcl2^{+/+}$) and *Bcl2*-deficient ($Bcl2^{-/-}$) MEFs were infected with retroviruses expressing each *MYC* allele and then subjected to serum starvation to trigger apoptosis. In the presence of *Bcl2*, wild-type *MYC* consistently induced higher levels of apoptosis than either of the *MYC* point mutants (Fig. 3f). However, in $Bcl2^{-/-}$ MEFs, P57S and T58A were as capable of inducing apoptosis as wild-type *MYC*. Interestingly, apoptosis induced by wild-type *MYC* was not affected by *Bcl2* ablation, perhaps because wild-type MYC already efficiently inhibits Bcl2. Similarly, recipients of $Bcl2^{+/+}$ HSCs infected with wild-type *MYC* showed significantly less GFP-positive bone marrow after reconstitution compared with those receiving mutant *MYC*-infected stem cells; however, all *MYC*-expressing cells were strongly selected against during haematopoietic reconstitution with $Bcl2^{-/-}$ HSCs (see Supplementary Fig. 3b). Thus, in the absence of *Bcl2*, wild-type and mutant *MYC* were equally capable of promoting cell death, implying that *MYC* mutants—through reduced Bim induction and perhaps other additional effects—are ultimately deficient in their ability to inhibit Bcl2.

Evasion of p53 tumour suppressor function

These data suggest that the increased oncogenicity of tumour-derived *MYC* point mutants results from their inability to efficiently couple proliferation to cell death. Certain targeted apoptotic defects can obviate the requirement for tumour suppressor inactivation during tumorigenesis. For example, both *Bcl2* overexpression and *Bim* deficiency facilitate *MYC*-induced lymphomagenesis in the absence of $p19^{ARF}$ or *p53* mutations[24,26]. To determine whether *MYC* point mutations could also bypass the requirement for *p53* loss, we examined tumour onset and *p53* status in lymphomas derived from $p53^{+/-}$ stem cells. In this context, both wild-type and mutant *MYC* rapidly produced lymphomas with complete penetrance (Fig. 4a), although those produced by wild-type *MYC* arose with slightly longer latency (56 ± 24 days for wild-type *MYC*, 35 ± 3 days for T58A and 40 ± 9 days for P57S; $P < 0.001$ for both mutants compared with wild type). However, although tumours induced by wild-type *MYC* invariably lost the wild-type *p53* allele during tumour development (0 out of 9 with wild-type *p53* allele), those induced by the *MYC* mutants typically retained the wild-type *p53* allele (7 out of 9 with wild-type *p53* allele for both P57S and T58A; Fig. 4a, inset).

The requirement to lose the wild-type *p53* allele in tumours initiated by wild-type MYC may account for their delayed onset relative to tumours in recipients of mutant *MYC*-infected stem cells. Indeed, mice reconstituted with $p53^{-/-}$ HSCs expressing either the wild-type or mutant *MYC* formed tumours rapidly, with an identical penetrance and latency (Fig. 4b; 40 ± 8 days for wild-type *MYC*, 40 ± 4 days for T58A and 41 ± 5 days for P57S). All of the $p53^{-/-}$ lymphomas were oligoclonal (see Supplementary Fig. 3b), and showed no obvious differences between wild-type and mutant MYC expression (Fig. 4b, inset). Thus *MYC* mutations can facilitate tumour development in the absence of p53 inactivation.

MYC, *p53* and Bim status in Burkitt's lymphoma

These data predict that Bim expression should be reduced in Burkitt's lymphomas expressing *MYC* mutations, and that *p53* mutations should be under-represented in Burkitt's lymphomas harbouring the P57S and T58A mutant *MYC* alleles. Indeed, seven out of seven Burkitt's lymphomas expressing wild-type *MYC* were positive for Bim expression by immunohistochemistry, displaying substantially higher Bim levels than several diffuse, large B cell lymphomas (DLBCL) without *MYC* translocations (three out of three; Fig. 4c). In contrast, six out of seven Burkitt's lymphomas harbouring *MYC* alleles with MYC box I mutations (including four out of five tumours with P57 and T58 mutations) were scored as Bim negative (Fig. 4c). Furthermore, none of the *p53* mutant Burkitt's lymphomas (0 out of 17) that we analysed or identified from the literature[6,27,28] had coincident mutations at *MYC* codons 57 and 58, despite their common occurrence in our data set (12 out of 71) (see Supplementary Table; and data not shown). Of note, two out of two lymphomas harbouring a Pro to Ser mutation at codon 60 of *MYC* had a *p53* mutation at codon 248, suggesting that not all *MYC* mutations are equivalent. Nevertheless, our data demonstrate that two mutant *MYC* alleles commonly observed in Burkitt's lymphoma can bypass p53 action during lymphomagenesis in mice, and provide evidence to suggest that this may also occur in humans.

Discussion

Our results reveal new insights into the biology of the MYC oncoprotein. For example, they suggest that tumour-derived MYC mutants are relevant to the pathology of Burkitt's lymphoma because they are unable to upregulate Bim and efficiently inhibit Bcl2. Furthermore, they show that the MYC mutants we examined retain their ability to activate the $p19^{ARF}$–p53 pathway and efficiently promote proliferation. Thus, *MYC* mutations alter MYC activity in both quantitative[12,14] and, as shown here, qualitative ways. These qualitative changes may be particularly important for the development of tumours such as Burkitt's lymphoma, which acquire independent alterations that quantitatively upregulate MYC. Such qualitative changes to MYC may also be important for normal cell growth; in this regard, some tumour-derived *MYC* mutations lead to constitutive phosphorylation of MYC at Ser 62, a target of mitogen-stimulated kinases[14]. Perhaps *MYC* mutations 'lock in' a qualitative state that normally uncouples proliferation from apoptosis in response to mitogens, allowing tissue expansion.

Mutations that are selected for in malignant tumours pinpoint genes and processes that must be altered during tumour evolution. Thus, the fact that certain mutant *MYC* alleles selected for in human lymphomas are specifically impaired in their ability to promote apoptosis *in vivo* provides compelling genetic evidence that the phenomenon of oncogene-induced apoptosis is relevant to human tumorigenesis. Furthermore, our results demonstrate that the MYC-induced apoptotic programme is not linear, but instead involves p53-dependent and -independent signals that act in parallel to promote cell death and suppress MYC-induced tumorigenesis. That a cell expressing mutant MYC can activate p53 in response to hyperproliferative signalling but evade p53 action implies that MYC-induced apoptosis involves a threshold phenomenon, such that inactivation of any one of several MYC effectors can cause apoptotic firing to drop below the threshold level and allow unabated proliferation. Consequently, certain mutations in *MYC* alleviate the requirement for a cooperating gene mutation during tumorigenesis.

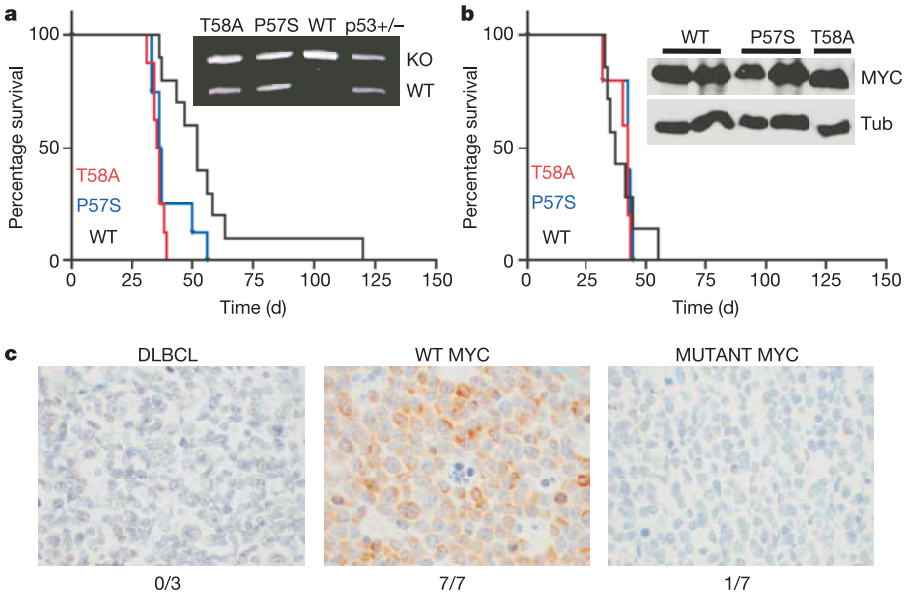


Figure 4 | Impact of mutations in *MYC* on p53 tumour suppressor action and Bim induction in human Burkitt's lymphomas. **a**, Kaplan–Meier curve showing mouse survival after adoptive transfer with $p53^{+/-}$ HSCs infected with the indicated *MYC* retroviral constructs. Inset, Representative allele-specific PCR assay showing retention of the wild-type *p53* allele in lymphomas produced by mutant but not wild-type *MYC*. **b**, Kaplan–Meier curve showing mouse survival after adoptive transfer with $p53^{-/-}$ HSCs infected as in **a**. Inset, western blot analysis of representative lymphomas harbouring the indicated *MYC* allele. **c**, Bim expression in DLBCL and wild-type *MYC*-expressing and mutant *MYC*-expressing Burkitt's lymphomas as assessed by immunohistochemisty using an anti-Bim antibody. The frequency of Bim-positive tumours in each category is indicated.

METHODS

Retroviral infection and cell culture. HAM (a methionine-rich variant of the widely used haemagglutinin tag)-tagged wild-type *MYC* and mutants were subcloned into MSCV-IRES-GFP[26] for *in vivo* studies and MSCVpuro (Clontech) for cell culture studies. Primary MEFs were generated from embryonic 13.5 (E13.5) day embryos from a wild-type C57BL/6J background or from a $Bcl2^{+/-}$ to $Bcl2^{+/-}$ cross. Retroviral-mediated gene transfer was performed using Phoenix packaging cells as previously described[29]. Cells for *in vivo* studies were infected with MSCV-GFP, MSCV-GFP-WT *MYC*, MSCV-GFP-P57S or MSCV-GFP-T58A. Infected cell populations of FL5.12 cells and MEFs used for *in vitro* studies were selected in puromycin ($2\,\mu g\,ml^{-1}$ for 2 days). Infected HSCs used for *in vitro* studies were diluted to equal infection efficiency (assessed by GFP) then sorted using magnetic beads to 70–80% purity with a lineage cell depletion kit (Miltenyi Biotec). For cell viability assays, MEFs were infected with *MYC*-expressing retroviruses as described, and were incubated for 24 h in either 10% fetal bovine serum (high serum) or 0.1% fetal bovine serum (low serum) and analysed for viability by trypan blue exclusion.

Stem cell isolation and adoptive transfer. Pregnant E14.5 C57BL/6 mice from wild-type, $p53^{+/-}$, $Bim^{+/-}$ or $Bcl2^{+/-}$ crosses were killed to obtain fetal livers. Stem cell infection and transplantation were performed as described[26]. For stem cell reconstitution and BrdU analysis, total bone marrow was isolated 21–28 days after tail vein injection. *In vivo* BrdU incorporation was assessed using an APC BrdU Flow Kit (BD Pharmingen). Bone marrow was harvested 2 h after an intraperitoneal injection of $200\,\mu l$ $10\,mg\,ml^{-1}$ BrdU, and BrdU incorporation was assessed for a population of GFP-positive cells at a defined GFP intensity. Flow cytometry analysis was performed on a Becton Dickinson LSRII cell analyser, with FACSVantage DiVa software.

Lymphoma monitoring and analysis. Reconstituted animals were monitored for illness by lymph node palpation, monitoring overall morbidity, and, in some cases, whole-body fluorescence imaging[26]. Overall survival was defined as the time from stem cell reconstitution until the animal reached a terminal stage and was killed. Statistical analysis was performed using a one-way ANOVA (analysis of variance) test using Graph Pad Prism version 3.0 (Graph Pad Software). Immunohistochemistry was performed using CD45R/B220–clone RA3-6B2 (BD Biosciences, Pharmingen) and rabbit anti-Bim (Stressgen) antibodies. Tumour cell DNA content was determined by FACS analysis with propidium iodide staining of ethanol-fixed cells[26]. Nested PCR analysis of variable, diversity and joining (V(D)J) recombination was performed as described[30].

Protein analysis. Immunoblots were performed from whole-cell lysates obtained by boiling cell pellets that were solubilized in Laemmli sample buffer. Twenty micrograms of protein samples (Bio-Rad protein assay) were separated by SDS polyacylamide electrophoresis (SDS–PAGE) and transferred to Immobilon-P membranes (Millipore). The antibodies used were anti-MYC (Ab1, Oncogene; 1:250), anti-$p19^{ARF}$ (ab80-100, Abcam; 1:1,000), anti-p21 (C-19, Santa Cruz; 1:500), anti-p53 (NCL-p53-505, Novocastra; 1:500), anti-phospho-p53 (Ser 15) (16G8, Cell Signaling; 1:1,000), anti-Bim/BOD (AAP-330, Stressgen; 1:1,000), anti-Bcl2 (N-19, Santa Cruz; 1:1,000) and anti-Bcl-XL (S-18, Santa Cruz; 1:1,000). Proteins were visualized using ECL (Amersham) or Lumi-light (Roche).

Received 1 April; accepted 23 May 2005.

1. Dang, C. V. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol. Cell. Biol.* **19,** 1–11 (1999).
2. Cole, M. D. Activation of the c-myc oncogene. *Basic Life Sci.* **38,** 399–406 (1986).
3. Spencer, C. A. & Groudine, M. Control of c-myc regulation in normal and neoplastic cells. *Adv. Cancer Res.* **56,** 1–48 (1991).
4. Dalla-Favera, R. *et al.* Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. *Proc. Natl Acad. Sci. USA* **79,** 7824–7827 (1982).
5. Davis, M., Malcolm, S. & Rabbitts, T. H. Chromosome translocation can occur on either side of the c-myc oncogene in Burkitt lymphoma cells. *Nature* **308,** 286–288 (1984).
6. Bhatia, K. *et al.* Point mutations in the c-Myc transactivation domain are common in Burkitt's lymphoma and mouse plasmacytomas. *Nature Genet.* **5,** 56–61 (1993).
7. Albert, T., Urlbauer, B., Kohlhuber, F., Hammersen, B. & Eick, D. Ongoing mutations in the N-terminal domain of c-Myc affect transactivation in Burkitt's lymphoma cell lines. *Oncogene* **9,** 759–763 (1994).
8. Clark, H. M. *et al.* Mutations in the coding region of c-MYC in AIDS-associated and other aggressive lymphomas. *Cancer Res.* **54,** 3383–3386 (1994).
9. Henriksson, M., Bakardjiev, A., Klein, G. & Luscher, B. Phosphorylation sites mapping in the N-terminal domain of c-myc modulate its transforming potential. *Oncogene* **8,** 3199–3209 (1993).
10. Hoang, A. T. *et al.* A link between increased transforming activity of lymphoma-derived MYC mutant alleles, their defective regulation by p107, and altered phosphorylation of the c-Myc transactivation domain. *Mol. Cell. Biol.* **15,** 4031–4042 (1995).
11. Westaway, D., Payne, G. & Varmus, H. E. Proviral deletions and oncogene base-substitutions in insertionally mutagenized c-myc alleles may contribute to the progression of avian bursal tumors. *Proc. Natl Acad. Sci. USA* **81,** 843–847 (1984).
12. Salghetti, S. E., Kim, S. Y. & Tansey, W. P. Destruction of Myc by ubiquitin-mediated proteolysis: cancer-associated and transforming mutations stabilize Myc. *EMBO J.* **18,** 717–726 (1999).
13. Chang, D. W., Claassen, G. F., Hann, S. R. & Cole, M. D. The c-Myc transactivation domain is a direct modulator of apoptotic versus proliferative signals. *Mol. Cell. Biol.* **20,** 4309–4319 (2000).
14. Sears, R. C. The life cycle of c-Myc: from synthesis to degradation. *Cell Cycle* **3,** 1133–1137 (2004).
15. Frykberg, L., Graf, T. & Vennstrom, B. The transforming activity of the chicken c-myc gene can be potentiated by mutations. *Oncogene* **1,** 415–422 (1987).
16. Yeh, E. *et al.* A signalling pathway controlling c-Myc degradation that impacts oncogenic transformation of human cells. *Nature Cell Biol.* **6,** 308–318 (2004).
17. Rabbitts, T. H., Hamlyn, P. H. & Baer, R. Altered nucleotide sequences of a translocated c-myc gene in Burkitt lymphoma. *Nature* **306,** 760–765 (1983).
18. Bemark, M. & Neuberger, M. S. The c-MYC allele that is translocated into the IgH locus undergoes constitutive hypermutation in a Burkitt's lymphoma line. *Oncogene* **19,** 3404–3410 (2000).
19. Adams, J. M. *et al.* The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice. *Nature* **318,** 533–538 (1985).
20. Evan, G. I. *et al.* Induction of apoptosis in fibroblasts by c-myc protein. *Cell* **69,** 119–128 (1992).
21. Lowe, S. W. & Sherr, C. J. Tumor suppression by Ink4a-Arf: progress and puzzles. *Curr. Opin. Genet. Dev.* **13,** 77–83 (2003).
22. Maclean, K. H., Keller, U. B., Rodriguez-Galindo, C., Nilsson, J. A. & Cleveland, J. L. c-Myc augments gamma irradiation-induced apoptosis by suppressing Bcl-XL. *Mol. Cell. Biol.* **23,** 7256–7270 (2003).
23. Seoane, J., Le, H. V. & Massague, J. Myc suppression of the p21(Cip1) Cdk inhibitor influences the outcome of the p53 response to DNA damage. *Nature* **419,** 729–734 (2002).
24. Egle, A., Harris, A. W., Bouillet, P. & Cory, S. Bim is a suppressor of Myc-induced mouse B cell leukemia. *Proc. Natl Acad. Sci. USA* **101,** 6164–6169 (2004).
25. O'Connor, L. *et al.* Bim: a novel member of the Bcl-2 family that promotes apoptosis. *EMBO J.* **17,** 384–395 (1998).
26. Schmitt, C. A. *et al.* Dissecting p53 tumour suppressor functions *in vivo*. *Cancer Cell* **1,** 289–298 (2002).
27. Gaidano, G. *et al.* p53 mutations in human lymphoid malignancies: association with Burkitt lymphoma and chronic lymphocytic leukemia. *Proc. Natl Acad. Sci. USA* **88,** 5413–5417 (1991).
28. Bhatia, K. G., Gutierrez, M. I., Huppi, K., Siwarski, D. & Magrath, I. T. The pattern of p53 mutations in Burkitt's lymphoma differs from that of solid tumors. *Cancer Res.* **52,** 4273–4276 (1992).
29. Serrano, M., Lin, A. W., McCurrach, M. E., Beach, D. & Lowe, S. W. Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell* **88,** 593–602 (1997).
30. Yu, D. & Thomas-Tikhonenko, A. A non-transgenic mouse model for B-cell lymphoma: *in vivo* infection of p53-null bone marrow progenitors by a Myc retrovirus is sufficient for tumorigenesis. *Oncogene* **21,** 1922–1927 (2002).

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

Acknowledgements We thank M. S. Jiao, C. Rosenthal, M. Yang, S. Ray, C. Yang and I. Linkov for technical assistance and J. Zilfou, R. Dickins, E. Cepero, M. Spector, J. Fridman and other members of the Lowe laboratory for advice and/or critical reading of the manuscript. We also thank A. Hunt and G. Evan for helpful advice, and T. Mak and A. Strasser for mice. This work was supported by a postdoctoral fellowship from the Helen Hay Whitney Foundation (M.T.H.), an NCI postdoctoral training grant (A.B.), a program project grant from the National Cancer Institute (W.P.T and S.W.L.) and a grant from the Irving Hansen Memorial Foundation (W.P.T). W.P.T. is a Leukemia and Lymphoma Society Scholar and S.W.L. is an AACR-NCRF Research Professor.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W.L. (lowe@cshl.org).

ARTICLES

A contractile nuclear actin network drives chromosome congression in oocytes

Péter Lénárt[1]†, Christian P. Bacher[2], Nathalie Daigle[1], Arthur R. Hand[3], Roland Eils[2], Mark Terasaki[4,5] & Jan Ellenberg[1]

Chromosome capture by microtubules is widely accepted as the universal mechanism of spindle assembly in dividing cells. However, the observed length of spindle microtubules and computer simulations of spindle assembly predict that chromosome capture is efficient in small cells, but may fail in cells with large nuclear volumes such as animal oocytes. Here we investigate chromosome congression during the first meiotic division in starfish oocytes. We show that microtubules are not sufficient for capturing chromosomes. Instead, chromosome congression requires actin polymerization. After nuclear envelope breakdown, we observe the formation of a filamentous actin mesh in the nuclear region, and find that contraction of this network delivers chromosomes to the microtubule spindle. We show that this mechanism is essential for preventing chromosome loss and aneuploidy of the egg—a leading cause of pregnancy loss and birth defects in humans.

In dividing eukaryotic cells, microtubules are believed to be involved in all aspects of spindle assembly[1]. In prophase, long interphase microtubules depolymerize[2,3] and centrosomes nucleate M-phase asters of short and dynamic microtubules. After nuclear envelope breakdown (NEBD), astral microtubules invade the nuclear region, 'search and capture' chromosomes and then align them on the spindle[4,5]. Rapid turnover of microtubules is essential for spindle assembly[5,6], but limits their maximal length to about 20 μm (refs 7, 8). Consequently, chromosome capture is inefficient at distances greater than 40 μm away from the centrosomes[9]. Chromosomes are not completely passive during this process but may facilitate capture by generating a high local concentration of RanGTP[10] and/or by directly nucleating microtubules on their surface[11]. Computer simulations have shown that with such bias towards chromosomes, 'search and capture' can indeed function in somatic cells at experimentally observed timescales[12]. However, the model also predicts microtubule capture to be inefficient in cells with nuclei larger than 30 μm in diameter[12].

Animal oocytes store large amounts of nuclear proteins and RNA for early embryonic divisions, and thus have much larger nuclei. Starfish oocytes are a typical example and are excellent specimens for live-cell confocal microscopy owing to their transparency and fast, reliable maturation. Oocytes arrested at meiotic prophase I are 170 μm in diameter and contain a large, 80-μm nucleus at the animal pole of the cell. Typically for oocytes, condensed chromosomes (1–2 μm in size) are randomly scattered across the nucleoplasm. The two centrosomes are located at the animal pole, sandwiched between the cell cortex and the nuclear envelope[13]. After addition of the maturation hormone 1-methyladenine, oocytes re-enter meiosis, leading to NEBD in 20 min (ref. 14). After NEBD, chromosomes move towards the centrosomes and form the first meiotic spindle under the cell cortex, 40 min after hormone addition. Extrusion of the first polar body 1 h after hormone treatment then completes meiosis I[13].

During meiosis I, chromosomes in the oocytes of starfish and many other animals have to travel much longer distances to the spindle than in somatic cells. Consequently, microtubule 'search and capture' is likely to be inefficient. Here we show that in starfish oocytes, chromosome congression is driven by a contractile actin network. This network forms and fills the nuclear region after NEBD, and by its contraction delivers the embedded chromosomes to within the ~30-μm capture range of microtubule asters. We show that this mechanism is essential, because actin-depolymerizing or stabilizing drugs cause chromosome loss and lead to the formation of aneuploid eggs.

Asters are too short to capture chromosomes

First, we analysed meiotic chromosome movement by imaging microtubules and chromosomes in live as well as in fixed oocytes during maturation, when chromosomes travel up to 80 μm to reach the meiotic spindle. Before NEBD, we visualized a large array of cytoplasmic microtubules anchoring the nucleus to the animal pole[15] (Fig. 1a, time −3:24 and Supplementary Video S1). These microtubules rapidly depolymerized at NEBD and centrosomes started to nucleate short M-phase asters (Fig. 1a, b) with a maximal length of 15–20 μm (Fig. 1c). Nevertheless, all chromosomes, including those more than 40 μm away from the closest microtubule end, moved towards the animal pole immediately after NEBD (Fig. 1c and Supplementary Video S1). This movement occurred 5–10 min before microtubules contacted those distal chromosomes, as visualized in live cells or by immunofluorescence (Fig. 1a–c), suggesting that a mechanism other than microtubule capture is responsible for chromosome congression in meiosis I. To test this directly, we depolymerized microtubules using nocodazole (Supplementary Fig. S1a). We found that chromosome congression proceeded almost normally in the absence of microtubules. All chromosomes gathered in a compact group under the cell cortex at the animal pole, as for untreated cells (Fig. 2a, b).

[1]Gene Expression and Cell Biology/Biophysics Programmes, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [2]Theoretical Bioinformatics B080, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany. [3]Departments of Orthodontics, Oral and Maxillofacial Surgery, Pediatric Dentistry and Advanced General Dentistry, and [4]Department of Physiology, University of Connecticut Health Center, Farmington, Connecticut 06030, USA. [5]Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA. †Present address: Research Institute of Molecular Pathology (IMP), Dr Bohr-Gasse 7, A-1030 Vienna, Austria.

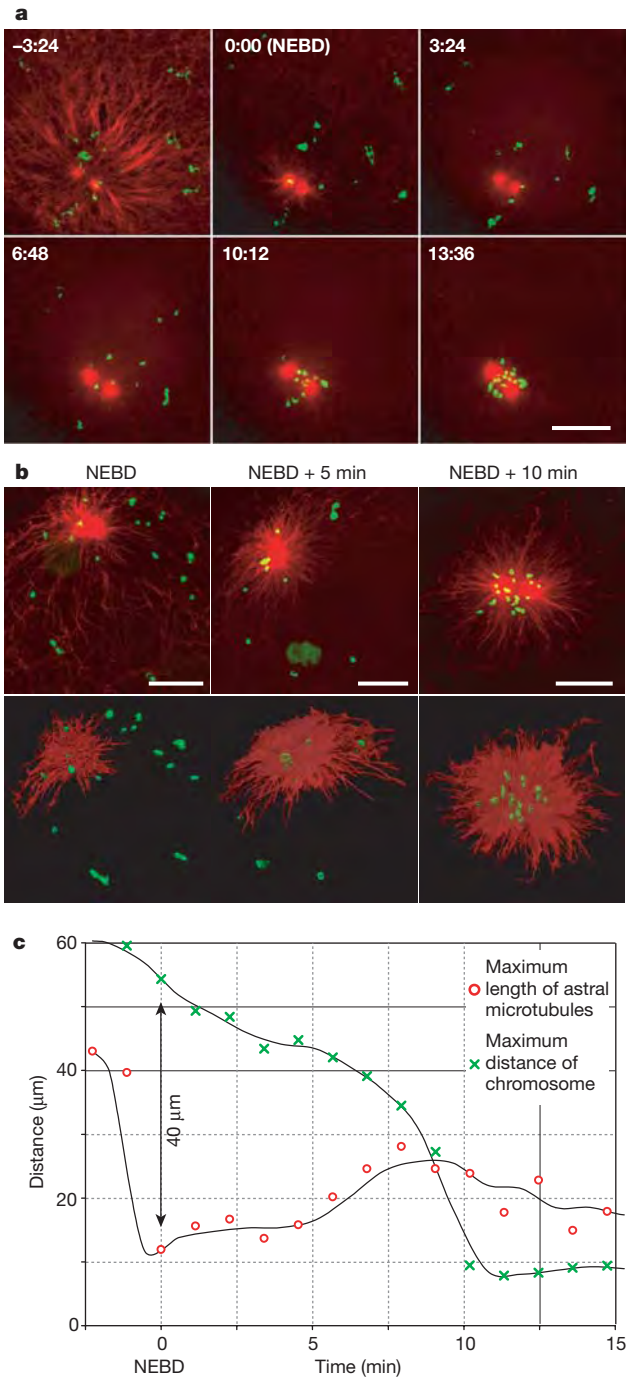


Figure 1 | Chromosomes congress without attachment to meiotic asters. **a**, 3D time-lapse (4D) imaging of a maturing oocyte expressing ensconsin–3EGFP (microtubules, red) and injected with Alexa-568-Ran (chromosomes, green) (20 sections every 4 μm, stacks of images taken every 68 s). Images show *Z*-projections of selected time points; a full series is shown in Supplementary Video S1. Time shown in minutes and seconds. Scale bar, 20 μm. **b**, Immunofluorescence of oocytes. *Z*-projections (top row) and isosurface reconstructions (bottom row). Microtubules shown in red, DNA in green. Scale bar, 20 μm. Residual interphase microtubules not connected to centrosomes, and nucleoli were removed in isosurface reconstructions. **c**, Quantification of data in **a**, showing the maximal distance of chromosomes from the midpoint between centrosomes (green crosses) and the maximal length of astral microtubules (red circles) over time. In fixed oocytes, microtubules 5–10 μm longer were occasionally found. Lines show running average fits.

Chromosome congression is actin-dependent

As microtubules were not required for long-range directional chromosome movement, we tested other cytoskeletal components for their involvement. Strikingly, when the actin cytoskeleton was disrupted during maturation by latrunculin B (Supplementary Fig. S1b), congression failed in 75% of the cells (Fig. 2a, b). It seems that only chromosomes close enough to centrosomes were captured by microtubules, while more distal chromosomes were lost (Fig. 2a). As chromosomes are positioned randomly in the nucleus, chromosomes in the remaining 25% of the cells were probably already in close proximity to the centrosomes before NEBD, and thus could all be captured by microtubules. To exclude the possibility that chromosomes in these oocytes congressed independently of both actin and microtubules, we also treated cells with a combination of nocodazole and latrunculin B. This caused complete failure of chromosome congression, and chromosomes remained scattered in the nuclear region in 100% of the cells (Fig. 2a, b).

To characterize congression defects caused by latrunculin B treatment, we quantified chromosome movement in live oocytes expressing histone H2B fused to a red fluorescent protein (H2B–diHcRed[16]). In three-dimensional time-lapse (4-dimensional) data sets, chromosomes were reconstructed by isosurface rendering and tracked (Fig. 2c, see Supplementary Videos S2–S5 for videos and reconstructions). In untreated cells, tracking revealed two distinct phases of movement. At NEBD, chromosomes started to move synchronously with a speed of $\sim 3\,\mu m\,min^{-1}$ in the general direction of the animal pole, as if connected to each other (Fig. 2d, e). Between 5 and 15 minutes after NEBD, when chromosomes entered the range of short astral microtubules 20–40 μm away from the animal pole, they individually switched to a faster speed ($>12\,\mu m\,min^{-1}$), changed direction and moved straight towards the centrosomes (arrows in Fig. 2d, e; Supplementary Video S2). Nocodazole completely prevented this second fast movement (Fig. 2d, e), which thus corresponds to chromosome capture by microtubules. In the absence of microtubules, chromosomes nevertheless continued to move at the slow speed, eventually reaching the animal pole (Fig. 2d, e and Supplementary Video S3), confirming our observations in fixed cells (Fig. 2a, b).

The initial slower phase of microtubule-independent chromosome movement was markedly affected by depolymerizing actin with latrunculin B. Long incubation periods with low doses (120 nM, which is still compatible with cell integrity) greatly decreased the initial movement (Supplementary Fig. S2). However, the residual motion was often sufficient to move chromosomes within 40 μm of the centrosomes and microtubules were still able to capture many chromosomes (Supplementary Video S9). Higher concentrations of latrunculin B (2 μM), which completely prevent actin polymerization, could be used if added just a few minutes before NEBD. This completely abolished movement of distal chromosomes (asterisks in Fig. 2d, e), which did not reach the animal pole even by the time the few proximal chromosomes that had been captured by microtubules entered anaphase (Supplementary Video S4). Despite chromosome loss, anaphase initiated without delay, indicating that the spindle checkpoint is easily bypassed or might not be active in meiosis I of starfish oocytes. Failure of the actin-dependent phase of chromosome congression thus led to formation of an aneuploid egg.

Treating oocytes with the combination of nocodazole and latrunculin B prevented any directional movement of chromosomes (Fig. 2d, e and Supplementary Video S5), demonstrating that microtubules and actin account for all chromosome dynamics observed in untreated cells. Results similar to those with latrunculin B were also obtained with 1 μM cytochalasin D (refs 17, 18 and data not shown).

NEBD triggers nuclear actin polymerization

We have shown that an actin-dependent mechanism is necessary for

delivering distal chromosomes to within the capture distance of centrosomal microtubules to prevent chromosome loss. To visualize the structures responsible for this actin-dependent congression in live oocytes we injected them with fluorescently labelled actin. Before NEBD, a significant fraction of fluorescent actin was localized to the nucleus, but we did not detect any nuclear actin structures (Fig. 3a) or observe staining of polymerized actin by phalloidin in nuclei of fixed cells (not shown). Additionally, before NEBD, actin

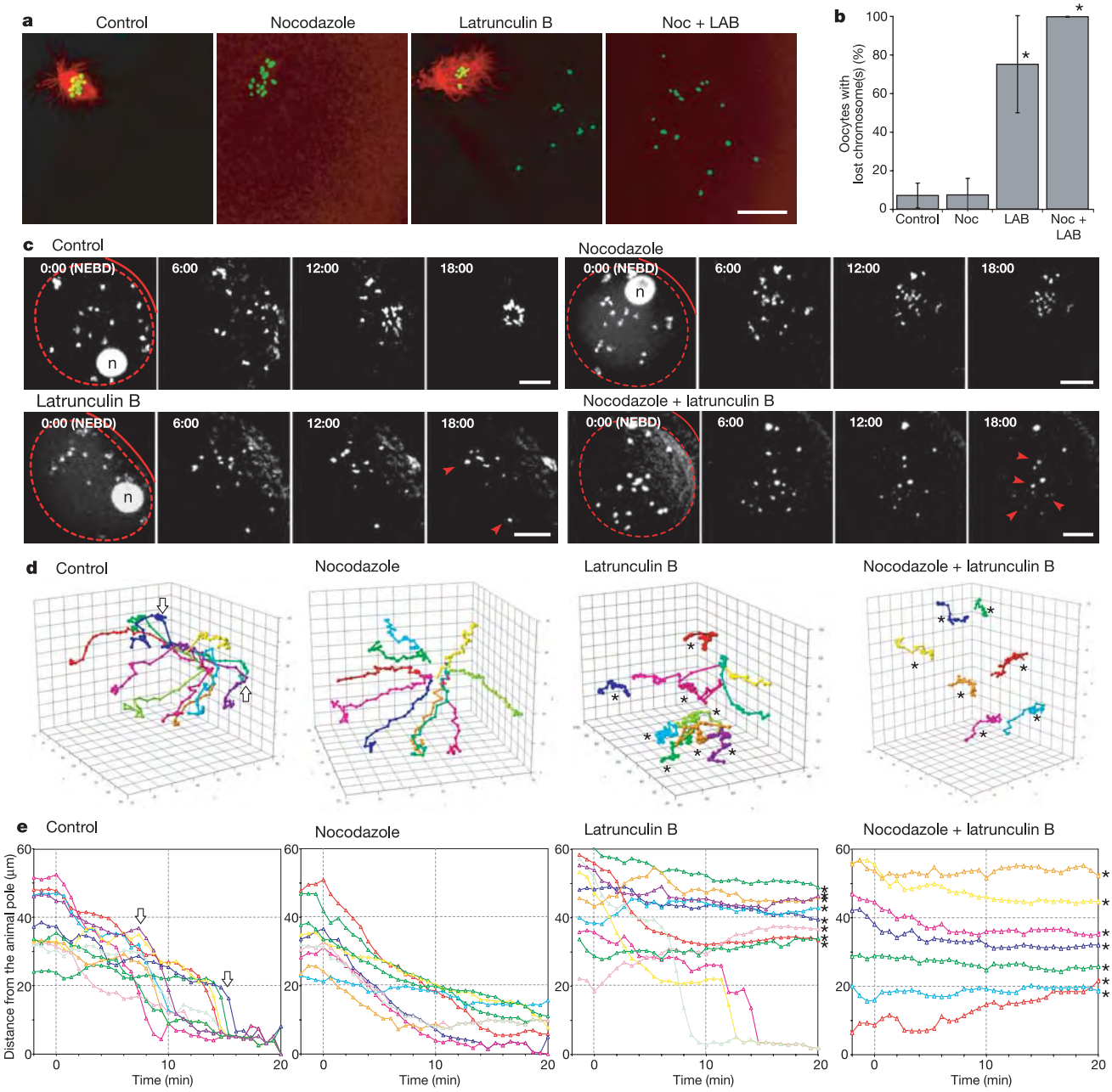


Figure 2 | Chromosome congression is independent of microtubules but relies on actin polymerization. **a**, Immunofluorescence of oocytes matured without treatment or in 3.3 μM nocodazole (Noc), 250 nM latrunculin B (LAB) or a combination of both. *Z*-projections of oocytes fixed 30 min after NEBD. Microtubules shown in red, DNA in green. Scale bar, 20 μm. **b**, Batches of oocytes were fixed 30 min after NEBD, stained for DNA and scored for the presence of all chromosomes at the animal pole. Averages and standard deviations from four independent experiments are shown ($n \geq 50$ each). Asterisks mark significant differences from control oocytes ($P < 0.05$, Student's *t*-test). Variations originate mainly as a result of differences in the ages of animals. **c**, 4D imaging of chromosomes in live oocytes (20 slices every 3 μm, image stacks taken every 40 s), showing *Z*-projections of selected time points. Solid line, cortex; dashed line, nucleus; n, nucleolus (disassembling shortly after NEBD). Red arrowheads show examples of lost chromosomes. Oocytes treated with nocodazole were pre-incubated for 1 h and matured in 3.3 μM nocodazole. For oocytes treated with latrunculin B, 2 μM latrunculin B was added 10 min before NEBD. For treatment with both nocodazole and latrunculin B, oocytes were matured in 3.3 μM nocodazole and 250 nM latrunculin B. Chromosomes were labelled by expressing H2B–diHcRed (which also labels the nucleolus) or by injecting Alexa-568-Ran (nocodazole + latrunculin B). Time shown in minutes and seconds. Scale bar, 20 μm. Full videos, reconstructions and tracking are available in Supplementary Videos S2–S5. **d**, 3D chromosome tracks for the data sets shown in **c**. **e**, Distances of chromosomes from the animal pole over time, calculated from **d**. Corrected for translational movement of the oocyte. Arrows in **d** and **e** highlight coinciding changes in speed and direction; lost chromosomes are labelled with asterisks. A high dose (2 μM) of latrunculin B caused rotation of the cell cortex at late times, distorting tracks of chromosomes close to the cell surface (for example, red, yellow and pale green tracks).

polymerization could not be induced by perforating the nucleus with a microneedle, by injecting phalloidin into oocyte nuclei or by overexpressing green fluorescent protein (GFP)-labelled actin artificially targeted to the nucleus by a nuclear localization signal (not shown), indicating that M-phase progression is required for actin polymerization in the nuclear region.

In contrast, exactly correlated in space and time with NEBD, we observed massive actin polymerization at the nuclear rim, which electron microscopy revealed to be directly underneath the nuclear envelope (Fig. 3c, d). The polymerization wave initiated 30–60 s before the fragmentation of nuclear membranes (visualized by the entry of a 160-kDa fluorescent dextran[14]) and also depleted the soluble nuclear actin pool (Fig. 3a, b). At this stage, nuclear pore complexes are largely disassembled and the nuclear envelope is permeable to smaller dextrans similar in size to actin as well as proteins up to several hundred kDa (Fig. 3a, b and ref. 14). We therefore conclude that after M-phase progression has been initiated, it is simply the mixing of nuclear and cytoplasmic compartments caused by NEBD that induces actin polymerization. That actin polymerization is a consequence, rather than a cause of NEBD, is also indicated by the fact that latrunculin B treatment did not affect NEBD (not shown).

An F-actin network moves chromosomes

The initial actin polymerization at the site of nuclear envelope disruption rapidly extended to a shell engulfing the entire nucleus (Figs 3a and 4a). The actin shell then depolymerized again within 1–2 min except for dense actin patches around chromosomes (Fig. 4a, b). We could observe filaments connecting these patches, but the contrast was low owing to the large soluble pool of fluorescent G-actin. To specifically detect polymerized actin we expressed an enhanced (E)GFP-tagged F-actin-binding domain of the actin-binding protein ABP-120 (EGFP–ABD, ref. 19). This enabled us to clearly visualize a network of actin filaments in the nucleoplasmic region connecting the patches around chromosomes (Fig. 4b, c and Supplementary Videos S6, S7). We confirmed identical structures by staining with phalloidin in fixed oocytes (Fig. 4e and also seen in part earlier, ref. 20) or by microinjecting phalloidin into live oocytes (Fig. 4f). High resolution time-lapse imaging of EGFP–ABD showed that contraction of the actin network delivered embedded chromosomes to the animal pole (Supplementary Videos S6, S7). The actin patches persisted around chromosomes during the low speed motion (Fig. 4d), and disappearance of the patch was shortly followed by fast microtubule capture (Fig. 4d and Supplementary Video S6). This provides a direct explanation for the actin-dependent phase of

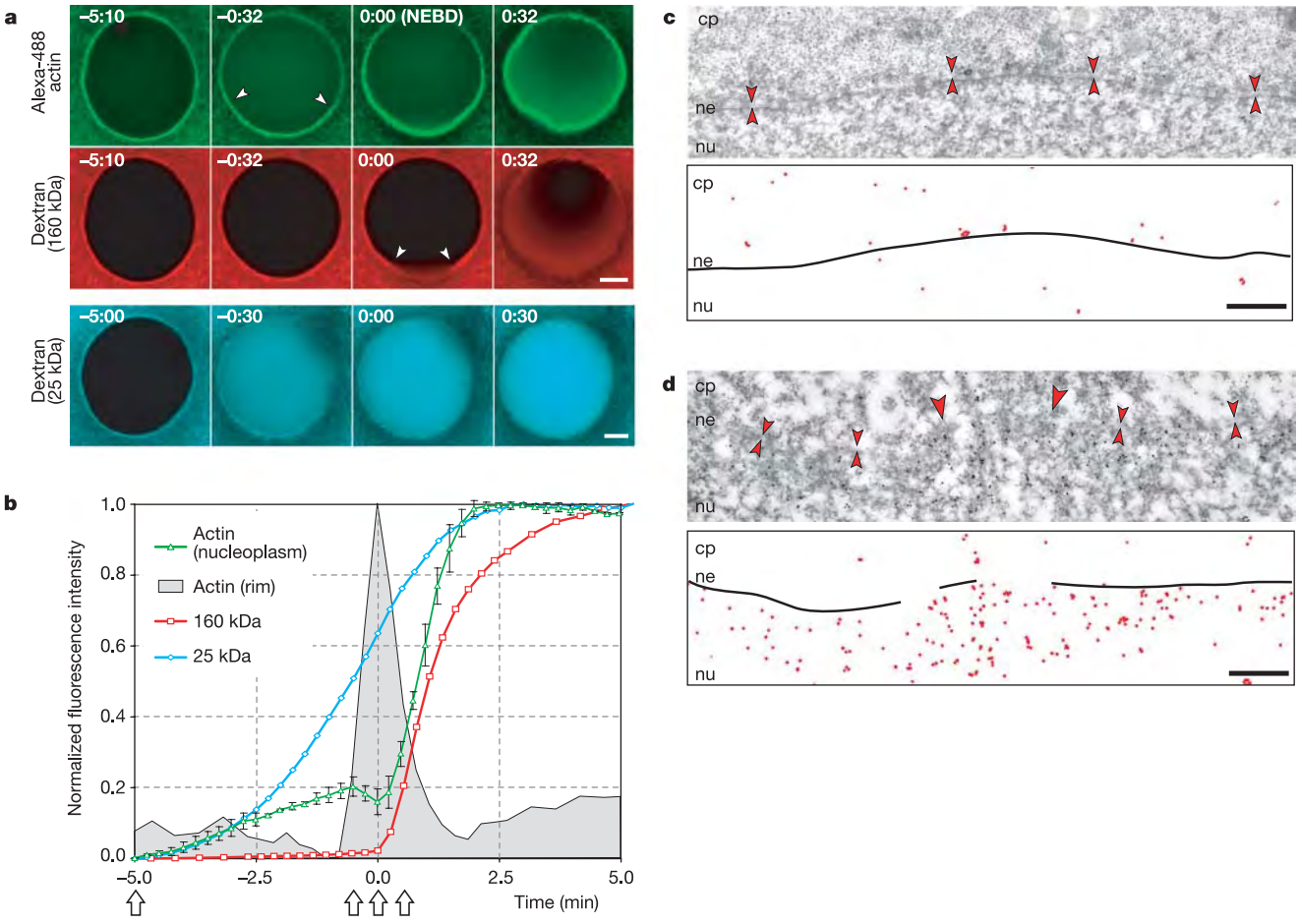


Figure 3 | Actin polymerization initiates at NEBD. **a**, Oocytes injected with Alexa-488-actin and 160 or 25 kDa TRITC-labelled dextran. Shown are single optical sections along the equator of the nucleus. Sequences were aligned to the initiation of the actin wave. The actin channel is omitted for the 25 kDa data set. Arrowheads denote region of initial actin polymerization or dextran entry. Time shown in minutes and seconds. Scale bar, 20 μm. **b**, Quantification of **a**, showing entry kinetics of actin and dextran as mean fluorescence intensities in the nuclear region (for actin average and standard deviation of the 160 and 25 kDa experiments). The nuclear rim was also quantified for actin polymerization. Data were normalized from minimum to maximum values. Arrows mark time points shown in **a**. **c**, **d**, Electron micrographs of immunogold labelling with an anti-actin antibody on a thin section fixed before (**c**) or at the time of (**d**) NEBD. Double arrowheads mark the nuclear envelope; cp, cytoplasm; ne, nuclear envelope; nu, nucleus. Scale bar, 1 μm. Schematics show gold particles and an outline of the nuclear envelope. Single arrowheads in **d** mark nuclear envelope fragments after fenestration of the nuclear envelope[14].

chromosome congression. Furthermore, beads coated with plasmid DNA injected into oocyte nuclei triggered the formation of actin patches similar to chromosomes but empty beads had no effect, indicating a chromatin-specific mechanism of actin nucleation (Fig. 4g, Supplementary Fig. S3 and Supplementary Videos S10, S11).

Depolymerization is required for contraction

Using any of the methods of visualization, the F-actin network collapsed towards the animal pole over time, suggesting that either bundling of filaments or depolymerization could drive its contraction. To test the latter hypothesis, we examined the effect of two drugs (phalloidin and jasplakinolide) that stabilize actin filaments[21]. Cells treated with either of the compounds showed severe chromosome loss, similar to latrunculin B (Fig. 5a, b). Importantly, both phalloidin and jasplakinolide were also effective if applied after NEBD, indicating that stabilization of actin filaments can prevent contraction of the already formed network (Fig. 5a, b). In control cells, the network completely depolymerized in less than 10 min (Fig. 4b and Supplementary Videos S6, S7). In contrast, in cells injected with rhodamine-phalloidin, the network was still largely intact even 20 min after NEBD (Fig. 5b and Supplementary Video S8). Thus, actin depolymerization is required for contraction of the network. Moreover, time-lapse imaging of the filament network (Supplementary Video S7) revealed streaming of thin filaments towards the animal pole. We therefore propose that the directionality of the contraction could be due to a local depolymerization of actin around the centrosomes and/or anchoring of the nuclear actin network to the cell cortex at the animal pole[22]. This would result in a 'fishnet' like mechanism, in which the three-dimensional actin net is 'pulled in' from the animal pole, carrying with it the attached chromosomes.

Although we could not observe filament thickening by live imaging of the actin network (Supplementary Video S7), we cannot rule out that myosin II-driven bundling contributes to the contraction, as it has been reported in morphologically similar actin filament networks in *Xenopus* egg extracts[23]. In such a model, the nucleoplasm of starfish oocytes could behave in a manner similar to

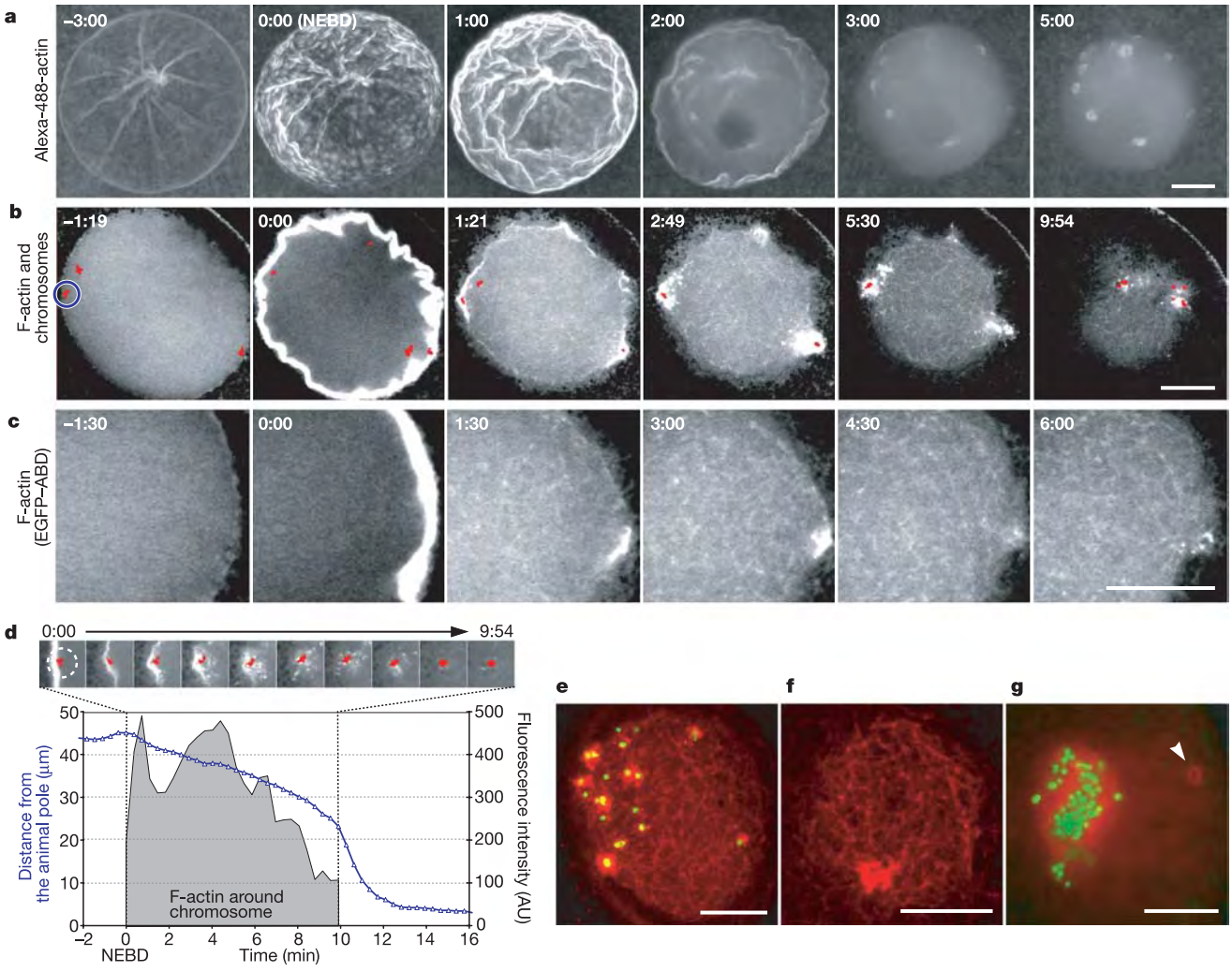


Figure 4 | Actin polymerizes into a contractile network in the nuclear region that moves chromosomes to the animal pole. **a**, 4D imaging of a maturing oocyte injected with Alexa-488-actin (15 sections every 2.75 μm, image stacks taken every 60 s), showing *Z*-projections of selected time points. **b**, 4D imaging of an oocyte expressing EGFP–ABD (F-actin, grey) injected with Alexa-568-Ran (chromosomes, red) (3 sections every 3.8 μm, image stacks every 22 s). Images show single confocal sections at selected time points. See Supplementary Video S6 for a full video. **c**, Enlarged region of an oocyte nucleus expressing EGFP–ABD, showing details of the actin network. Images show single confocal sections. See Supplementary Video S7 for a full video. **d**, Quantification of the chromosome circled in blue in **b**. Distance from animal pole (blue line) and mean EGFP–ABD fluorescence intensity (grey shading) over time. **e**, *Z*-projection of an oocyte fixed 5 min after NEBD, showing F-actin (TMR-phalloidin) staining in red and DNA in green. **f**, Single confocal section showing TMR-phalloidin injected into a live oocyte 5 min after NEBD. **g**, DNA-coated beads (green) were microinjected into the nucleus and the oocyte was imaged during maturation. Actin (Alexa-488-actin, red) forms patches around beads similar to chromosomes. Arrowhead shows an actin patch around a chromosome. Time shown in minutes and seconds (**a**–**d**). Scale bar, 20 μm.

in vitro-formed contractile actomyosin gels[24,25]. If such a gel were anchored to the animal pole, myosin-driven contraction could also provide the observed directionality. We attempted to test a role for myosin II by treating oocytes with blebbistatin, a specific inhibitor of myosin II or the Rho-associated protein kinase inhibitor Y-27632. Although neither compound affected chromosome congression, they failed to block myosin II-dependent polar body extrusion (not shown), making these experiments inconclusive.

Actin dependence in oocytes of other species

In summary, we have shown that in starfish oocytes, microtubules alone are unable to capture chromosomes more than 40 μm away from centrosomes. Instead, an actin 'fishnet' is essential for delivering chromosomes within the capture distance of microtubule asters. A similar architectural problem exists in many animal species, as oocytes have large nuclei for storing nuclear proteins and RNA. For example, *Xenopus* oocytes are morphologically similar to those of starfish and contain a microtubule array of 50–80-μm radius that has to capture chromosomes in a 500-μm diameter nucleus[26]. Consistent with our observations, disrupting the actin cytoskeleton with cytochalasin B has been reported to cause defects in spindle formation in *Xenopus* oocytes[27,28]. Moreover, in meiosis I of mouse oocytes, chromosomes are known to move to the cell cortex in an actin-dependent manner, even in the absence of microtubules[29–31]. This strongly suggests that actin nucleation on M-phase chromatin, which we directly visualized for the first time in starfish oocytes, is an evolutionarily conserved mechanism that is also functional in vertebrates and mammals. In starfish oocytes, failure of actin-dependent chromosome congression did not delay anaphase and consequently led to chromosome loss and aneuploidy. Such failure should have similar deleterious effects in *Xenopus* oocytes, because the spindle checkpoint is inactive in meiosis I (ref. 32), and although the checkpoint is active in mouse, oocytes escape checkpoint arrest after few hours[33]. At least 5% of human oocytes are aneuploid, and aneuploidy is the leading cause of pregnancy loss and birth defects[34]. It will therefore be important to test whether a similar actin-dependent mechanism is also involved in chromosome congression in humans.

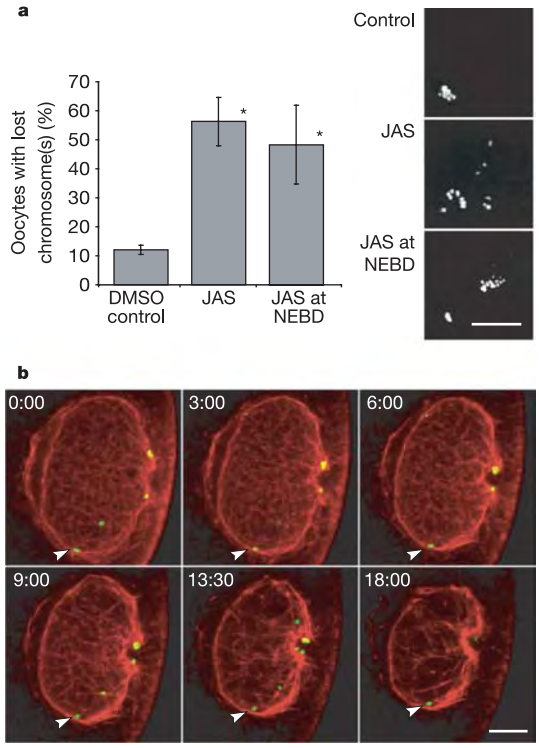


Figure 5 | F-actin-stabilizing drugs prevent chromosome movement and contraction of the actin network. a, Oocytes were matured in 0.1% DMSO or 1 μM jasplakinolide (JAS+0.01% DMSO), or 10 μM JAS (+0.1% DMSO) was added 2 min after NEBD. Oocytes were fixed 30 min after NEBD, stained for DNA and scored for the presence of all chromosomes at the animal pole. Data show average ± s.d. from two independent experiments ($n \geq 50$ each). Asterisks mark significant differences from control oocytes ($P < 0.05$, Student's *t*-test). **b**, 4D imaging of an oocyte expressing H2B–EGFP (green) and injected with TMR-phalloidin (red) approximately 3 min after NEBD (25 sections every 2.75 μm, image stacks taken every 45 s). Single confocal sections of selected time points are shown; see Supplementary Video S8 for a full video. Arrowhead indicates a lost chromosome. Time shown in minutes and seconds. Scale bar, 20 μm (**a**, **b**).

METHODS

Oocyte injection, maturation and drug treatments. Starfish (*Asterina miniata*) were obtained from Marinus Scientific and maintained as described in ref. 14. Oocytes were isolated and injected with mercury-filled needles using methods described elsewhere (ref. 14; see also http://155.37.3.143/panda/injection/index.html). Injection and microscopy were done at 20 °C. mRNA-injected oocytes were incubated overnight at 16 °C, or in the case of H2B–DiHcRed-expressing oocytes, for 36 h to allow incorporation of core histones. Maturation was triggered by the addition of 1–10 μM 1-methyladenine (Sigma). NEBD typically started 20 min after hormone addition, and only oocytes starting NEBD between 15 and 35 min were analysed.

For drug treatments, oocytes were matured in 3.3 μM nocodazole or 250 nM latrunculin B (both from EMD Biosciences, diluted from 1 mg ml^{-1} stocks), which efficiently depolymerized microtubules and actin filaments, respectively (Supplementary Fig. S1). In some cases, oocytes were also preincubated with the drugs for 30 min before maturation, which gave identical results. In live-cell imaging experiments, latrunculin B was added at 100–2,000 nM concentrations 10 min after hormone addition to prevent shape changes of oocytes caused by disruption of cortical actin. Jasplakinolide (Molecular Probes) was dissolved at 0.7 mg ml^{-1} in dimethylsulphoxide (DMSO) and used at a final concentration of 1 or 10 μM.

GFP constructs and fluorescent markers. H2B–DiHcRed[16], ensconsin–3EGFP (3EGFP–EMTB)[35] and EGFP–ABD[19] were transferred from constructs for mammalian expression to pGEMHE for *in vitro* transcription as described in ref. 14. Capped mRNA was synthesized from linearized templates using the mMessage mMachine kit (Ambion), dissolved in 10 μl water (typically 1–2 μg μl^{-1}) and injected to 1–5% oocyte volume.

Alexa-488-labelled rabbit muscle actin (Molecular Probes; 11 mg ml^{-1}) was diluted 1:4 in G-buffer (2 mM Tris pH 8.0, 0.1 mM ATP, 0.1 mM dithiothreitol, 0.1 mM $CaCl_2$) and injected to 5% oocyte volume. In some experiments, Alexa-488- or Alexa-568-labelled Ran was used to label chromosomes[36]. From the concentrated stocks (50 and 60 μM, respectively), amounts corresponding to 0.5–1% of the egg volume (2–5% of the endogenous Ran protein level[36]) were injected. To inject rhodamine-phalloidin (Molecular Probes), 50 μl of methanol stock was dried, dissolved in 1 μl water and injected into oocytes.

DNA-coated beads were prepared as previously published[37], except that DNA was fluorescently labelled by incorporating Cy3-dUTP (Amersham) and 1-μm diameter beads (Dynabeads MyOne Streptavidin, Dynal Biotech) were used.

Immunofluorescence. We have used the methods developed in ref. 38, with minor modifications. Briefly, oocytes were fixed in buffer containing 100 mM HEPES buffer, 50 mM EGTA, 10 mM $MgSO_4$, 2–500 mM maltose, 2% formaldehyde, 0.2% glutaraldehyde and 0.2% Triton-X100 for 2 h at 20 °C. Oocytes were extracted overnight in 0.2% Triton-X100 in PBS, blocked with PBS + 5% BSA and incubated with a monoclonal anti-α-tubulin antibody (DM 1A, Sigma) and Alexa-488-labelled secondary antibody (Molecular Probes). Rhodamine-phalloidin was from Molecular Probes, the methanol stock was dried then dissolved in PBS, and oocytes were stained for 1 h at 20 °C. DNA was stained by incubating oocytes for 1 h with 5 μg ml^{-1} Hoechst 33342 (Molecular Probes).

Confocal microscopy and image quantification. Microscopy was done either with a customized Zeiss LSM510 Axiovert confocal microscope as described[14], using an ×40 C-Apochromat 1.2 NA water immersion objective lens (Zeiss) (Figs 1b, 2a, 3a, 4a, e–g, 5b); a Nikon C1 E-800 confocal microscope with a ×20 SuperFluar 0.7 NA dry lens (Fig. 2c–e) or a Leica TCS SP2 confocal microscope equipped with an HCX Plan Apochromat λBlue 1.2 NA ×63 water immersion objective lens (Figs 1a and 4b, c).

Time series were analysed using LSM510 software (Zeiss) and ImageJ (http://rsb.info.nih.gov/ij/). Figures were assembled using Adobe Photoshop and Illustrator. A gaussian blur filter (0.3–1 pixel) was generally applied to reduce shot noise. Three-dimensional (3D) reconstructions and tracking were done

with in-house developed software, Tikal[39]. Images were filtered using an anisotropic diffusion filter and thresholded. Binary images were reconstructed in three dimensions by isosurface rendering, and chromosomes were tracked manually. Tracks were analysed using the graphing software Excel (Microsoft) and Graphis (Kylebank Software). For rendering Fig. 1b, Amira 2.3 (TGS) was used.

Electron microscopy. Oocytes were fixed in 2% formaldehyde plus 0.2–0.5% glutaraldehyde in sea water for 2 h, then stored at 4 °C in 1% formaldehyde in sea water. After embedding in agarose, oocytes were dehydrated in cold methanol and embedded in LR Gold resin (Polysciences) at −20 °C. The LR Gold was polymerized under ultraviolet light at −20 °C. Thin sections were collected on Formvar-coated nickel grids and immunolabelled as previously described[40]. A mouse monoclonal antibody against actin (clone C4, Boehringer Mannheim) was used at a dilution of 1:10 or 1:20. Goat anti-mouse IgG labelled with 10 nm gold was obtained from Amersham Biosciences and used at a dilution of 1:20. After labelling, the sections were stained and examined under a transmission electron microscope (CM-10, Philips).

Received 22 February; accepted 3 May 2005.
Published online 13 July 2005.

1. Gadde, S. & Heald, R. Mechanisms and molecules of the mitotic spindle. *Curr. Biol.* **14,** R797–R805 (2004).
2. Rusan, N. M., Fagerstrom, C. J., Yvon, A. M. & Wadsworth, P. Cell cycle-dependent changes in microtubule dynamics in living cells expressing green fluorescent protein-α tubulin. *Mol. Biol. Cell* **12,** 971–980 (2001).
3. Zhai, Y., Kronebusch, P. J., Simon, P. M. & Borisy, G. G. Microtubule dynamics at the G2/M transition: abrupt breakdown of cytoplasmic microtubules at nuclear envelope breakdown and implications for spindle morphogenesis. *J. Cell Biol.* **135,** 201–214 (1996).
4. Kirschner, M. & Mitchison, T. Beyond self-assembly: from microtubules to morphogenesis. *Cell* **45,** 329–342 (1986).
5. Kline-Smith, S. L. & Walczak, C. E. Mitotic spindle assembly and chromosome segregation: refocusing on microtubule dynamics. *Mol. Cell* **15,** 317–327 (2004).
6. Holy, T. E. & Leibler, S. Dynamic instability of microtubules as an efficient way to search in space. *Proc. Natl Acad. Sci. USA* **91,** 5682–5685 (1994).
7. Belmont, L. D., Hyman, A. A., Sawin, K. E. & Mitchison, T. J. Real-time visualization of cell cycle-dependent changes in microtubule dynamics in cytoplasmic extracts. *Cell* **62,** 579–589 (1990).
8. Piehl, M. & Cassimeris, L. Organization and dynamics of growing microtubule plus ends during early mitosis. *Mol. Biol. Cell* **14,** 916–925 (2003).
9. Rieder, C. L. & Alexander, S. P. Kinetochores are transported poleward along a single astral microtubule during chromosome attachment to the spindle in newt lung cells. *J. Cell Biol.* **110,** 81–95 (1990).
10. Carazo-Salas, R. E. & Karsenti, E. Long-range communication between chromatin and microtubules in *Xenopus* egg extracts. *Curr. Biol.* **13,** 1728–1733 (2003).
11. Tulu, U. S., Rusan, N. M. & Wadsworth, P. Peripheral, non-centrosome-associated microtubules contribute to spindle formation in centrosome-containing cells. *Curr. Biol.* **13,** 1894–1899 (2003).
12. Wollman, R. *et al.* Efficient chromosome capture requires a bias in the 'search-and-capture' process during mitotic-spindle assembly. *Curr. Biol.* **15,** 828–832 (2005).
13. Zhang, Q. Y., Tamura, M., Uetake, Y., Washitani-Nemoto, S. & Nemoto, S. Regulation of the paternal inheritance of centrosomes in starfish zygotes. *Dev. Biol.* **266,** 190–200 (2004).
14. Lenart, P. *et al.* Nuclear envelope breakdown in starfish oocytes proceeds by partial NPC disassembly followed by a rapidly spreading fenestration of nuclear membranes. *J. Cell Biol.* **160,** 1055–1068 (2003).
15. Miyazaki, A., Kamitsubo, E. & Nemoto, S. I. Premeiotic aster as a device to anchor the germinal vesicle to the cell surface of the presumptive animal pole in starfish oocytes. *Dev. Biol.* **218,** 161–171 (2000).
16. Gerlich, D. *et al.* Global chromosome positions are transmitted through mitosis in mammalian cells. *Cell* **112,** 751–764 (2003).
17. Stricker, S. A. & Schatten, G. The cytoskeleton and nuclear disassembly during germinal vesicle breakdown in starfish oocytes. *Dev. Growth Differ.* **33,** 163–171 (1991).
18. Terasaki, M. Redistribution of cytoplasmic components during germinal vesicle breakdown in starfish oocytes. *J. Cell Sci.* **107,** 1797–1805 (1994).
19. Pang, K. M., Lee, E. & Knecht, D. A. Use of a fusion protein between GFP and an actin-binding domain to visualize transient filamentous-actin structures. *Curr. Biol.* **8,** 405–408 (1998).
20. Heil-Chapdelaine, R. A. & Otto, J. J. Characterization of changes in F-actin during maturation of starfish oocytes. *Dev. Biol.* **177,** 204–216 (1996).
21. Visegradi, B., Lorinczy, D., Hild, G., Somogyi, B. & Nyitrai, M. The effect of phalloidin and jasplakinolide on the flexibility and thermal stability of actin filaments. *FEBS Lett.* **565,** 163–166 (2004).
22. Weber, K. L., Sokac, A. M., Berg, J. S., Cheney, R. E. & Bement, W. M. A microtubule-binding myosin required for nuclear anchoring and spindle assembly. *Nature* **431,** 325–329 (2004).
23. Waterman-Storer, C. *et al.* Microtubules remodel actomyosin networks in *Xenopus* egg extracts via two mechanisms of F-actin transport. *J. Cell Biol.* **150,** 361–376 (2000).
24. Szent-Gyorgyi, A. *Chemistry of Muscular Contraction* (Academic, New York, 1951).
25. Spicer, S. S. The clearing response of actomyosin to adenosinetriphosphate. *J. Biol. Chem.* **199,** 289–300 (1952).
26. Gard, D. L. Microtubule organization during maturation of *Xenopus* oocytes: assembly and rotation of the meiotic spindles. *Dev. Biol.* **151,** 516–530 (1992).
27. Ryabova, L. V., Betina, M. I. & Vassetzky, S. G. Influence of cytochalasin B on oocyte maturation in *Xenopus laevis*. *Cell Differ.* **19,** 89–96 (1986).
28. Gard, D. L., Cha, B. J. & Roeder, A. D. F-actin is required for spindle anchoring and rotation in *Xenopus* oocytes: a re-examination of the effects of cytochalasin B on oocyte maturation. *Zygote* **3,** 17–26 (1995).
29. Longo, F. J. & Chen, D. Y. Development of cortical polarity in mouse eggs: involvement of the meiotic apparatus. *Dev. Biol.* **107,** 382–394 (1985).
30. Leader, B. *et al.* Formin-2, polyploidy, hypofertility and positioning of the meiotic spindle in mouse oocytes. *Nature Cell Biol.* **4,** 921–928 (2002).
31. Maro, B. & Verlhac, M. H. Polar body formation: new rules for asymmetric divisions. *Nature Cell Biol.* **4,** E281–E283 (2002).
32. Peter, M. *et al.* The APC is dispensable for first meiotic anaphase in *Xenopus* oocytes. *Nature Cell Biol.* **3,** 83–87 (2001).
33. Wassmann, K., Niault, T. & Maro, B. Metaphase I arrest upon activation of the Mad2-dependent spindle checkpoint in mouse oocytes. *Curr. Biol.* **13,** 1596–1608 (2003).
34. Hassold, T. & Hunt, P. To err (meiotically) is human: the genesis of human aneuploidy. *Nature Rev. Genet.* **2,** 280–291 (2001).
35. Faire, K. *et al.* E-MAP-115 (ensconsin) associates dynamically with microtubules *in vivo* and is not a physiological modulator of microtubule dynamics. *J. Cell Sci.* **112,** 4243–4255 (1999).
36. Hinkle, B. *et al.* Chromosomal association of Ran during meiotic and mitotic divisions. *J. Cell Sci.* **115,** 4685–4693 (2002).
37. Heald, R. *et al.* Self-organization of microtubules into bipolar spindles around artificial chromosomes in *Xenopus* egg extracts. *Nature* **382,** 420–425 (1996).
38. Strickland, L. *et al.* Light microscopy of echinoderm embryos. *Methods Cell Biol.* **74,** 371–409 (2004).
39. Bacher, C. P., Reichenzeller, M., Athale, C., Herrmann, H. & Eils, R. 4-D single particle tracking of synthetic and proteinaceous microspheres reveals preferential movement of nuclear particles along chromatin-poor tracks. *BMC Cell Biol.* **5,** 45 (2004).
40. Hand, A. R. in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research* (eds Glasel, J. & Deutscher, M.) 205–260 (Academic, New York, 1995).

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

Acknowledgements Part of this work was performed at the Marine Biological Laboratory (MBL) in Woods Hole, supported by summer research fellowships from Nikon Inc. and the E. and M. Spiegel, F.B. and B.G. Bang, L.B. Lehmann, R.D. Allen and H.W. Rand foundations to J.E. Martin Hoppe and Leica Microsystems in Mannheim are gratefully acknowledged for providing equipment at the MBL. P.L. was supported by a predoctoral fellowship from the Louis-Jeantet Foundation. We would like to thank J. C. Bulinski for providing p3EGFP-EMTB, K. Weijer for EGFP-ABD and K. Ribbeck for fluorescently labelled Ran.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.E. (jan.ellenberg@embl.de).

LETTERS

A dark jet dominates the power output of the stellar black hole Cygnus X-1

Elena Gallo[1], Rob Fender[1,2], Christian Kaiser[2], David Russell[2], Raffaella Morganti[3], Tom Oosterloo[3] & Sebastian Heinz[4]

Black holes undergoing accretion are thought to emit the bulk of their power in the X-ray band by releasing the gravitational potential energy of the infalling matter[1]. At the same time, they are capable of producing highly collimated jets of energy and particles flowing out of the system with relativistic velocities[2]. Here we show that the 10-solar-mass ($10M_{\odot}$) black hole in the X-ray binary Cygnus X-1 (refs 3–5) is surrounded by a large-scale (~5 pc in diameter) ring-like structure that appears to be inflated by the inner radio jet[6]. We estimate that in order to sustain the observed emission of the ring, the jet of Cygnus X-1 has to carry a kinetic power that can be as high as the bolometric X-ray luminosity of the binary system. This result may imply that low-luminosity stellar-mass black holes as a whole dissipate the bulk of the liberated accretion power in the form of 'dark', radiatively inefficient relativistic outflows, rather than locally in the X-ray-emitting inflow.

Relativistic jets are a common feature of accreting black holes on all mass scales, ranging from supermassive black holes at the centres of active galactic nuclei[7,8] to stellar-mass black holes in X-ray binary systems within our own Galaxy[9–11]. Whereas the inflow of hot gas can be very efficient in producing light (up to 40% of the accreted material may be transformed into energy and radiated away in the form of optical/ultraviolet/X-ray photons), the same is not true for the synchrotron-emitting outflow, whose efficiency might be lower than a few per cent. Estimating the total—radiated plus kinetic—power content of the jets, and hence their importance with respect to the accretion process in terms of energetics, is a primary aim of high energy astrophysics.

We observed the field of the $10M_{\odot}$ black hole and Galactic jet source Cygnus X-1 at 1.4 GHz for 60 h with the Westerbork Synthesis Radio Telescope (WSRT), yielding the deepest low frequency radio observation of that region to date[12]. A ring of radio emission—with a diameter of $\sim 10^6$ AU—appears northeast of Cygnus X-1 (Fig. 1), and seems to draw an edge between the tail of the nearby H II nebula Sh2-101 (ref. 13; whose distance is consistent with that to Cygnus X-1 (ref. 14)) and the direction of the radio jet powered by Cygnus X-1 (ref. 6). As Cygnus X-1 moves in the sky along a trajectory that is roughly perpendicular to the jet[6,15,16], and thus can not possibly be traced back to the ring centre, this rules out the possibility that the ring might be the low-luminosity remnant of the natal supernova of the black hole. In analogy with extragalactic jet sources, the ring of Cygnus X-1 could be the result of a strong shock that develops at the

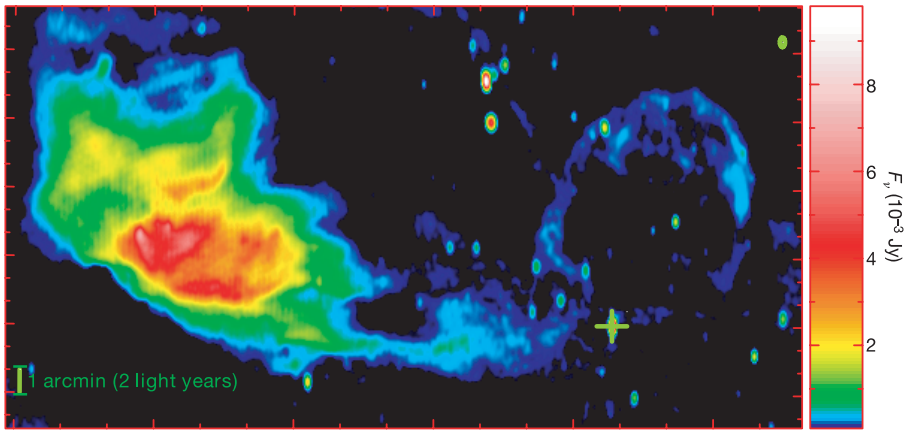


Figure 1 | A jet-blown ring around the Galactic stellar black hole in Cygnus X-1. The interstellar gas around a Galactic stellar black hole is stirred by the pressure of a highly collimated relativistic jet of energy and particles, resulting in a 15-light-yr-wide ring of radio emission. The field of view of the $10M_{\odot}$ black hole in Cygnus X-1 (marked by a cross) was observed by the Westerbork Synthesis Radio Telescope for 60 h at 1.4 GHz: the ring appears to draw an edge between the tail of Sh2-101, the nearby H II nebula on the left hand side, and the direction of the inner radio jet of Cygnus X-1 (shown in the inset of Fig. 2). The spatial resolution in this map (25×14 arcsec^2) is illustrated by the green open ellipse at the top right corner. Flux F_ν is colour-coded on a logarithmic scale, between 0 and 0.01 Jy (see colour scale); the average ring monochromatic flux is 0.2 mJy beam^{-1}. At a distance of 2.1 kpc, the separation L between Cygnus X-1 (coincident with the jet base) and the ring's outermost point is $1.9 \times 10^{19} \sin\theta^{-1}$ cm, where θ is the jet inclination to the line of sight ($\theta \approx 35°$). Because of limb-brightening, we observe a ring whose thickness ΔR in the plane of the sky equals the effective length we are looking through into the bubble. At 2.1 kpc, $\Delta R \approx 1.6 \times 10^{18}$ cm.

[1]Astronomical Institute 'Anton Pannekoek' and Center for High Energy Astrophysics, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, The Netherlands. [2]School of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK. [3]Netherlands Foundation for Research in Astronomy, Postbus 2, NL-7990 AA, Dwingeloo, The Netherlands. [4]Center for Space Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.

location where the collimated jet strikes the ambient interstellar medium (ISM; Fig. 2). The jet particles inflate a radio lobe that is over-pressured with respect to the surroundings, and thus the lobe expands sideways, forming a spherical bubble of shock-compressed ISM, which we observe as a ring because of limb brightening effects. The collisionally ionized gas behind the bow shock would produce the observed bremsstrahlung radiation; in addition, if the shock is radiative, significant line emission is expected from hotter gas at the bow shock front.

Structures similar to the ring of Cygnus X-1 have been found at the edges of radio lobes inflated by the jets of supermassive black holes at the centre of powerful radio galaxies[17], where the much higher temperatures of the intra-cluster medium compared to the ISM shift the bremsstrahlung emission to X-ray frequencies. Striking confirmation of this interpretation comes from follow-up optical observations of the field of Cygnus X-1 with the Isaac Newton Telescope Wide Field Camera: the ring is clearly detected using a Hα filter in an exposure of only 1,200 s (Fig. 3). The estimated flux of

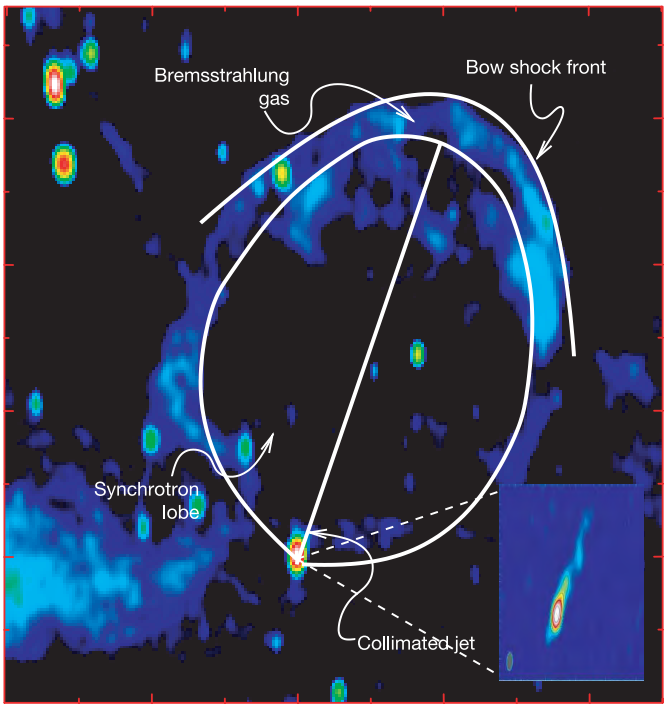


Figure 2 | The ring of Cygnus X-1: sketch of the model. The ring is the result of a strong shock that develops at the location where the pressure exerted by the collimated milliarcsec-scale jet, shown in the inset, is balanced by the interstellar medium. The jet particles start to inflate a synchrotron-emitting lobe that is over-pressured with respect to the surrounding gas, so the lobe expands sideways forming a spherical bubble of shock-compressed bremsstrahlung-emitting gas. The monochromatic luminosity of the ring, $L_{1.4\,\mathrm{GHz}} \approx 10^{18}\,\mathrm{erg\,s^{-1}\,Hz^{-1}}$, equals the product $(V \times \varepsilon_\nu)$, where the source unit volume V is given by the beam area times the measured ring thickness ($V \approx 4 \times 10^{53}\,\mathrm{cm^3}$), and ε_ν is the expression of the bremsstrahlung emissivity for a pure hydrogen gas emitting at a temperature T: $\varepsilon_\nu = 6.8 \times 10^{-38}\, g(\nu,T)\, T^{(-1/2)}\, n_e^2 \exp(h\nu/k_B T)\,\mathrm{erg\,cm^{-3}\,s^{-1}\,Hz^{-1}}$ (where h and k_B are Planck's constant and Boltzmann's constant, respectively). For $T \approx 10^4$ K and a Gaunt factor $g \approx 6$, the density n_e of the ionized particles in the ring is about $25\,\mathrm{cm^{-3}}$. The ionization fraction at 10^4 K is[23] $x \approx 0.02$, resulting in a total particle density $n_t \approx 1{,}300\,\mathrm{cm^{-3}}$. The minimum pressure inside the lobe predicted by the model is about $5 \times 10^{-11}\,\mathrm{erg\,cm^{-3}}$. If this pressure is solely due to a magnetized relativistic pair plasma in equipartition, then the strength of the magnetic field is about 40 μG. Assuming minimum energy conditions[22], this yields an expected lobe synchrotron surface brightness of $\sim$35 mJy beam^{-1} at 1.4 GHz, more than 150 times brighter than the observed ring. The much lower upper limit for the lobe surface brightness means that either the system is far from equipartition, or the most of the energy is stored in non-radiating particles, presumably baryons.

the ring at Hα frequencies exceeds the measured radio flux by a factor ≥20, indicating that the collisionally ionized gas in the ring is indeed emitting bremsstrahlung radiation and also that a significant amount of the measured Hα flux is due to line emission, as expected in the case of radiative shock.

Acting as an effective jet calorimeter, the ISM allows an estimate of the jet's power × lifetime product[18], which is, in principle, independent of the uncertainties associated with the jet spectrum and radiative efficiency. Following a self-similar fluid model developed for extragalactic jet sources[19–21], we assume that the jet of Cygnus X-1 is supplying energy at a constant rate P_{jet} and is expanding in a medium of constant density. We set the minimum temperature of the thermal gas to $T_{shock} \geq 10^4$ K, a typical temperature above which the cooling time becomes critically short, and below which the ionization fraction becomes too low for the ring to emit observable bremsstrahlung radiation. Given the average ring monochromatic luminosity, we are able to estimate a density of $\sim$1,300 cm^{-3} for the shock-compressed particles in the ring from the expression for the bremsstrahlung emissivity[22,23]. By balancing the interior pressure exerted by the lobe and the ram pressure of the shocked ISM, it can be shown[20] that the jet length within the lobe grows with time t in such a way that $t = (L/2)^{5/3}\,(\rho_0/P_{jet})^{1/3}$, where L is the separation between Cygnus X-1 and the ring's outermost point, and ρ_0 is the mass density of the unshocked gas. By writing the time derivative of this equation, there follows a simple relation between the jet lifetime, its length within the lobe, and the ring velocity: $t = (3/5)\,(L/v_{ring})$.

For a strong shock in a mono-atomic gas, the expansion velocity is set by the temperature of the shocked gas. If the shock is radiative, then the initial post-shock temperature can be higher than that of the thermalized, bremsstrahlung-emitting gas. A stringent constraint comes from X-ray observations: from the non-detection of soft X-ray emission in a 12,000-s observation of Cygnus X-1 taken with the Chandra X-ray Observatory, we can place an upper limit of $T_{shock} \leq 3 \times 10^6$ K by modelling the emission as a radiative shock. This, combined with the lower limit of 10^4 K, gives a ring velocity $v_{ring} \approx 20$–$360\,\mathrm{km\,s^{-1}}$. The resulting jet lifetime is $t \approx 0.02$–0.32 Myr, which can be compared with the estimated age of the progenitor of the black hole in Cygnus X-1—a few Myr (ref. 16).

Adopting a mass density of the unshocked gas that is four times lower than the initial post-shock density, we infer a time-averaged energy emission rate from the jet of between 8×10^{35} and $10^{37}\,\mathrm{erg\,s^{-1}}$. On the basis of daily X-ray and radio monitoring of Cygnus X-1 over the past 10 yr, we know that Cygnus X-1 is in a hard X-ray state[24]—and hence powers a collimated jet[11]—for about 90% of its lifetime. Taking this duty cycle into account, the total power carried by the jet of Cygnus X-1 is $\sim 9 \times 10^{35} \leq P_{jet} \leq 10^{37}\,\mathrm{erg\,s^{-1}}$, up to two orders of magnitude higher than the existing estimate based on the flat radio spectrum[25]. The fact that the jet of Cygnus X-1 switches off for short periods of time (typically for a few months over timescales of years) does not violate the model assumptions: the condition of constant power supply is met as long as the jet is intermittent over timescales that are short compared to its lifetime. The total power carried by the jet is a significant fraction $f \approx 0.03$–0.5 of the bolometric (0.1–200 keV) X-ray luminosity of Cygnus X-1 while in the hard state[26]; the total energy deposited by the jet into the surrounding ISM over its lifetime is about 7×10^{48} erg.

The particle density of the ISM through which the ring is expanding is constrained between $\sim$1 and $\sim$300 cm^{-3}, at most three orders of magnitude higher than the average ISM density in the Galaxy. The lack of a 'counter-ring' can be explained in terms of a much lower particle density in the opposite direction to Cygnus X-1. Such large density inhomogeneities are not unusual for dense star forming regions, such as the Cygnus association, and would support the hypothesis that the ring is the result of the interaction between the radio lobe and the tail of the H II nebula. If so, the counter-jet of Cygnus X-1 is travelling undisturbed to much larger distances,

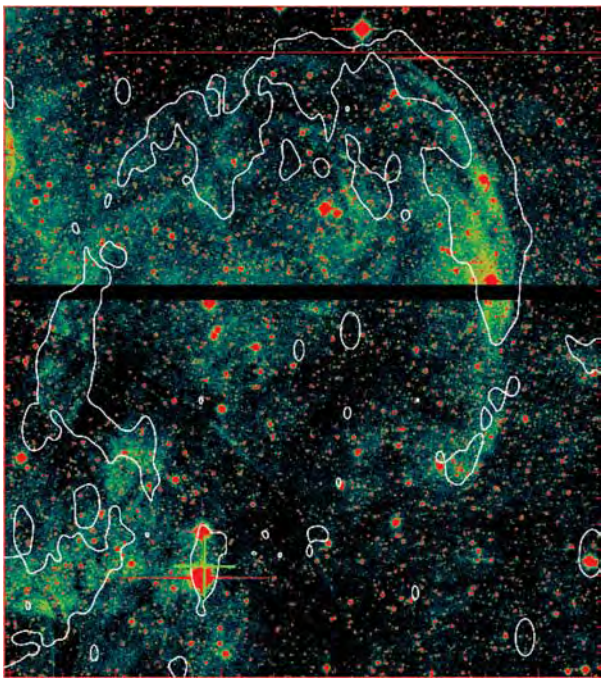


Figure 3 | Optical counterpart of the radio ring of Cygnus X-1. The optical image, taken with the Isaac Newton Telescope Wide Field Camera using an Hα filter, is shown with the 3σ radio contours over-plotted in white. As no calibration was taken during the observation, an absolute flux scale can not be set; red and blue colours correspond respectively to higher and lower relative fluxes. Given that the ring is clearly detected in a 1,200-s exposure, and taking into account the atmospheric and sky conditions during the observation, this translates into a minimum unabsorbed Hα flux of 0.02 mJy arcsec^{-2}. The corresponding radio-optical spectral index a (defined such that the monochromatic flux F_ν scales as ν^a) is greater than 0.2. This implies an emission mechanism with a flat spectrum, such as bremsstrahlung, plus excess flux possibly due to line emission, as expected in the case of radiative shock. For comparison, if the ring emitted optically thin synchrotron radiation with a spectral index $a = -0.7$, the expected flux at Hα frequencies would be 1.2×10^{-7} mJy arcsec^{-2}, by no means detectable by the Isaac Newton Telescope Wide Field Camera in 1,200 s.

gradually expanding and releasing its enormous kinetic energy. This could mean that the ring of Cygnus X-1 is a rather exceptional detection for this class of objects, made possible by its proximity to the H II nebula. Taking into account the contribution of the counter-jet as well, the total power dissipated by the jets of Cygnus X-1 in the form of kinetic energy can be as high as the bolometric X-ray luminosity of the system ($f \approx 0.06$–1).

The results presented here have important consequences for low-luminosity stellar-mass black holes as a whole: several works[27–29] have suggested that hard-state stellar black holes below a critical X-ray luminosity dissipate most of the liberated gravitational power in the form of radiatively inefficient outflows, rather than locally in the accretion flow. This is because in hard-state black hole binaries, the total jet power $P_{\rm jet}$ and the observed X-ray luminosity $L_{\rm X}$ follow a nonlinear relation[30] of the form $P_{\rm jet} \propto L_{\rm X}^{0.5}$ (both $P_{\rm jet}$ and $L_{\rm X}$ are expressed in Eddington units). Using the ring of Cygnus X-1 as an effective calorimeter for the jet power, we have constrained the normalization factor of the above equation, and find that $P_{\rm jet} = fL_{\rm X}$, with $f \approx 0.06$–1, when $L_{\rm X} \approx 0.02$. Thus the critical X-ray luminosity below which $P_{\rm jet} > L_{\rm X}$ is no lower than a few times 10^{-5} Eddington units, and could even be as high as the peak luminosity of the hard X-ray state. This radically alters our concepts of the accretion process and of the feedback of accretion power into the surroundings. The new observations presented here provide strong evidence that the power output of low-luminosity (that is, the overwhelming majority of) stellar black holes is dominated by the kinetic energy of 'dark' outflows, whose key signature is the eventual energization of the ambient medium.

Received 11 April; accepted 25 May 2005.

1. Frank, J., King, A. R. & Raine, D. J. *Accretion Power in Astrophysics* (Cambridge Univ. Press, Cambridge, 2002).
2. Hughes, P. A. *Beams and Jets in Astrophysics* (Cambridge Astrophysics Series, Cambridge Univ. Press, Cambridge, 1991).
3. Bowyer, S., Byram, E. T., Chubb, T. A. & Friedman, H. Cosmic X-ray sources. *Science* **147,** 394–398 (1965).
4. Gies, D. R. & Bolton, C. T. The optical spectrum of HDE 226868 = Cygnus X-1. II—Spectrophotometry and mass estimates. *Astrophys. J.* **304,** 371–393 (1986).
5. Herrero, A., Kudritzki, R. P., Gabler, R., Vilchez, J. M. & Gabler, A. Fundamental parameters of galactic luminous OB stars. II. A spectroscopic analysis of HDE 226868 and the mass of Cygnus X-1. *Astron. Astrophys.* **297,** 556–566 (1995).
6. Stirling, A. M. *et al.* A relativistic jet from Cygnus X-1 in the low/hard X-ray state. *Mon. Not. R. Astron. Soc.* **327,** 1273–1278 (2001).
7. Urry, C. M. & Radovani, P. Unified schemes for radio-loud active galactic nuclei. *Publ. Astron. Soc. Pacif.* **107,** 803–845 (1995).
8. Blandford, R. D. Black holes and relativistic jets. *Prog. Theor. Phys. Suppl.* **143,** 182–201 (2001).
9. Mirabel, I. F. & Rodríguez, L. F. A superluminal source in the galaxy. *Nature* **371,** 46–48 (1994).
10. Mirabel, I. F. & Rodríguez, L. F. Sources of relativistic jets in the galaxy. *Annu. Rev. Astron. Astrophys.* **37,** 409–443 (1999).
11. Fender, R. P. Jets from X-ray binaries. in *Compact Stellar X-Ray Sources* (eds Lewin, W. H. G. & van der Klis, M.) (Cambridge Univ. Press, Cambridge, in the press).
12. Martí, J., Rodríguez, L. F., Mirabel, I. F. & Paredes, J. M. A search for arcminute-scale radio jets in Cygnus X-1. *Astron. Astrophys.* **306,** 449–454 (1996).
13. Sharpless, S. A Catalogue of HII regions. *Astrophys. J. Suppl. Ser.* **4,** 257–279 (1959).
14. Hunter, D. A. & Massey, P. Small galactic H II regions. I—Spectral classifications of massive stars. *Astron. J.* **99,** 846–856 (1990).
15. Lestrade, J.-F. *et al.* High-precision VLBI astrometry of radio-emitting stars. *Astron. Astrophys.* **344,** 1014–1026 (1999).
16. Mirabel, I. F. & Rodrigues, I. Formation of a black hole in the dark. *Science* **300,** 1119–1120 (2003).
17. Smith, D. A., Wilson, A. S., Arnaud, K. A., Terashima, Y. & Young, A. J. A Chandra X-Ray study of Cygnus A. III. The cluster of galaxies. *Astrophys. J.* **565,** 195–207 (2002).
18. Burbidge, G. R. Estimates of the total energy in particles and magnetic field in the non-thermal radio sources. *Astrophys. J.* **129,** 849–851 (1959).
19. Castor, J., McCray, R. & Weaver, R. Interstellar bubbles. *Astrophys. J.* **200,** L107–L110 (1975).
20. Kaiser, C. R. & Alexander, P. A self-similar model for extragalactic radio sources. *Mon. Not. R. Astron. Soc.* **286,** 215–222 (1997).
21. Heinz, S., Reynolds, C. S. & Begelman, M. C. X-ray signatures of evolving radio galaxies. *Astrophys. J.* **501,** 126–136 (1998).
22. Longair, M. S. *High Energy Astrophysics* (Cambridge Univ. Press, Cambridge, 1992).
23. Lotz, W. Electron-impact ionization cross-sections and ionization rate coefficients for atoms and ions. *Astrophys. J. Suppl. Ser.* **14,** 207–238 (1967).
24. McClintock, J. E. & Remillard, R. A. Black hole binaries. in *Compact Stellar X-Ray Sources* (eds Lewin, W. H. G. & van der Klis, M.) (Cambridge Univ. Press, Cambridge, in the press).
25. Fender, R. P., Pooley, G. G., Durouchoux, P., Tilanus, R. P. J. & Brocksopp, C. The very flat radio-millimetre spectrum of Cygnus X-1. *Mon. Not. R. Astron. Soc.* **312,** 853–858 (2000).
26. Di Salvo, T., Done, C., Zycki, P. T., Burderi, L. & Robba, N. R. Probing the inner region of Cygnus X-1 in the low/hard state through its X-ray broadband spectrum. *Astrophys. J.* **547,** 1024–1033 (2001).
27. Gallo, E., Fender, R. P. & Pooley, G. G. A universal radio/X-ray correlation in low/hard state black hole binaries. *Mon. Not. R. Astron. Soc.* **34,** 60–72 (2003).
28. Livio, M., Pringle, J. E. & King, A. R. The disk-jet connection in microquasars and AGN. *Astrophys. J.* **593,** 184–188 (2003).
29. Malzac, J., Merloni, A. & Fabian, A. C. Jet-disc coupling through a common energy reservoir in the black hole XTE J1118+480. *Mon. Not. R. Astron. Soc.* **351,** 253–264 (2004).
30. Fender, R. P., Gallo, E. & Jonker, P. G. Jet-dominated states: an alternative to advection across black hole event horizons in 'quiescent' X-ray binaries. *Mon. Not. R. Astron. Soc.* **343,** L99–L103 (2003).

Acknowledgements We thank D. Mislis and R. Corradi for the Hα observation presented in this work. The Westerbork Synthesis Radio Telescope is operated by ASTRON (the Netherlands Foundation for Research in Astronomy) with support from the Netherlands Foundation for Scientific Research (NWO).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to E.G. (egallo@science.uva.nl).

LETTERS

Discovery of the triple asteroidal system 87 Sylvia

Franck Marchis[1], Pascal Descamps[2], Daniel Hestroffer[2] & Jérome Berthier[2]

After decades of speculation[1], the existence of binary asteroids has been observationally confirmed[2,3], with examples in all minor planet populations[4]. However, no triple systems have hitherto been discovered. Here we report the unambiguous detection of a triple asteroidal system in the main belt, composed of a 280-km primary (87 Sylvia) and two small moonlets orbiting at 710 and 1,360 km. We estimate their orbital elements and use them to refine the shape of the primary body. Both orbits are equatorial, circular and prograde, suggesting a common origin. Using the orbital information to estimate its mass and density, 87 Sylvia appears to have a rubble-pile structure with a porosity of 25–60 per cent. The system was most probably formed through the disruptive collision of a parent asteroid, with the new primary resulting from accretion of fragments, while the moonlets are formed from the debris, as has been predicted previously[5].

From the study of a companion's orbit, unique information about the internal properties (mass, density and porosity) of an asteroid, as well as their role in the formation of our Solar System, can be inferred. During the past two years, our group has focused on the study of main-belt asteroids already identified as binaries to carry out high accuracy astrometric follow-up and then provide ephemerides after assessing long-term dynamical effects such as precession[6–8]. Starting in August 2004 and for a period of 6 months, we initiated an observing campaign using the Adaptive Optics (AO) system and its infrared camera (NACO) available on the Very Large Telescope (VLT) at Cerro Paranal of the European Southern Observatory. Our list of 6 targets included 87 Sylvia, an X-type main-belt asteroid close to opposition (phase $<14°$) whose doubleness was discovered[9] and confirmed[10] in 2001. The primary asteroid with an angular size of ~ 0.17 arcsec is resolved in each frame. The presence of the moonlet, called S/2001(87)1, is confirmed in all frames (with a difference in brightness of magnitude $\Delta m \approx 3.8$, and maximum separation of 0.84 arcsec). We report here the detection of a second companion (called S/2004(87)1), fainter ($\Delta m \approx 4.2$) and closer (maximum separation of 0.44 arcsec) to the primary, and which is seen only at 12 epochs (Figs 1 and 2).

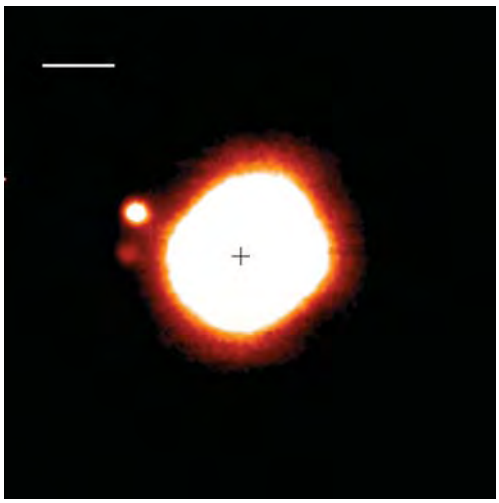


Figure 1 | One of the discovery images taken on 9 August 2004, showing 87 Sylvia and its two satellites, S/2001(87)1 and S/2004(87)1, in the K band (2.2 μm). This image is the result of shift-adding four frames with 2 s integration time reduced by standard methods. Scale bar, 0.25 arcsec. The cross indicates the position of the primary asteroid. The faintest moonlet is seen only in the frames taken under good seeing conditions (<1.0 arcsec in the visible). Twenty-seven observations were taken using the 13.3 mas per pixel scale over $\sim$2 months. AO systems, which provide in real time images close to the diffraction limit of the telescope ($\sim$0.06 arcsec at 2.2 μm), represent a very robust technique for detecting faint companions in proximity to a large asteroid and thus characterize their orbits. The peak signal-noise ratio on the main asteroid is relatively high ($\sim$2,000) because of its brightness (ephemeris predicted apparent magnitude $m_v \approx 11.5$).

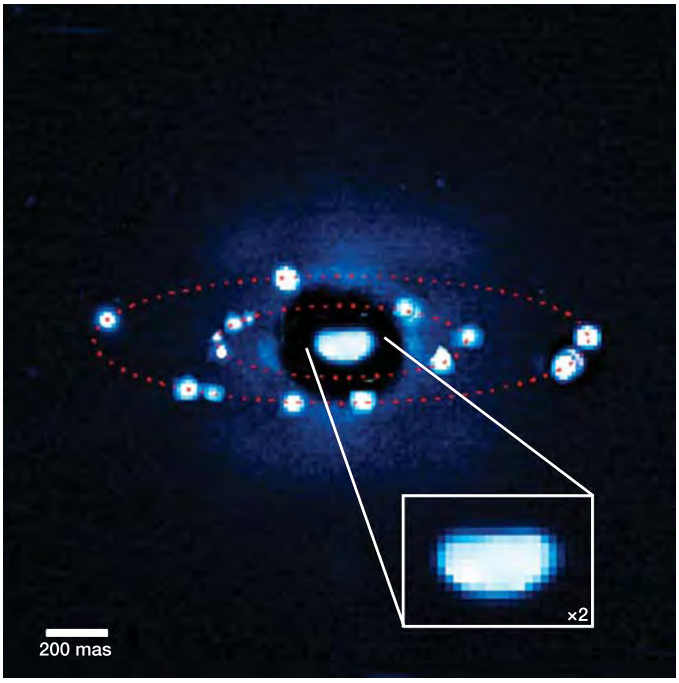


Figure 2 | Positions of S/2001(87)1 and S/2004(87)1 around 87 Sylvia. The dashed lines correspond to the moonlet orbits. This image is a composite of nine individual observations taken on nine nights and filtered with an unsharp mask (a sharpening digital process). North is up and east is left. The inset shows the shape of Sylvia's primary after applying a filtering process. The dark ring around the primary is an artefact due to the filtering. The centre of light of each moonlet is fitted by a Moffat–Gauss profile after subtraction of the residual background due to the imperfect phase correction of the AO system.

[1]University of California at Berkeley, Department of Astronomy, 601 Campbell Hall, Berkeley, California 94720, USA. [2]Institut de Mécanique Céleste et Calculs d'Éphémérides, Observatoire de Paris, 77 Avenue Denfert-Rochereau, F-75014 Paris, France.

Table 1 | Orbital elements of Sylvia's satellites S/2001(87)1 and S/2004(87)1, and physical properties of 87 Sylvia

	S/2001(87)1	S/2004(87)1
Properties of Sylvia's satellites		
Period (d)	3.6496 ± 0.0007	1.3788 ± 0.0007
Semi-major axis (km)	1,356 ± 5	706 ± 5
Eccentricity	0.001 ± 0.001	0.016 ± 0.011
Inclination w.r.t. equator (degrees)	1.7 ± 1.0	2.0 ± 1.0
Pole solution in ECJ2000 (degrees)	$\lambda = 72.4$, $\beta = 62.6$ (fixed)	$\lambda = 72.4 \pm 0.5$, $\beta = 62.6 \pm 0.5$
Inclination in J2000 (degrees)	7	7
Mean longitude at epoch (degrees)	112	63
Pericentre longitude (degrees)	14	51
Longitude of ascending node (degrees)	101	97
Physical properties of 87 Sylvia		
Mass (kg)	1.48×10^{19}	1.47×10^{19}
Density of primary* ($g\,cm^{-3}$)	1.2 ± 0.1	1.2 ± 0.1
J_2	0.17 ± 0.05	0.18 ± 0.01

Both orbits, which were fitted independently (with the exception of the pole solution derived from the S/2004(87)1 orbit), give extremely consistent physical characteristics for Sylvia's primary. The epoch of reference is JD 2453249.5 UTC for S/2001(87)1 and JD 2453246.5 UTC for S/2004(87)1. λ, longitude; β, latitude; J_2, gravity field term.
*Considering the shape of Sylvia's primary to be approximated by an ellipsoid with $a = 192$ km, $b = 132$ km and $c = 116$ km, and 4% error.

Supplementary Table 1 summarizes the epoch of observations and the relative positions and brightness of each moonlet. The observations were taken over a period of time covering ~2 months, enough to obtain a good estimate of the satellite orbit. Neglecting any mutual perturbations between both moonlets, we can determine their orbit individually. We used an algorithm specially suited for the analysis of binary asteroids[11] and already successfully applied to well-constrained binary systems, such as Pluto–Charon or 121 Hermione[7,8]. Table 1 summarizes the orbital elements for each satellite. Both solutions are extremely accurate, with a residual mean square error of 10 milliarcseconds (10 mas; corresponding to 16 km). Both orbits, displayed in Fig. 3, are quasi-circular, nearly equatorial (inclination $i \approx 2°$) and prograde with respect to the spin axis orientation of the primary (which is itself prograde). The satellite S/2001(87)1 orbits around the primary at ~1,360 km, about twice the distance of the S/2004(87)1 moonlet (710 km). The period and semi-major axes of the already known moonlet are in agreement with the analysis by our colleagues based on five observations taken shortly after its discovery[12].

The angular size of both moonlets is below the VLT-NACO angular resolution. Their sizes can only be estimated by comparing their integrated brightness with that of the primary and assuming the same albedo. Our analysis showed that the relative intensities measured on the reduced frames remain constant, suggesting that they are roughly spheroidal. Their diameters are estimated as 18 ± 4 km for S/2001(87)1 and 7 ± 2 km for S/2004(87)1, assuming the same surface composition (and thus the same albedo) for the primary and the moonlets. These sizes are typical of main-belt asteroid satellites discovered so far[3,4,6–8]. However, this measurement is biased by the quality of AO systems. Smaller, and therefore fainter, moonlets

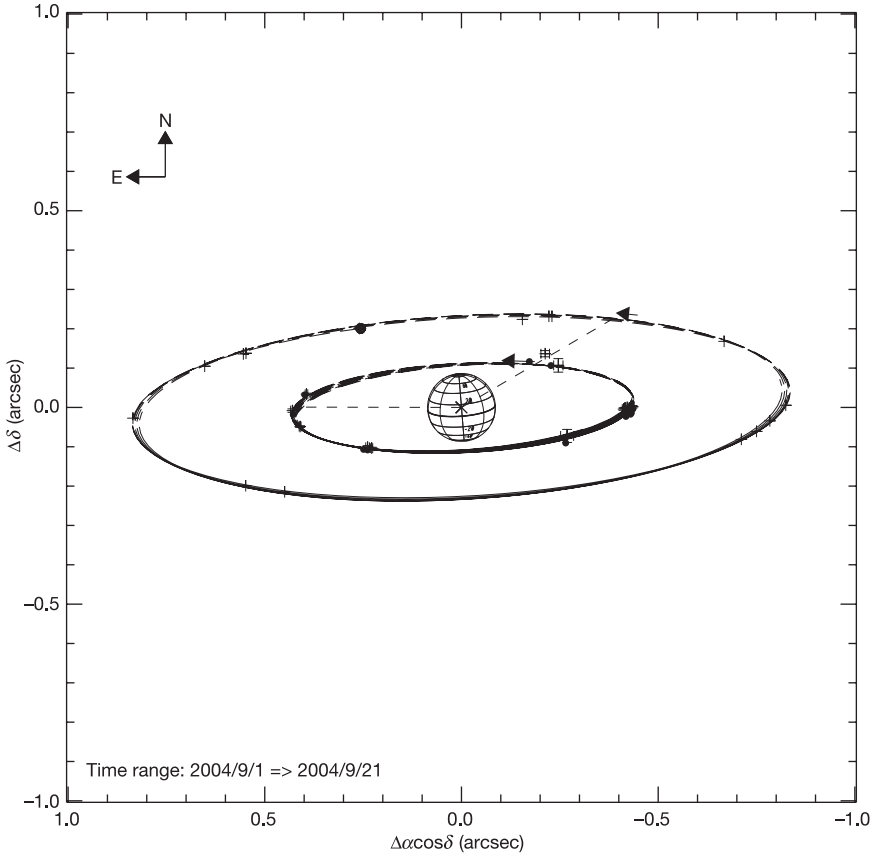


Figure 3 | The apparent orbit of Sylvia's companions projected on the plane of the sky at different epochs in September 2004. The measured positions (crosses) and their error bars (lower than 8 mas) are labelled. The primary is displayed using the pole solution of our model, and the mean diameter measured on our data set. The positions of the pericentres at their epoch of reference JD 2453249.5 UTC for S/2001(87)1 and JD 2453246.5 UTC for S/2004(87)1 are indicated (dashed lines). North is up and east is left. α, right ascension; δ, declination.

cannot be detected because of the glow of the primary produced by the residual uncorrected phase. Considering their period of revolution and neglecting the mass of the moonlets, the mass of the primary (M_{Sylvia}) is accurately derived using Kepler's third law on both moonlet orbits ($M_{\mathrm{Sylvia}} \approx (1.478 \pm 0.006) \times 10^{19}$ kg).

Since Sylvia's primary is well resolved in the AO observations, its shape can be directly estimated by fitting the image with an ellipsoid (see Supplementary Table 1). Using this technique, which was validated for other main-belt asteroids and compared with the lightcurve inversion shape[7], we can refine the size and shape of the primary. Considering the pole orientation determined by the analysis of the moonlet orbits, which is also in agreement with lightcurve inversion pole solution[13], the shape of Sylvia's primary can be approximated by an ellipsoid with axes $a = 192$ km, $b = 132$ km and $c = 116$ km (that is, $a/b = 1.46$, $b/c = 1.14$, with a mean radius Re = 143 km) with a 4% error, very close to the shape model derived by lightcurve inversion[13]. The averaged size is 10% larger than that determined by radiometric IRAS measurement in the far-infrared[14]. Interestingly, the second degree zonal coefficient of the gravitational potential of the primary (J_2) suggested[6] by our orbit analysis (~ 0.18) is compatible with a theoretical $J_2 = 0.14$ calculated considering the shape of Sylvia's primary and a homogeneous mass distribution. Even if the observations spanned a short period of time, the consideration of the precession effect is pertinent, especially for S/2004(87)1, since this satellite is closer to the primary and thus more sensitive to the precession effect[6,11]. Forcing $J_2 = 0$ in the orbit model for this moonlet will lead to a solution which is incompatible with the mass estimate derived from the orbit of S/2001(87)1. The results from the pole solution, the precessions of the orbits, and the mass values from both satellite orbits are highly consistent. We believe these data strongly support our discovery.

The bulk density (ρ) of the primary calculated using the mass from the orbit analyses and the size directly estimated from the AO images ($\rho = 1.2 \pm 0.1$ g cm^{-3}) is similar to that measured for C-type binary asteroids, such as 45 Eugenia[3,7], 121 Hermione[7,8] or 90 Antiope[3,7]. 87 Sylvia is classified, however, as an X-type asteroid[15], corresponding to an asteroid with a subtle absorption feature at 0.9 μm (tentatively interpreted as being due to the presence of orthopyroxene), adjoining both the C- and S-types. Since the composition of X-type asteroids remains unclear and their relation to mineralogical type is unknown, Sylvia's bulk porosity is not well constrained. The porosity could range from 25% up to 60%, depending on whether one considers the interior of Sylvia to be composed of hydrated carbonaceous chondrites[16] or a mixture of anhydrous silicates and water-ice. The reality might be somewhere between these two extreme values, but in all cases it is implying that 87 Sylvia's primary is made of a significant portion of void, suggesting a (loose) rubble-pile internal structure. Further analysis of X-class asteroids[17] should shed light on the actual composition of Sylvia.

Rubble-pile asteroids, as the outcomes of catastrophic collisions, are likely to be cohesionless bodies of gravitationally bound aggregates. Studies using the lightcurve inversion technique have already suggested the irregular shape of 87 Sylvia[13], implying that this asteroid may be the result of a gravitational re-accumulation process subsequent to a catastrophic collision[5]. The satellite moonlets, which revolve in a prograde manner describing quasi-equatorial and circular orbits, could be impact-captured debris. The very small eccentricity is also consistent with a well-relaxed system well damped by tidal effects. The existence of multiple moonlets formed by collisional events was suggested in several collisional simulations in which a large fragment is surrounded by a swarm of fragments[18,19], but little work has been done to study their stability over a large period of time. These moonlets (S/2001(87)1 and S/2004(87)1) orbit far enough inside the Hill sphere of the primary (1/50 and $1/100 \times r_{\mathrm{Hill}}$) to be insensitive to solar tides, but probably they are far enough from synchronous rotation with the primary (9.5 and $4.9 \times$ Re) to have stable orbits despite its ellipsoid shape[20,21]. It has been shown that the stability of Ida's moonlet[22], the first binary asteroid discovered[2], is mostly dependent on the distance to the primary; however, these authors failed to model a binary system that is stable over tens of millions of years. Our detection of the first multiple asteroidal system provides basic data that can be used to explore several long term stability scenarios for asteroidal satellites.

Received 14 April; accepted 6 July 2005.

1. Weidenschiling, S. J., Paolicchi, P. & Zappala, V. in *Asteroids II* (eds Binzel, R. P., Gehrels, T. & Matthews, M. S.) 643–658 (Univ. Arizona Press, Tucson, 1989).
2. Chapman, C. R. *et al.* Discovery and physical properties of Dactyl, a satellite of asteroid 243 Ida. *Nature* **374,** 783–785 (1995).
3. Merline, W. J. L. *et al.* Discovery of a moon orbiting the asteroid 45 Eugenia. *Nature* **401,** 565–568 (1999).
4. Merline, W. J. L., *et al.* in *Asteroids III* (eds Bottke, W. F., Cellino, A., Paolicchi, P. & Binzel, R. P.) 289–312 (Univ. Arizona Press, Tucson, 2002).
5. Michel, P., Benz, W., Tanga, P. & Richardson, D. C. Collision and gravitational reaccumulation: forming asteroid families and satellites. *Science* **294,** 1696–1700 (2001).
6. Marchis, F. *et al.* A three-dimensional solution for the orbit of the asteroidal satellite of 22 Kalliope. *Icarus* **165,** 112–120 (2003).
7. Marchis, F., Descamps, P., Berthier, J., Hestroffer, D. & de Pater, I. Fine analysis of 121 Hermione, 45 Eugenia, and 90 Antiope binary asteroid systems with AO observations. *Bull. Am. Astron. Soc. Div. Planet. Sci.* **36,** 33.23 (2004).
8. Marchis, F. *et al.* Mass and density of asteroid 121 Hermione from an analysis of its companion orbit. *Icarus* (in the press).
9. Brown, M. E., Margot, J. L., de Pater, I. & Roe, H. S/2001(87)1. *IAU Circ. No.* 7588 (2001).
10. Storrs, A. *et al.* S/2001 (87) 1. *IAU Circ. No.* 7590 (2001).
11. Descamps, P. Orbit of a visual binary system. *Bull. Am. Astron. Soc. Div. Dynam. Astron.* **35,** 04.01 (2005).
12. Margot, J. L. & Brown, M. E. Discovery and characterization of binary asteroids 22 Kalliope and 87 Sylvia. *Bull. Am. Astron. Soc. Div. Planet. Sci.* **33,** 52.02 (2001).
13. Kaasalainen, M., Torppa, J. & Piironen, J. Models of twenty asteroids from photometric data. *Icarus* **159,** 369–395 (2002).
14. Tedesco, E. F., Noah, P. V., Noah, M. & Price, S. D. The supplemental IRAS minor planet survey. *Astron. J.* **123,** 1056–1085 (2002).
15. Bus, S. J. & Binzel, R. P. Phase II of the small main-belt asteroid spectroscopic survey. *Icarus* **158,** 146–177 (2002).
16. Britt, D. T., Yeomans, D., Housen, K. & Consolmagno, G. in *Asteroids III* (eds Bottke, W. F., Cellino, A., Paolicchi, P. & Binzel, R. P.) 485–500 (Univ. Arizona Press, Tucson, 2002).
17. Clark, B. E. *et al.* Spectroscopy of X-type asteroids. *Icarus* **128,** 3070–3081 (2004).
18. Durda, D. D. The formation of asteroidal satellites in catastrophic collisions. *Icarus* **120,** 212–219 (1996).
19. Doressoundiram, A., Paolicchi, P., Verlicchi, A. & Cellino, A. The formation of binary asteroids as outcomes of catastrophic collisions. *Planet. Space Sci.* **45,** 757–770 (1997).
20. Chauvineau, B., Farinella, P. & Mignard, F. Planar orbits about a triaxial body—Application to asteroidal satellites. *Icarus* **105,** 370–384 (1993).
21. Scheeres, D. J. The effect of C22 on orbit energy and angular momentum. *Celest. Mech. Dynam. Astron.* **73,** 339–348 (1999).
22. Petit, J.-M., Durda, D. D., Greenberg, R., Hurford, T. A. & Geissler, P. E. The long-term dynamics of Dactyl's orbit. *Icarus* **130,** 177–197 (1997).

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

Acknowledgements The observations reported here are based on data collected at the European Southern Observatory, Chile. This work was partially supported by the National Science Foundation Science and Technology Center for Adaptive Optics, managed by the University of California at Santa Cruz under a cooperative agreement, and by the Chretien International Research Grant of the American Astronomical Society.

Author Contributions All authors contributed equally to this work.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.M. (fmarchis@berkeley.edu).

LETTERS

In situ multi-satellite detection of coherent vortices as a manifestation of Alfvénic turbulence

David Sundkvist[1,2], Vladimir Krasnoselskikh[1], Padma K. Shukla[3], Andris Vaivads[2], Mats André[2], Stephan Buchert[2] & Henri Rème[4]

Turbulence in fluids[1] and plasmas[2–5] is a ubiquitous phenomenon driven by a variety of sources—currents, sheared flows, gradients in density and temperature, and so on. Turbulence involves fluctuations of physical properties on many different scales, which interact nonlinearly to produce self-organized structures in the form of vortices[2–5]. Vortex motion in fluids and magnetized plasmas is typically governed by nonlinear equations[2–5], examples of which include the Navier–Stokes equation[1,2], the Charney–Hasegawa–Mima equations[2–5] and their numerous generalizations[6–9]. These nonlinear equations admit solutions[2–5] in the form of different types of vortices that are frequently observed in a variety of contexts: in atmospheres, in oceans and planetary systems[2,4], in the heliosphere[10,11], in the Earth's ionosphere and magnetosphere[12–17], and in laboratory plasma experiments[18]. Here we report the discovery by the Cluster satellites[19] of a distinct class of vortex motion—short-scale drift-kinetic Alfvén (DKA) vortices[8,9]—in the Earth's magnetospheric cusp region. As is the case for the larger Kelvin–Helmholtz vortices observed previously[17], these dynamic structures should provide a channel for transporting plasma particles and energy through the magnetospheric boundary layers.

The DKA vortices are result of nonlinear interaction of coupled finite-amplitude, low-frequency (in comparison with the proton gyrofrequency) drift and kinetic Alfvén waves[20], which accompany density, potential and sheared magnetic field perturbations, and which are propagating obliquely to the external magnetic field and density gradient directions. We present clear evidence that the DKA vortex structures exist on several spatial scales, contrasting them to those Alfvénic structures that have been deduced from simple magnetohydrodynamic (MHD) models[2,11]. The observed DKA vortices have a characteristic scale of the order of the ion (sound) gyroradius, $\rho_p \approx 25$ km. Hence, the electromagnetic DKA vortices, as discovered here, are in sharp contrast to a recent observation[17] of Kelvin–Helmholtz (KH) vortices which are typically 40,000 km across in space. Both the DKA and KH vortices provide new channels for transporting plasma particles and energy through the magnetospheric boundary layers, but on different spatio-temporal scales. Furthermore, it should be stressed that the DKA vortices are of fundamental interest also in the heliosphere[11] and in fusion plasmas[18] with regard to cross-field transport processes.

The formation of vortices is generally attributed to inverse cascade[1–5] of energy between two or three-dimensional (3D) dispersive modes that are nonlinearly interacting. For the DKA vortices in a non-uniform magnetoplasma, nonlinear equations[8,9] governing their dynamics include the evolution of nonlinearly coupled density perturbations, the ion vorticity, and the parallel (with respect to the external magnetic field direction) component of the vector potential. Owing to anisotropic dispersion characteristics of the DKA waves in a non-uniform magnetoplasma (with the plasma β—the ratio kinetic pressure/magnetic pressure—being larger than the electron-to-proton mass ratio), the above mentioned nonlinear equations admit inverse cascade, but the energy spectrum in the inertial range may not follow a Kolmogorov-like power law[1] because the turbulence now contains the characteristic scale ρ_p. This makes a clear distinction between the nonlinear dispersive DKA mode related vortices and those localized Alfvénic structures that are based on the nonlinear MHD equations[2,11]. Although extensive theoretical work has been carried out to investigate the properties of the DKA vortices[8,9,13], there are no direct microscale observations of these coherent structures in magnetized space plasmas. Earlier space-borne observations, which suggested vortex-like structures[13–16] in the Earth's ionosphere and magnetosphere, suffered from the inherent problem of separating temporal from spatial variations owing to limitations of single point measurements in space. Hence, up to now, the existence of the DKA vortices and many theoretical predictions regarding their properties could not be experimentally verified—for example, the characteristic length scale perpendicular to the ambient magnetic field direction, which is a crucial parameter for identifying the size of a plasma vortex.

The Cluster multi-spacecraft mission, with *in situ* simultaneous multi-point measurements[19], provides a unique opportunity to investigate the existence of the DKA vortices inside the cusp region of the Earth's magnetosphere, where there exist sheared plasma flows and density gradients sufficient to excite finite amplitude DKA fluctuations in a non-uniform magnetoplasma with $\beta \approx 0.02$. In the following, we present multipoint microscale measurements that unambiguously show the existence of short scale DKA vortices.

The cusp is an important region in the Earth's magnetosphere where the solar wind can directly access the ionosphere, and where large amounts of the plasma as well as kinetic and electromagnetic energies are transported. The cusp often exhibits the characteristics of low-frequency electromagnetic turbulence[16,21]. On 9 March 2002, the Cluster satellites passed outward through the high-altitude northern cusp during a time of northward interplanetary magnetic field[21]. Under such conditions, high-latitude magnetic reconnection takes place tailward from the cusp. A turbulent boundary layer on the border between the cusp and the polar cap is then formed, characterized by reconnection induced field-aligned proton jets. The four Cluster satellites entered the region of the turbulent boundary layer at 02:49 UT (Fig. 1). The cusp is identified as a high-density region with the presence of intense ion flux (Fig. 2a, b). Associated with these proton injections, there are broadband low-frequency electromagnetic fluctuations of Alfvénic nature ($E_\perp/B_\perp \approx v_A$, where $\perp$ denotes perpendicular components and v_A is the Alfvén speed) around and below the proton gyrofrequency, $f_{cp} \approx 1.6$ Hz, giving rise to the turbulent spectra shown in Fig. 2c (only the magnetic

[1]Laboratoire de Physique et Chimie de l'Environnement, CNRS, 45071 Orléans, France. [2]Swedish Institute of Space Physics, SE-751 21 Uppsala, Sweden. [3]Institut für Theoretische Physik IV, Ruhr-Universität Bochum, D-44780 Bochum, Germany. [4]Centre d'Etude Spatiale des Rayonnements, Toulouse 31029, France.

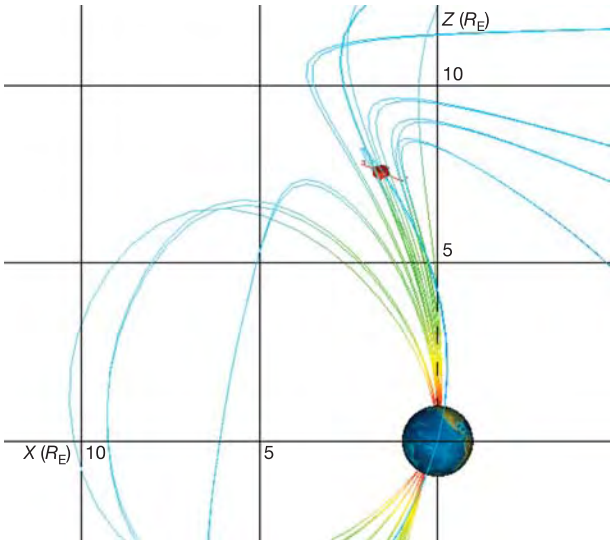


Figure 1 | Location of the Cluster spacecraft and the Earth's magnetic field lines. The red line marks the trajectory of the spacecraft for the time interval in Fig. 2, while traversing outbound through a boundary layer to the cusp. The interplanetary magnetic field was northward and reconnection occurred on the same field lines at higher altitudes. R_E, Earth radius.

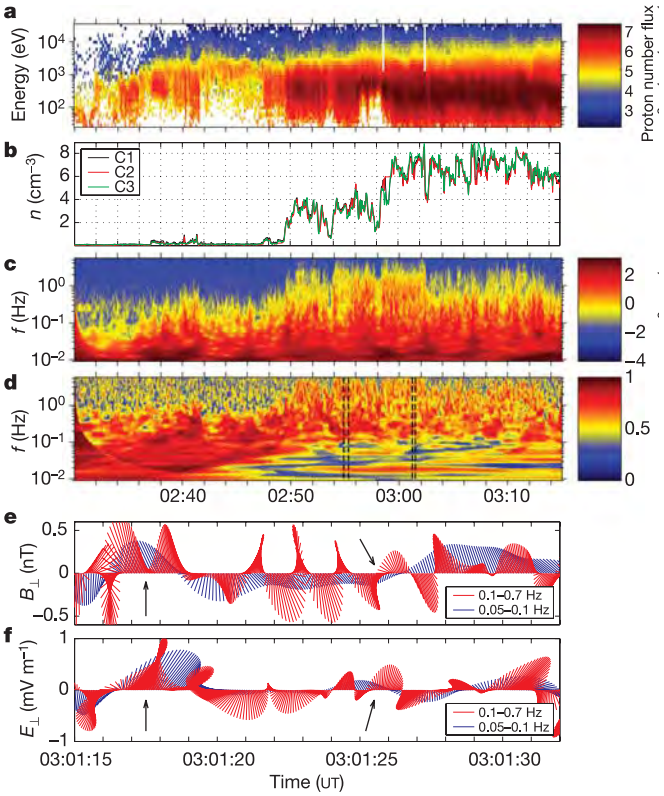


Figure 2 | Proton flux, density and low frequency fields for a cusp passage. **a**, Proton number flux. **b**, Plasma density, n, measured by three of the Cluster spacecraft, C1, C2 and C3. **c**, Power spectral density of magnetic field strength B as a function of frequency f and time. A wavelet transform has been used to better resolve the low-frequency structures. **d**, Degree of polarization (DOP) of the three components of the magnetic field, which can be viewed as a test for plane waves. Note how the DOP decreases dramatically for emissions below ~0.1–0.2 Hz when entering the boundary layer to the cusp, implicitly indicating the presence of structures. **e**, Hodogram representations of perpendicular magnetic field, $B_\perp$, for two frequency intervals, corresponding to different spatial scales through a Doppler shift. The arrows shows where rotational reversal occurs. **f**, Same as **e**, here for the perpendicular electric field.

spectrogram is shown here). The Alfvén speed is typically around 1,000 km s^{-1} in this region. Direct multi-point measurements[21] revealed the presence of coupled drift and kinetic Alfvén waves (KAWs) as well as electromagnetic ion-cyclotron waves whose characteristic perpendicular scale-lengths are of the order of (2–4)ρ_p. The boundary layer plasma is strongly inhomogeneous (density gradients up to 25% at distances of a few proton gyroradii), with prevalent shear velocity flows[21]. These gradients are well above the theoretical threshold values for spontaneously exciting finite amplitude DKA waves which self-organize in the form of vortical structures. Below, we study the low-frequency electromagnetic waves in detail, and show that a significant part of the frequency spectrum can be attributed to coherent DKA vortex structures.

The result of a polarization analysis of the three magnetic field components in the frequency domain around and below the proton gyrofrequency is displayed in Fig. 2d. The degree of polarization[22] (DOP) can be regarded as a test for the plane wave ansatz. A high value of DOP implies that the fluctuations considered as linear plane waves are coherent over several wavelengths. The sudden drop in DOP at frequencies below 0.1–0.2 Hz, when the spacecraft enters the cusp boundary layer, indicates that the electromagnetic field fluctuations can be structures, rather than plane wavepackets. Hodogram representations of the perpendicular magnetic and electric fields along the spacecraft trajectory (Fig. 2e, f) indeed show rotation reversal behaviour reminiscent of vortices. Moreover, we notice that the vortex structures exist nested on several different scales (Fig. 2e, f). To demonstrate that the observed variations are coherent vortex structures, it is necessary to show their characteristics and their spatial origin. The latter was not possible with earlier single-spacecraft measurements. We now demonstrate the spatial occurrence of these structures by using the fact that we have four simultaneous measurement points in space.

Figure 3e exhibits the configuration of the Cluster spacecraft in the plane perpendicular to the ambient magnetic field direction, together with the projection of the convective plasma flow velocity $V_\perp$ at about 02:55 UT on 9 March 2002. The relative velocity between the spacecraft and the bulk plasma is ~28 km s^{-1}. The relative alignment of the C1 and C2 spacecrafts along the plasma flow vector (Fig. 3e) predicts that a structure fixed in the plasma frame should be visible on both spacecraft with a time lag of approximately one second. Moreover, a quasi-static structure of the size of a few ρ_p will be Doppler-shifted in frequency domain to $f \approx V_\perp/\lambda_\perp \approx 0.1$–0.3 Hz, where $\lambda_\perp$ is the typical perpendicular scale size. The time series of the perpendicular magnetic field components, which are bandpass filtered at 0.1–0.2 Hz, are depicted in Fig. 3a, b. The time lag between the spacecraft is clearly visible and is about 1.2 s, which is consistent with the spacecraft observing a quasi-stationary structure moving with the plasma flow. Observations by the C3 and C4 satellites, which are not aligned along the flow with any other satellites, do not correlate with the observations from other spacecraft (Fig. 3d, e), thus giving an estimate of the transverse radial scale size of the structure of the order of (2–3)ρ_p.

To confirm the vortex nature of the coherent structures, we perform an analysis of the temporal evolution of the instantaneous direction of the magnetic field vectors (Fig. 3c, d). The angle between the magnetic field and the fixed direction of the spacecraft trajectory shows the polarization reversal feature, which is a unique characteristic of a vortex: polarization change on its spatial scale. At 02:55:03 UT and 02:55:16 UT (indicated by arrows) the sense of rotation changes direction from left-handed to right-handed and then back again.

We have demonstrated that the variations are spatially coherent vortex structures. Taken together with their Alfvénic nature in a strongly inhomogeneous magnetoplasma, we conclude that we have

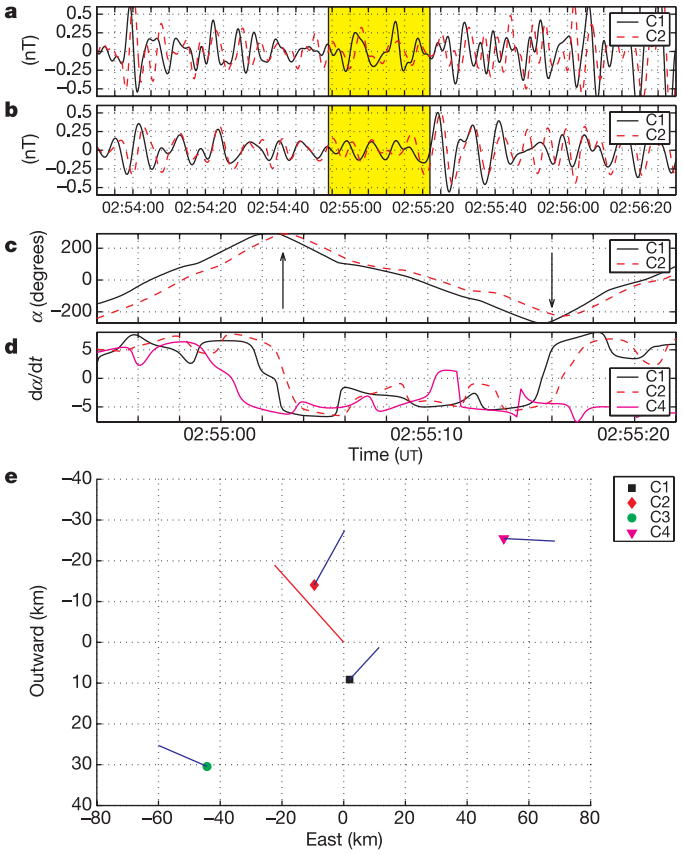


Figure 3 | Magnetic field data from the Cluster satellites. The time series are filtered to correspond to the length scales of the Doppler shifted structures. **a**, Eastward perpendicular component of the magnetic field. The highlighted area is used for panels **c** and **d**. **b**, Similarly, the outward perpendicular component. **c**, Angle α between the instantaneous magnetic field vector in the perpendicular plane and a fixed direction in this plane, corresponding to the highlighted area in **a** and **b**. It should be pointed out that the sense of rotation changes direction from left-handed to right-handed and then back again during this time interval. **d**, Angular velocity of the same angle as in **c**. **e**, Relative positions of the satellites in the plane perpendicular to the ambient magnetic field. The plasma flow vector was mainly in this plane and is shown by the red line. The length of the vector indicates the distance the plasma convects with respect to the satellites during one second. The blue lines represent the instantaneous magnetic field measured on board respective spacecraft. Note that C1 and C2 are almost exactly aligned along the flow vector.

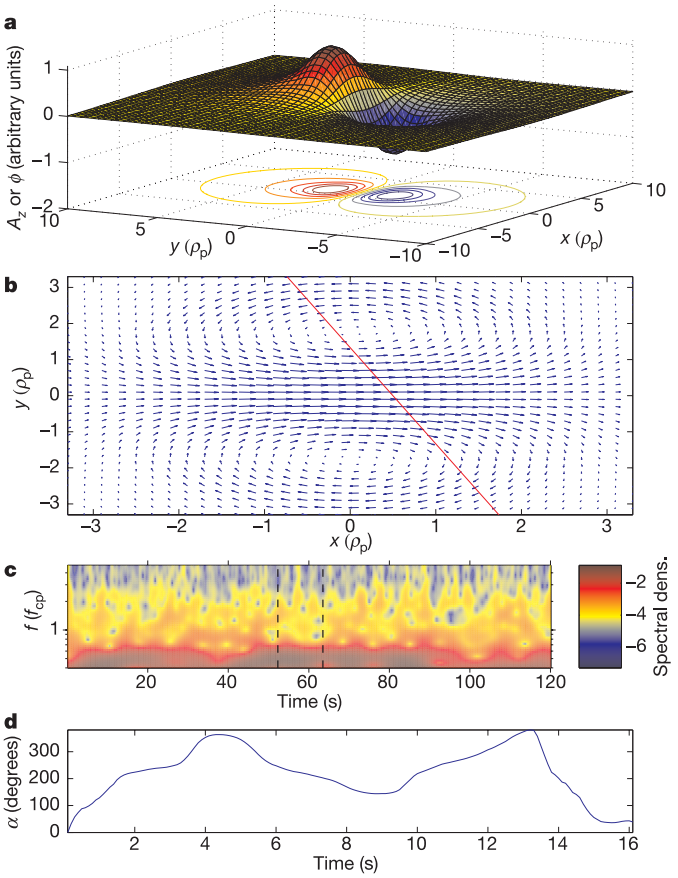


Figure 4 | Analytical DKA vortex solution and a vortex model for the turbulence in the cusp. **a**, Magnetic vector potential A_z or scalar potential Φ of a dipole Alfvénic vortex. **b**, Magnetic field in perpendicular plane calculated from A_z. Note the rotation of the field vectors. The red line is an example of a spacecraft trajectory through the vortex field. **c**, Magnetic field spectral intensity from a model time series for Alfvénic turbulence consisting of a series of vortices in the plasma. Compare with Fig. 2c. **d**, the angle α between instantaneous perpendicular magnetic field and a simulated spacecraft trajectory through the model field for the time period marked with vertical lines in **c**. Compare with Fig. 3c.

indeed observed short scale coherent DKA vortices. The angular rotation in Fig. 3c, d is consistent with a vortex chain convected past the spacecraft. We can also estimate the characteristic size of the DKA vortex from these time series. It is difficult to say when one vortex ends and the next begins, but we note (Fig. 3c, d) that the typical time period is of the order of 10–20 s, which corresponds to spatial radial scales of the order of $(3–6)\rho_p$. Thus, we have two independent methods for calculating the transverse spatial scale that are in agreement with each other—one from the spatial configuration of the spacecraft at the same instant of time, and the other from the time series as the vortices are convected past the spacecraft. This is also in good agreement with the theoretical estimate[8,9], which gives the radial scale-size as $\sim 3\rho_p$. The vortices propagate along (across) the geomagnetic field lines with a speed comparable to the Alfvén speed, v_A ($(0.01–0.1)v_A$).

Strong electromagnetic wave turbulence in a non-uniform magnetoplasma can be considered as an ensemble of nonlinear dispersive Alfvén and drift-waves, DKA vortices, and quasi-magnetostatic structures[23]. We have constructed an isolated DKA vortex model[8,9] for conditions applicable to the cusp plasma, and used it to extract the main characteristic features to be compared with our observational data (Fig. 4). An example of the magnetic field from such a model is shown in Fig. 4c. The spectrogram has been obtained from simulating a spacecraft trajectory through a model field, the temporal fluctuations depending solely upon a Doppler shift of quasi-static DKA vortices (compare Fig. 4b). The spectra obtained are broadband with the largest intensities in the lowest frequencies, and can be compared to the observed field in Fig. 2c. The angular rotation of the magnetic field, while passing through a vortex chain (Fig. 4d), is in perfect agreement with our observations (Fig. 3c).

Received 15 March; accepted 16 June 2005.

1. Frisch, U. *Turbulence: The Legacy of A.N. Kolmogorov* (Cambridge Univ. Press, Cambridge, 1995).
2. Petviashvili, V. I. & Pokhotelov, O. A. *Solitary Waves in Plasmas and in the Atmosphere* (Gordon and Breach Science Publishers, Philadelphia, 1992).
3. Horton, W. & Hasegawa, A. Quasi-two-dimensional dynamics of plasmas and fluids. *Chaos* **4**, 227–251 (1994).
4. Pokhotelov, O. A., Stenflo, L. & Shukla, P. K. Nonlinear structures in the earth's magnetosphere and atmosphere. *Plasma Phys. Rep.* **22**, 852–863 (1996).
5. Horton, W. Drift waves and transport. *Rev. Mod. Phys.* **71**, 735–778 (1999).
6. Shukla, P. K., Yu, M. Y. & Varma, R. K. Formation of kinetic Alfvén vortices. *Phys. Lett. A* **109**, 322–324 (1985).
7. Petviashvili, V. I. & Pokhotelov, O. A. Dipole Alfvén vortices. *JETP Lett.* **42**, 54–56 (1985).

8. Shukla, P. K., Yu, M. Y. & Stenflo, L. Electromagnetic drift vortices. *Phys. Rev. A* **34,** 3478–3480 (1986).
9. Liu, J. & Horton, W. The intrinsic electromagnetic solitary vortices in magnetized plasma. *J. Plasma Phys.* **36,** 1–24 (1986).
10. Burlaga, L. F. A heliospheric vortex street. *J. Geophys. Res.* **95,** 4333–4336 (1990).
11. Zhou, Y., Matthaeus, W. H. & Dmitruk, P. Magnetohydrodynamic turbulence and time scales in astrophysical and space plasmas. *Rev. Mod. Phys.* **76,** 1015–1035 (2004).
12. Hones, E. W. *et al.* Further determination of the characteristics of magnetospheric plasma vortices with ISEE 1 and 2. *J. Geophys. Res.* **86,** 814–820 (1981).
13. Chmyrev, V. M. *et al.* Alfvén vortices and related phenomena in the ionosphere and the magnetosphere. *Physica Scripta* **38,** 841–854 (1988).
14. Chmyrev, V. M. *et al.* Vortex structures in the ionosphere and magnetosphere of the earth. *Planet. Space Sci.* **39,** 1025–1030 (1991).
15. Stasiewicz, K. *et al.* Small scale Alfvénic structure in the aurora. *Space Sci. Rev.* **92,** 423–533 (2000).
16. Savin, S. *et al.* Turbulent boundary layer at the border of geomagnetic trap. *JETP Lett.* **74,** 547–551 (2001).
17. Hasegawa, H. *et al.* Transport of solar wind into Earth's magnetosphere through rolled-up Kelvin-Helmholtz vortices. *Nature* **430,** 755–758 (2004).
18. Spolaore, M. *et al.* Vortex-induced diffusivity in reversed field pinch plasmas. *Phys. Rev. Lett.* **93,** 215003 (2004).
19. Escoubet, C. P., Schmidt, R. & Goldstein, M. L. Cluster—science and mission overview. *Space Sci. Rev.* **79,** 11–32 (1997).
20. Weiland, J. *Collective Modes in Inhomogeneous Plasma* (IoP Publishing, Bristol, 2000).
21. Sundkvist, D. *et al.* Multi-spacecraft determination of wave characteristics near the proton gyrofrequency in high-altitude cusp. *Ann. Geophys.* **23,** 983–995 (2005).
22. Samson, J. C. Some comments on the descriptions of the polarization states of waves. *Geophys. J. R. Astr. Soc* **61,** 115–129 (1980).
23. Volokitin, A. S. & Dubinin, E. M. The turbulence of Alfvén waves in the polar magnetosphere of the earth. *Planet. Space Sci.* **37,** 761–765 (1989).

Acknowledgements The authors thank the FGM, EFW and CIS Cluster instrument teams for supplying data for this study. Gratitude goes to T. D. de Wit for his help with the wavelet calculations. The research of D.S., V.K. and P.K.S. was partially supported by the European Commission. The research of A.V. was supported by the Swedish Research Council.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to D.S. (davids@irfu.se).

LETTERS

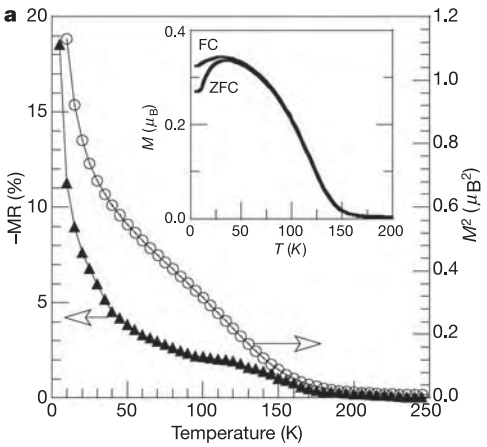
Negative lattice expansion from the superconductivity–antiferromagnetism crossover in ruthenium copper oxides

A. C. Mclaughlin[1], F. Sher[2,3] & J. P. Attfield[2]

The mechanism of high-transition-temperature (high-T_c) superconductivity in doped copper oxides is an enduring problem. Antiferromagnetism is established as the competing order[1,2], but the relationship between the two states in the intervening 'pseudogap' regime has become a central puzzle[3]. The role of the crystal lattice, which is important in conventional superconductors, also remains unclear. Here we report an anomalous increase of the distance between copper oxide planes on cooling, which results in negative thermal volume expansion, for layered ruthenium copper oxides[4,5] that have been doped to the boundary of antiferromagnetism and superconductivity. We propose that a crossover between these states is driven by spin ordering in the ruthenium oxide layers, revealing a novel mechanism for negative lattice expansion in solids. The differences in volume and lattice strain between the distinct superconducting and antiferromagnetic states can account for the phase segregation phenomena found extensively in low-doped copper oxides, and show that Cooper pair formation is coupled to the lattice. Unusually large variations of resistivity with magnetic field are found in these ruthenium copper oxides at low temperatures through coupling between the ordered Ru and Cu spins.

Layered ruthenium copper oxides are most easily stabilized when the copper oxide planes are in the low doped ('underdoped') superconducting regime. The Ru spins order at 100–140 K, and copper oxide superconductivity is observed for $T_c \lesssim 50$ K (ref. 6). We have investigated how the magnetic order in the RuO_2 layers influences the CuO_2 planes at 5% hole-doping, which is the lower limit for superconductivity in copper oxides. Polycrystalline 1222-type $RuSr_2R_{1.1}Ce_{0.9}Cu_2O_{10}$ ceramics (R is a mixture of Nd and other trivalent rare earth elements) were prepared by sintering oxide pellets at 1,000–1,100 °C in air. The pellets are semiconducting down to 4 K, with resistivities at 300 K of 50–100 mΩ cm, and do not show superconducting transitions or resistive anomalies at the magnetic transitions described below.

Two spin ordering transitions are observed in magnetoresistance and magnetization measurements on $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ (Fig. 1a) and are assigned from magnetic neutron scattering (Fig. 2a). Additional neutron diffraction peaks from a (1/2 1/2 1/2) magnetic superstructure appear at temperatures below the Ru spin ordering transition at $T_{Ru} = 140$ K (Fig. 2a). The magnetic intensities down to 60 K are fitted by a model of antiferromagnetically ordered Ru moments aligned in the c direction, as found in the related 1212-type ruthenium copper oxide $RuSr_2GdCu_2O_8$ (ref. 7). (The 1212 materials have only one rare earth layer between ruthenium copper oxide slabs, whereas two layers are present in the 1222 homologues.) The slight canting of Ru moments, which creates a small in-plane ferromagnetism in this and other layered ruthenium copper oxides, is not seen by neutron diffraction but is evidenced in



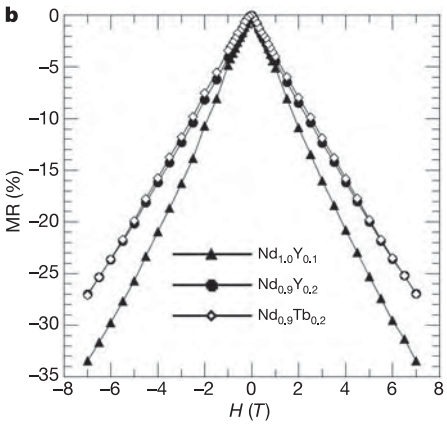


Figure 1 | Magnetization and electronic transport measurements for $RuSr_2R_{1.1}Ce_{0.9}Cu_2O_{10}$. Measurements were made using a 7-T physical properties measurement system. Magnetoresistance, the change of electrical resistance ρ in applied magnetic field H, is defined as $MR_H = (\rho(H) - \rho(0))/\rho(0)$. **a**, The correlation between negative magnetoresistance (−MR) and the square of the magnetization (M^2) of $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ (both measured at $H = 5$ T), which is typical for ceramic magnetoresistive oxides. μ_B, Bohr magneton. The increase in both quantities below 140 K reveals the Ru spin ordering transition (T_{Ru}), and the further sharp rises below 60 K coincide with the ordering of Cu spins (T_{Cu}). Inset, the divergence of field-cooled (FC) and zero field-cooled (ZFC) magnetizations measured in a low field (0.1 T) also corresponds to T_{Cu}. The transitions are corroborated by the neutron scattering results in Fig. 2c. **b**, The field dependences of magnetoresistance for several ceramic $RuSr_2R_{1.1}Ce_{0.9}Cu_2O_{10}$ samples (key shows R). These are again typical for ceramic spin polarized materials, with a high field $\sim|H|$ dependence and additional low field contributions from spin-polarized tunnelling across domain[29] or phase boundaries.

[1]Department of Chemistry, University of Aberdeen, Meston Walk, Aberdeen AB24 3UE, UK. [2]Centre for Science at Extreme Conditions and School of Chemistry, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, UK. [3]Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK.

magnetization measurements. Below $T_{Cu} = 60$ K, the Cu spins order antiferromagnetically in the *a*–*b* plane with a (1/2 1/2 0) superstructure—the fitted Ru and Cu spin model is shown in Fig. 2b. Ordering of both Cu and Ru spins has not been previously reported in layered ruthenium copper oxides, and is also surprising because long range Cu spin order persists only up to 2% doping in simple copper oxides such as $La_{2-x}Sr_xCuO_4$ (although short range antiferromagnetism is observed up to 5% doping). The same two spin ordering transitions are observed by neutron diffraction in $RuSr_2NdY_{0.1}Ce_{0.9}Cu_2O_{10}$ and a nominally zero-doped sample, $RuSr_2Nd_{0.9}Y_{0.1}CeCu_2O_{10}$. Both transition temperatures increase with reduced hole-doping, up to $T_{Ru} = 170$ K and $T_{Cu} = 120$ K in the latter sample; these and other results are shown in Supplementary Information.

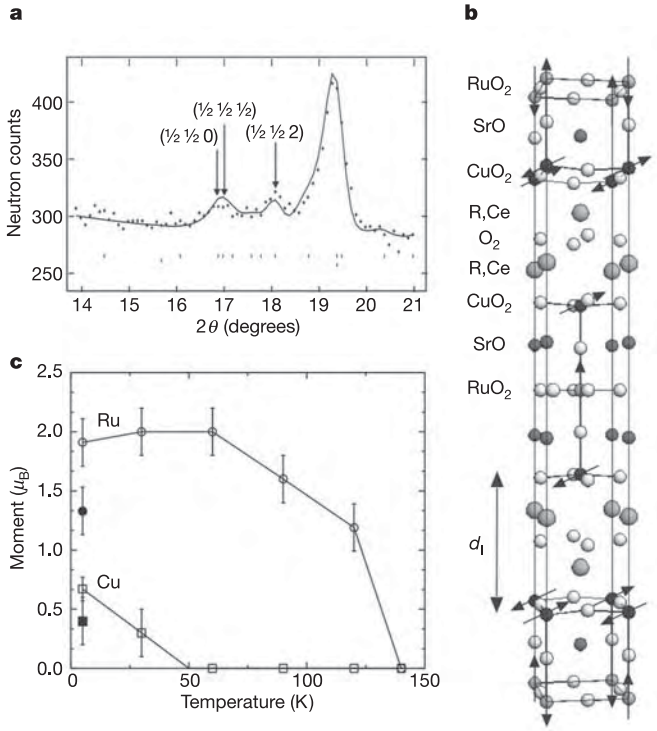


Figure 2 | Magnetic order in $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ at low temperatures. This was determined from high resolution powder neutron diffraction data, collected at a wavelength of 1.5943 Å on instrument SuperD2B at ILL, Grenoble. Magnetic superstructure peaks are labelled on the 4 K diffraction plot (**a**); the fit is calculated from the refined model for structural and antiferromagnetic Ru and Cu spin ordering (**b**). The stacking sequence of metal oxide layers and the interplanar separation of copper oxide planes, d_I, are labelled. Refinement of the tetragonal crystal structure (space group *I*4/*mmm*) revealed no significant oxygen deficiency (<1%) in the R oxide layers, and the nominal doping level of 0.05 is also corroborated by a bond valence sum calculation using parameters derived previously for layered ruthenium copper oxides[30] which gives an estimate of 0.06 ± 0.01 holes per Cu. These analyses confirm that $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ is hole-doped to the normal lower limit for superconductivity. **c**, The temperature variation of the antiferromagnetically ordered Ru and Cu moments (from the slow-cooled experiment described in Fig. 3 legend). Applying a 5-T field at 4 K diminishes the intensities of the antiferromagnetic superstructure peaks (see also Supplementary Information) and the refined moment values (filled symbols in **c**) are consistent with a ~45° canting of both Ru and Cu moments towards a parallel spin structure. The magnetic alignment axis is unclear from our data but is likely to be in the *a*–*b* plane, as found for $RuSr_2GdCu_2O_8$ at high field strengths[9]. Error bars, one estimated s.d., calculated from the Rietveld fits to the neutron diffraction profiles.

Crystal structure refinements of $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ from neutron diffraction data show that tetragonal lattice symmetry is preserved down to 4 K, but the cell parameters reveal an unexpected and remarkable change of behaviour at T_{Ru}. In both slow cooling and slow warming experiments, the unit cell shows a normal (positive) thermal expansion above $T_{Ru} = 140$ K but below this, the cell shrinks with increasing temperature (Fig. 3a). This negative volume expansion has not been found in other copper oxides (or ruthenium copper oxides) and is unprecedented in magnetic transition metal oxides. No unusual changes in the metal-to-oxygen bond distances are observed, but an anomalous expansion in the separation between copper oxide planes (d_I) below T_{Ru} is seen to drive the negative lattice expansion (Fig. 3b). The negative expansion is not observed when the same sample is quenched to 4 K and then warmed rapidly (Fig. 3a).

Negative expansion at temperatures below an electronic transition results from a crossover between two states of differing volume. In ferromagnets, this is the well-known Invar effect[8], driven by a crossover from small to large spin states below the Curie transition, for example, in the Invar alloy, $Fe_{0.64}Ni_{0.36}$, and in the conducting ferromagnetic oxide $SrRuO_3$, although here the volume expansion remains positive below the Curie transition[9]. A distinct, intervalence, mechanism[10] has recently been identified in rare earth compounds such as the fulleride $Sm_{2.75}C_{60}$ (ref. 11) and YbGaGe (ref. 12) that change between large, low temperature R^{2+} and small, high temperature R^{3+} states. Neither of these mechanisms explains the negative expansion of $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$, most obviously because the structure, the Ru and Cu valences, and the Ru spin ordering here are not noticeably different to those in previously-characterized superconducting ruthenium copper oxides, none of which show unusual lattice expansion below T_{Ru} (refs 8, 13, 14). Instead, the proximity of $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ to the boundary between antiferromagnetism and superconductivity suggests that a crossover between these two states is responsible for the negative expansion. However, as the change from positive to negative expansion occurs at 140 K, well above the possible ordering temperatures for Cu antiferromagnetism or superconductivity, the crossover indicates that two high-temperature correlated states are available for the CuO_2 planes in the pseudogap regime. For convenience, we label these as 'para-antiferromagnetic' and 'parasuperconducting'.

We propose that the 5% doped CuO_2 planes in $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ are parasuperconducting above 140 K. This precursor state for superconductivity in underdoped copper oxides shows high temperature pairing fluctuations[15] that correspond to pre-formed Cooper pairs in strong coupling theories[16], but without the phase stiffness between pairs that characterizes the superconducting phase below T_c. When cooled below 140 K, the Ru spin ordering switches the copper oxide state from parasuperconductivity to para-antiferromagnetism—the state of a paramagnet above a conventional antiferromagnetic transition. Fluctuating clusters of antiparallel copper spins in the para-antiferromagnetic phase condense into the observed long range (1/2 1/2 0) order on cooling to $T_{Cu} = 60$ K. The switching action of the Ru spin order on the copper oxide planes could be through ferromagnetic pair-breaking from the small in-plane moment, which destabilizes the parasuperconducting state, or through magnetic Ru–Cu exchange, which induces antiferromagnetic Cu spin correlations and so stabilizes para-antiferromagnetism. The latter seems plausible, given the unusual stabilization of long range Cu spin ordering up to 5% doping.

In the above interpretation, the increase in the inter-copper plane distance d_I below T_{Ru} (Fig. 3b) shows that the pairing correlations in the parasuperconducting state result in a slight attraction between the copper oxide planes, and the release of this stress below T_{Ru} drives the negative lattice expansion. This is corroborated by a previous study of a superconducting $RuSr_2GdCu_2O_8$ sample[17], in which a small discontinuity in d_I was found at $T_{Ru} = 133$ K, although no lattice parameter anomalies or negative expansion resulted. Here, the Ru spin ordering again leads to a slight release of inter-

CuO_2 plane stress, but the pair correlations are stronger in this more highly-doped material so parasuperconductivity survives to low temperatures leading to the zero-resistance state below $T_c = 35$ K.

The proposed transformation between parasuperconducting and para-antiferromagnetic states is first order because of the difference in their lattice volumes. This is confirmed by the volume hysteresis in the slow cooling and warming data around T_{Ru} (Fig. 3a). Furthermore, when the same sample was quenched to 4 K and then warmed rapidly, a low volume state that expands positively up to T_{Ru} was found, although antiferromagnetic Cu spin ordering was still observed below T_{Cu} (see Supplementary Figures). This shows that the crossover between parasuperconducting and para-antiferromagnetic spin states occurs rapidly, but the consequent structural relaxation is slower, requiring several hours in our ceramic samples. Similar timescales are observed for low temperature lattice relaxation in phase-separated manganese oxides[18]. From the difference between quenched and slow cooled $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ lattices, we estimate that the zero temperature superconducting phase has a >0.1% smaller volume and a >0.04% smaller c/a ratio than the antiferromagnetic phase.

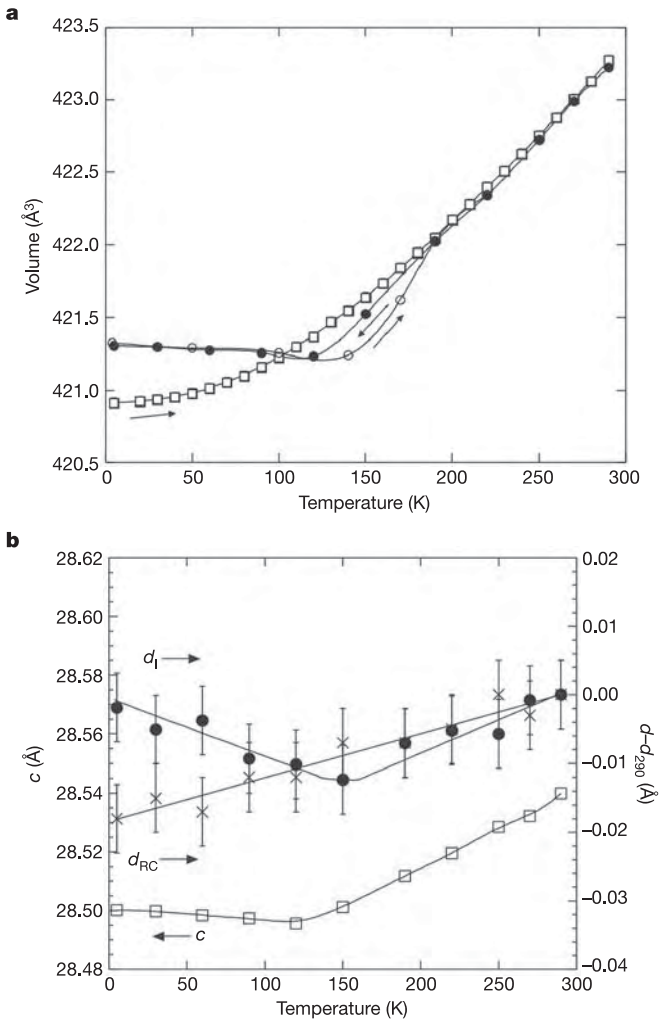


Figure 3 | Temperature variation of the $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ structure, revealing negative thermal expansion below the $T_{Ru} = 140$ K Ru spin ordering temperature. **a**, Volume evolution from three separate powder neutron diffraction experiments on the same sample. 'Slow warming' (open circles; sample was held at 4 K for 5 h then warmed to 290 K at an average rate of 18 K h^{-1}) and 'slow cooling' (filled circles; average cooling rate, 16 K h^{-1}) data were collected on diffractometer SuperD2B. In the 'fast warming' experiment (open squares) on ILL diffractometer D20 (wavelength 2.42 Å), the sample was quickly cooled to 4 K (at 430 K h^{-1}), and diffraction patterns were recorded while warming at 60 K h^{-1}. A negative volume expansion is seen below T_{Ru} in both the slow warming and the slow cooling experiments. The a and c cell parameters show the same trends as the volumes. The negative volume (V) expansion of $\alpha_V = d(\ln V)/dT = -1.43 \times 10^{-6}$ K^{-1} in the 4–140 K interval comes principally from the c-axis contribution ($\alpha_c = -1.39 \times 10^{-6}$ K^{-1}), although the a axis also shows a small negative expansion ($\alpha_a = -0.02 \times 10^{-6}$ K^{-1}). **b**, The c-axis length is usefully split into contributions from the thickness of the CuO_2.SrO.RuO_2.SrO.CuO_2 ruthenium copper oxide slabs (d_{RC}) and the interplanar separation of copper oxide layers through the $(R,Ce)_2O_2$ blocks (d_I; marked on Fig. 2b) such that $c/2 = d_{RC} + d_I$. The changes in these distances relative to their 290 K values ($d_{RC} = 8.174$, $d_I = 6.095$ Å) are shown on the right-hand scale of **b**. d_{RC} has a monotonic temperature variation, but the separation of the CuO_2 layers d_I shows an anomalous expansion below T_{Ru} that mirrors the overall negative expansion of c. Error bars, one estimated s.d., calculated from the Rietveld fits to the neutron diffraction profiles.

Ru spin ordering in layered ruthenium copper oxides significantly changes the electronic behaviour of low-doped copper oxide planes by extending the limit of long range antiferromagnetic Cu spin order up to the boundary of superconductivity near 5% hole-doping. Another of the consequences of this 'alternative physics' for low-doped copper oxides are large values of negative magnetoresistances (−MR), observed in all of our $RuSr_2R_{1.1}Ce_{0.9}Cu_2O_{10}$ samples (Fig 1b). Figure 1a shows that −MR initially rises to ~2% below T_{Ru}, as is found in superconducting ruthenium copper oxides[19–21], but increases dramatically on cooling below T_{Cu}, unlike its behaviour in the superconducting analogues. The $-MR_{7T} = 34\%$ value for $RuSr_2Nd_{1.0}Y_{0.1}Ce_{0.9}Cu_2O_{10}$ is the largest reported for copper oxides at this field strength. Anisotropic magnetoresistances with $-MR_{7T}$ up to 20% are found in single crystals of 1% doped antiferromagnetic copper oxides[22,23] (and can increase in higher fields, up to 80% for $La_{1.99}Sr_{0.01}CuO_4$ at 14 T). Magnetoresistance results from a field-induced canting of antiferromagnetically-ordered Cu spins towards a parallel alignment in the 1% doped crystals. We find that the Ru spin order in $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ enhances this effect, as neutron diffraction shows that both Cu and Ru spins are canted towards a parallel alignment by an applied field (Fig. 2c). The large negative magnetoresistance effects observed in our 1222 ruthenium copper oxides are comparable to those in spin-polarized conductors such as colossal magnetoresistance manganese oxide perovskites[24,25] and Sr_2FeMoO_6 (ref. 26). Phase segregation enhances magnetoresistance in the manganese oxides, and we speculate that parasuperconducting–antiferromagnetic phase coexistence could also contribute to the large magnetoresistance in ruthenium copper oxides.

Our results and our proposed two-state model have several implications for superconductivity in copper oxides. They show that Ru spin ordering creates an alternative physics for low-doped copper oxides that helps to clarify the relationship between antiferromagnetism and superconductivity. The volume difference between the two states demonstrates that Cooper pair formation is coupled to the lattice through a slight attraction between CuO_2 planes, and so a charge–lattice coupling may be necessary for the mechanism of superconductivity. Coupling between carriers and out-of-plane phonons has been demonstrated in recent studies of superconducting bismuth copper oxides[27]. The superconducting and antiferromagnetic correlations are competitive, with different lattice volumes and c/a ratios for their para-states at a given doping concentration, although the structural transformation between phases is slow. The lattice strains between the two phases may account for the many observations of phase separation, coexistence, and heterogeneities in the pseudogap regime, without requiring electronic segregation into hole-rich and hole-poor regions. A recent theoretical analysis has found that strains resulting from electron–lattice coupling are sufficient to create phase-separated textures over a range of length scales in transition metal oxides[28].

Finally, we note more general materials science implications. $RuSr_2Nd_{0.9}Y_{0.2}Ce_{0.9}Cu_2O_{10}$ has revealed a novel mechanism for

negative thermal expansion through a crossover between two different types of antiparallel spin correlation, and other crossovers between competing exotic states may lead to new classes of electronic material showing this useful phenomenon. The ordered Ru spins also enhance spin polarized transport in the doped antiferromagnetic copper oxide planes, leading to magnetoresistances comparable to those of spintronic materials.

Received 10 December 2004; accepted 17 May 2005.

1. Kang, H. J. *et al.* Antiferromagnetic order as the competing ground state in electron-doped $Nd_{1.85}Ce_{0.15}CuO_4$. *Nature* **423,** 522–525 (2003).
2. Hinkov, V. *et al.* Two-dimensional geometry of spin excitations in the high-transition-temperature superconductor $YBa_2Cu_3O_{6+x}$. *Nature* **430,** 650–654 (2004).
3. Orenstein, J. & Millis, A. J. Advances in the physics of high-temperature superconductivity. *Science* **288,** 468–474 (2000).
4. Felner, I., Asaf, U., Levi, Y. & Millo, O. Coexistence of magnetism and superconductivity in $R_{1.4}Ce_{0.6}RuSr_2Cu_2O_{10-delta}$ (R = Eu and Gd). *Phys. Rev. B* **55,** R3374–R3377 (1997).
5. Noce, C. & Vecchione, A., Cuoco, M. & Romano, A. (eds) *Ruthenate and Rutheno-Cuprate Materials* (Springer, Berlin, 2002).
6. Bernhard, C. *et al.* Coexistence of ferromagnetism and superconductivity in the hybrid ruthenate-cuprate compound $RuSr_2GdCu_2O_8$ studied by muon spin rotation and dc magnetization. *Phys. Rev. B* **59,** 14099–14107 (1999).
7. Lynn, J. W., Keimer, B., Ulrich, C., Bernhard, C. & Tallon, J. L. Antiferromagnetic ordering of Ru and Gd in superconducting $RuSr_2GdCu_2O_8$. *Phys. Rev. B* **61,** 14964–14967 (2000).
8. van Schilfgaarde, M., Abrikosov, I. A. & Johansson, B. Origin of the Invar effect in iron-nickel alloys. *Nature* **400,** 46–49 (1999).
9. Kiyama, T., Yoshimura, K., Kosuge, K., Ikeda, Y. & Bando, Y. Invar effect of $SrRuO_3$: itinerant electron magnetism of Ru 4*d* electrons. *Phys. Rev. B* **54,** R756–R759 (1996).
10. Sleight, A. Zero-expansion plan. *Nature* **425,** 674–676 (2003).
11. Arvanitidis, J., Papagelis, K., Margadonna, S., Prassides, K. & Fitch, A. N. Temperature-induced valence transition and associated lattice collapse in samarium fulleride. *Nature* **425,** 599–602 (2003).
12. Salvador, J. R., Guo, F., Hogan, T. & Kanatzidis, M. G. Zero thermal expansion in YbGaGe due to an electronic valence transition. *Nature* **425,** 702–705 (2003).
13. McLaughlin, A. C., Attfield, J. P. & Tallon, J. L. A variable temperature structural study of the ferromagnetic superconductor $RuSr_2GdCu_2O_8$. *Int. J. Inorg. Mater.* **2,** 95–99 (2000).
14. Shi, L., Li, G., Fan, X. J., Feng, S. J. & Li, X.-G. Structural, transport and magnetic properties of $RuSr_2Sm_{1.4}Ce_{0.6}Cu_2O_{10-\delta}$. *Physica C* **399,** 69–74 (2003).
15. Emery, V. J. & Kivelson, S. A. Importance of phase fluctuations in superconductors with small superfluid density. *Nature* **374,** 434–437 (1995).
16. Yanase, Y. & Yamada, K. Theory of pseudogap phenomena in high-T-c cuprates based on the strong coupling superconductivity. *J. Phys. Soc. Jpn* **68,** 2999–3015 (1999).
17. Chmaissem, O., Jorgensen, J. D., Shaked, H., Dollar, P. & Tallon, J. L. Crystal and magnetic structure of ferromagnetic superconducting $RuSr_2GdCu_2O_8$. *Phys. Rev. B* **61,** 6401–6407 (2000).
18. Matsukawa, M. *et al.* Stretched exponential behaviour in remanent lattice striction of a $(La,Pr)_{1.2}Sr_{1.8}Mn_2O_7$ bilayer manganite single crystal. *Phys. Rev. B* **70,** 132402 (2004).
19. Chen, X. H. *et al.* Transport properties and specific heat of $RuSr_2GdCu_2O_8$ and $RuSr_2Gd_{1.4}Ce_{0.6}Cu_2O_y$ in magnetic fields. *Phys. Rev. B* **63,** 064506 (2001).
20. Awana, V. P. S., Ichihara, S., Karppinen, M. & Yamauchi, H. Comparison of magneto-superconductive properties of $RuSr_2GdCu_2O_{8-\delta}$ and $RuSr_2Gd_{1.5}Ce_{0.5}Cu_2O_{10-\delta}$. *Physica C* **378,** 249–254 (2002).
21. McCrone, J. E. *et al.* Magnetotransport properties of doped $RuSr_2GdCu_2O_8$. *Phys. Rev. B* **68,** 064514 (2003).
22. Ando, Y., Lavrov, A. N. & Komiya, S. Anisotropic magnetoresistance in lightly doped $La_{2-x}Sr_xCuO_4$: Impact of antiphase domain boundaries on the electron transport. *Phys. Rev. Lett.* **90,** 247003 (2003).
23. Lavrov, A. N. *et al.* Spin-flop transition and the anisotropic magnetoresistance of $Pr_{1.3-x}La_{0.7}Ce_xCuO_4$: unexpectedly strong spin-charge coupling in the electron-doped cuprates. *Phys. Rev. Lett.* **92,** 227003 (2004).
24. Rao, C. N. R. & Raveau, B. (eds) *Colossal Magnetoresistance, Charge Ordering and Related Properties of Manganese Oxides* (World Scientific, Singapore, 1998).
25. Tokura, Y. (ed.) *Colossal Magnetoresistive Oxides* (Gordon and Breach Science, New York, 2000).
26. Kobayashi, K. L., Kimura, T., Sawada, H., Terakura, K. & Tokura, Y. Room-temperature magnetoresistance in an oxide material with an ordered double-perovskite structure. *Nature* **395,** 677–680 (1998).
27. Cuk, T. *et al.* Coupling of the B_{1g} phonon to the antinodal electronic states of $Bi_2Sr_2Ca_{0.92}Y_{0.08}Cu_2O_{8+\delta}$. *Phys. Rev. Lett.* **93,** 117003 (2004).
28. Ahn, K. H., Lookman, T. & Bishop, A. R. Strain-induced metal–insulator phase coexistence in perovskite manganites. *Nature* **428,** 401–404 (2004).
29. Mclaughlin, A. C., Zhou, W., Attfield, J. P., Fitch, A. N. & Tallon, J. L. The structure and microstructure of the ferromagnetic superconductor $RuSr_2GdCu_2O_8$. *Phys. Rev. B* **60,** 7512–7516 (1999).
30. Mclaughlin, A. C., Attfield, J. P., Asaf, U. & Felner, I. Chemical control of hole-doped superconductivity and magnetism in $Gd_{2-x}Ce_xRuSr_2Cu_2O_{10-\delta}$. *Phys. Rev. B* **68,** 014503 (2003).

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

Acknowledgements We thank P. Littlewood, P. Monthoux and N. Mathur for discussions, and P. Henry and E. Suard for assistance with the neutron experiments. We also acknowledge the Royal Society of Edinburgh for a SEELLD research fellowship (A.C.M.), the Ministry of Science and Technology, Government of Pakistan for a studentship (F.S.), and the UK EPSRC for beam time provision and financial support.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.P.A. (j.p.attfield@ed.ac.uk).

LETTERS

A late Eemian aridity pulse in central Europe during the last glacial inception

F. Sirocko[1], K. Seelos[1], K. Schaber[1], B. Rein[1], F. Dreher[1], M. Diehl[1], R. Lehne[1], K. Jäger[1], M. Krbetschek[2] & D. Degering[2]

Investigating the processes that led to the end of the last interglacial period is relevant for understanding how our ongoing interglacial will end, which has been a matter of much debate (see, for example, refs 1,2). A recent ice core from Greenland demonstrates climate cooling from 122,000 years ago[3] driven by orbitally controlled insolation, with glacial inception at 118,000 years ago[4–8]. Here we present an annually resolved, layer-counted record of varve thickness, quartz grain size and pollen assemblages from a maar lake in the Eifel (Germany), which documents a late Eemian aridity pulse lasting 468 years with dust storms, aridity, bushfire and a decline of thermophilous trees at the time of glacial inception. We interpret the decrease in both precipitation and temperature as an indication of a close link of this extreme climate event to a sudden southward shift of the position of the North Atlantic drift, the ocean current that brings warm surface waters to the northern European region. The late Eemian aridity pulse occurred at a 65° N July insolation of 416 W m^{-2}, close to today's value of 428 W m^{-2} (ref. 9), and may therefore be relevant for the interpretation of present-day climate variability.

Greenland ice cores[3] and marine sediments from the North Atlantic[10], for example, have produced a very consistent picture of the last glacial cycle, with a succession of 26 stadials and interstadials, which have also been observed on the European continent[11–14].The warm interstadials were characterized by a strong North Atlantic drift (Fig. 1) with warm waters reaching far into the north[15], which provided sensible heat, latent heat and precipitation for the central European continent. However, the cold stadials were associated with ice-rafted debris from surging continental glaciers[16]. The first of the large ones of these cold stadials was C25 at 115 kyr BP (Figs 2 and 3). Marine evidence from the North Atlantic indicated an even earlier stadial (C26 of ref. 17), which we now also observe on the central European continent as a 468-year-long cold and arid event that we describe in detail in Fig. 4 and call the late Eemian aridity pulse (LEAP). We avoid the C-nomenclature because this event can hardly be explained by surging ice; it occurred during the inception, when ice sheets just started to grow.

The interglacial immediately before the LEAP was characterized by temperate forests for a duration of 10 kyr (ref. 18) (the Eemian *sensu stricto*), which was recently U/Th-dated from about 127 to 117 kyr BP (ref. 19). The end of the Eemian was associated with a global sea level drop, dated at 118 ± 1 kyr BP (ref. 7). This was the last glacial inception (LGI), when continental glaciers first started to grow. Fully coupled ocean and atmosphere climate models have recently studied this initial growth of ice and corroborated that the severe decrease in Northern Hemisphere insolation is fully sufficient to initiate this inception and growth of ice at 118 kyr BP (refs 5,6). In particular the model of ref. 4 finds a response to the insolation change by a strong cooling of the Norwegian–Greenland Sea which enhances the temperature gradient across the North Atlantic, leading to lower temperatures and precipitation over central Europe. These hydrological changes are matched by pollen-based temperature reconstructions that reveal a 3 °C summer cooling and a 10 °C winter cooling at the end of the Eemian in central Europe[20]. In what follows we will now evaluate a new annually laminated lake sediment record from a dry maar lake in the Eifel (Germany), which covers the entire time from the C24 event into the Eemian and allows us to extract more detailed information on the environmental changes and their rate of change during the inception and LEAP.

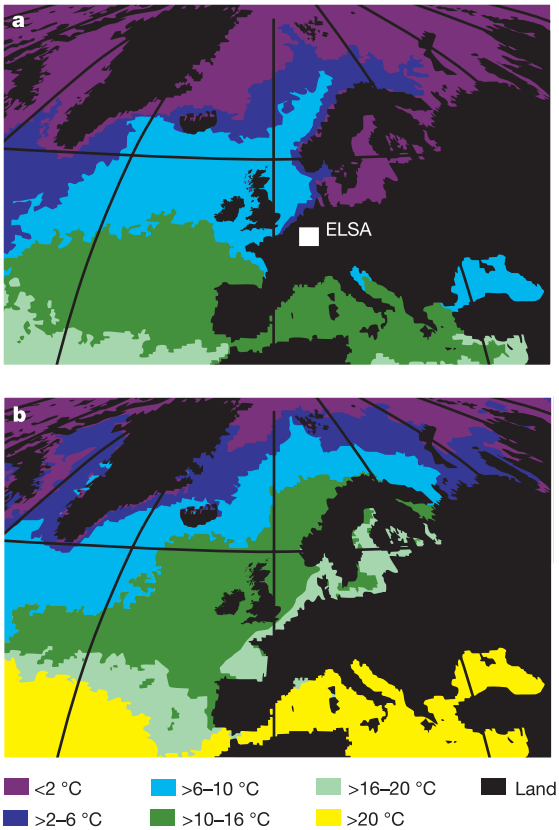


Figure 1 | Equal-area projection of North Atlantic sea surface temperatures derived from the satellite-borne Advanced Very High Resolution Radiometer. Images taken during the winter (**a**) and the summer (**b**) of 2002, together with the locations of the Eifel and ELSA cores.

[1]Institute of Geoscience, Johannes Gutenberg-University, 55099 Mainz, Germany. [2]Saxon Academy of Sciences in Leipzig, Bernhard-von-Cotta-Strasse 4, 09596 Freiberg/Sa, Germany.

The west Eifel volcanic field shows 68 maar structures[21], 6 of which are Holocene lakes with varved (annually laminated) sediments; 62 maar lakes have dried out at some time during the middle and late Pleistocene. The Eifel Laminated Sediment Archive (ELSA) Project cored 30 of these dry maar lakes during the past 5 years. These lakes were 800–1,500 m wide with a maximum palaeo-waterdepth of 150 m; bottom water in such small and deep lakes with no large inflow or outflow is at least seasonally anoxic[22], which inhibits bioturbation and allows layer preservation.

Accelerator mass spectrometry (AMS) ^{14}C and luminescence dates were used for absolute age control, but the final stratigraphy came from tuning the ELSA greyscale curve to the NorthGRIP $\delta^{18}O$ record (see Methods). The undisturbed sections of four ELSA cores were combined to the 'ELSA greyscale stack 2005' (Fig. 2). The tuned stratigraphy was applied down to the C24 event. To study the time earlier than C24 we switched to a varve-counted chronology and used core HL2 in particular. The continuous annual layer counts below C24 were 14.4 kyr before a massive coarse grained slump that apparently destroyed parts of the earlier Eemian section of core HL2.

To study the LEAP and LGI with annual resolution we used thin sections 10 cm in length that were prepared with standard methods of resin impregnation for soft sediments. Thin sections were ground down to 30 μm thickness and used for annual layer (varve) counting, grain size analysis and determination of charcoal content. Quartz grain size was determined with sub-millimetre resolution by application of the RADIUS method[23]. Charcoal content was estimated from the thin sections as absent, present or frequent. Varve classification and photos of the LEAP are documented in Supplementary Fig. 1. Preparation of pollen samples and full pollen spectra for two cores, which show an identical pollen succession, are given in Supplementary Fig. 2. Turbidites, a blockslide, and tephra layers were detected on the split cores and thin sections and removed from the records when plotted against age (Figs 2 and 3).

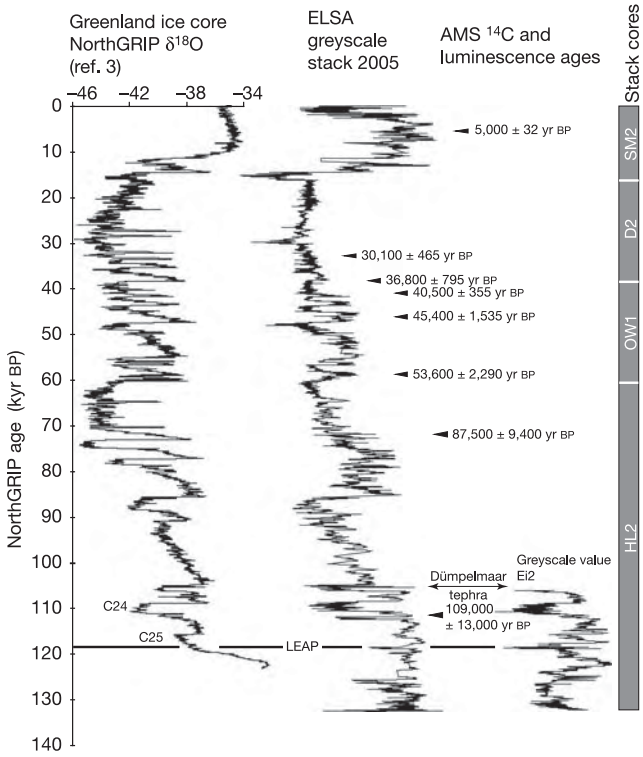


Figure 2 | Climate change during the last glacial cycle. NorthGRIP ice-core $\delta^{18}O$ (temperature) variations[3]. ELSA greyscale stack, constructed from the individual cores from the dry maar lakes west of Hoher List (HL2), Oberwinkel (OW1), Dehner (D2) and the maar lake of Schalkenmehren (SM2). The sections used for the stack all come from a palaeo-waterdepth of 20–60 m. Dark sediment in these cores is caused by a high organic carbon content during warm phases; light sections are coloured by abundant silt-sized quartz. The LEAP is incorporated into the figure at its varve-counted age of 118 kyr BP. Information on the stratigraphic control points is given in Methods. The ELSA greyscale stack can be downloaded from http://www.uni-mainz.de/FB/Geo/Geologie/sedi/index.html

Figure 3 shows the varve thickness, quartz grain size and pollen for the entire last interglacial in the Eifel area. According to the varve counts the Eemian *sensu stricto* ended at 118 kyr BP with the final deterioration of *Abies*, immediately followed by the LEAP with an increase in varve thickness, loess content and charcoal. The deterioration of the temperate forest lasted for several centuries; broadleaf forest taxa such as *Carpinus* reappeared soon after the cooling and continued to spread until the C24 event, when climate and environment changed drastically to a shrub tundra landscape. The reappearance of *Carpinus* after the decline in *Abies* was also observed in the Ribains dry maar record from the French Massif Central[24] about 600 km south of the Eifel. In general, the pollen succession of the entire last interglacial in the Eifel is much more similar to the French Ribains pollen record than to records from northern and southern Germany, where *Carpinus* did not reappear after the end of the *Abies* zone (see a compilation of records in ref. 25). Apparently, the gradient in sea surface temperature over the North Atlantic was paralleled by a temperature gradient also over land, today (Fig. 1) and also during the end of the last interglacial.

A magnification of the HL2 data for the time of the LEAP is shown in Fig. 4 and documents 52 dust layers in 468 years. The event starts and ends abruptly within less than 20 years. Dust storms became perennial only during the middle of the event. The thermophilous trees *Alnus*, *Quercus*, *Corylus*, *Tilia* and *Carpinus* disappeared within a century during the early part of the LEAP. Charcoal layers indicate that the vegetation change was caused not only by a decrease in temperature but in particular by an increase in aridity. Grass pollen abundance was high during the entire LEAP, which indicates an opening of the forests, but trees apparently continued to grow somewhere, most probably on stands with good access to groundwater and high summer temperatures, as in the nearby Mosel and Rhine valley. The end of the LEAP started with an increase in the abundance of *Picea*, and terminated with a 30-year interval of *Betula* and *Pinus* dominance, which apparently formed a pioneer vegetation in the partly deforested parts of the high Eifel. The areas covered by the grasses and pioneer vegetation were then afforested again by interglacial tree taxa, a process of a few decades, which appears as the very abrupt change in the pollen record at the very end of the LEAP in Fig. 4.

The LEAP might very well represent the abrupt 6 b anomaly in the Grand Pile/Vosges/France record[26], the transition between pollen zone 20/21 in the French Ribains maar/Massif Central[24], the first late Eemian increase of grasses at 39.3 m in the north German Rederstall core[25] and the H6 zone at Hollerup/Denmark[27]. In particular the Hollerup record clearly shows the severe aridity during the LEAP. However, this event is not visible in several other central European pollen records. This could probably be explained because the LEAP was very short and associated with sand/silt layers (our grain size maxima), which are often systematically avoided when sampling for pollen analysis. The time between the LEAP and the start of C24 (118–111 kyr BP) is accordingly badly documented in numerous records. Following our above interpretation, one must assume that for this time there were rather colder conditions in northern Germany and Denmark, but a still temperate forest in the Eifel and France. Accordingly, the duration of the last interglacial in the north (the Eemian *sensu stricto*) would be indeed about 10 kyr (ref. 18) but at least 16 kyr in the southwest of Europe, which we interpret to reflect a shift in the position of the warm north Atlantic drift waters (Fig. 1).

The parallel decrease in precipitation and temperature indicate the most likely cause of the LEAP. The area of the North Atlantic drift and

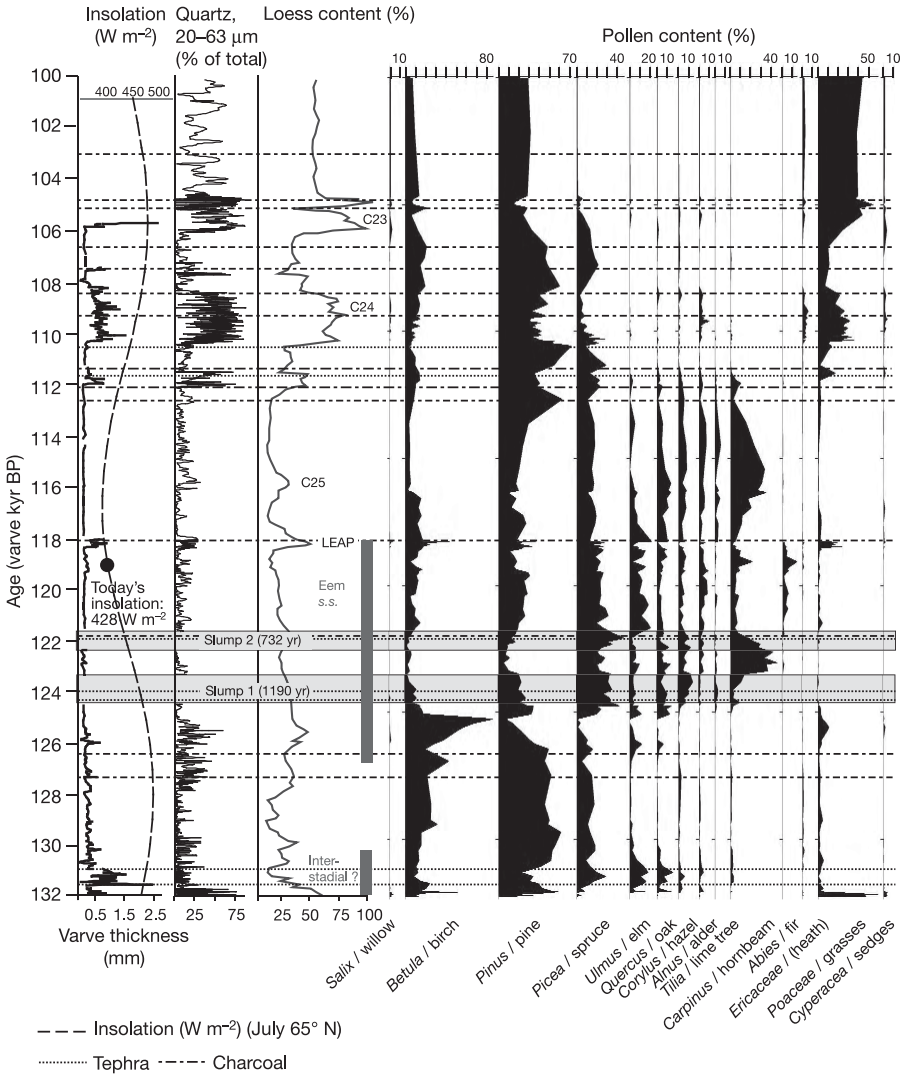


Figure 3 | Climate change in the Eifel between 100 and 132 kyr BP. Varve thickness, abundance of quartz grains in the loess fraction 20–63 μm (applying a 200-point running mean), loess detected by the RADIUS method[23], ash/tuff and charcoal layers and pollen content in core HL2 are plotted against varve-counted age. The start of the Eemian at 127 kyr BP is taken from the dating of last-interglacial stalagmites[19], which we use to determine the number of years lost in the slump. The chronology of the pre-Eemian is again based on varve counts.

its associated temperature anomaly has about the same size as the whole of Europe (Fig. 1). A sudden southward change in the extent or position of the warm ocean surface waters is therefore a very likely cause of the cold and arid climate anomaly. This interpretation is in accordance with the model results of Khodri[4], who observed a decrease in oceanic northward heat transport in response to a lowering of insolation at the end of the last interglacial. Following this interpretation, the end of the LEAP would represent a new northward shift of the North Atlantic drift, but not to latitudes as high in the north as before the LEAP. The onset of the LEAP occurred within less than two decades (see the core photograph in Fig. 4), demonstrating the existence of a sharp threshold, which must be near 416 W m^{-2}, which is the 65° N July insolation for 118 kyr BP (ref. 9). This value is only slightly below today's value of 428 W m^{-2}. Insolation will remain at this level slightly above the inception for the next 4,000 years before it then increases again. Accordingly the model results predict a continuation of our present interglacial for the next 55 kyr (ref. 5), when insolation will decrease for the first time again to the LGI level. However, the Earth will be in a fragile state for the whole of the next 4000 years, and one can only hope that the expected climate extremes of the Anthropocene will not lead to conditions that cross the threshold to glaciation.

METHODS

AMS ^{14}C dates are on the bulk organic fraction after an acid–alkali–acid extraction, measured at the Leibniz Laboratory at Kiel University, Germany. All AMS ^{14}C dates were converted to calendar years as described[28]. Luminescence dates were determined by the Freiberg group of the Saxon Academy of Sciences by infrared radiofluorescence (IR-RF) with a single aliquot regeneration technique[29]. The method is based on the dose dependence of the infrared fluorescence emission at 1.43 eV (865 nm) from potassium feldspar during irradiation and is similar to other optically or thermally stimulated luminescence dating methods. Radionuclide contents of the sediments were analysed by low-level spectrometry. Additional stratigraphic information comes from the occurrence of the phonolitic Dümpelmaar tephra, which was Ar/Ar dated[30] to 116 ± 16 kyr BP and is now tuned to 105 kyr BP, very shortly before the start of DO23 (Fig. 2).The fine tuning of the stratigraphy is based on the greyscale record of Fig. 2 but also includes significant lithological changes in grain size, organic carbon content, lamination thickness and occurrence of light diatom/calcite layers. The similarity between the ELSA greyscale and NorthGRIP δ^{18}O is very high during the last termination, the early MIS3, MIS4 and down to the C24 event. However, the similarity is low for the Holocene and last interglacial because the ELSA interglacial laminae consist of both dark organic carbon-rich layers and light calcite and diatom layers. Similarity is also low during the late MIS3 (that is, DO8 to DO2), which have an organic carbon content too low to be visible in the greyscale.

The stratigraphy of the last interglacial is not performed by any tuning but strictly by varve counting. The entire varve-counted section reaches from 106 kyr BP down to 132 kyr BP. The connection of the floating ELSA varve chronology to the NorthGRIP record is done at the beginning of C24, which is most clearly visible in both records.

Received 22 February; accepted 10 June 2005.

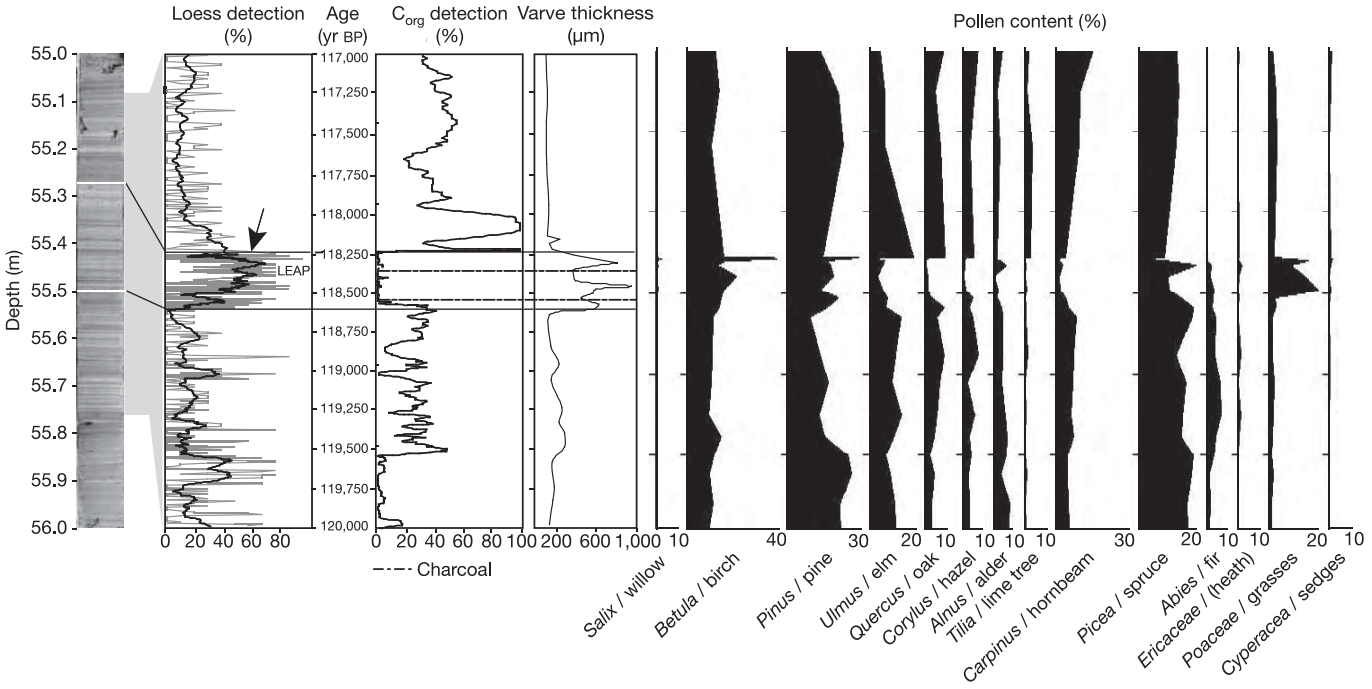


Figure 4 | Climate change in the Eifel during the last glacial inception from 117–120 kyr BP. Core photo of HL2 55–56 m, with varve thickness, abundance of quartz grains in the loess fraction 20–63 μm (applying a 30-point running mean), loess (the arrow indicates a region of 52 loess events in 468 years, which is one event about every nine years) and C_{org} detected by the RADIUS method[23], ash/tuff and charcoal layers and pollen content, plotted against varve-counted age.

1. Kukla, G. J. & Went, E. *Start of a Glacial* (Springer, Berlin, 1992).
2. Kukla, G., McManus, J. F., Rousseau, D.-D. & Chuine, I. How long and how stable was the last Interglacial? *Quat. Sci. Rev.* **16,** 605–612 (1997).
3. NorthGrip, High resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **43,** 147–151 (2004).
4. Khodri, M. *et al.* Simulating the amplification of orbital forcing by ocean feedbacks in the last glaciation. *Nature* **410,** 570–574 (2001).
5. Berger, A. & Loutre, M. F. An exceptionally long interglacial ahead? *Science* **297,** 1287–1288 (2002).
6. Calov, R., Ganopolski, A., Claussen, M., Petukhov, V. & Greve, R. Transient simulation of the last glacial inception. Part I: Glacial inception as a bifurcation in the climate system. *Climate Dyn.* **24,** doi:10.1007/s00382-005-0007-6 (2005).
7. Lambeck, K. & Chappell, J. Sea level change through the last glacial cycle. *Science* **292,** 679–685 (2001).
8. Sánchez-Goñi, M. F. *et al.* Synchroneity between marine and terrestrial responses to millennial scale climatic variability during the last glacial period in the mediterranean region. *Climate Dyn.* **19,** 95–105 (2002).
9. Berger, A. & Loutre, M. F. Insolation values for the climate of the last 10 million years. *Quat. Sci. Rev.* **10,** 297–317 (1991).
10. Bond, G. & Lotti, R. Iceberg discharges into the North Atlantic on millennial time scales during the last glaciation. *Science* **267,** 1005–1010 (1995).
11. Thouveny, N. *et al.* Climate variations in Europe over the past 140 kyr deduced from rock magnetism. *Nature* **371,** 503–506 (1994).
12. Allen, J. R. M. *et al.* Rapid environmental changes in southern Europe during the last glacial period. *Nature* **400,** 740–743 (1999).
13. Lehman, S. J., Sachs, J. P., Rotwell, A. M., Keigwin, L. D. & Boyle, E. A. Relation of subtropical Atlantic temperature, high-latitude ice rafting, deep water formation, and European climate 130,000–60,000 years ago. *Quat. Sci. Rev.* **21,** 1917–1924 (2002).
14. Müller, U. C. & Kukla, G. J. North Atlantic Current and European environments during the declining stage of the last interglacial. *Geology* **32,** 1009–1012 (2004).
15. Van Kreveld, S. *et al.* Potential links between surging ice sheets, circulation changes, and the Daansgaard–Oeschger cycles in the Irminger Sea, 60–18 kyr. *Paleoceanography* **15,** 425–442 (2000).
16. McManus, J. F. *et al.* High resolution climate records from the North Atlantic during the last interglacial. *Nature* **371,** 326–329 (1994).
17. Chapman, M. R. & Shackleton, N. J. Global ice-volume fluctuations, North Atlantic ice-rafting events, and deep-ocean circulation changes between 130 and 70 ka. *Geology* **27,** 795–798 (1999).
18. Müller, H. Pollenanalytische Untersuchungen und Jahresschichtenzählungen an der eem-zeitlichen Kieselgur von Bispingen/Luhe. *Geol. Jb. A* **21,** 149–169 (1974).
19. Holzkämper, S., Mangini, A., Spötl, C. & Mudelsee, M. Timing and progression of the Last Interglacial derived from a high alpine stalagmite. *Geophys. Res. Lett.* **31,** L07201 (doi:10.1029/2003GL019112) (2004).
20. Kühl, N. & Litt, T. Quantitative time series reconstruction of Eemian temperature at three European sites using pollen data. *Veget. Hist. Archaeobot.* **12,** 205–214 (2003).
21. Büchel, G. *Vulkanologische Karte der West- und Hocheifel, Map 1:50,000* (Institut für Geowissenschaften, Universität Mainz, 1994).
22. Larsen, C. P. S. *et al.* Relations between lake morphometry and the presence of laminated lake sediments: a re-examination of Larsen and MacDonald (1993). *Quat. Sci. Rev.* **17,** 711–717 (1998).
23. Seelos, K. & Sirocko, F. RADIUS—rapid particle analysis of digital images by ultra-high resolution scanning of thin sections. *Sedimentology* **52,** 669–681 (2005).
24. de Beaulieu, J. L. & Reille, M. Long Pleistocene pollen sequences from the Velay Plateau (Massif Central, France). *Veget. Hist. Archaeobot.* **1,** 233–242 (1992).
25. Menke, B. & Tynni, R. Das Eeminterglazial und das Weichselfrühglazial von Rederstall/Dithmarschen und ihre Bedeutung für die mitteleuropäische Jungpleistozän-Gliederung. *Geol. Jb. A* **76,** 3–120 (1984).
26. Woillard, G. Abrupt end of the last interglacial s.s. in north-east France. *Nature* **281,** 558–565 (1979).
27. Björck, S. *et al.* Eemian lake development, hydrology and climate: a multi-stratigraphic study of the Hollerup site in Denmark. *Quat. Sci. Rev.* **19,** 509–536 (2000).
28. Voelker, A. H. L. Global distribution of centennial-scale records for Marine Isotope Stage (MIS) 3: a database. *Quat. Sci. Rev.* **21,** 1185–1212 (2001).
29. Erfurt, G., Krbetschek, M. R., Bortolot, V. J. & Preusser, F. A fully automated multi-spectral radioluminescence reading system for geochronometry and dosimetry. *Nucl. Instrum. Methods Phys. Res. B* **207,** 487–499 (2003).
30. van den Bogaard, P., Hall, C. M., Schmincke, H.-U. & York, D. Precise single-grain 40Ar/39Ar dating of a cold to warm climate transition in Central Europe. *Nature* **342,** 523–525 (1989).

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

Acknowledgements Drilling of the ELSA cores was done by Stölben GmbH (www.stoelbenbohr.de), and we thank J. Schmitz for his continuous effort to increase the quality of the core material. We thank the Stiftung Rheinland-Pfalz für Innovation, which initiated the first ELSA drillings, and the DEKLIM program of the German Ministry for Education and Research, which financed ELSA over the past 5 years.

Author Contributions K. Seelos, grain size analysis; K. Schaber, tuning; B.R., varve counting; F.D. and M.D., pollen analysis; K.J., varve classification; M.K. and D.D., luminescence dating.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.S. (sirocko@uni-mainz.de).

LETTERS

The chemical structure of the Hawaiian mantle plume

Zhong-Yuan Ren[1], Stephanie Ingle[1]†, Eiichi Takahashi[1,2], Naoto Hirano[1] & Takafumi Hirata[1]

The Hawaiian–Emperor volcanic island and seamount chain is usually attributed to a hot mantle plume, located beneath the Pacific lithosphere, that delivers material sourced from deep in the mantle to the surface[1–5]. The shield volcanoes of the Hawaiian islands are distributed in two curvilinear, parallel trends (termed 'Kea' and 'Loa'), whose rocks are characterized by general geochemical differences[2–5]. This has led to the proposition that Hawaiian volcanoes sample compositionally distinct, concentrically zoned, regions of the underlying mantle plume[4,5]. Melt inclusions, or samples of local magma 'frozen' in olivine phenocrysts during crystallization, may record complexities of mantle sources[6], thereby providing better insight into the chemical structure of plumes. Here we report the discovery of both Kea- and Loa-like major and trace element compositions in olivine-hosted melt inclusions in individual, shield-stage Hawaiian volcanoes—even within single rock samples. We infer from these data that one mantle source component may dominate a single lava flow, but that the two mantle source components are consistently represented to some extent in all lavas, regardless of the specific geographic location of the volcano. We therefore suggest that the Hawaiian mantle plume is unlikely to be compositionally concentrically zoned. Instead, the observed chemical variation is probably controlled by the thermal structure of the plume.

Most of our current understanding of the geochemistry of mantle plumes originates from data on whole rocks. Yet, whole rocks record only an 'average' composition produced by complex petrogenesis involving partial melting in the mantle, subsequent aggregation of melts on their way to the surface, and mixing of melts in shallow magma chambers[6]. In contrast, melt inclusions can provide important information on compositions of magmas before extensive magma mixing[6–9]. To test the homogeneity of the mantle plume source sampled in the Kea and Loa trends by Hawaiian shield volcanoes, we examined major and trace element compositions of olivine-hosted melt inclusions from Hawaiian shield lavas, using electron microprobe and laser ablation inductively coupled plasma source mass spectrometry (see Methods). The data are presented in Supplementary Table 1. We selected lava samples from the submarine Hana ridge, Haleakala volcano (Maui), and the submarine exposures of the Makapuu stage, Koolau volcano (Oahu), because these two volcanoes are believed to represent the sampling of the two distinct compositions, having erupted above the Kea trend and Loa trend, respectively[4] (Fig. 1). Detailed sample information, including specific locality and whole rock geochemistry, has been presented elsewhere[10–12].

Primary information about the mantle source composition beneath Hawaii can, in theory, be obtained from data on melt inclusions, provided the host olivines are derived exclusively from Hawaiian plume melts. All host olivines have CaO contents (0.16–0.29 wt%) that are typical of magmatic olivine and significantly higher than those of mantle (that is, xenocrystic) olivine[10,13]. Furthermore, we exclusively use major element ratios (such as Al_2O_3/CaO and TiO_2/Na_2O) and incompatible trace element ratios (such as Zr/Nb and Sr/Nb) in order to circumvent problems associated with olivine fractionation in the magmas, possible modification of inclusions from exchange with olivine during eruption, and re-equilibration between the host olivine and inclusion before analysis[8,14].

Major element ratios (such as Al_2O_3/CaO and TiO_2/Na_2O), together with incompatible trace element ratios (such as Zr/Nb and Sr/Nb), of melt inclusion from submarine Hana ridge lavas overlap the fields for both Kilauea and Mauna Loa lavas (Figs 2a, 3a). Most major and trace element compositions of melt inclusions from the Makapuu stage of Koolau olivines are similar to data for whole rocks from Koolau (Makapuu stage, extreme Loa) lavas[3,12], however, other melt inclusions extend well into the fields for Mauna Loa and Kilauea lavas (Figs 2d, 3d). Most importantly, melt inclusions from individual rock samples at both Hana ridge and Makapuu display not only significant heterogeneity, but also span the compositional fields for both Kilauea and Mauna Loa lavas (for example, K214-9, K214-15A, S500-5B, S500-1; Figs 2b–d, 3b–d). Multiple melt inclusions from individual olivine phenocrysts show much more limited variability than observed in the population from whole rocks.

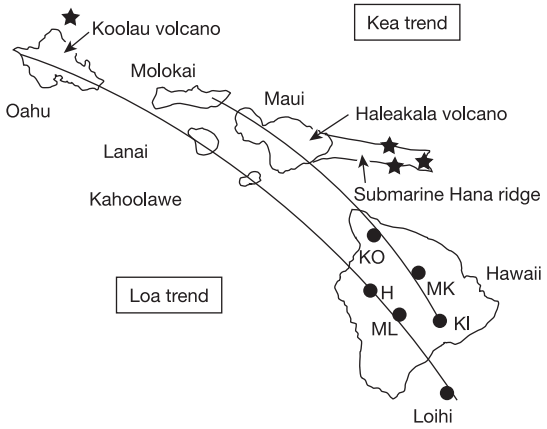


Figure 1 | Map of the Hawaiian islands. Prominent volcanoes, and the Kea and Loa volcano trends, are shown. KI, Kilauea; ML, Mauna Loa; MK, Mauna Kea; H, Hualalai; KO, Kohala. Sample locations from the submarine Hana ridge, Haleakala volcano and the submarine Makapuu stage of Koolau volcano (on the Kea and Loa trends, respectively) are labelled with stars.

[1]Department of Earth and Planetary Sciences, Tokyo Institute of Technology 2-12-1 Ookayama, Meguro-ku, 152-8551, Japan. [2]Institute for Frontier Research on Earth Evolution, Japan Agency for Marine-Earth Science and Technology, Natsushima-cho 2-15, Yokosuka, Kanagawa 237-0061, Japan. †Present address: School of Ocean and Earth Science and Technology, University of Hawaii, 1680 East-West Road, POST 615A, Honolulu, Hawaii 96822, USA.

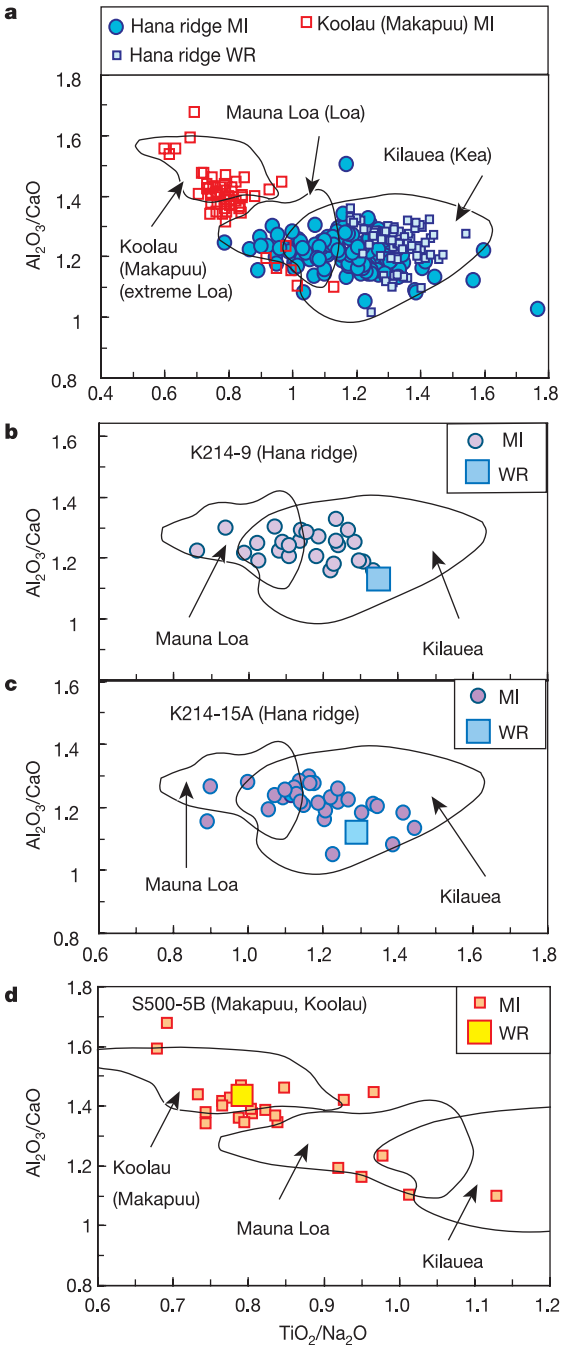


Figure 2 | Variation of major element ratio in bulk rocks compared to melt inclusions. **a**, TiO_2/Na_2O versus Al_2O_3/CaO for the melt inclusions from the submarine Hana ridge (Haleakala shield) and the submarine Makapuu stage of Koolau volcano span both Kea- and Loa-like compositions in single, shield-stage volcanoes (Haleakala and Koolau). Even melt inclusions from some individual samples (for example, K214-9 (**b**), K214-15A (**c**) and S500-5B (**d**)) overlap both Kilauea and Mauna Loa fields. Bulk rock data for Hana ridge and Koolau Makapuu are from refs 10 and 12, respectively. Fields for Kilauea and Mauna Loa whole rocks are from literature sources available at the GEOROC database (http://georoc.mpch-mainz.gwdg.de). The field for Koolau (Makapuu stage) lavas is from refs 3 and 12. Only whole rock data from fresh lavas ($K_2O/P_2O_5 > 1$) are used for the fields. Data for submarine Hana ridge bulk lavas are from ref. 10. All whole rock and melt inclusion major element data have been normalized to 100% without volatiles and with all iron as FeO. MI, melt inclusion; WR, whole rock.

Potential processes that might cause the observed major and trace element ratio variations among the melt inclusions include the assimilation of ocean crust basalt by Hawaiian plume-derived magmas, or mixing of evolved magmas with primitive magmas. Ocean crust basalt typically has low Sm/Yb values as melting occurs extensively and at shallow depths, whereas Hawaiian shield lavas usually have higher Sm/Yb values because of deep melting of a garnet-bearing source. The melt inclusions have relatively constant Sm/Yb values that are high, much higher than those associated with ocean crust basalt (Fig. 4), making assimilation an unlikely scenario[11]. Mixing of primitive and evolved (for example, plagioclase-fractionated) magmas in the magma plumbing system could also cause variations in Al_2O_3, CaO and Sr, because of the affinity of these elements for plagioclase. However, Zr and Nb should not be affected as they are highly incompatible in both plagioclase and clinopyroxene. Melt inclusions reported here have Al_2O_3/CaO and Sr/Nb that are well correlated with Zr/Nb (not shown), indicating that mixing of primitive and evolved magmas is unlikely to have occurred.

Previous work has consistently demonstrated that trace element ratios, such as Zr/Nb and Sr/Nb, are correlated with radiogenic isotope ratios in Hawaiian shield lavas, and it is widely accepted that these trace element ratios reflect mantle source compositions[2,3,15,16]. Also some major element ratios, such as Na/Ti, vary systemically with isotope compositions[17]. Major element ratios (Al_2O_3/CaO, TiO_2/Na_2O) and incompatible trace element ratios (Zr/Nb, Sr/Nb) are correlated in our melt inclusion data, implying that these ratios reflect source geochemical characteristics and the variations, therefore, must reflect source heterogeneity. Therefore, in the same way that isotopic compositions (for example, Pb isotopic ratios) are used, these major and trace element ratios may be sufficient in distinguishing the Kea and Loa signatures of the Hawaiian shield lavas.The Kea and Loa geochemical trends among the Hawaiian shield volcanoes are commonly believed to reflect melting above a compositionally concentrically zoned[4,5] or compositionally left-right asymmetrically zoned[18] mantle plume. Here, we demonstrate that both Kea and Loa components co-exist in a single shield at both Haleakala and Koolau; therefore, a compositionally zoned mantle plume is inconsistent with our data. Overall, general geochemical differences exist in whole rocks between Kea- and Loa-trend volcanoes, but in detail these differences can also exist within a single shield[19], and can vary relative to eruption age[11,12,16,20], suggesting a more complicated structure for the Hawaiian mantle plume[20,21]. The presence of both Kea and Loa compositions in melt inclusions from individual shield volcanoes, and even in single lava samples, indicates that whole rocks may sample one mantle source component dominantly, but that the two (or more) source components are consistently represented to some extent, regardless of the specific geographic locations of the volcano (for example, above the 'Kea' zone).

To best explain both the overall geochemical trends and the newer, more detailed whole rock and melt inclusion data, we propose a Hawaiian mantle plume characterized by more random heterogeneity than would be present in a compositionally zoned mantle plume. The compositional variations in melt inclusions in individual rock samples from Haleakala (Kea trend) and Koolau (Loa trend) suggest that the length-scale of chemical heterogeneity is remarkably smaller than that estimated on the basis of bulk rock geochemistry. The plume may be envisioned as having a matrix characterized by one composition (for example, Kea) with streaks or ribbons of another composition (for example, Loa) distributed throughout the entire plume. Experimental petrology shows that, for a given degree of partial meting, depleted peridotite generates melts having higher Ca/Al, lower SiO_2 and higher FeO contents than does fertile peridotite[22]. Therefore, changes in major element compositions of melt inclusions from Kilauea- to Mauna Loa- to Koolau (Makapuu)-like probably reflect changes in the source available for melting (that is, from relatively depleted to relatively fertile peridotite). A final

implication is that melt compositions are 'fingerprinted' by the source and transported independently before final mixing and homogenization in the magma chambers of individual volcanoes[23].

The geochemistry of Hawaiian lavas, particularly isotopically 'enriched' lavas from Koolau, has been explained by the presence of a recycled eclogite (or pyroxenite) component in the Hawaiian plume[4,24]. Partial melting at 2–3 GPa of eclogite with a composition like that of mid-ocean-ridge basalt produces silica-rich basaltic andesites low in FeO and with low CaO/Al_2O_3 (refs 25, 26). Major element compositions of melts derived from recycled eclogite would be expected to be associated with enriched isotopic signatures[4]. Lavas suspected of derivation from a source containing recycled oceanic gabbroic crust have Th/La and Th/Ba values less than the primitive mantle[27], and excesses of Eu and Sr[28]. These characteristics are also present in melt inclusions examined in this study. Likewise, the presence of Sr-enriched, light rare earth element-depleted melt inclusions from Mauna Loa also suggests a recycled lower oceanic crustal component in the Hawaiian plume[29]. This recycled ancient oceanic crust may remain distinct geochemically[4], forming streaks or ribbons that are deformed and stirred by stretching and folding as the plume rises through the mantle[30].

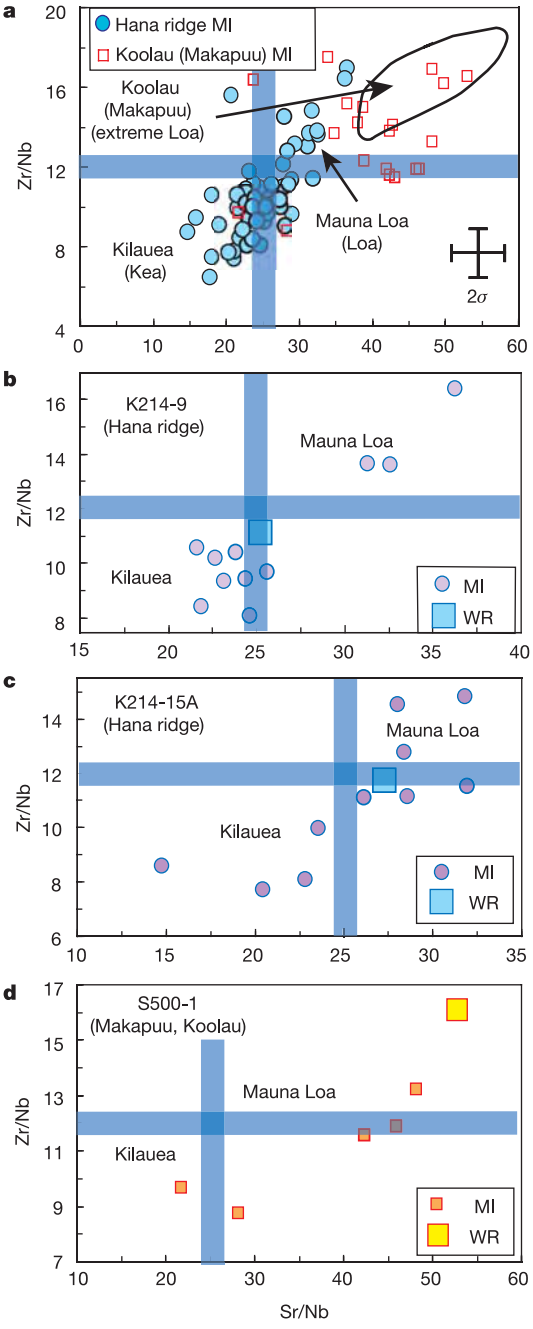


Figure 3 | Variation of trace element ratios in bulk rocks compared to melt inclusions. **a**, Melt inclusions from Hana ridge and Koolau Makapuu have Sr/Nb and Zr/Nb values that overlap those of Kilauea and Mauna Loa lavas. Melt inclusions from some individual samples (for example, K214-9 (**b**), K214-15A (**c**) and S500-1 (**d**)) overlap both the Kilauea field and Mauna Loa field. Delimitations for Kilauea and Mauna Loa are from ref. 2, and for Koolau (Makapuu stage) lavas are from refs 3 and 12. The 2σ error bars are indicated in **a**.

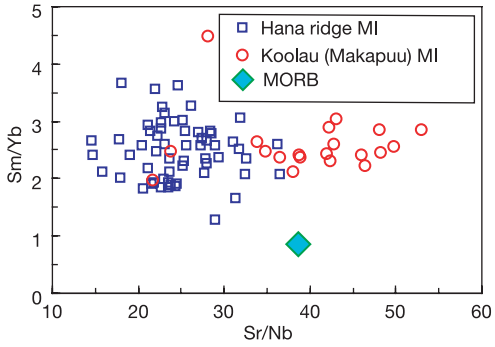


Figure 4 | Variation of trace element ratios in melt inclusions compared to oceanic crust basalt. The wide range of Sr/Nb, and relatively constant and higher Sm/Yb, in the melt inclusions implies that their compositions are unlikely to have been derived from assimilation of ocean crust basalt by plume magmas. Data for ocean crust basalt are from ref. 33. MORB, mid-ocean-ridge basalt.

The dominant component sampled at a given shield volcano may reflect melting processes related to the position of the volcano relative to the hot centre of the plume. As the volcano grows, it migrates away from the hot plume axis with plate motion. At higher temperatures, basaltic eclogite melts will react with peridotite, producing Kilauea-like melts with a composition higher in MgO and lower in SiO_2 than melts formed at lower temperatures[31]. In other words, in the plume core, higher temperatures are able to generate melts from the more refractory component (that is, depleted peridotite), but at plume margins, lower temperatures permit melting of only the more fertile material. Bulk lavas from some younger stages of the Hawaiian shields have enriched isotopic characteristics (for example, Mauna Loa, the Haleakala shield (Honomanu), and the Koolau shield Makapuu stage), suggesting that the proportion of fertile pyroxenite is higher than the proportion of depleted peridotite sampled by the melts. This is most easily explained if melting zones for lavas from the younger stages of Hawaiian shields are located farther from the hot mantle plume axis. Far from the plume's centre, temperatures would be lower, and the more fertile, isotopically enriched material (for example, eclogite) may be sampled to a greater extent.

METHODS

Homogenization of melt inclusions. Olivine-hosted melt inclusions were subjected to experiments designed to homogenize them by melting and quenching, using a 1-atm gas mixing furnace following accepted procedures and conditions[7]. The oxygen fugacity was kept at the quartz–fayalite–magnetite (QFM) buffer. Olivines from single samples were loaded into Pt capsules that were gradually lowered, over 10 min, from the top of the furnace (100 °C) to the hottest place in the furnace (1,250 °C) and kept at that temperature for 10 min before quenching.

EPMA and LA-ICP-MS analyses. We measured major element compositions for 147 melt inclusions from 13 rock samples of submarine Hana ridge, Haleakala volcano and 63 melt inclusions from 3 rock samples of submarine Koolau volcanoes. Inclusions and host olivine were analysed by electron probe micro-analysis (EPMA) with a JEOL-8800 instrument at the Tokyo Institute of

Technology, following procedures described elsewhere[10]. Each melt inclusion was measured using 3–20 points, the compositions of which were then averaged. Analytical uncertainty is 1–2% for major elements, and 5–10% for minor elements. We measured trace element compositions for 58 melt inclusions from 13 Hana ridge lavas and 21 melt inclusions from 3 Koolau lavas. Trace element abundances were measured using laser ablation inductively coupled plasma source mass spectrometry (LA-ICP-MS) coupled with a ThermoElemental PQ2 Ω instrument at the Tokyo Institute of Technology, following procedures described elsewhere[32]. The resulting precision, determined by analysing a standard as an unknown, is better than about 10% for most elements.

Received 8 December 2004; accepted 20 May 2005.

1. Morgan, W. J. Convection plumes in the lower mantle. *Nature* **230,** 42–43 (1971).
2. Frey, F. A. & Rhodes, J. M. Intershield geochemical differences among Hawaiian volcanoes: implications for source compositions, melting process and magma ascent paths. *Phil. Trans. R. Soc. Lond. A* **342,** 121–136 (1993).
3. Frey, F. A., Garcia, M. O. & Roden, M. F. Geochemical characteristics of Koolau Volcano: implications of intershield differences among Hawaiian volcanoes. *Geochim. Cosmochim. Acta* **58,** 1441–1462 (1994).
4. Hauri, E. H. Major-element variability in the Hawaiian mantle plume. *Nature* **382,** 415–419 (1996).
5. Lassiter, J. C., DePaolo, D. J. & Tatsumoto, M. Isotopic evolution of Mauna Kea Volcano: results from the initial phase of the Hawaii Scientific Drilling Project. *J. Geophys. Res.* **101,** 11769–11780 (1996).
6. Sobolev, A. V. Melt inclusions in minerals as a source of principle petrological information. *Petrology* **4,** 209–220 (1996).
7. Hauri, E. SIMS analysis of volatiles in silicate glasses, 2: isotopes and abundances in Hawaiian melt inclusions. *Chem. Geol.* **183,** 115–141 (2002).
8. Norman, M. D., Garcia, M. O., Kamenetsky, V. S. & Nielson, R. L. Olivine-hosted melt inclusions in Hawaiian picrites: equilibration, melting, and plume source characteristics. *Chem. Geol.* **183,** 143–168 (2002).
9. Maclennan, J., McKenzie, D., Hilton, F., Gronvöld, K. & Shimizu, N. Geochemical variability in a single flow from northern Iceland. *J. Geophys. Res.* **108,** 2007, doi:10.1029/2000JB000142 (2003).
10. Ren, Z.-Y., Takahashi, E., Orihashi, Y. & Johnson, K. M. T. Petrogenesis of tholeiitic lavas from the submarine Hana Ridge, Haleakala volcano, Hawaii. *J. Petrol.* **45,** 2067–2099 (2004).
11. Ren, Z.-Y., Shibata, T., Yoshikawa, M., Johnson, K. T. M. & Takahashi, E. Isotope compositions of the submarine Hana ridge lavas, Haleakala volcano, Hawaii: Implications for source compositions, melting process and the structure of Hawaiian plume. *J. Petrol.* (in the press).
12. Tanaka, R., Nakamura, E. & Takahashi, E. in *Hawaiian Volcanoes: Deep Underwater Perspectives* (eds Takahashi, E., Lipman, P. W., Garcia, M. O., Naka, J. & Aramaki, S.) 311–332 (Vol. 128, Geophysical Monograph Series, AGU, Washington DC, 2002).
13. Garcia, M. O., Hulsebosch, T. P. & Rhodes, J. M. in *Mauna Loa Revealed: Structure, Composition, History, and Hazards* (eds Rhodes, J. M. & Lockwood, J. P.) 219–239 (Vol. 92, Geophysical Monograph Series, AGU, Washington DC, 1995).
14. Gaetani, G. A. & Watson, E. B. Open system behaviour of olivine-hosted melt inclusions. *Earth Planet. Sci. Lett.* **183,** 27–41 (2000).
15. Huang, S. & Frey, F. A. Trace element abundances of Mauna Kea basalt from phase 2 of the Hawaii Scientific Drilling Project: petrogenetic implications of correlations with major element content and isotopic ratios. *Geochem. Geophys. Geosyst.* **6,** doi:10.1029/2002GC000322 (2004).
16. Rhodes, J. M. & Vollinger, M. J. Composition of basaltic lavas sampled by phase-2 of the Hawaii Scientific Drilling Project: Geochemical stratigraphy and magma types. *Geochem. Geophys. Geosyst.* **5,** doi:10.1029/2002GC000434 (2004).
17. Putirka, K. Melting depth and mantle heterogeneity beneath Hawaii and the East Pacific Rise: constraints from Na/Ti and rare earth elements ratios. *J. Geophys. Res.* **104,** 2817–2829 (1999).
18. Abouchami, W. *et al.* Lead isotopes reveal bilateral asymmetry and vertical continuity in the Hawaiian mantle plume. *Nature* **434,** 851–856 (2005).
19. West, H. B. & Leeman, W. P. Isotopic evolution of lavas from Haleakala Crater, Hawaii. *Earth Planet. Sci. Lett.* **84,** 211–225 (1987).
20. Eisele, J., Abouchami, W., Galer, S. J. G. & Hofmann, A. W. The 320 kyr Pb isotope evolution of Mauna Kea lavas recorded in the HSDP-2 drill core. *Geochem. Geophys. Geosyst.* **4,** doi:10.1029/2002GC000339 (2003).
21. Blichert-Toft, J., Weis, D., Maerschalk, C., Agranier, A. & Albarède, F. Hawaiian hot spot dynamics as inferred from the Hf and Pb isotope evolution of Mauna Kea volcano. *Geochem. Geophys. Geosyst.* **4,** doi:10.1029/2002GC000340 (2003).
22. Hirose, K. & Kushiro, I. Partial melting of dry peridotites at high pressures: determination of compositions of melts segregated from peridotite using aggregates of diamond. *Earth Planet. Sci. Lett.* **114,** 477–489 (1993).
23. Slater, L., McKenzie, D., Gronvold, K. & Shimizu, N. Melt generation and movement beneath Theistareykir, NE Iceland. *J. Petrol.* **42,** 321–354 (2001).
24. Lassiter, J. C. & Hauri, E. H. Osmium-isotopic variations in Hawaiian lavas and evidence for recycled oceanic lithosphere in the Hawaiian plume. *Earth Planet. Sci. Lett.* **164,** 483–496 (1998).
25. Takahashi, E., Nakajima, K. & Wright, T. L. Origin of the Columbia River basalts: melting model of a heterogeneous plume head. *Earth Planet. Sci. Lett.* **162,** 63–80 (1998).
26. Pertermann, M. & Hirschmann, M. M. Partial melting experiments on a MORB-like pyroxenite between 2 and 3 GPa: constraints on the presence of pyroxenite in basalt source regions from solidus location and melting rate. *J. Geophys. Res.* **108,** doi:10.1029/2000JB000118 (2003).
27. Hofmann, A. W. & Jochum, K. P. Source characteristics derived from very incompatible trace elements in Mauna Loa and Mauna Kea basalt, Hawaiian Scientific Drilling Project. *J. Geophys. Res.* **101,** 11831–11839 (1996).
28. Gasperini, D. *et al.* Evidence from Sardinian basalt geochemistry for recycling of plume heads into the Earth's mantle. *Nature* **408,** 701–704 (2000).
29. Sobolev, A. V., Hofmann, A. W. & Nikogosian, I. K. Recycled oceanic crust observed in 'ghost plagioclase' within the source of Mauna Loa lavas. *Nature* **404,** 986–989 (2000).
30. Farnetani, C. G., Legras, B. & Tackley, P. J. Mixing and deformation in mantle plumes. *Earth Planet. Sci. Lett.* **196,** 1–15 (2002).
31. Takahashi, E. & Nakajima, K. in *Hawaiian Volcanoes: Deep Underwater Perspectives* (eds Takahashi, E., Lipman, P. W., Garcia, M. O., Naka, J. & Aramaki, S.) 403–418 (Vol. 128, Geophysical Monograph Series, AGU, Washington DC, 2002).
32. Norman, M. D., Griffin, W. L., Pearson, N. J., Garcia, M. O. & O'Reilly, S. Y. Quantitative analysis of trace element abundances in glasses and minerals: a comparison of laser ablation ICPMS, solution ICPMS, proton microprobe, and electron microprobe data. *J. Anal. Atom. Spectrom.* **13,** 477–482 (1998).
33. Sun, S.-S. & McDonough, W. F. in *Magmatism in the Ocean Basins* (eds Saunders, A. D. & Norry, M. J.) 313–345 (Special Publication 42, Geological Society, London, 1989).

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

Acknowledgements We thank F. A. Frey for comments and suggestions, and M. F. Coffin, L. Danyushevsky, A. W. Hofmann, J. Lassiter, D. A. Clague, S. Escrig, D. Weis and K. Putirka for discussions, comments and technical advice. C. Herzberg, J. M. Rhodes and M. Kurz provided constructive criticism that led to improvements in the manuscript. Z.-Y.R., S.I. and N.H. are grateful to the JSPS programme for funding.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to Z.-Y.R. (zyren@jamstec.go.jp).

LETTERS

A wide depth distribution of seismic tremors along the northern Cascadia margin

Honn Kao[1], Shao-Ju Shan[1], Herb Dragert[1], Garry Rogers[1], John F. Cassidy[1] & Kumar Ramachandran[1]

The Cascadia subduction zone is thought to be capable of generating major earthquakes with moment magnitude as large as $M_w = 9$ at an interval of several hundred years[1–3]. The seismogenic portion of the plate interface is mostly offshore and is currently locked, as inferred from geodetic data[4–6]. However, episodic surface displacements—in the direction opposite to the long-term deformation motions caused by relative plate convergence across a locked interface—are observed about every 14 months with an unusual tremor-like seismic signature[7–9]. Here we show that these tremors are distributed over a depth range exceeding 40 km within a limited horizontal band. Many occurred within or close to the strong seismic reflectors above the plate interface where local earthquakes are absent, suggesting that the seismogenic process for tremors is fluid-related. The observed depth range implies that tremors could be associated with the variation of stress field induced by a transient slip along the deeper portion of the Cascadia interface or, alternatively, that episodic slip is more diffuse than originally suggested.

In anticipation of an episodic tremor-and-slip (ETS) event[7–9] in February–March 2003, the Geological Survey of Canada augmented the existing seismographic network by deploying temporary broadband instruments to create a dense seismic array in southern Vancouver Island (Fig. 1). Clear tremor signals were first observed on 25 February 2003 (defined as 'day 0') and lasted for more than three weeks (Fig. 1). Concurrently, the global positioning system (GPS) site near Victoria, at the southern tip of Vancouver Island, showed a steady westward movement during the first ten days with a total displacement of ~3.7 mm. The GPS site located ~175 km towards the northwest near Ucluelet, however, showed little or no corresponding movement, implying that the slip was mostly beneath southern Vancouver Island.

Because of the emergent arrivals and the clustering nature of tremor signals, determination of tremor origin times and locations using standard seismological procedures is extremely difficult. The newly developed source-scanning algorithm (SSA) method is used instead to overcome such problems[10]. The SSA method identifies the

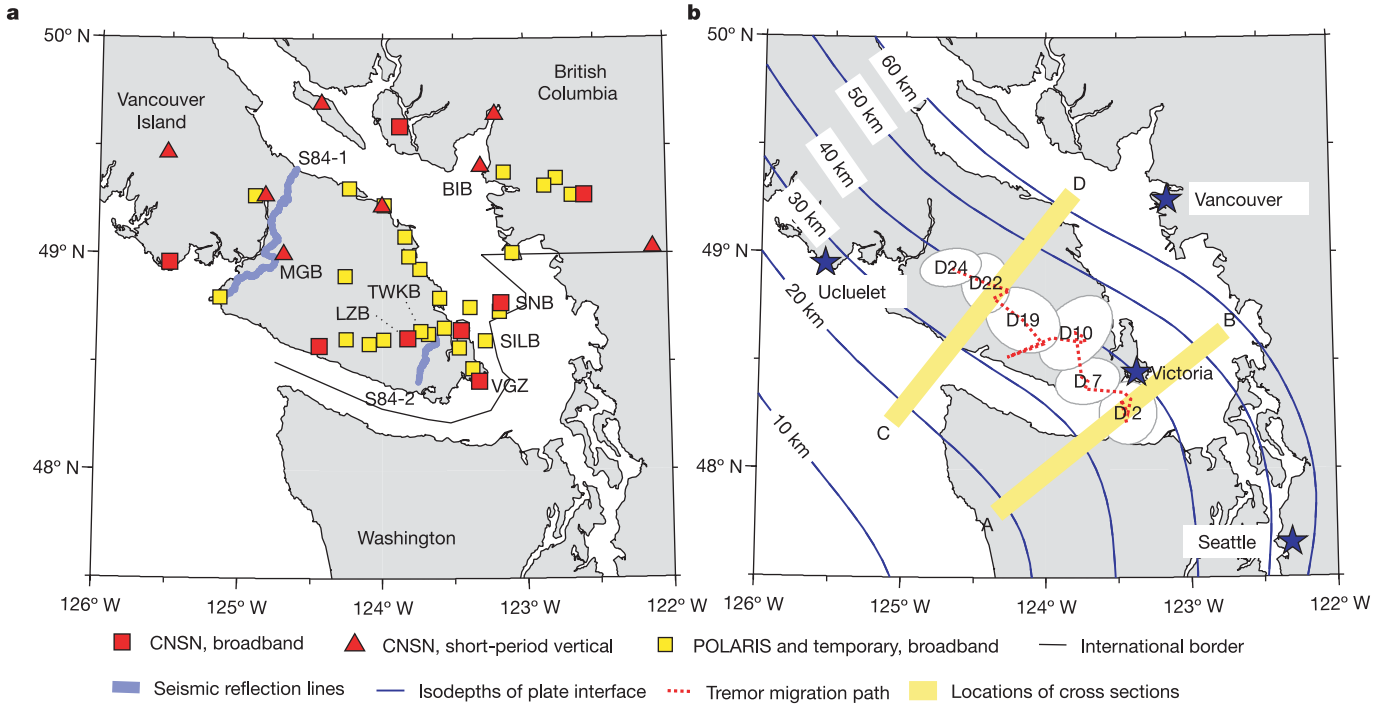


Figure 1 | Maps showing the distribution of seismograph stations (a) and the gradual migration path of seismic tremors (b). CNSN, Canadian National Seismograph Network; POLARIS, Portable Observatories for Lithospheric Analysis and Research Investigating Seismicity. This tremor sequence began on 25 February 2003 (day 0, D0). Thus D2 means 27 February 2003 and so on. Each white oval represents one standard deviation (from the geometric centre) of the spatial distribution of tremors occurring in the annotated day. Acronyms indicate seismograph stations. Stars mark major cities in the region and the town mentioned in the text.

[1]Geological Survey of Canada, Pacific Geoscience Centre, 9860 West Saanich Road, Sidney, British Columbia V8L 1B9, Canada.

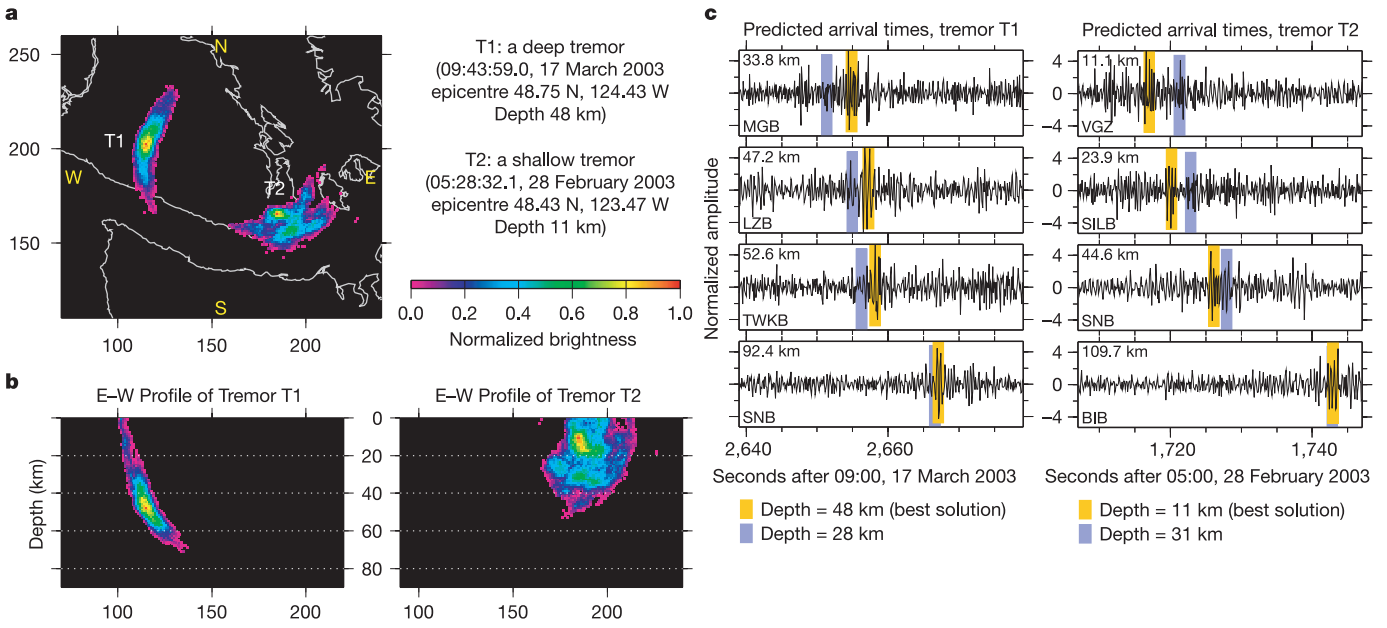


Figure 2 | Map and profiles showing two seismic tremors occurring at different locations and depths. a, Location map. **b**, The normalized brightness function, as determined by SSA, is displayed in colour. Higher brightness means a better consistency between the source location and the observed arrivals of tremor signals. The large depth difference cannot be an artefact because the predicted arrivals from erroneous depths are inconsistent with the observed seismograms (**c**). Waveforms are plotted in normalized amplitudes, which are used in brightness calculation. The station name and its corresponding distance from the epicentre are marked at the beginning of each trace. Notice that the predicted arrival time difference varies with epicentral distance, thus allowing no trade-off between the depth and the origin time.

existence of seismic sources in space and time by calculating the so-called 'brightness' function for all grid points inside the model space. The larger the value of the brightness function, the better the consistency between the identified sources (that is, epicentre, depth and origin time) and observed waveforms. A more detailed description of our SSA analysis and a list of determined tremor locations and origin times are presented as Supplementary Information.

The most striking result of our analysis is that ETS tremors are distributed across a wide depth range of over 40 km, with a peak at the depth of 25–35 km. Two representative examples, one centred at a shallow depth of 11 km and another at 48 km, are shown in Fig. 2 to demonstrate that this wide span cannot be attributed to analysis uncertainties (which are estimated to be ± 3 and ± 5 km for epicentral location and depth, respectively). Forward calculations indicate that there is virtually no trade-off between the determined depth and origin time because of the optimal station coverage (Figs 1 and 2). We also conducted a series of experiments to confirm that the depth difference did not arise from a systematic bias in the velocity model and/or different choices of controlling parameters used in the SSA analysis (see the Supplementary Information). For a few relatively large and isolated tremors, the best solutions are further verified by conventional earthquake-location methods and three-dimensional relocation techniques. All results consistently suggest the existence of ETS tremors over a wide depth range.

The epicentres of most tremors are confined to a limited band bounded approximately by the surface projections of 30- and 45-km depth contours of the subducting plate interface, with a clear northwestward migration pattern (Fig. 1). We constructed two cross-sections to show the depth distribution of ETS tremors in Fig. 3. The southern section (A–B) corresponds to the initial stage of this ETS sequence (days 0–6), whereas the northern section (C–D) represents the ending phase. Background seismicity since 1990 is also plotted for comparison. It is clear that ETS tremors do not occur strictly along the plate interface (or any other planar structures) as has been proposed for the slip[7]. Instead, they are distributed in both the overriding crust and within the subducted slab (as demonstrated by examples in Fig. 2), similar to the pattern of local earthquakes (Fig. 3).

However, ETS tremors tend to occur in places where local earthquakes are sparse. This is confirmed by simple 'nearness' tests on both the tremor and earthquake data sets. On average, the distance from one tremor to the nearest earthquake hypocentre is ~10 km, about double of that to the nearest tremor (~5 km). In contrast, the average earthquake-to-earthquake 'nearness' is only 3 km, suggesting that the difference between the spatial distributions of tremors and earthquakes is not an artefact.

This difference becomes more apparent when the cross-sections are overlaid with available tomographic images[11] and seismic reflection profiles (Fig. 3). About 50%–55% of ETS tremors are located within or close to the strong reflectors above the plate interface. In contrast, >90% of local earthquakes tend to be located away from the reflectors[12,13]. Previous studies have shown that these strong reflectors coincide with an electrically conductive layer[14–16] and a low shear-wave velocity zone[17]. It has been argued that these regions are where fluids released from the dehydration reactions in the subducted oceanic slab are trapped[13–16]. Theoretical calculations of the amplitude contrast also suggest that the reflectors may be associated with extensive shearing[13–15]. Consequently, the observed tremor distribution suggests that the underlying physical processes, which might be closely related to both shear deformation and fluids, are distinct from those involved in earthquake rupture.

Another noticeable difference can be found in the frequency contents of ETS tremors and local earthquakes (Fig. 4). To ensure that the calculated frequency spectra are representative of the source characteristics with minimum path effects, we only include seismograms recorded within 50 km from the corresponding earthquake or tremor epicentres. The two tremor data sets contain seismograms corresponding to the ten largest tremors observed in this ETS sequence and the ten best-located tremors in this study, respectively, whereas the three earthquake data sets are for different magnitude ranges ($M_L = 2.5$–3.5, 1.0–1.5 and 0.0–0.2). It is clear that the stacked Fourier spectra of ETS tremors show relatively larger amplitudes in the frequency range of 1–5 Hz with a rapid decrease at higher

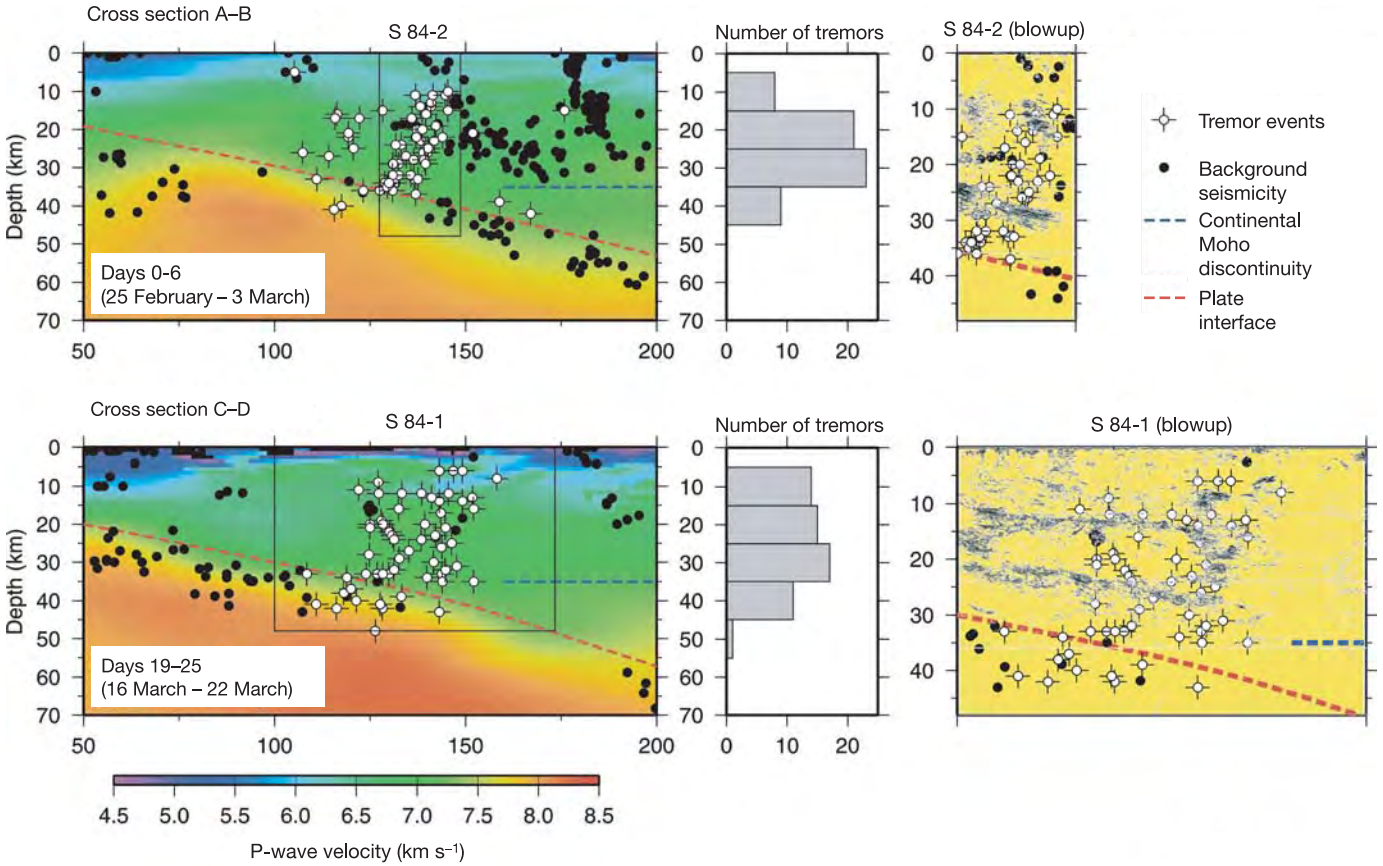


Figure 3 | Two cross-sections showing the spatial distribution of seismic tremors with respect to background seismicity, regional tomography, and strong reflectors identified from seismic reflection surveys. Histograms of tremor numbers at various depths are shown to the right of each cross-section. Many tremors occurred within (or close to) the reflectors where local earthquakes are absent (blow-ups of the seismic profiles S84-1 and S84-2, whose locations are shown in Fig. 1). The symbol size of tremors roughly corresponds to the hypocentral uncertainty. To avoid any distortion from projecting over a large distance, only tremors and earthquakes within 15 km from the location of each cross-section are used.

(5–30 Hz) frequencies (Fig. 4). The sharp contrast in the frequency characteristics between ETS tremors and earthquakes again hints at different seismogenic behaviour.

Our results provide important insights into the physical and mechanical processes along the Cascadia margin that subsequently control the seismogenic and rheological behaviour of the subduction system. If fluid released from dehydration of subducted materials plays an important role in forming the strong seismic reflectors in Cascadia's forearc, as suggested previously[14,15,18], it may also play a critical role in facilitating the occurrence of ETS tremors. One possible scenario is that ETS tremors manifest the hydro-seismic processes in both the overriding continental crust and the subducting oceanic crust in response to the temporal stress/strain field induced by the transient slip along the deeper (aseismic) portion of the plate interface. In this case, the existence of fluids provides a favourable condition for tremors to occur, whereas the stress/strain variation due to the transient slip actually triggers their occurrences. Alternatively, if the tremors are interpreted as the de facto seismic part of the episodic slip, then the slip may have occurred in a zone much more diffuse than previous studies have suggested[7]. Although the inversion of surface GPS measurements cannot rule out the latter scenario, it is not only more complex but could also introduce a mass balance problem in the forearc region if the diffused slip above the plate interface continues over geological timescales and permanently displaces material through shearing[19]. However, many other processes can be involved in the mass balance issue in the forearc region, including frontal accretion that transfers materials into the wedge[20], subduction erosion or underplating at the base of the forearc[21], and regional N–S crustal shortening[22]. From our

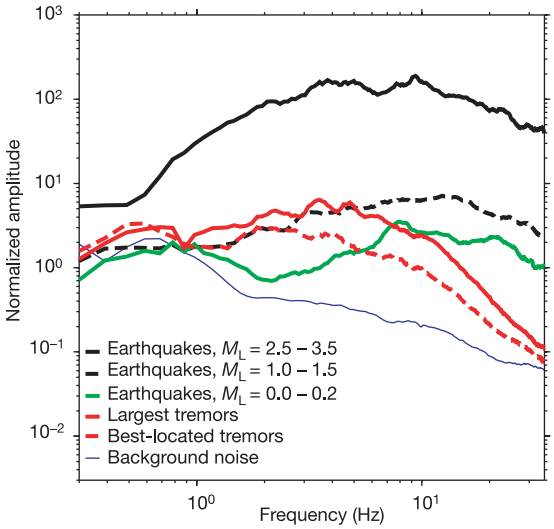


Figure 4 | A comparison of frequency spectra between local earthquakes and ETS tremors in northern Cascadia. At frequencies between 1 and 5 Hz, the amplitudes of seismic tremors are comparable to that of $M_L \approx 1.5$ earthquakes. However, the high frequency (>5 Hz) content of tremors is much smaller, even less than that of $M_L = 0.2$ earthquakes. Such a dramatic difference in source spectra suggests distinct physical processes for tremors versus local earthquakes.

observations alone, we cannot conclusively reject either interpretation at this stage.

Whichever interpretation is chosen, a concurrence of tremors and slip in an ETS event is always implied. However, the original ETS report acknowledged the existence of minor tremor activities without corresponding GPS signatures[9]. Similar events have also been documented for other regions with irregular minor tremors, including southern Cascadia[23], Costa Rica[24] and southwestern Japan[25,26]. One possible explanation is that the corresponding slip may be too small to be detected by surface GPS measurements. The installation of borehole strain meters in the region or increasing the resolution of GPS analysis may be able to provide more diagnostic constraints.

Previous studies have shown that aseismic slips on the interplate thrust zone at depth may slightly increase the stress along the locked portion of the interface located up-dip[7,27–29], and thereby increase the probability of triggering a large subduction thrust earthquake[30]. The wide depth distribution of tremors reported here suggests that detailed modelling incorporating various slip scenarios is needed to properly address the possible effect of ETS events on the occurrence of megathrust earthquakes.

Finally, we note that the extraordinarily high signal-to-noise ratio of the Japanese borehole seismic networks enables the identification of P and S phases of some relatively large seismic tremors. Using more conventional earthquake-location methods, the depths of Japanese tremors are determined to be 25–35 km (refs 25, 26), a range remarkably consistent with the peak of tremor depths reported here (Fig. 3). If ETS events in various regions are associated with similar physical processes, then a wide distribution of tremor depths with a peak in the lower crust can be expected. The resolution of this extended depth range will depend on the accurate locations of both small and large tremors.

Received 24 November 2004; accepted 7 June 2005.

1. Atwater, B. F. Evidence for great Holocene earthquakes along the outer coast of Washington state. *Science* **236,** 942–944 (1987).
2. Atwater, B. F. & Hemphill-Haley, E. Recurrence intervals for great earthquakes of the past 3500 years at northeastern Willapa Bay, Washington. *US Geol. Surv. Prof. Pap.* **1576,** 1–108 (1997).
3. Satake, K., Wang, K. & Atwater, B. F. Fault slip and seismic moment of the 1700 Cascadia earthquake inferred from Japanese tsunami descriptions. *J. Geophys. Res.* **108,** doi:10.1029/2003JB002521 (2003).
4. Savage, J. C. & Lisowski, M. Strain measurements and the potential for a great subduction earthquake off the coast of Washington. *Science* **252,** 101–103 (1991).
5. Dragert, H. & Hyndman, R. D. Continuous GPS monitoring of elastic strain in the northern Cascadia subduction zone. *Geophys. Res. Lett.* **22,** 755–758 (1995).
6. Wang, K., Wells, R., Mazzotti, S., Hyndman, R. D. & Sagiya, T. A revised dislocation model of interseismic deformation of the Cascadia subduction zone. *J. Geophys. Res.* **108,** 2026, doi:10.1029/2001JB001227 (2003).
7. Dragert, H., Wang, K. & James, T. S. A silent slip event on the deeper Cascadia subduction interface. *Science* **292,** 1525–1528 (2001).
8. Miller, M. M., Melbourne, T., Johnson, D. J. & Sumner, W. Q. Periodic slow earthquakes from the Cascadia subduction zone. *Science* **295,** 2423 (2002).
9. Rogers, G. & Dragert, H. Episodic tremor and slip on the Cascadia subduction zone: The chatter of silent slip. *Science* **300,** 1942–1943 (2003).
10. Kao, H. & Shan, S.-J. The source-scanning algorithm: mapping the distribution of seismic sources in time and space. *Geophys. J. Int.* **157,** 589–594 (2004).
11. Ramachandran, K., Dosso, S. E., Spence, G. D., Hyndman, R. D. & Brocher, T. M. Forearc structure beneath southwestern British Columbia: A three-dimensional tomographic velocity model. *J. Geophys. Res.* **110,** B02303, doi: 10.1029/2004JB003258 (2005).
12. Calvert, A. J. Seismic reflection imaging of two megathrust shear zones in the northern Cascadia subduction zone. *Nature* **428,** 163–167 (2004).
13. Nedimovic, M. R., Hyndman, R. D., Ramachandran, K. & Spence, G. D. Reflection signature of seismic and aseismic slip on the northern Cascadia subduction interface. *Nature* **424,** 416–420 (2003).
14. Calvert, A. J. & Clowes, R. M. Deep, high-amplitude reflections from a major shear zone above the subducting Juan de Fuca plate. *Geology* **18,** 1091–1094 (1990).
15. Hyndman, R. D. Dipping seismic reflectors, electrically conductive zones, and trapped water in the crust over a subducting plate. *J. Geophys. Res.* **93,** 13391–13405 (1988).
16. Clowes, R. M. *et al.* LITHOPROBE-southern Vancouver island: Cenozoic subduction complex imaged by deep seismic reflections. *Can. J. Earth Sci.* **24,** 31–51 (1987).
17. Cassidy, J. F. & Ellis, R. M. S wave velocity structure of the northern Cascadia subduction zone. *J. Geophys. Res.* **98,** 4407–4421 (1993).
18. Peacock, S. M. Fluid processes in subduction zones. *Science* **248,** 329–337 (1990).
19. Wang, K. & Hyndman, R. D. Challenges in defining seismogenic zone using geodetic and structural observations. *Eos* (AGU Fall Meeting Suppl.) **85**, abstr. S43D–05 (2004).
20. Adam, J., Klaeschen, D., Kukowski, N. & Flueh, E. Upward delamination of Cascadia Basin sediment infill with landward frontal accretion thrusting caused by rapid glacial age material flux. *Tectonics* **23,** TC3009, doi:10.1029/2002TC001475 (2004).
21. von Huene, R., Ranero, C. R. & Vannucchi, P. Generic model of subduction erosion. *Geology* **32,** 913–916 (2004).
22. Hyndman, R. D., Mazzotti, S., Weichert, D. & Rogers, G. C. Frequency of large crustal earthquakes in Puget Sound-Southern Georgia Strait predicted from geodetic and geological deformation rates. *J. Geophys. Res.* **108,** 2033, doi:10.1029/2001JB001710 (2003).
23. Szeliga, W., Melbourne, T. I., Miller, M. M. & Santillan, V. M. Southern Cascadia episodic slow earthquakes. *Geophy. Res. Lett.* **31,** L16602, doi:10.1029/2004GL020824 (2004).
24. Broan, K. M., DeShon, H., Tryon, M., Dorman, L. & Schwartz, S. Transient fluid pulsing and noise in the Costa Rican subduction zone: Nearly silent slip events? *Eos* (AGU Fall Meet. Suppl.) **84**, abstr. T41E–06 (2003).
25. Obara, K. Nonvolcanic deep tremor associated with subduction in southwest Japan. *Science* **296,** 1679–1681 (2002).
26. Katsumata, A. & Kamara, N. Low-frequency continuous tremor around the Moho discontinuity away from volcanoes in the southwest Japan. *Geophys. Res. Lett.* **30,** 1020, doi:10.1029/2002GL015981 (2003).
27. Linde, A. T. & Silver, P. G. Elevation changes and the great 1960 Chilean earthquake: Support for aseismic slip. *Geophys. Res. Lett.* **16,** 1305–1308 (1989).
28. Thatcher, W. Seismic triggering and earthquake prediction. *Nature* **299,** 12–13 (1982).
29. Shimazaki, K. & Nakata, T. Time predictable recurrence model for large earthquakes. *Geophys. Res. Lett.* **7,** 279–282 (1980).
30. Mazzotti, S. & Adams, J. Variability of near-term probability for the next great earthquake on the Cascadia subduction zone. *Bull. Seismol. Soc. Am.* **94,** 1954–1959 (2004).

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

Acknowledgements We thank I. Al-Khoubbi, T. Christie, T. Claydon and R. Hall for their efforts in deploying and maintaining the seismic network; A. Bird, T. Mulder and W. Bentkowski for their assistance in seismic data processing; R. Baldwin for his software support; A. Calvert, R. Hyndman and K. Wang for stimulating discussion; and the POLARIS consortium for providing data from the BC POLARIS array. Most SSA computation is done on the Mercury cluster of the University of Victoria. This work is supported in part by a USGS NEHRP research grant.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.K. (hkao@nrcan.gc.ca).

LETTERS

The *hangover* gene defines a stress pathway required for ethanol tolerance development

Henrike Scholz[1,3], Mirjam Franz[3] & Ulrike Heberlein[1,2]

Repeated alcohol consumption leads to the development of tolerance, simply defined as an acquired resistance to the physiological and behavioural effects of the drug. This tolerance allows increased alcohol consumption, which over time leads to physical dependence and possibly addiction[1–3]. Previous studies have shown that *Drosophila* develop ethanol tolerance, with kinetics of acquisition and dissipation that mimic those seen in mammals. This tolerance requires the catecholamine octopamine, the functional analogue of mammalian noradrenaline[4]. Here we describe a new gene, *hangover*, which is required for normal development of ethanol tolerance. *hangover* flies are also defective in responses to environmental stressors, such as heat and the free-radical-generating agent paraquat. Using genetic epistasis tests, we show that ethanol tolerance in *Drosophila* relies on two distinct molecular pathways: a cellular stress pathway defined by *hangover*, and a parallel pathway requiring octopamine. *hangover* encodes a large nuclear zinc-finger protein, suggesting a role in nucleic acid binding. There is growing recognition that stress, at both the cellular and systemic levels, contributes to drug- and addiction-related behaviours in mammals. Our studies suggest that this role may be conserved across evolution.

When flies are exposed to ethanol vapour, they become hyperactive, uncoordinated and eventually sedated. These effects of ethanol cause loss of postural control, which can readily be quantified in an inebriometer[5]. Naive wild-type flies emerge from the inebriometer with a mean elution time (MET) of ~20 min at standard ethanol vapour concentrations[6,7]. A single exposure to ethanol in the inebriometer leads to the development of tolerance: flies re-introduced into the apparatus 4 h after their initial exposure elute with a MET of ~28 min (ref. 4; Fig. 1a). This acquired resistance, or tolerance, correlates with an increase in the absorbed ethanol levels required to induce loss of postural control, and is measured as the percentage increase (~35–40% at standard ethanol concentrations) in MET between the first and second ethanol exposures[4].

To identify molecules and pathways involved in tolerance development, we carried out a screen for P-element-induced mutants with aberrant tolerance (see Methods). As the degree of tolerance is proportional to the length of initial ethanol exposure[4], we limited our screen to strains that reacted normally to their first ethanol exposure. One mutant strain, AE10, showed a normal initial MET but a reduced ability to develop tolerance (14 ± 3% increase in MET compared with 35 ± 2% for controls), and was named *hangover* (*hang*) (Fig. 1a). The tolerance defect in $hang^{AE10}$ flies is not simply due to a change in the rate of tolerance acquisition, as the mutant flies were also impaired when tested in a paradigm that induces maximal tolerance by means of several consecutive ethanol exposures[4] (Fig. 1b). *hang* flies show normal ethanol absorption and metabolism (Supplementary Fig. 1a), so their phenotype is therefore not due to altered drug pharmacokinetics.

The P-element in $hang^{AE10}$ flies is inserted in the predicted open reading frame of a new gene, CG32575 (http://www.fruitfly.org/; Fig. 2a), encoding a protein with 15 nucleic-acid-binding zinc-finger domains, two of which are of the U1 subclass found in RNA-binding proteins[8] (Fig. 2b). The P-element insertion causes the mutant phenotype, as precise excision of the transposon causes reversion to the wild-type phenotype (Supplementary Fig. 1b). *hang* is expressed ubiquitously in the nervous system (Supplementary Fig. 2a, b) and HANG protein is localized to the nuclei of neurons (Fig. 2c and Supplementary Fig. 2c). $hang^{AE10}$ appears to be a null allele, as the ~7 kb messenger RNA encoded by *hang* (Fig. 2d) and the HANG protein (Fig. 2c) are undetectable in the mutant.

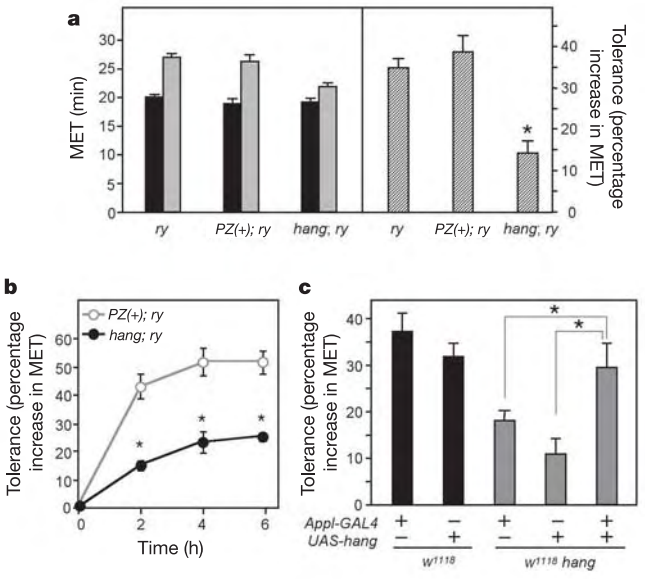


Figure 1 | $hang^{AE10}$ flies have abnormal ethanol tolerance development. **a**, Left panel shows mean elution times (MET) from the inebriometer of naive flies (first exposure, black bars) and flies pre-exposed to ethanol (second exposure, grey bars). Right panel shows development of ethanol tolerance, expressed as a percentage increase in MET between the two exposures. $hang^{AE10}$ mutants (*hang; ry*) show significantly reduced tolerance compared to two controls, (*ry*) and (*PZ(+); ry*). Asterisk, $P < 0.0001$; $n = 18$ experiments. **b**, $hang^{AE10}$ flies show defective ethanol tolerance in a chronic tolerance protocol in which flies are exposed to ethanol in the inebriometer four times at 2-h intervals. Asterisk, $P < 0.001$; $n = 8$ experiments. **c**, The $hang^{AE10}$ tolerance defect can be rescued by expression of a *UAS-hang* transgene in the nervous system under the control of *Appl-GAL4* driver. Mutant w^{1118} *hang* flies carrying either transgene alone show a reduced tolerance (similar to that of $hang^{AE10}$ flies), whereas flies carrying both transgenes show normal tolerance. Asterisk, $P < 0.005$; $n = 5–10$ experiments. In all panels, error bars represent s.e.m.

[1]Department of Anatomy and [2]Program in Neuroscience, University of California at San Francisco, California 94143-0452, USA. [3]Biozentrum, University of Würzburg, Am Hubland, D-97074 Würzburg, Germany.

Consistent with a role for *hang* in the nervous system, expression of a *UAS–hang* transgene under the control of the pan-neuronal *Appl-GAL4* driver[9] restores normal ethanol tolerance to $hang^{AE10}$ flies (Fig. 1c).

At high concentrations, ethanol is known to induce cellular responses similar to those elicited by heat shock[10,11]. Indeed, ethanol exposure in adult flies induces the expression of the heat-shock protein Hsp70 (data not shown). To assess whether these cellular stress responses might mediate the development of ethanol tolerance, we tested whether a heat pulse could mimic the effects of ethanol pre-exposure. Heat exposure of control flies (37 °C for 30 min) led to a 46 ± 4% increase in MET (compared to untreated flies) when measured in the inebriometer 4 h later. This heat–ethanol cross-tolerance is substantially reduced in $hang^{AE10}$ flies (26 ± 4%, Fig. 3a). The fact that $hang^{AE10}$ flies are deficient in both forms of tolerance leads us to conclude that the cellular changes induced by ethanol and heat overlap. However, $hang^{AE10}$ flies retain some capacity for developing tolerance (Figs 1a and 3a, b), suggesting that other pathways are also involved.

We have previously found that flies lacking the neuromodulator octopamine, owing to a mutation in the gene encoding tyramine β hydroxylase (*Tbh*)[12], show a reduction in ethanol tolerance similar to that seen with $hang^{AE10}$ flies[4] (22 ± 5% increase in MET compared with 34 ± 7% for control flies, Fig. 3b). To test whether the residual tolerance seen with $hang^{AE10}$ mutant flies might be mediated by octopaminergic systems (and vice versa), we tested flies lacking both octopamine and the *hang* gene product. Ethanol tolerance is almost completely abolished in *Tbh* $hang^{AE10}$ double mutant flies (6 ± 2% increase in MET compared with 34 ± 7% for control flies, Fig. 3b). As both mutations cause complete loss of gene function (ref. 12 and Fig. 2), these data suggest that the development of ethanol tolerance relies on two parallel molecular pathways, one involving octopaminergic systems and the other pathway involving the *hang* gene. The octopamine pathway is specific to ethanol tolerance, as *Tbh* flies develop normal heat–ethanol cross-tolerance (Fig. 3a). In contrast, the *hang* pathway involves cellular stress responses that are shared with those induced by heat. Finally, the fact that some heat–ethanol cross-tolerance remains in *hang* and *Tbh hang* double mutant flies (Fig. 3b) suggests that heat treatment engages a third, unidentified pathway involved in ethanol resistance.

To test whether $hang^{AE10}$ flies are more generally deficient in their ability to deal with environmental stressors, we exposed the mutant flies to paraquat, an agent that induces the formation of reactive oxygen species and that has been used to test for oxidative-stress responses[13]. When fed paraquat-containing food, $hang^{AE10}$ flies show a dose- and time-dependent reduction in survival rate compared with control flies (Fig. 3c, d). This reduced survival was not seen with *Tbh* flies, which appear to be even slightly resistant to the effects of the drug (Fig. 3c, d). The generation of endogenous reactive oxygen species is believed to contribute to the ageing of organisms[14]. We find that $hang^{AE10}$, but not *Tbh* flies, have a reduced life span (Supplementary Fig. 3a, b), which may be caused by an impaired ability to cope with reactive oxygen species. Mutations that cause neurodegeneration often result in reduced lifespan[15]. This does not appear to be the case in $hang^{AE10}$ flies, as their overall brain structure is normal, even in old flies (Supplementary Fig. 3c). This suggests a more direct involvement of *hang* in the functional response to environmental stressors. *hang* expression was not altered by ethanol exposure (data not shown). However, the predicted HANG protein contains, in addition to multiple zinc-fingers, a calcium-binding EF hand (Fig. 2b), suggesting that HANG function might be regulated by calcium. As calcium levels are sensitive to ethanol, reactive oxygen species and heat exposure[10,16–18], this provides a potential mechanism by which these environmental stressors could regulate HANG function.

In summary, the development of ethanol tolerance in *Drosophila* engages two systems that function in parallel, one involving a cellular stress pathway defined by the *hang* gene and the other involving

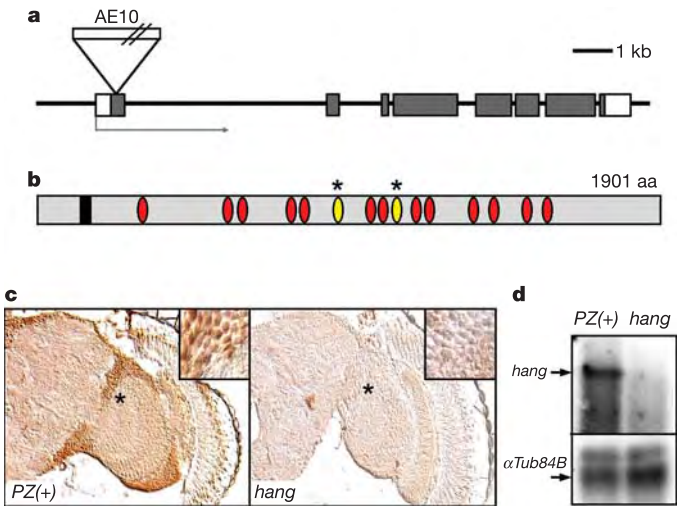


Figure 2 | The $hang^{AE10}$ mutation disrupts a gene encoding a zinc-finger protein. **a**, Genomic map of the *hang* locus (CG32575). The 8 exons are shown as boxes; grey shading indicates protein-coding regions. The position of the P-element insertion AE10 is shown. **b**, Conceptual translation predicts a protein of 1,901 amino acids (aa) (http://www.fruitfly.org/). The positions of an EF hand (black box) and zinc-finger domains (ovals) are indicated; U1-like zinc-fingers are marked with asterisks. **c**, HANG protein is broadly expressed in the nervous system (left panel) and localized to the nuclei of neurons (inset, see also Supplementary Fig. 2). Expression is undetectable in the $hang^{AE10}$ mutant (right panel). **d**, Northern blot analysis shows that expression of the full-length *hang* transcript (~7 kb, top arrow) is absent (or severely reduced) in the $hang^{AE10}$ mutant compared with the *PZ(+)* control strain. The lower panel shows expression of *α-Tubulin at 84B* (*αTub84B*), used as a control.

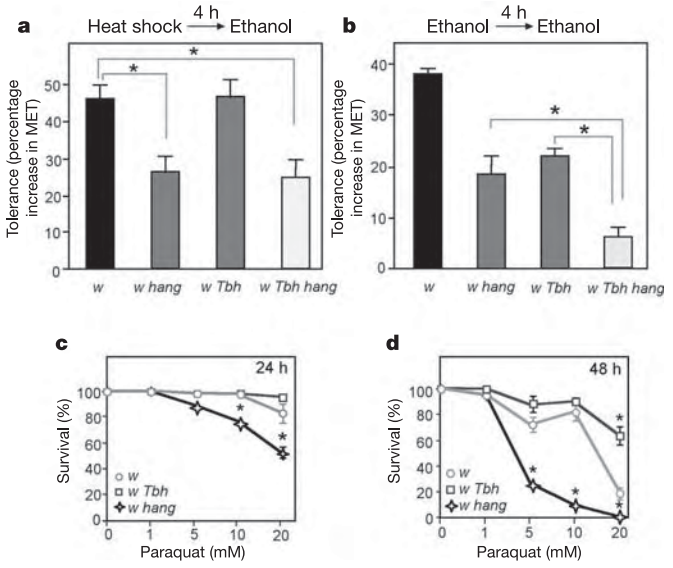


Figure 3 | $hang^{AE10}$ results in impaired heat–ethanol cross-tolerance. **a**, Heat–ethanol cross-tolerance. Control flies (*w*) and flies lacking octopamine (*w Tbh*) show normal cross-tolerance, whereas $hang^{AE10}$ mutants (*w hang*) and the double mutant (*w Tbh hang*) show strongly reduced cross-tolerance. Asterisk, $P < 0.0001$; $n = 8$–10 experiments. **b**, Ethanol tolerance, although reduced in $hang^{AE10}$ and *Tbh* flies (significantly different from control flies, not indicated in panel), is essentially absent in the *Tbh hang* double mutant flies. Asterisk, $P < 0.001$; $n = 4$–6 experiments. **c**, **d**, $hang^{AE10}$ flies show increased sensitivity to paraquat. Survival data obtained 24 (**c**) and 48 h (**d**) after exposure to different concentrations of paraquat. Asterisk, $P < 0.001$; $n = 5$ experiments. For all panels, error bars represent s.e.m.

octopaminergic systems. Octopamine has recently been implicated in the formation of appetitive (sugar-reinforced) memories in *Drosophila*[19]. The contribution of learned behaviours and stress, at both the cellular and systemic levels, to drug- and addiction-related behaviours in mammals is being increasingly recognized[20,21]. Our studies in *Drosophila* suggest that these pathways are conserved, allowing further analysis in this genetically tractable model organism.

METHODS

Fly stocks and genetics. We generated and tested 404 homozygous viable X-linked lines carrying *PZ(ry$^+$)*-element insertions[22] for their ability to develop ethanol tolerance. Lines were generated[6] and tested as described previously[4]. We isolated two lines with altered ethanol tolerance, one of them being the insertion line *hang^{AE10}*. To ensure that the phenotype was not caused by chromosomal alterations unlinked to the *PZ(ry$^+$)*-element, *hang^{AE10}* mutants were crossed for five generations to *ry^{506}* flies and re-tested for tolerance. The *hang^{AE10}* flies retained their ethanol tolerance defect. The *PZ(ry$^+$)* control line shows normal ethanol sensitivity and tolerance as measured in the inebriometer[4,6]. Excisions of the *hang^{AE10} PZ(ry$^+$)*-insertion were generated as described previously[6]. Four independent excision stocks were tested for ethanol tolerance and analysed by polymerase chain reaction (PCR) and DNA sequencing. A perfect correlation between precise excision of the P-element and phenotype reversion was observed (Supplementary Fig. 1b).

The *w^{1118} Tbh^{nM18} hang^{AE10}* double mutant stock was generated by recombination, as *Tbh* and *hang* both reside on the X chromosome. Stocks that showed female sterility, associated with the *Tbh^{nM18}* mutation[12], were tested for β-galatosidase expression to determine the presence of the *hang^{AE10} PZ(ry$^+$)*-insertion. A fertile *w^{1118} Tbh$^+$ hang^{AE10}* strain was also generated as a control. Two independent recombinant double mutant strains with the same genotype (*w^{1118} Tbh^{nM18} hang^{AE10}*) were tested for ethanol tolerance, with identical results.

Behavioural analyses. For all behavioural tests, flies were generated as described before[4,6]. Approximately 100 3–4-day-old males were tested in the inebriometer (MET1)[6,7]. After the initial exposure, flies were collected in vials and allowed to recover in a humidified incubator at 25 °C. The second exposure in the inebriometer (MET2) was initiated exactly 4 h after the start of the first exposure. Tolerance was calculated as described previously[4], using the formula (MET2–MET1/MET1 × 100). Ethanol absorption was quantified in extracts prepared from adult flies as described previously[6,7]. Significance was established using either Student's paired *t*-tests assuming equal variance or one-way analysis of variance (ANOVA) with Newman-Keuls post-hoc tests.

For heat-shock–ethanol cross-tolerance experiments, flies were incubated in a vial at 37 °C for 30 min in a water bath. After a recovery period of 3.5 h at 25 °C in a humidified incubator, the flies were exposed to ethanol in the inebriometer (MET^{hs+}). Tolerance was calculated with respect to flies that were not heat-treated (MET^{hs-}), using the formula (MET^{hs+}–MET^{hs-}/MET^{hs-} × 100).

To measure sensitivity to paraquat, groups of 20 3–4-day-old male flies were transferred into a glass vial containing a filter paper saturated with aqueous solutions of paraquat (0, 1, 5, 10 or 20 mM) in 5% sucrose. Flies were kept in an incubator with 70% humidity at 25 °C. Every 12 h, dead flies were counted and live flies were transferred into new vials.

Molecular biology. The genomic DNA flanking the *hang^{AE10} PZ(ry$^+$)* insertion was isolated by plasmid rescue. Comparisons with the published genome sequence of *Drosophila* revealed that the insertion was located in the first exon of CG32575, a finding that was confirmed by PCR. Several cDNA clones (GH14331, LD34334 and GH19829) were sequenced. The GH14331 cDNA is approximately 7 kb long, and its sequence corresponds to the predicted structure of CG32575 (http://www.fruitfly.org/). The sequence of GH14331 does not differ from the genomic sequence, but differs from the annotated transcripts CG32575RA and CG32575RB. The full-length cDNA of GH14331 was cloned into pUAST[23] to generate the *UAS-hang* transgene. For northern blots, 1 μg polyA$^+$ RNA from *PZ(ry$^+$)* and *hang^{AE10}* was used. A 3.7-kb *Eco*RI/*Kpn*I-digested fragment from the GH14331 cDNA clone, spanning from exon 1 (3′ of the insertion site) to exon 5, was used as a hybridization probe. A probe that recognizes *α-Tubulin at 84B* (*αTub84B*) mRNA (nucleotides 10–340; http://flybase.bio.indiana.edu/) was used as a control.

Histology. To generate antibodies to the HANG protein, a His-tagged antigen was made. The sequence encoding amino acids 62–319 of HANG was cloned into the pET28-b vector (Novagene). Protein was purified from *Escherichia coli* extracts using a nickel column, according to the manufacturer's instructions, and injected into rabbits. For HANG immunohistochemistry, flies were fixed for 2.5 h in 5% paraformaldehyde at 4 °C, washed and infiltrated with a 25% sucrose solution overnight at 4 °C. Whole flies were sectioned into 10-μm slices using a cryostat. Sections were washed twice for 5 min in PBT (PBS plus 0.3% Triton X-100) and preincubated with 10% normal goat serum in PBT for 2 h at 20 °C. Incubation with anti-HANG serum (1:5,000 in PBT plus 10% normal goat serum) was carried out overnight at 4 °C. After two 5-min washes with PBT, signal detection was carried out using the Vectastain ABC elite kit with a peroxidase-conjugated mouse anti-IgG antibody.

Received 15 April; accepted 27 May 2005.

1. Tabakoff, B., Cornell, N. & Hoffman, P. L. Alcohol tolerance. *Ann. Emerg. Med.* **15,** 1005–1012 (1986).
2. Lê, A. D. & Mayer, J. M. in *Pharmacological Effects of Ethanol on the Nervous System* (eds Deitrich, R. A. & Erwin, V. G.) 251–268 (CRC Press, Boca Raton, 1996).
3. Fadda, F. & Rossetti, Z. L. Chronic ethanol consumption: from neuroadaptation to neurodegeneration. *Prog. Neurobiol.* **56,** 385–431 (1998).
4. Scholz, H., Ramond, J., Singh, C. M. & Heberlein, U. Functional ethanol tolerance in *Drosophila*. *Neuron* **28,** 261–271 (2000).
5. Weber, K. E. & Diggins, L. T. Increased selection response in larger populations. II. Selection for ethanol vapor resistance in *Drosophila melanogaster* at two population sizes. *Genetics* **125,** 585–597 (1990).
6. Moore, M. S. *et al.* Ethanol intoxication in *Drosophila*: genetic and pharmacological evidence for regulation by the cAMP signalling pathway. *Cell* **93,** 997–1007 (1998).
7. Singh, C. M. & Heberlein, U. Genetic control of acute ethanol-induced behaviors in *Drosophila*. *Alcohol. Clin. Exp. Res.* **24,** 1127–1136 (2000).
8. Legrain, P. & Choulika, A. The molecular characterization of PRP6 and PRP9 yeast genes reveals a new cysteine/histidine motif common to several splicing factors. *EMBO J.* **9,** 2775–2781 (1990).
9. Torroja, L., Chu, H., Kotovsky, I. & White, K. Neuronal overexpression of APPL, the *Drosophila* homologue of the amyloid precursor protein (APP), disrupts axonal transport. *Curr. Biol.* **9,** 489–492 (1999).
10. Wilke, N., Sganga, M., Barhite, S. & Miles, M. F. Effects of alcohol on gene expression in neural cells. *EXS* **71,** 49–59 (1994).
11. Piper, P. W. The heat shock and ethanol stress responses of yeast exhibit extensive similarity and functional overlap. *FEMS Microbiol. Lett.* **134,** 121–127 (1995).
12. Monastirioti, M., Linn, C. E. J. & White, K. Characterization of *Drosophila tyramine β-hydroxylase* gene and isolation of mutant flies lacking octopamine. *J. Neurosci.* **16,** 3900–3911 (1996).
13. Arking, R., Buck, S., Berrios, A., Dwyer, S. & Baker, G. T. III Elevated paraquat resistance can be used as a bioassay for longevity in a genetically based long-lived strain of *Drosophila*. *Dev. Genet.* **12,** 362–370 (1991).
14. Finkel, T. & Holbrook, N. J. Oxidants, oxidative stress and the biology of ageing. *Nature* **408,** 239–247 (2000).
15. Min, K. T. & Benzer, S. Spongecake and eggroll: two hereditary diseases in *Drosophila* resemble patterns of human brain degeneration. *Curr. Biol.* **7,** 885–888 (1997).
16. Suzuki, Y. J., Forman, H. J. & Sevanian, A. Oxidants as stimulators of signal transduction. *Free Radic. Biol. Med.* **22,** 269–285 (1997).
17. Walter, H. J. & Messing, R. O. Regulation of neuronal voltage-gated calcium channels by ethanol. *Neurochem. Int.* **35,** 95–101 (1999).
18. Sun, A. Y. *et al.* Ethanol and oxidative stress. *Alcohol. Clin. Exp. Res.* **25,** 237S–243S (2001).
19. Schwaerzel, M. *et al.* Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in *Drosophila*. *J. Neurosci.* **23,** 10495–10502 (2003).
20. Hyman, S. E. & Malenka, R. C. Addiction and the brain: the neurobiology of compulsion and its persistence. *Nature Rev. Neurosci.* **2,** 695–703 (2001).
21. Nestler, E. J. Molecular basis of long-term plasticity underlying addiction. *Nature Rev. Neurosci.* **2,** 119–128 (2001).
22. Mlodzik, M. & Hiromi, Y. Enhancer trap method in *Drosophila*: Its application to neurobiology. *Methods Neurosci.* **9,** 397–414 (1992).
23. Brand, A. H. & Perrimon, N. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118,** 401–415 (1993).

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

Acknowledgements We thank R. Threlkeld for the northern blot, N. Funk and I. Schwenkert for the production of rescue constructs and anti-HANG antibody, H.Saumweber and G.Krohne for gifts of antibodies, and B. Poeck, A. Corl, A. Rothenfluh, D. Guarnieri, F. Wolf and C. Kenyon for comments on the manuscript. This work was supported by the NIH/NIAAA (U.H.), ABMRF (U.H.), the Sandler Foundation (U.H.), the Wheeler Center (H.S. and U.H.) and the German Science Foundation DFG (H.S.)

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.S. (henrike.scholz@biozentrum.uni-wuerzburg.de) or U.H. (ulrike@itsa.ucsf.edu).

LETTERS

Small vertical movement of a K^+ channel voltage sensor measured with luminescence energy transfer

David J. Posson[1], Pinghua Ge[1], Christopher Miller[2], Francisco Bezanilla[3] & Paul R. Selvin[1]

Voltage-gated ion channels open and close in response to voltage changes across electrically excitable cell membranes[1]. Voltage-gated potassium (Kv) channels are homotetramers with each subunit constructed from six transmembrane segments, S1–S6 (ref. 2). The voltage-sensing domain (segments S1–S4) contains charged arginine residues on S4 that move across the membrane electric field[2,3], modulating channel open probability. Understanding the physical movements of this voltage sensor is of fundamental importance and is the subject of controversy. Recently, the crystal structure of the KvAP[4] channel motivated an unconventional 'paddle model' of S4 charge movement, indicating that the segments S3b and S4 might move as a unit through the lipid bilayer with a large (15–20-Å) transmembrane displacement[5]. Here we show that the voltage-sensor segments do not undergo significant transmembrane translation. We tested the movement of these segments in functional *Shaker* K^+ channels by using luminescence resonance energy transfer to measure distances between the voltage sensors and a pore-bound scorpion toxin. Our results are consistent with a 2-Å vertical displacement of S4, not the large excursion predicted by the paddle model. This small movement supports an alternative model in which the protein shapes the electric field profile, focusing it across a narrow region of S4 (ref. 6).

Conformational changes in proteins can be studied in great detail by using fluorescence energy transfer as a spectroscopic ruler[7,8]. Luminescence resonance energy transfer (LRET) is a modified version that employs a lanthanide donor complex with a long excited-state lifetime[9,10]. This unconventional probe can donate energy to a conventional fluorescent acceptor in the standard distance-dependent manner of Förster theory[10], and energy transfer efficiency and distances are calculated from the time constants of acceptor fluorescence emission (see Methods). LRET is capable of accurately measuring distances on *Shaker* channels *in vivo* because only donor–acceptor pairs produce sensitized acceptor emission, which is measured after a brief time-gate rejects fast background fluorescence. Further advantages arise from the minimal spectral overlap of donor and acceptor, the zero intrinsic anisotropy of the donor lanthanide[11], and the accuracy with which donor quantum yield[10] and R_o (the characteristic distance of 50% energy transfer) can be estimated.

Gating-driven protein movements have previously been measured on the *Shaker* channel by using LRET[12] and conventional fluorescence resonance energy transfer[13] with both the donor and the acceptor labelled at sites on the voltage-sensing domains. This configuration measured distance changes parallel to the membrane between S4 helices in the same tetrameric channel but did not directly measure transmembrane movements. Here we have attached the acceptor dye to the channel by means of a scorpion toxin that binds to the pore from the external solution[14,15]. Toxin binding is insensitive to the channel's open or closed status and does not alter the movement of the voltage sensor[15,16]. Lanthanide donors were attached to several sites on S4, S3b and the S3–S4 linker region near S4, to test directly *in vivo* whether the voltage-sensing segments undergo a large transmembrane movement (Fig. 1). Any paddle-type mechanism by definition requires a transmembrane movement of 15–20 Å (ref. 5), equivalent to a change in distance of about 10 Å from S4 to toxin in the configuration used here, as estimated from conservative structural assumptions (Fig. 1). Testing the vertical translation of S3b–S4 is of central importance in evaluating the validity of the paddle mechanism, because the model's other unusual feature, the location of S4 at the lipid interface, has been shown experimentally to be plausible[17,18].

For LRET, ionic currents of *Shaker* expressed in *Xenopus* oocytes were blocked with 100 nM fluorescent charybdotoxin (CTX) or agitoxin-2 (AgTX) such that almost all channels were blocked and residual unblocked current was limited to 10–30 μA to minimize voltage-clamping errors[19]. The charge–voltage relations for donor-labelled channels were measured separately from cells blocked with a saturating level of unlabelled toxin (see Supplementary Information). Lanthanide-chelate donors (terbium) were attached to site-directed cysteine residues on the voltage-sensor domain (see Methods). The labelling of voltage-sensing segments with fluorescent probes does not disrupt the movement of gating charge. The effect of fluorescein (and rhodamine) was tested, but not terbium-chelate (see Supplementary Information). Toxin binding brought acceptor fluorophores into proximity to the labelled donor sites on the channel,

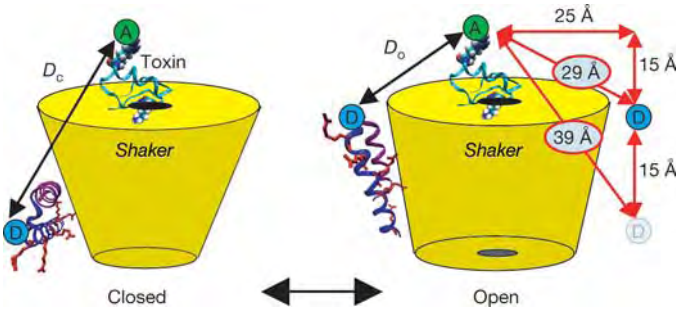


Figure 1 | Diagram of the paddle model. LRET measures distances from donor labelled sites (blue circles) on the S3b–S4 paddle (structure taken from the isolated voltage sensor[4]). The voltage-sensing arginine residues are shown in red. The energy-transfer acceptors (green circles) are attached to the top of the channel with a scorpion toxin. The paddle model predicts a change in distance, $D_c - D_o$, of 10 Å, estimated by a conservative geometric calculation assuming a 15-Å vertical translation (red arrows).

[1]Department of Physics and Biophysics Center, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA. [2]Department of Biochemistry, Howard Hughes Medical Institute, Brandeis University, Waltham, Massachusetts 02454, USA. [3]Department of Physiology and Department of Anesthesiology, UCLA School of Medicine, Los Angeles, California 90095, USA.

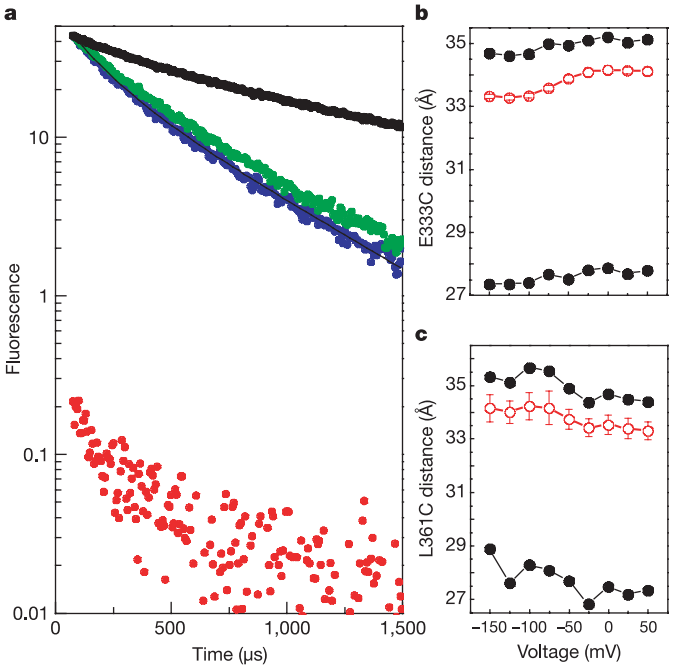


Figure 2 | LRET raw data and distance calculations. **a**, Acceptor-sensitized emission at two extreme voltages is shown for the E333C mutant near the top of the paddle. Black, donor only ($\tau_D = 1{,}600\,\mu s$); green, acceptor +50 mV; blue, acceptor −150 mV ($\tau_{AD} = 160\,\mu s$, 509 μs); red, average control data ($n = 6$). **b**, The time constants displayed a small voltage dependence corresponding to a small movement 0.8 Å away from the toxin. The distances calculated from the two lifetime components (black) and the weighted average (red; see Methods) are shown. **c**, L361C showed voltage-dependent movement of 0.8 Å towards the toxin. Error bars on the average distance represent s.e.m. ($n = 13$ in **b**; $n = 8$ in **c**).

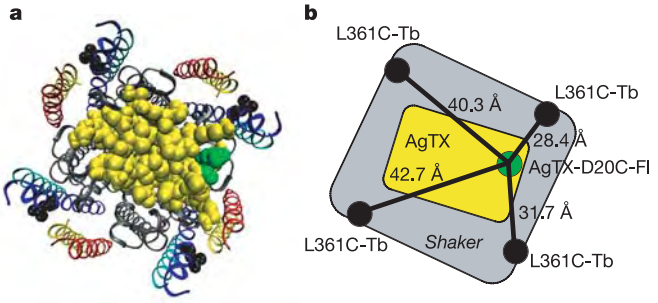


Figure 3 | A model of *Shaker* with docked AgTX predicts four distances for each LRET experiment. **a**, The coordinates provide an opportunity to compare our measurements with a picture of *Shaker* that has S3 and S4 placed against the pore domain. **b**, Distances for L361C on S4, measured from α-carbons.

and the distances were calculated as a function of voltage from the measured LRET time constants. Acceptor-sensitized emission data from E333C on S3b and background controls (see Methods) are shown in Fig. 2. LRET signals fitted well to two time constants that reflect the asymmetry of the toxin–acceptor position with respect to the central axis of the channel (Fig. 3). Distances from both time constants were calculated, as was a population-weighted average distance versus voltage (see Methods). These distances are shown for E333C on S3b and L361C on S4 (Fig. 2). The results show that sites homologous to the KvAP voltage-sensor paddle move less than 1 Å with respect to the toxin when going from the closed to the open channel positions. If the S4 segment moves in a purely vertical direction, a change in LRET distance of 1 Å corresponds to a 2 Å vertical displacement, as estimated by a conservative geometric calculation similar to that shown in Fig. 1.

Small but unambiguous voltage-dependent movements were seen at many sites (Fig. 4), with S3 moving about 1 Å away from the toxin, S4 moving about 1 Å towards the toxin, and the sites in the linker moving up to 2.5 Å in a manner consistent with a change in linker tilt[12]. We note that S3b and S4 move in opposite directions, instead of translating together as a rigid unit. For three sites, N353C, E335C and L361C, both AgTX and CTX gave similar calculated distances. Two sites on the S3–S4 linker were studied with two different acceptors, CTX-Lucifer Yellow and CTX-Atto465, which are useful for measuring distances as short as 15 Å (see Methods). For S346C, the calculated distances differed by only 2.5 Å, which may be attributed to differences in dye size and linker lengths. For S351C the distances obtained using CTX-Atto465 and CTX-4,4-difluoro-4-bora-3a,4a-diaza-*s*-indacene (BODIPY) Fl differed by 5 Å, but the gating-induced change in distance was unaffected by the choice of acceptor. Thus, the absolute distances are slightly uncertain, but the changes in distance are reproducible. As a further control, we switched the donor and acceptor for one experiment, labelling E333C with fluorescein acceptors and attaching a CTX-Tb donor to the top of the channel. The resulting distance measurements were nearly identical to those in Fig. 2 (see Supplementary Information). In every experiment, minimal voltage-dependent changes in both LRET amplitudes and time constants showed that only small changes in energy transfer occurred. These small changes refute the most central feature of the paddle model: substantial physical movement of gating charge transverse to the membrane plane.

We are confident that this LRET technique estimates distances faithfully, on proteins in general and in K^+ channels in particular. Beyond the technique's agreement with known structures in soluble proteins[10], distances measured here agree well with independent estimates of distances from the *Shaker* voltage sensors to the pore made with the use of tethered tetraethylammonium blockers[20]. For example, in the tether experiments, Q348C, D349C and K350C were found to be 17–18 Å from the pore in the open state, very similar to our measurements of 17–19 Å and 21–25 Å for S3–S4 linker residues S346C and S351C, respectively. Similarly, E334C and E335C were found to be about 30 Å from the pore in the tether experiments, close to our measurements for these same residues at the end of S3b, 32–34 Å. Although tethered blockers and LRET measure distances to two different points near the central pore, the close agreement between the approaches shows their power for constraining structural distances on the *Shaker* channel. Furthermore, tethered blockers measure distances only for the open state whereas LRET has the advantage of probing both open and closed states.

LRET measures absolute distances with less systematic error than traditional energy-transfer techniques[8,10] and can therefore be used to evaluate and constrain structural models. Recently, a structural model was proposed for the *Shaker* open state based on a combination of experimental data and molecular dynamics[21,22]. This model was supplemented with a computationally docked AgTX[23] so that theoretical distances from the toxin labelling site to sites on the

Table 1 | Comparison of LRET distance analysis with distances taken from a model of the *Shaker* open state

Donor site	Acceptor used	Average model *D* (Å)	Average simulation *D* (Å)	LRET average *D*
(S4) L361	BODIPY Fl-CTX	35.8	33.4	33.3
(S4) R365	BODIPY Fl-CTX	39.8	35.5	33.2
(S3b) V330	BODIPY Fl-CTX	30.9	30.2	34.8
(S3–S4 Linker) S351	BODIPY Fl-CTX	30.2	30.5	25.5
(S3–S4 Linker) S351	Atto465-CTX	30.2	25.0	21.0

Model *D* is the mean of four distances measured from the Cα of D20 on a docked AgTX to the Cα of specified sites on each subunit (Fig. 3). For simulation data see Methods and Supplementary Information.

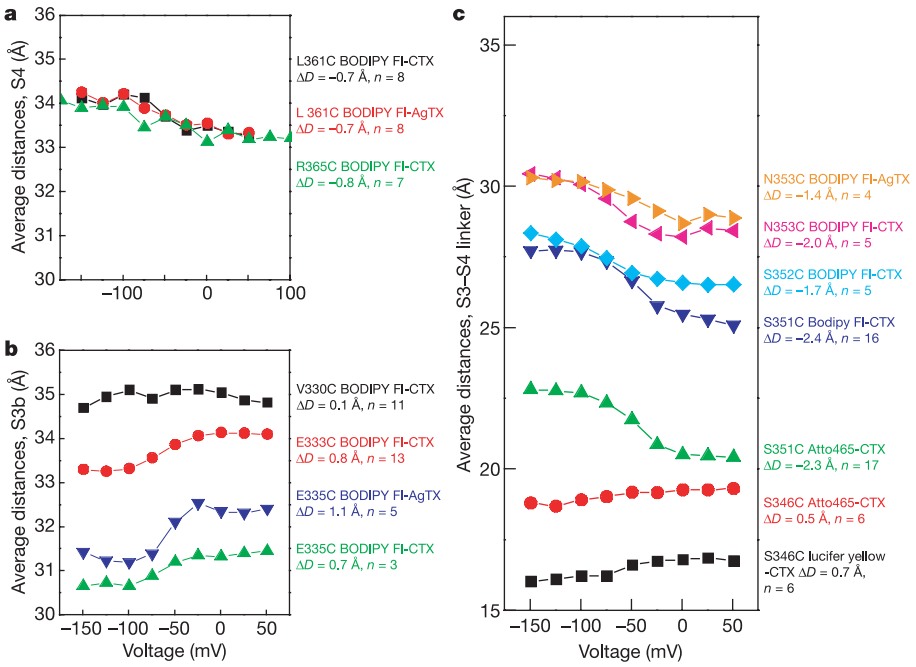


Figure 4 | Average distances for many *Shaker* sites. a, **b**, The S4 (**a**) and S3b (**b**) sites are homologous to sites on the KvAP voltage-sensing paddle. The distances for S4 change, 0.8 Å, consistent with a roughly 2-Å vertical translation. S3b moves in the direction opposite to S4, moving just 0.8 Å away from the toxin. **c**, Sites in the S3–S4 linker are clearly closer to the toxin than the transmembrane segments, as expected, and move no more than a few ångströms.

Shaker voltage sensor could be compared directly with our LRET measurements (Fig. 3). The model predicts four theoretical distances and we have used simulations to test how well LRET experiments can measure the average distance for situations of such geometric complexity (Table 1; see Methods and Supplementary Information). The simulations reproduce average distances in close agreement with model values; the exception was S351 using CTX-Atto465, for which the small R_o caused an underestimation. The LRET experimental results for two sites on S4 show very close agreement between model and data. However, the LRET measurements may systematically underestimate distances slightly because the position of the probes can wobble around their linker attachment points, weighting the measurement towards the distance of closest approach. The model prediction for S3b was unique in that it predicted a shorter distance (about 4 Å) than the distance measured experimentally. Nevertheless, the distance values obtained with LRET are consistent with the general structural view that S3 and S4 are transmembrane segments at all voltages.

The small vertical S4 movements presented here supplement the even smaller lateral movements between voltage sensors obtained previously[12] and indicate that the conformational changes underlying gating charge movement are subtle rather than substantial. The paddle model could be altered to account for our data by allowing the paddle segments to swing laterally outwards while undergoing vertical movement such that distances to the toxin remain constant. However, this kind of movement would be flatly inconsistent with the small lateral displacements observed in previous LRET measurements[12]. The small physical movements of voltage-sensing segments indicate that the membrane electric field must be focused over a very tight region of the voltage sensor, as if aqueous crevices penetrate the protein and thereby shape the field profile[6,24]. Small S4 movements relative to these crevices and voltage-induced changes in crevice shape can produce the large gating charge that must traverse the field to account for the steep voltage dependence of voltage-gated channels.

METHODS

Distance calculations. The lifetime of acceptor-sensitized emission was used to calculate energy transfer efficiency by using the relation $E = 1 - \tau_{AD}/\tau_D$, where τ_D is the lifetime of the donor in the absence of acceptor. τ_D was measured on channel sites in the absence of acceptors. On a few sites τ_D displayed a slight voltage dependence (S346C and S351C less than 10%; E335C less than 5%) and these changes were included in the analysis (distances changed were less than 1 Å). Sensitized emission data were fit to two exponentials using four parameters; A_1, τ_1, A_2, τ_2. Multiple time constants indicate that the acceptor molecule is not an equal distance from all four labelled donors to the voltage-sensing domains. An average lifetime was calculated by normalizing the sensitized emission lifetimes by the rate of energy transfer to obtain a 'population average'[25]

$$\bar{\tau} = \frac{(A_1/k_1)\tau_1 + (A_2/k_2)\tau_2}{(A_1/k_1) + (A_2/k_2)}$$

where $k_n = \tau_n^{-1} - \tau_D^{-1}$. Distances from τ_1, τ_2 and $\bar{\tau}$ were calculated by finding E (above) and using $R = R_o(E^{-1} - 1)^{1/6}$. Most data were taken by using BODIPY Fl-maleimide acceptors (Molecular Probes) for which $R_o = 39$ Å. Other data were taken with the use of Atto465-maleimide (Atto-Tec), $R_o = 27$ Å, and Lucifer Yellow-iodoacetamide (Molecular Probes), $R_o = 23$ Å.

Toxin biochemistry, *Shaker* expression and block. CTX-R19C and AgTX-D20C were prepared, labelled and purified as described previously[26]. The mass of each fluorescent toxin was verified by mass spectrometry, and high-affinity block with *Shaker* was evaluated qualitatively by examining the slow rate of toxin dissociation. The channel construct was the fast inactivation-removed, conducting *Shaker* H4IR, with the mutations F425G, K427D that increase the toxin binding to sub-nanomolar affinities[27]. *Xenopus* ooctye preparation, channel mutagenesis (Stratagene) and mRNA synthesis (Ambion) were performed with standard procedures. Experiments were performed typically 3–5 days after microinjection of 20 ng of *Shaker* mRNA. Voltage clamping was performed with a two-electrode setup (CA-1B; Dagan). Charge–voltage relations were measured with a saturating wild-type CTX block (2 μM), and LRET measurements were recorded with a nearly complete block with 100 nM fluorescent toxin (see Supplementary Information).

LRET protocols and controls. The optical setup consisted of an Olympus inverted IX-70 microscope with a 40 × quartz objective (numerical aperture 0.8; Partec). The lanthanide was excited with a pulsed 337-nm nitrogen laser source (Oriel), reflected by a 400DCLP dichroic filter (Chroma). Donor and acceptor

fluorescence were collected simultaneously with D490/10 and HQ520/20 filters, respectively (Chroma). Fluorescence was detected with two water-cooled R943-02 photomultiplier tubes (Hamamatsu) operated at −1,760 V. Prompt fluorescence was rejected by using an electronic gate (Products for Research) with a dead-time of 70 μs. The detector current was converted to voltage with a C7319 preamplifier (10^6 V/A; Hamamatsu) and filtered at 50 kHz (eight-pole Bessel filter; Dagan). The laser pulse was given exactly 40 ms after the initiation of a voltage step to ensure that the channels had reached conformational equilibrium before the LRET signals were measured.

Background cysteine residues on oocytes were generally prelabelled with β-maleimidopropionic acid (Sigma) for 1 h after a 2–3-day incubation at 12 °C to increase the specificity of donor labelling[28]. Oocytes were then incubated for 24–30 h at 18 °C to allow the surface expression of *Shaker*. Cells were placed in depolarizing solution for 30 min with 100 μM dithiothreitol to reduce cysteine thiols for reaction with maleimide. Dithiothreitol was washed away before the cells were placed in depolarizing solution containing 80 μM maleimide-lanthanide chelate. LRET signal to background was estimated for every LRET experiment by recording acceptor-sensitized emission signals from oocytes that were expressing high levels of the background *Shaker* construct without the experimental cysteine mutation. Controls were labelled identically to the LRET experiments.

LRET simulations with a *Shaker* model. Coordinates for the *Shaker* open-state model with docked AgTX[21,23] provide predictions for four different distances between the AgTX-D20 α-carbon and the four α-carbons of selected sites on the voltage sensors. Assuming these four distances, LRET signals were simulated by assuming a bi-exponential donor with a dominant component, 75% at 1,600 μs, and a minor component, 25% at 300 μs. The minor component adds a systematic error that slightly underestimates distances (less than 5%). The multiple components can be well described by fitting to two exponentials (see Supplementary Information), as were the experimental data. These calculations show how the complicated geometry of the model can be reduced to distance estimations in close agreement with actual LRET measurements (Table 1).

Received 6 March; accepted 18 May 2005.

1. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (Lond.)* **117,** 500–544 (1952).
2. Bezanilla, F. The voltage sensor in voltage-dependent ion channels. *Physiol. Rev.* **80,** 555–592 (2000).
3. Armstrong, C. M. & Bezanilla, F. Currents related to movement of the gating particles of the sodium channels. *Nature* **242,** 459–461 (1973).
4. Jiang, Y. *et al.* X-ray structure of a voltage-dependent K^+ channel. *Nature* **423,** 33–41 (2003).
5. Jiang, Y., Ruta, V., Chen, J., Lee, A. & MacKinnon, R. The principle of gating charge movement in a voltage-dependent K^+ channel. *Nature* **423,** 42–48 (2003).
6. Starace, D. M. & Bezanilla, F. A proton pore in a potassium channel voltage sensor reveals a focused electric field. *Nature* **427,** 548–553 (2004).
7. Clegg, R. M. Fluorescence resonance energy transfer. *Curr. Opin. Biotechnol.* **6,** 103–110 (1995).
8. Selvin, P. R. The renaissance in fluorescence resonance energy transfer. *Nature Struct. Biol.* **7,** 730–734 (2000).
9. Selvin, P. R., Rana, T. M. & Hearst, J. E. Luminescence resonance energy transfer. *J. Am. Chem. Soc.* **116,** 6029–6030 (1994).
10. Selvin, P. R. Principles and biophysical applications of luminescent lanthanide probes. *Annu. Rev. Biophys. Biomol. Struct.* **31,** 275–302 (2002).
11. Reifenberger, J. G., Snyder, G. E., Baym, G. & Selvin, P. R. Emission polarization of europium and terbium chelates. *J. Phys. Chem. B* **107,** 12862–12873 (2003).
12. Cha, A., Snyder, G. E., Selvin, P. R. & Bezanilla, F. Atomic scale movement of the voltage sensing region in a potassium channel measured via spectroscopy. *Nature* **402,** 809–813 (1999).
13. Glauner, K. S., Mannuzzu, L. M., Gandhi, C. S. & Isacoff, E. Y. Spectroscopic mapping of voltage sensor movement in the Shaker potassium channel. *Nature* **402,** 813–817 (1999).
14. MacKinnon, R. & Miller, C. Mechanism of charybdotoxin block of the high-conductance, Ca^{2+}-activated K^+ channel. *J. Gen. Physiol.* **91,** 335–349 (1988).
15. Goldstein, S. A. & Miller, C. Mechanism of charybdotoxin block of a voltage-gated K^+ channel. *Biophys. J.* **65,** 1613–1619 (1993).
16. Aggarwal, S. K. & MacKinnon, R. Contribution of the S4 segment to gating charge in the Shaker K^+ channel. *Neuron* **16,** 1169–1177 (1996).
17. Hessa, T., White, S. H. & von Heijne, G. Membrane insertion of a potassium-channel voltage sensor. *Science* **307,** 1427 (2005).
18. Cuello, L. G., Cortes, D. M. & Perozo, E. Molecular architecture of the KvAP voltage-dependent K^+ channel in a lipid bilayer. *Science* **306,** 491–495 (2004).
19. Baumgartner, W., Islas, L. & Sigworth, F. J. Two-microelectrode voltage clamp of *Xenopus* oocytes: voltage errors and compensation for local current flow. *Biophys. J.* **77,** 1980–1991 (1999).
20. Blaustein, R. O., Cole, P. A., Williams, C. & Miller, C. Tethered blockers as molecular 'tape measures' for a voltage-gated K^+ channel. *Nature Struct. Biol.* **7,** 309–311 (2000).
21. Laine, M. *et al.* Atomic proximity between S4 segment and pore domain in Shaker potassium channels. *Neuron* **39,** 467–481 (2003).
22. Laine, M., Papazian, D. M. & Roux, B. Critical assessment of a proposed model of Shaker. *FEBS Lett.* **564,** 257–263 (2004).
23. Eriksson, M. A. & Roux, B. Modeling the structure of agitoxin in complex with the Shaker K^+ channel: a computational approach based on experimental distance restraints extracted from thermodynamic mutant cycles. *Biophys. J.* **83,** 2595–2609 (2002).
24. Asamoah, O. K., Wuskell, J. P., Loew, L. M. & Bezanilla, F. A fluorometric approach to local electric field measurements in a voltage-gated ion channel. *Neuron* **37,** 85–97 (2003).
25. Heyduk, T. & Heyduk, E. Luminescence energy transfer with lanthanide chelates: interpretation of sensitized acceptor decay amplitudes. *Anal. Biochem.* **289,** 60–67 (2001).
26. Shimony, E., Sun, T., Kolmakova-Partensky, L. & Miller, C. Engineering a uniquely reactive thiol into a cysteine-rich peptide. *Protein Eng.* **7,** 503–507 (1994).
27. Goldstein, S. A., Pheasant, D. J. & Miller, C. The charybdotoxin receptor of a Shaker K^+ channel: peptide and channel residues mediating molecular recognition. *Neuron* **12,** 1377–1388 (1994).
28. Mannuzzu, L. M., Moronne, M. M. & Isacoff, E. Y. Direct physical measure of conformational rearrangement underlying potassium channel gating. *Science* **271,** 213–216 (1996).

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

Acknowledgements We thank B. Roux for putting together coordinates for a combined model of the AgTX–*Shaker* complex[23] with the model for the *Shaker* open state[21], and L. Kolmakova-Partensky and T. Lawrecki for technical assistance. This work was supported by grants from the NIH, NSF, the Carver Foundation and the Cottrell funds of the Research Corp to P.R.S., from an NIH grant to F.P., and from the Howard Hughes Medical Institute to C.M. P.R.S. also thanks J. Ackland, J. Stenehjem and the other members of the Sharp Rehabilitation Center of San Diego for their care, which made this study possible.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.R.S. (selvin@uiuc.edu).

LETTERS

Gating charge displacement in voltage-gated ion channels involves limited transmembrane movement

Baron Chanda[1], Osei Kwame Asamoah[1]†, Rikard Blunck[1], Benoît Roux[2] & Francisco Bezanilla[1,3]

Voltage-gated ion channels are responsible for generating electrical impulses in nerves and other excitable cells. The fourth transmembrane helix (S4) in voltage-gated channels is the primary voltage-sensing unit that mediates the response to a changing membrane electric field[1,2]. The molecular mechanism of voltage sensing, particularly with respect to the magnitude of the transmembrane movement of S4, remains controversial[3–5]. To determine the extent of this transmembrane movement, we use fluorescent resonance energy transfer between the S4 domain and a reference point in the lipid bilayer. The lipophilic ion dipicrylamine distributes on either side of the lipid bilayer depending on the membrane potential, and is used here as a resonance-energy-transfer acceptor from donor molecules attached to several positions in the Shaker K^+ channel. A voltage-driven transmembrane movement of the donor should produce a transient fluorescence change because the acceptor also translocates as a function of voltage. In Shaker K^+ channels no such transient fluorescence is observed, indicating that the S4 segment does not translocate across the lipid bilayer. Based on these observations, we propose a molecular model of voltage gating that can account for the observed 13*e* gating charge with limited transmembrane S4 movement.

Molecular models of voltage gating that invoke a large translocation of the voltage-sensing S4 segment, such as the 'helical screw'[6,7] (in its original form) or the 'paddle'[5,8] models, account for the observed 13*e* gating charges intuitively by incorporating a complete translocation of the charged S4 residues across the full thickness of the lipid membrane. For example, in the paddle model the gating charges of the S4 segment are buried in the lipid bilayer and traverse approximately 20 Å in response to a transmembrane potential change. However, accessibility data[9–12], resonance energy transfer[13,14] and potentiometric studies[15] suggest that the voltage sensor undergoes relatively small transmembrane excursions. Previous measurements of the extent of the voltage sensor's translocation using fluorescent resonance energy transfer (FRET) did not resolve this issue because the distance estimates were obtained by measuring relative distances between mobile voltage sensors[13,14]. Our aim is to determine whether, during the gating process, the main voltage-sensing element (S4) undergoes a small or a large transmembrane movement relative to the membrane plane. To address this issue, we use the lipid bilayer as an independent reference plane and a novel strategy based on FRET.

The non-fluorescent hydrophobic anion dipicrylamine (DPA) (Fig. 1a, inset), which localizes at the lipid-aqueous interface, can be used as a classical Förster-energy-transfer acceptor from a variety of donor fluorophores owing to its absorbance in the blue region of the visible spectrum. In response to a change in membrane potential, DPA generates robust gating currents as it moves between the inner and outer membrane leaflets (Fig. 1b). The voltage- and time-dependent FRET signal ($\Delta F/F$) that arises from donor fluorescence quenching in the presence of DPA reflects the magnitude and direction of movement that is normal to the bilayer plane. If the S4 segment remains relatively static during the gating process (Fig. 1c, left), the donor fluorescence at the beginning of the pulse is quenched by the presence of DPA molecules that accumulate in the outer leaflet. Upon depolarization, the fluorescence increases monotonically because only DPA migrates to the inner leaflet of the lipid bilayer. Conversely, if the S4 segment also translocates across the membrane, the FRET signal is more complex (Fig. 1c, right). At hyperpolarized potentials, the S4 segment resides near the inner leaflet while DPA molecules populate the outer leaflet. Upon depolarization, the DPA molecules quickly migrate to the inner leaflet and quench the donor fluorescence. As the S4 segment subsequently moves towards the outer membrane the donor fluorescence recovers. This interplay results in a characteristic biphasic time trace with the donor fluorescence returning to baseline on sustained depolarization. We note that the biphasic time trace is observed even when both donor and acceptor groups have equal charge and mobility in the membrane, as long as they translocate from one side of the bilayer to the other (Supplementary Discussion). Thus, the fluorescence profile (biphasic versus monophasic) generated by such FRET pairs can provide a stringent qualitative test to discriminate between the two contrasting gating models.

We performed a series of control experiments in native *Xenopus laevis* oocyte membranes to validate this FRET technique. To reproduce the model of a large S4 translocation, the lipophilic anionic fluorophore oxonol (bis-(1,3-diethylthiobarbituric acid)trimethine oxonol; also called $DiSBAC_2$) was employed as a FRET donor. Because oxonol translocates across the bilayer with a time constant of 300 ms at 0 mV (ref. 16), 350-fold slower than DPA, we expect a transient FRET signal owing to this large difference in translocation time. Upon depolarization, the oxonol fluorescence shows an initial increase that is consistent with reduced FRET efficiency as DPA migrates away from oxonol. As oxonol molecules slowly relocate to the inner membrane leaflet, a decrease in fluorescence is observed. Upon repolarization, an analogous optical signal is observed as the hydrophobic anions re-establish their original positions (Fig. 1d). In contrast, a monotonic FRET signal is observed upon depolarization from di-8-ANEPPS (see Methods), which is a potentiometric dye that is exclusively localized to the outer leaflet

[1]Departments of Physiology and Anesthesiology, David Geffen School of Medicine, UCLA, 650 Charles E. Young Dr. South, Los Angeles, California 90025, USA. [2]Department of Physiology and Biophysics, Weill Medical College, Cornell University, 1300 York Avenue, New York, New York 10021, USA. [3]Centro de Estudios Cientificos, Valdivia, Chile.
†Present address: Department of Emergency Medicine, University of New Mexico Health Sciences Campus, Albuquerque, New Mexico 87106, USA.

(Fig. 1e, top). A small intrinsic electrochromic signal from di-8-ANEPPS (Fig. 1e, bottom) is overwhelmed by the FRET signal in the presence of DPA. As a further test, we used fluorescently labelled melittin peptide[17] for which there is evidence of voltage-activated insertion or translocation across the membrane upon hyperpolarization; akin to the paddle model[18]. When the potential was stepped to inside negative (see Methods), the fluorescence showed an early decrease corresponding to the translocation of DPA molecules (Fig. 1f). On sustained hyperpolarization, the fluorescence recovers with a time course that correlates with increasing membrane conductance, and reflects the formation of functional melittin channels. Thus, the transient FRET signal supports the hypothesis that the amino terminus of melittin moves across the width of the bilayer in response to a voltage pulse (Supplementary Discussion). Taken together, these results confirm that this FRET approach can distinguish a large transmembrane movement of protein segments from a minimal movement.

The extent of voltage-sensor translocation in Shaker K^+ channels was determined by measuring FRET signals from different positions in the S4 segment relative to DPA in the membrane. Our measurements were limited to residues D349, Q354, V363, V367 and F425, as these sites generate small (if any) intrinsic voltage-dependent fluorescence signals (Supplementary Discussion). The fluorescent label sulphorhodamine was attached using a methanethiosulphonate (MTS) linker to the cysteine-substituted residues in the S4 segment of the cysteine-less Shaker channels. FRET signals were recorded in response to a single voltage pulse to +50 mV (from −120 mV). This protocol is expected to move all gating charge in the Shaker channels, together with the majority of DPA molecules, across the bilayer. FRET signals from D349C and Q354C, close to the N terminus of S4, show a clear monophasic fluorescence increase consistent with their static position (Fig. 2a). FRET signals from V363C (close to the charge-carrying residues) is also monophasic, but inverted in orientation relative to the signals from the previous sites, suggesting that this site lies closer to the inner leaflet. This was initially surprising, because the neighbouring R362 residue is the first charge-carrying residue and is expected to be close to the outer membrane surface. However, the inverted signal is probably due to the long length of the sulphorhodamine dipole midpoint (~10 Å). This results in a shorter DPA–chromophore distance when DPA populates the inner (as compared to outer) membrane (Fig. 2b and Supplementary Discussion). The energy-transfer midline, as defined by FRET orientation, is unlikely to coincide with the physical midline of the lipid bilayer. For instance, if the footprint of the channel protein on the outer leaflet is larger than that on the inner leaflet, the energy-transfer midline is biased towards the outer leaflet. Thus, inverted FRET signals from V363C may reflect the long chromophore length, channel geometry and/or orientation of the voltage sensor. Nonetheless, it clearly does not produce the transient fluorescence signal expected from a group that translocates across the bilayer. The V367C residue, located before the third gating charge in the S4 segment and below V363C, also shows a monophasic decrease in fluorescence. As a control for fixed position, F425C (close to outer pore) was tested; it showed a monophasic fluorescence increase.

To further constrain the movement of S4 during gating, we labelled residue V363C with a 7-fluorobenz-2-oxa-1,3-diazole-4-sulphonamide (ABD) group using an MTS linker. The ABD chromophore is extremely sensitive to its environment and its electrical dipole midpoint is only 6 Å from the thiol group. The FRET signals between ABD–V363C and DPA in the membrane is monophasic, but reversed in direction relative to sulphorhodamine at the same position (Fig. 2c). This indicates that the ABD chromophore, unlike sulpho-

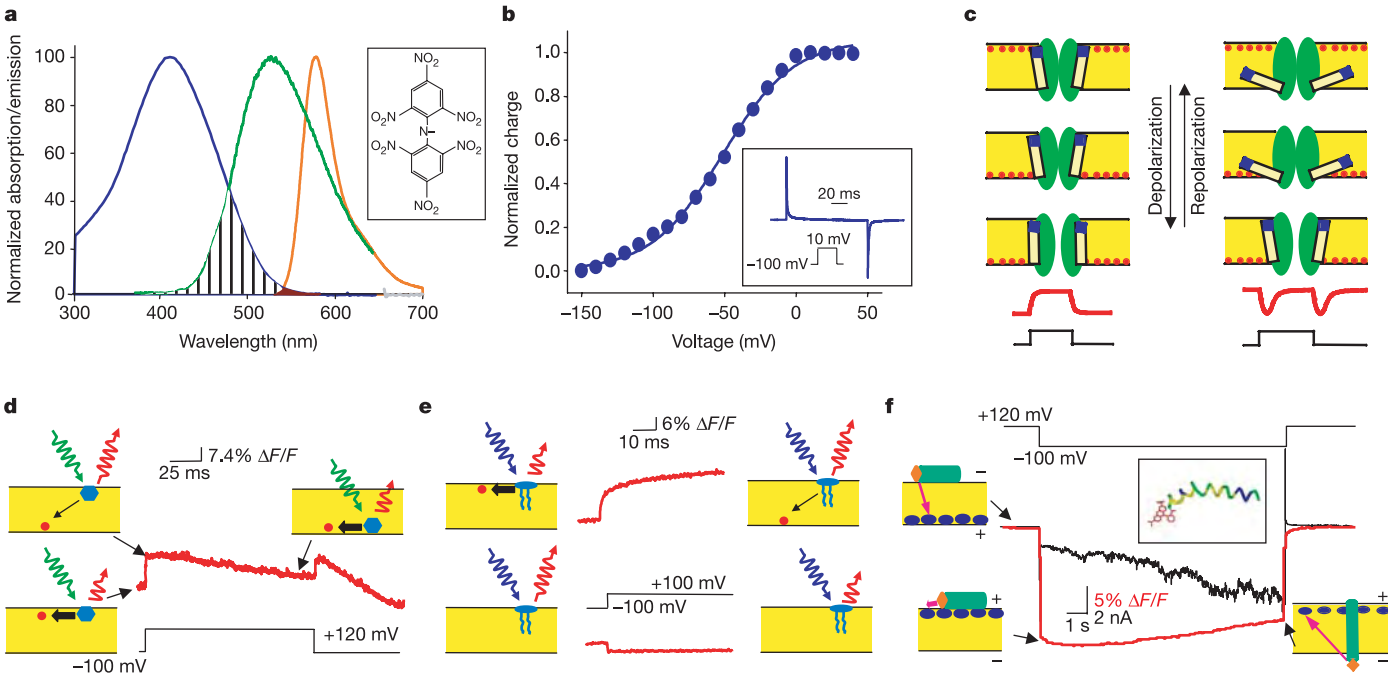


Figure 1 | Voltage-dependent FRET between dipicrylamine and model membrane probes. a, Dipicrylamine (structure shown in inset) absorbance spectra (blue) along with tetramethylrhodamine (orange) and ABD (green) emission spectra. **b**, Charge–voltage relationship of dipicrylamine with a representative gating-current trace (inset). **c**, Schematic representation of molecular movements of dipicrylamine (red circles) and labelled (blue) S4 segments (yellow rectangles) in the channel (green) embedded in the membrane (yellow) based on contrasting models of voltage-gating (left panel, no transmembrane movement; right panel, large transmembrane movement). The time course of the FRET signal (red traces) and the voltage pulses (black traces) in each gating model are shown at the bottom. **d**, The time trace of voltage-dependent FRET (red trace) between the membrane-translocating probe oxonol (blue hexagons) and dipicrylamine (red circles). **e**, Upper panel, the voltage-dependent FRET signal (red trace) between di-8-ANEPPS (blue) and dipicrylamine (red circles). Lower panel, the intrinsic di-8-ANEPPS electrochromic signal (red trace). **f**, The time trace of FRET signal (red trace) between N-TMR-melittin (green cylinder) and dipicrylamine (blue ovals) is overlaid on the current recordings (black trace). Inset, structure of N–TMR–melittin.

rhodamine, is located closer to the outer leaflet, as would be expected from their structural differences. However, like sulphorhodamine, the FRET signal of the ABD–DPA pair is monophasic, confirming that there is no evidence of transmembrane crossing. Our measurements indicate that the residues in the S3b–S4 loop and the S4 segment do not translocate across the bilayer during the gating process.

The lack of significant transmembrane displacement of the S4 segment strongly constrains the type of molecular motions that are able to generate a gating charge of 13*e*. To illustrate the counter-intuitive concept of voltage gating in the context of limited transmembrane movement, we have constructed a molecular model of the Shaker K^+ channel using the X-ray structure of the bacterial channel KvAP (voltage-dependent K^+ channel from *Aeropyrum pernix*) together with the available experimental data (Fig. 3). The channel was modelled conservatively by preserving the overall packing of all α-helical elements of the voltage sensor. The structural details are obviously uncertain owing to the lack of experimental data; however, the most important features of the model are strongly supported by the currently available information. In particular, the R362 residue in S4 is in proximity to the A419 residue in S5 near the extracellular side of the Shaker channel in the activated state[19]. In order to generate 13*e* gating charges upon hyperpolarization in the context of limited inward S4 transmembrane movement, it is necessary for the intracellular electrical potential to extend near the extracellular surface when the channel is in the resting state; presumably through a high dielectric aqueous crevice. This notion—that the transmembrane potential is 'focused' near the extracellular side in the region of the first four gating charges in S4 through an intracellular aqueous cavity—is supported by experimental evidence of a proton-conducting pore in the R362H mutant of the Shaker channel in the closed state[12], the strong dependence of gating charge quantity on intracellular ionic strength[20], and the measurement of an amplified membrane electric field near the second gating charge residue[15].

Figure 3a shows the models of the Shaker K^+ channel in the open and closed state where the out-of-plane subunits have been removed

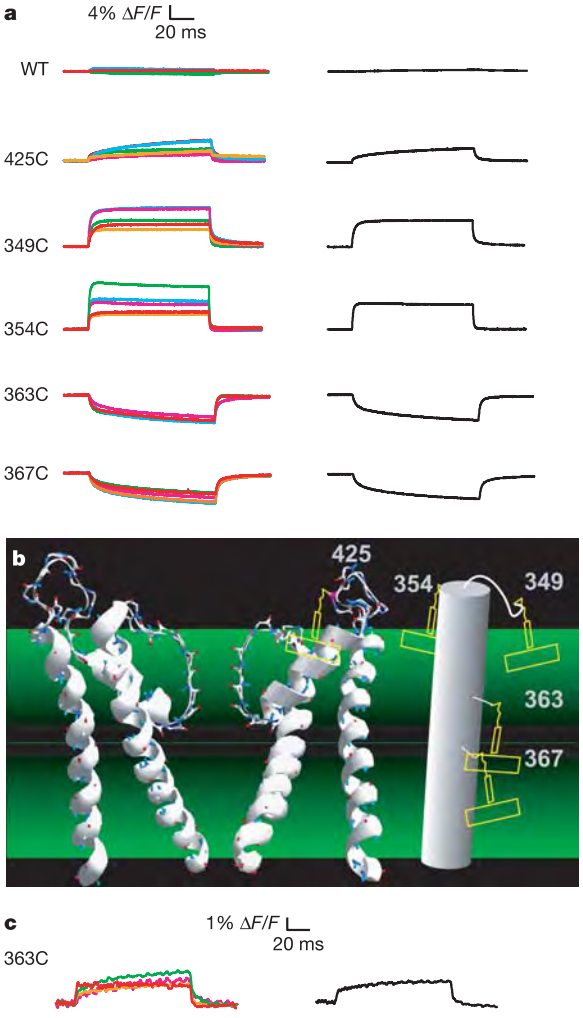


Figure 2 | FRET measurements between labelled Shaker potassium channel and dipicrylamine. a, Time trace of voltage-dependent FRET change of sulphorhodamine measured from different positions in the Shaker channel in response to a voltage pulse to +50 mV from a pre-pulse and a post-pulse of −120 mV. Left, each colour corresponds to a measurement in an individual oocyte; right, average of voltage-dependent FRET from all oocytes. **b**, Cartoon representation of the labelled positions in the Shaker S4 segment and the pore turret. **c**, Time trace of voltage-dependent FRET measured from position V363C labelled with ABD–MTS in response to a pulse to +50 mV from a pre-pulse and post-pulse of − 120 mV. (Left and right panels have the same meaning as in **a**.) WT, wild type.

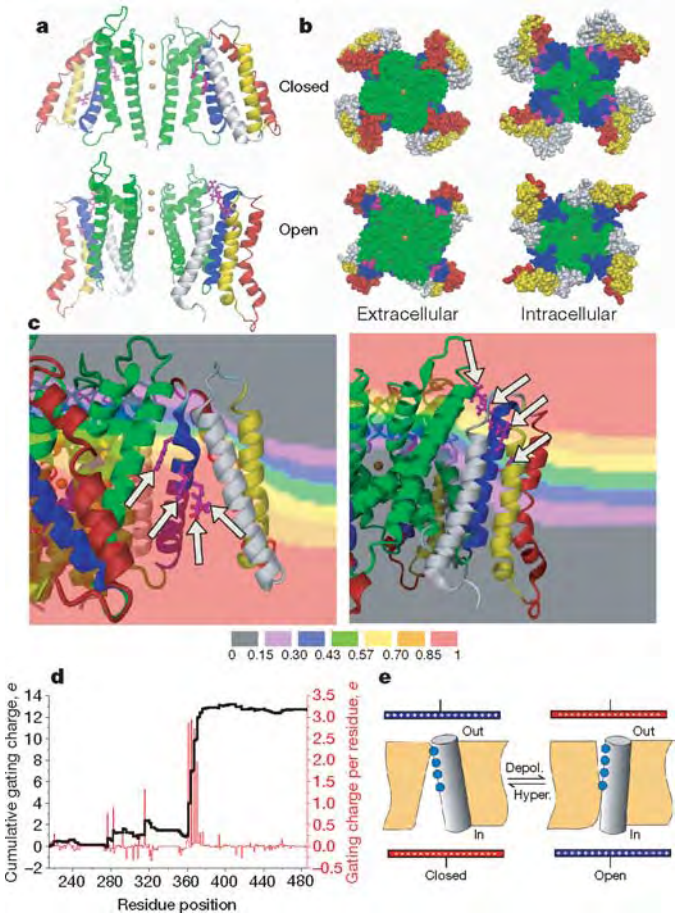


Figure 3 | Molecular model of voltage-gating in Shaker potassium channel. a, Axial view showing two of the four subunits of the Shaker potassium channel embedded in the lipid bilayer in a closed (upper panel) and an open state (lower panel). S1, white; S2, yellow; S3, red; S4, blue; pore region, green; charge-carrying arginine residues, magenta. **b**, Intracellular and extracellular views of the space-filled model of the Shaker potassium channel. **c**, Close-up view of voltage-sensor charges in the closed (left panel) and open state (right panel) overlaid with the isopotential surface. The colour bars (below) correspond to dimensionless fractions of the total transmembrane potential. Arrows highlight the arginine residues. **d**, Gating charge per residue (red) and cumulative gating charge (black) transported across the membrane field as the channel goes from a closed to an open state as a function of residue position. **e**, Schematic of the transporter model of voltage-gating where depolarization alters the accessibility of the charged residues from internal to external water-filled crevices.

for clarity. Figure 3b shows the intra- and extracellular views of the structure, where it is apparent that the arginine residues are visible from the outside in the open state and from the inside in the closed state. The molecular models of the open and closed states of the channel allow for the computation of the gating charge by solving the modified Poisson–Boltzmann equation[21]. A close-up view of the voltage sensor and the iso-potential lines (Fig. 3c) shows that the voltage-sensing arginine residues move across most of the focused electric field by re-orienting the side chains with very little translocation of the voltage-sensor backbone (Fig. 3c). The cumulative charge that is translocated in the Shaker channel, computed as a function of residue position, is shown in Fig. 3d. The structural model is able to quantitatively reproduce the displacement of 12–13 elementary charges contributed by the first four charges on S4, as determined experimentally in the Shaker channel[1,2], without the requirement for large transmembrane movements of S4. In this model, the S4 backbone undergoes a 45° change in tilt upon gating but translates only a short distance (less than 2 Å) perpendicular to the membrane, while the charges are relocated from a deep internally facing aqueous crevice in the closed state to an external crevice when it opens (Fig. 3). In the open state, the intracellular crevice that is apparent in the closed state essentially disappears, as the S1–S4 segments pack into a tight helical bundle. The residues near the S3–S4 loop (I325–R371) that form the structural helix–turn–helix paddle motif in the X-ray structure of KvAP move together as a stable dynamic unit (a difference of only 3 Å root mean square deviation, r.m.s.d., for the Cα between the modelled open and closed state), with the carboxy terminus of S3 serving to electrically insulate S4 from the extracellular side in the closed state.

These results suggest that the operation of the voltage sensor could be qualitatively similar to a transporter, in which the accessibility to a binding site alternates between the inside and outside with each transport cycle[22]. In the case of voltage-gated ion channels, the charged arginine residues are analogous to 'tethered substrates' whose accessibility alternates as a function of membrane voltage (Fig. 3e). Thus, an evolutionarily conserved transporter-like mechanism may be sufficient to move the large amount of charge across the electric field that is required for the steep voltage dependence of voltage-gated channels, without translocating the S4 segment across the membrane.

Note added in proof: The atomic structure of Kv1.2 determined by X-ray crystallography appeared online while this publication was in press[31]. Despite some differences, the current model of Shaker in the open state is broadly consistent with the structure of Kv1.2, and the calculated gating charge is not expected to change (see Supplementary Discussion).

METHODS

Molecular biology and labelling. All the Shaker-channel constructs used in this work were generated in an inactivation-removed (Δ6–46 ShH4) non-conducting (W434F) cysteine-less (C96S/C245V/C286V/C301S/C308S/C462A/C505S) channel background[23]. Site-directed mutagenesis, *in vitro* transcription, complementary RNA injections and labellings were performed as described previously[24] with small modifications. Labelling with methanethiosulphonate–(MTS)–sulphorhodamine (Toronto Research Chemicals) and 7-fluorobenz-2-oxa-1,3-diazole-4-sulphonamide (ABD)–MTS (Toronto Research Chemicals) was performed for 5 min on ice, whereas *N*-(3-sulphopropyl)-4-(2-(6-*N*,*N*-dioctyl-beta-aminonaphthalene)ethenyl)pyridinium (di-8-ANEPPS; Molecular Probes) was at room temperature (20–22 °C). During recording, the external solution was perfused with 20 μM of DPA. FRET measurements of oxonol–DPA were in presence of 2 μM bis-(1,3-diethylthiobarbituric acid)trimethine oxonol (Molecular Probes).

FRET measurements in cut-open oocytes. Time-resolved FRET measurements under voltage-clamp conditions between labelled Shaker channels and DPA were carried out in a modified cut-open oocyte set-up using customized acquisition and analysis systems described previously[25]. Sulphorhodamine and oxonol fluorescence was measured using standard tetramethylrhodamine (TMR) optics[24], whereas di-8-ANEPPS fluorescence measurements used a 425 nm/50 nm band pass (BP) excitation filter. ABD fluorescence was excited with a filter cube consisting of a 425 nm/50 nm BP excitation filter, a 475 nm long pass dichroic filter with no emission filters. Voltage-dependent FRET signals from the expressed channels were generated by subtracting the average FRET signal obtained in uninjected oocytes within the same batch. All FRET changes ($\Delta F/F$) were normalized to the background fluorescence before adding DPA. The R_o in FRET measurements is defined as the distance corresponding to 50% efficiency for a specific donor–acceptor pair. The calculated R_o for the sulphorhodamine–dipicrylamine pair was 22 Å and the ABD–dipicrylamine pair was 23 Å. The Q_D (the quantum yield of the donor) of ABD in ethanol was 0.18.

FRET measurements in planar lipid bilayers. Bilayer fluorescence was recorded using a horizontal bilayer set-up mounted on an inverted microscope. Fluorophores were excited with de-focused laser light ($\lambda = 514$ nm) and fluorescence (emission filter 610 nm/75 nm BP) recorded with a Cascade charge-coupled device (CCD) camera (Roper Scientific). Bilayers were formed from a 3:1mixture of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine:1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPC:POPE; Avanti Polar Lipids) in decane (25 mg ml^{-1}). Melittin labelled at the N-terminus with tetramethylrhodamine (N–TMR–melittin; Alpha Diagnostics International) was added to the top (outside) chamber (concentration, 10 μM). Voltage convention is inside potential minus outside potential.

Molecular modelling and Poisson–Boltzmann calculations. The Shaker model was constructed as described previously[19]. Briefly, the structure of the S1 through S4 crystal was docked to the pore region of the full crystal structure of KvAP (ref. 8) for the open state, and to the KcsA (ref. 26) for the S5–S6 pore structure of the closed state. To avoid ambiguous insertions in the homology modelling, the Shaker S3–S4 loop was shortened to eight residues[27]. The r.m.s.d. relative to the X-ray structure of the isolated voltage sensor of KvAP is 6 Å and 9 Å (Cα) for the open and closed states, respectively. Models were refined by energy minimization and molecular dynamics simulations using the program CHARMM (ref. 28). Gating charge was calculated from the structural models using the modified Poisson–Boltzmann voltage equation[21] implemented in the finite-difference PBEQ module of CHARMM. A dielectric of 2 was assigned to the protein, and the protein–solvent interface was determined using the set of optimized atomic Born radii[29]. The membrane was represented as a 24 Å slab of low dielectric 2, assembled by packing neutral spheres around the channel; all aqueous crevices of the pore and voltage sensor were represented as high dielectric regions of 80. A 1 Å grid was used. S4 is significantly exposed to the membrane (~20%), although the arginine residues are exposed to aqueous regions in agreement with electron paramagnetic resonance results in KvAP (ref. 30). The volume of the intracellular crevice of the closed state, which is partly bounded by the membrane, corresponds to 30–40 water molecules, in accord with experimental estimates[20]. Figure 3c was created using the DINO3D software (www.dino3d.org).

Received 6 April; accepted 6 June 2005.

1. Aggarwal, S. K. & MacKinnon, R. Contribution of the S4 segment to gating charge in the Shaker K^+ channel. *Neuron* **16,** 1169–1177 (1996).
2. Seoh, S. A., Sigg, D., Papazian, D. M. & Bezanilla, F. Voltage-sensing residues in the S2 and S4 segments of the Shaker K^+ channel. *Neuron* **16,** 1159–1167 (1996).
3. Bezanilla, F. The voltage sensor in voltage-dependent ion channels. *Physiol. Rev.* **80,** 555–592 (2000).
4. Ahern, C. A. & Horn, R. Stirring up controversy with a voltage sensor paddle. *Trends Neurosci.* **27,** 303–307 (2004).
5. Jiang, Y., Ruta, V., Chen, J., Lee, A. & MacKinnon, R. The principle of gating charge movement in a voltage-dependent K(+) channel. *Nature* **423,** 42–48 (2003).
6. Guy, H. R. & Seetharamulu, P. Molecular model of the action potential sodium channel. *Proc. Natl Acad. Sci. USA* **83,** 508–512 (1986).
7. Catterall, W. A. Molecular properties of voltage-sensitive sodium channels. *Annu. Rev. Biochem.* **55,** 953–985 (1986).
8. Jiang, Y. *et al.* X-ray structure of a voltage-dependent K(+) channel. *Nature* **423,** 33–41 (2003).
9. Yang, N., George, A. L. Jr & Horn, R. Molecular basis of charge movement in voltage-gated sodium channels. *Neuron* **16,** 113–122 (1996).
10. Larsson, H. P., Baker, O. S., Dhillon, D. S. & Isacoff, E. Y. Transmembrane movement of the shaker K^+ channel S4. *Neuron* **16,** 387–397 (1996).
11. Starace, D. M., Stefani, E. & Bezanilla, F. Voltage-dependent proton transport by the voltage sensor of the Shaker K^+ channel. *Neuron* **19,** 1319–1327 (1997).
12. Starace, D. M. & Bezanilla, F. A proton pore in a potassium channel voltage sensor reveals a focused electric field. *Nature* **427,** 548–553 (2004).
13. Cha, A., Snyder, G. E., Selvin, P. R. & Bezanilla, F. Atomic scale movement of the voltage-sensing region in a potassium channel measured via spectroscopy. *Nature* **402,** 809–813 (1999).
14. Glauner, K. S., Mannuzzu, L. M., Gandhi, C. S. & Isacoff, E. Y. Spectroscopic

mapping of voltage sensor movement in the Shaker potassium channel. *Nature* **402,** 813–817 (1999).
15. Asamoah, O. K., Wuskell, J. P., Loew, L. M. & Bezanilla, F. A fluorometric approach to local electric field measurements in a voltage-gated ion channel. *Neuron* **37,** 85–97 (2003).
16. Gonzalez, J. E. & Tsien, R. Y. Voltage sensing by fluorescence resonance energy transfer in single cells. *Biophys. J.* **69,** 1272–1280 (1995).
17. Tosteson, M. T. & Tosteson, D. C. The sting. Melittin forms channels in lipid bilayers. *Biophys. J.* **36,** 109–116 (1981).
18. Kempf, C. *et al.* Voltage-dependent trans-bilayer orientation of melittin. *J. Biol. Chem.* **257,** 2469–2476 (1982).
19. Laine, M. *et al.* Atomic proximity between S4 segment and pore domain in Shaker potassium channels. *Neuron* **39,** 467–481 (2003).
20. Islas, L. D. & Sigworth, F. J. Electrostatics and the gating pore of Shaker potassium channels. *J. Gen. Physiol.* **117,** 69–89 (2001).
21. Roux, B. Influence of the membrane potential on the free energy of an intrinsic protein. *Biophys. J.* **73,** 2980–2989 (1997).
22. Abramson, J. *et al.* Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301,** 610–615 (2003).
23. Boland, L. M., Jurman, M. E. & Yellen, G. Cysteines in the Shaker K^+ channel are not essential for channel activity or zinc modulation. *Biophys. J.* **66,** 694–699 (1994).
24. Chanda, B. & Bezanilla, F. Tracking voltage-dependent conformational changes in skeletal muscle sodium channel during activation. *J. Gen. Physiol.* **120,** 629–645 (2002).
25. Cha, A. & Bezanilla, F. Structural implications of fluorescence quenching in the Shaker K^+ channel. *J. Gen. Physiol.* **112,** 391–408 (1998).
26. Zhou, Y., Morais-Cabral, J. H., Kaufman, A. & MacKinnon, R. Chemistry of ion coordination and hydration revealed by a K^+ channel-Fab complex at 2.0Å resolution. *Nature* **414,** 43–48 (2001).
27. Gonzalez, C., Rosenman, E., Bezanilla, F., Alvarez, O. & Latorre, R. Periodic perturbations in Shaker K^+ channel gating kinetics by deletions in the S3–S4 linker. *Proc. Natl Acad. Sci. USA* **98,** 9617–9623 (2001).
28. Brooks, B. R. *et al.* CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4,** 187–217 (1983).
29. Nina, M., Beglov, D. & Roux, B. Atomic Born radii for continuum electrostatic calculations based on molecular dynamics free energy simulations. *J. Phys. Chem. B* **101,** 5239–5248 (1997).
30. Cuello, L. G., Cortes, D. M. & Perozo, E. Molecular architecture of the KvAP voltage-dependent K^+ channel in a lipid bilayer. *Science* **306,** 491–495 (2004).
31. Long, S. B., Campbell, E. B. & MacKinnon, R. Crystal structure of a mammalian voltage-dependent *Shaker* family K^+ channel. *Science* 7 July 2005 (doi:10.1126/science.1116269).

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

Acknowledgements We thank M. J. Hahn for technical assistance, M. Holmgren for the cysteine-less Shaker clone, W. Hubell for the gift of dipicrylamine and the members of Bezanilla and Correa laboratories for their comments. This work was supported by funds from an AHA postdoctoral fellowship to B.C., NRSA funding to O.K.A, DFG funding to R.B. and an NIH grant to F.B.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.B. (fbezanil@ucla.edu)

LETTERS

Voltage-sensor activation with a tarantula toxin as cargo

L. Revell Phillips[1]*, Mirela Milescu[1]*, Yingying Li-Smerin[1], Joseph A. Mindell[2], Jae Il Kim[3] & Kenton J. Swartz[1]

The opening and closing of voltage-activated Na^+, Ca^{2+} and K^+ (Kv) channels underlies electrical and chemical signalling throughout biology, yet the structural basis of voltage sensing is unknown. Hanatoxin is a tarantula toxin that inhibits Kv channels by binding to voltage-sensor paddles[1–5], crucial helix-turn-helix motifs within the voltage-sensing domains that are composed of S3b and S4 helices[6]. The active surface of the toxin is amphipathic[7,8], and related toxins have been shown to partition into membranes[9–12], raising the possibility that the toxin is concentrated in the membrane and interacts only weakly and transiently with the voltage sensors. Here we examine the kinetics and state dependence of the toxin–channel interaction and the physical location of the toxin in the membrane. We find that hanatoxin forms a strong and stable complex with the voltage sensors, far outlasting fluctuations of the voltage sensors between resting (closed) conformations at negative voltages and activated (open) conformations at positive voltages. Toxin affinity is reduced by voltage-sensor activation, explaining why the toxin stabilizes the resting conformation. We also find that when hanatoxin partitions into membranes it is localized to an interfacial region, with Trp 30 positioned about 8.5 Å from the centre of the bilayer. These results demonstrate that voltage-sensor paddles activate with a toxin as cargo, and suggest that the paddles traverse no more than the outer half of the bilayer during activation.

The scheme below describes a scenario wherein hanatoxin (Fig. 1a) partitions from the extracellular aqueous phase (Tx_{aq}) into the membrane (Tx_{mem}) and then binds the channel (Ch) to form a toxin–channel complex (Tx–Ch). Kv channels are tetramers and four toxins bind to each channel[1–3], one per voltage sensor.

$$Tx_{aq} \underset{k_{out}}{\overset{k_{in}}{\rightleftharpoons}} \left(Tx_{mem} + Ch \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} Tx\text{–}Ch \right)^4$$

If partitioning serves to concentrate the toxin near its target, the protein–protein interaction could have low affinity (for example, in the millimolar range) and toxin unbinding (k_{off}) could be as fast as channel gating (milliseconds)[9]. In this case, the conformational changes that underlie voltage-dependent gating in the presence of toxin (Fig. 1b) might only occur when the toxin is not bound. The very slow recovery of channel activity that is observed after removal of the toxin from the aqueous solution (~1,000 s for wild type; Fig. 1f) would then necessarily arise from slow departitioning (low k_{out}) rather than slow unbinding (low k_{off})[1]. Thus, membrane partitioning has fundamental implications for the inhibitory mechanisms of tarantula toxins, and for movements of the voltage sensors.

To evaluate the contribution of membrane partitioning we examined the roles of departitioning and unbinding in the slow recovery kinetics, taking advantage of mutations to F274 within the voltage-sensor paddle that weaken toxin binding affinity[4,5]. When the relationship between $[Tx_{aq}]$ and occupancy of the Kv2.1 channel is examined, the F274A mutation increases the apparent equilibrium dissociation constant (K_d) for hanatoxin by 28-fold (Fig. 1c). Simulations of the above scheme (Fig. 1d, e) demonstrate that slow recovery can result either from slow departitioning ($k_{out} = 5 \times 10^{-3}\,s^{-1}$) in the context of low affinity ($K_d = 1$ mM) and fast unbinding ($k_{off} = 10\,s^{-1}$), or slow unbinding ($k_{off} = 4 \times 10^{-3}\,s^{-1}$) in the context of high affinity ($K_d = 100$ nM) and fast departitioning ($k_{out} = 0.04\,s^{-1}$). Notably, weakening the binding affinity by 28-fold has minimal effects on recovery in the case of slow departitioning (Fig. 1d, grey and blue-dashed curves), but marked effects on recovery in the case of slow unbinding (Fig. 1e, grey and blue curves). We observe that the F274A mutation speeds the kinetics of recovery to an extent similar to that predicted by slow toxin unbinding (Fig. 1f). Similar effects on recovery are seen with other mutants with weakened toxin affinity (for example, F274G, F274K and F274R, and E277A, E277G, E277W, E277K and E277R; data not shown), and comparable effects of Na^+ channel mutations on α-scorpion toxin dissociation have been reported[13]. Because the partitioning step should be unaffected by channel mutations, k_{out} must be $0.04\,s^{-1}$ or higher to explain the fast recovery seen in the F274A mutant ($\tau = 25$ s; Fig. 1f). In the context of this relatively fast departitioning, the slow recovery observed in the wild-type channel suggests that k_{off} is low ($\sim 4 \times 10^{-3}\,s^{-1}$) (ref. 2) and the binding affinity is high (~100 nM; Fig. 1c). An alternative possibility that k_{on} is very fast is ruled out by further experiments (see below) and theoretical considerations (see Methods).

We next explored how the conformation of the voltage sensor affects the kinetics of toxin binding. Hanatoxin stabilizes the resting conformation of the voltage sensor[1], raising the possibility that the toxin binds less tightly to the activated conformation, as observed with other toxins[14–16]. The above recovery experiments (Fig. 1f) are dominated by toxin unbinding from resting voltage sensors because the cell is held at −80 mV, where this conformation predominates, and toxin occupancy of the channel is assayed using short depolarizations to −10 mV. In the context of a low k_{off} for the wild-type channel, the fast (millisecond timescale) opening observed during strong depolarizations (Fig. 1b) corresponds to the opening of toxin-bound channels. If the toxin binds less tightly to the activated voltage sensor, we might expect to see unbinding during strong depolarizations in F274A, which exhibits weaker binding and faster recovery (Fig. 1c, f). To explore this possibility we used a voltage protocol consisting of two weak depolarizing steps to −10 mV (insufficient to open toxin-bound channels; Fig. 1b), separated by a strong depolarization to +60 mV (Fig. 2a). In the wild-type channel, strong depolarizations do not alter the extent of inhibition observed

[1]Molecular Physiology and Biophysics Section, and [2]Membrane Transport Biophysics Unit, Porter Neuroscience Research Center, National Institute of Neurological Disorders and Stroke, National Institutes of Health, 35 Convent Drive, MSC 3701 Bethesda, Maryland 20892-3701, USA. [3]Department of Life Science, Gwangju Institute of Science and Technology, Gwangju 500-712, Korea.
*These authors contributed equally to this work.

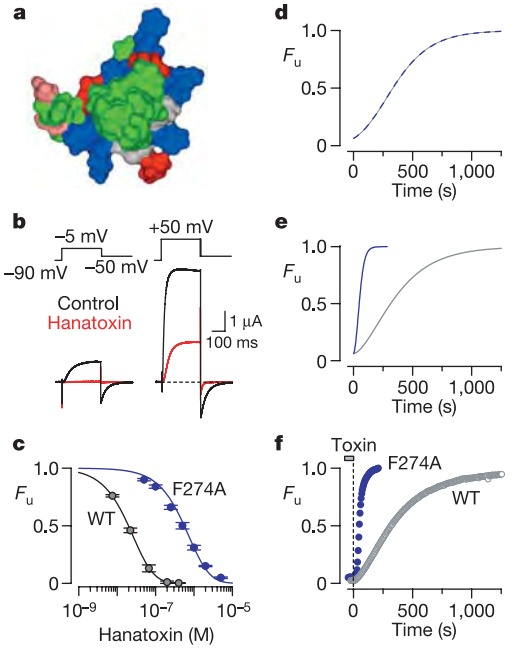


Figure 1 | Voltage-sensor paddle mutations affecting hanatoxin affinity and recovery kinetics. **a**, Surface rendering of hanatoxin (Protein Data Bank 1D1H)[7] with hydrophobic residues coloured green, basic residues blue, acidic residues red and Ser/Thr pink. **b**, Voltage activation of wild-type Kv2.1 channels and inhibition by hanatoxin. **c**, Concentration dependence for fractional occupancy of the Kv2.1 channel by hanatoxin. The fraction of unbound channels (F_u) was estimated from tail currents elicited after weak depolarizations (−20 to −10 mV)[2,4,8]. Data are mean ± s.e.m. (n = 3–7). Solid lines are fits of $F_u = (1 - P)^4$ where $P = [\text{toxin}]/([\text{toxin}] + K_d)$ and K_d = 101 nM for wild type and K_d = 2.8 µM for F274A. **d**, Simulated recovery kinetics for low-affinity toxin binding in the context of slow departitioning ($k_{out} = 5 \times 10^{-3}\,s^{-1}$) and fast unbinding. Grey curve: K_d = 1 mM, $k_{off} = 10\,s^{-1}$, $k_{on} = 10^4\,M^{-1}\,s^{-1}$. Blue dashed curve: K_d = 28 mM, $k_{off} = 280\,s^{-1}$, $k_{on} = 10^4\,M^{-1}\,s^{-1}$. Indistinguishable results are obtained when K_d is increased by lowering k_{on} 28-fold. [Tx_{mem}] was initially set to the K_d. **e**, Simulated recovery kinetics for high-affinity toxin binding in the context of slow unbinding and fast departitioning ($k_{out} = 0.04\,s^{-1}$). Grey curve: K_d = 100 nM, $k_{off} = 4 \times 10^{-3}\,s^{-1}$, $k_{on} = 4 \times 10^4\,M^{-1}\,s^{-1}$. Blue curve: K_d = 2.8 µM, $k_{off} = 0.112\,s^{-1}$, $k_{on} = 4 \times 10^4\,M^{-1}\,s^{-1}$. Speeding up of recovery is not observed when K_d is increased by lowering k_{on} 28-fold. **f**, Recovery kinetics for wild-type and F274A channels after removal of the toxin from the recording chamber. Test depolarizations (400 ms) to −10 mV were elicited from a holding voltage of −80 mV every 5 s. Best fit of single exponential functions to the final 70% of recovery yields τ = 25 ± 1 s (mean ± s.e.m.; n = 4) for F274A and 390 ± 50 s (n = 3) for wild type. Concentrations of toxin were chosen to produce equivalent inhibition at −10 mV (wild type, 200 nM; F274A, 5 µM)[2].

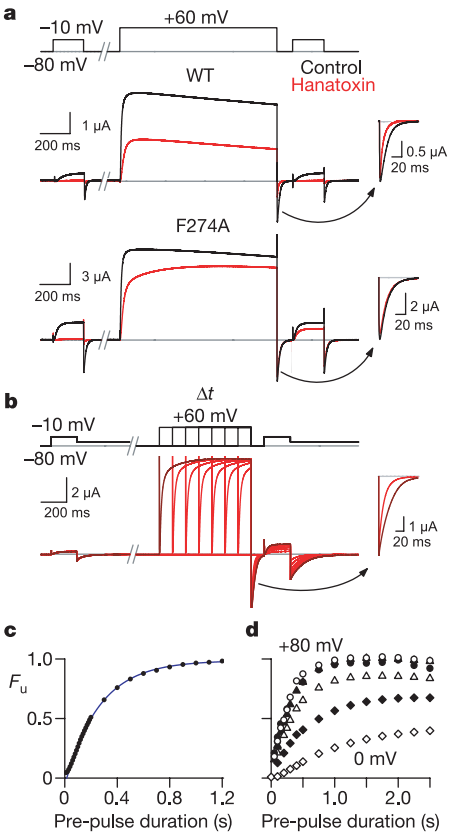


Figure 2 | Voltage-dependent unbinding of hanatoxin from F274A channels. **a**, Protocols for distinguishing between opening of toxin-bound channels in wild type and rapid dissociation in F274A. The break in the axis is 1.3 s in duration. Toxin concentration was 2 µM in both wild-type and F274A channels. **b**, Relationship between duration of the strong depolarization and relief of toxin inhibition during subsequent weak depolarizations. All traces were elicited in the presence of 2 µM toxin. Inset to right shows tail currents after either a 100-ms (red trace) or 700-ms (maroon trace) strong depolarization. The break in the axis is 900 ms in duration. Delays of 30 s between consecutive pulses were used to allow for rebinding of toxin. **c**, F_u, estimated from inhibition of currents elicited by the second weak depolarization, plotted against duration of strong depolarizations to +80 mV. Blue curve, exponential fit to the final 70% of recovery (τ = 261 ms). **d**, F_u plotted against pre-pulse duration for depolarizations to +80 mV (filled circle), +70 mV (open circles), +60 mV (filled triangles), +40 mV (open triangles), +20 mV (filled diamonds) and 0 mV (open diamonds).

during the weak depolarizations (Fig. 2a), suggesting that toxin occupancy of the channel remains unchanged. In contrast, strong depolarizations in the F274A channel result in a marked change in the extent of inhibition observed during the two weak depolarizations (Fig. 2a), pointing to a large change in toxin occupancy of the channel. The time dependence for this change in the extent of inhibition (Fig. 2b) correlates with a slow current relaxation at +60 mV (Fig. 2a; F274A), reflecting unbinding on the timescale of several hundred milliseconds (Fig. 2b, c). The interpretation of these phenomena as toxin unbinding is supported by an inspection of deactivation (closing) upon repolarization from +60 mV to −80 mV, which identifies toxin-bound channels by their faster deactivation kinetics (Fig. 2a, b, insets). The voltage dependence for the kinetics of depolarization-induced unbinding (Fig. 2d) is similar to that observed for gating charge movement in the presence of the toxin[1], suggesting that the toxin has higher affinity for the resting conformation of the voltage sensor. The weakening of toxin binding affinity by voltage-sensor activation provides a thermodynamic explanation for toxin stabilization of the resting sensor—the toxin binds this conformation more tightly and thus biases the equilibrium towards the resting state. These results are consistent with the results of the recovery experiments (Fig. 1f) in that they distinguish the kinetics of opening for toxin-bound channels (milliseconds for the wild-type and F274A channels) from the kinetics of toxin dissociation (hundreds of milliseconds for the F274A channel at +60 mV, whereas even at this voltage it takes >5 s for the wild-type channel).

The unbinding of hanatoxin observed during strong depolarizations in the F274A channel allows an assessment of the kinetics of the toxin-binding step without contributions from the partitioning step, because the membrane is already equilibrated with the toxin. Although it is unknown whether the toxin unbinds into the membrane or into the aqueous solution, it is unlikely that toxin unbinding influences the partitioning equilibrium because [Tx_{mem}] >> [Ch] (see Methods). We examined the kinetics of the binding step by varying the interval between a strong and subsequent weak depolarization and measuring how long it takes the toxin to rebind (Fig. 3c).

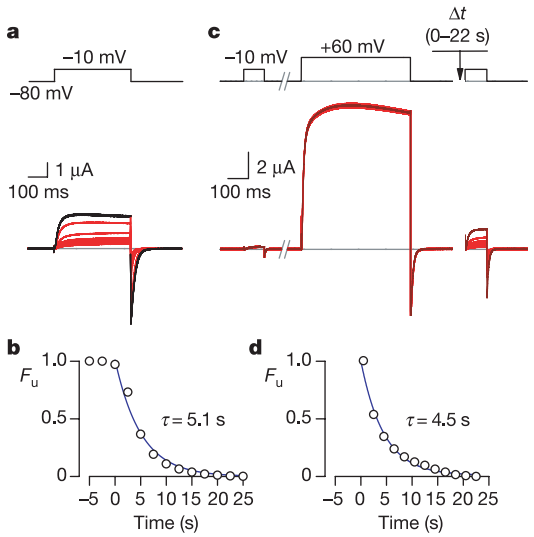


Figure 3 | Rebinding of hanatoxin to F274A channels. **a**, Onset of inhibition when hanatoxin (2 μM; red traces) is added to the recording chamber (pulse frequency is 0.4 Hz). **b**, F_u plotted versus time after the application of hanatoxin (2 μM). The solid line is an exponential fit with $\tau = 5.1$ s ($\tau = 5.3 \pm 0.2$ s, mean ± s.e.m., $n = 6$). **c**, Rebinding of hanatoxin after dissociation induced by a strong depolarization to +60 mV. The interval between the strong depolarization (+60 mV) and subsequent weak depolarization (−10 mV) was varied between 0.5 and 22.5 s in 2-s increments. **d**, F_u plotted versus interval duration. The solid line is an exponential fit with $\tau = 4.5$ s ($\tau = 3 \pm 0.1$ s, mean ± s.e.m., $n = 6$).

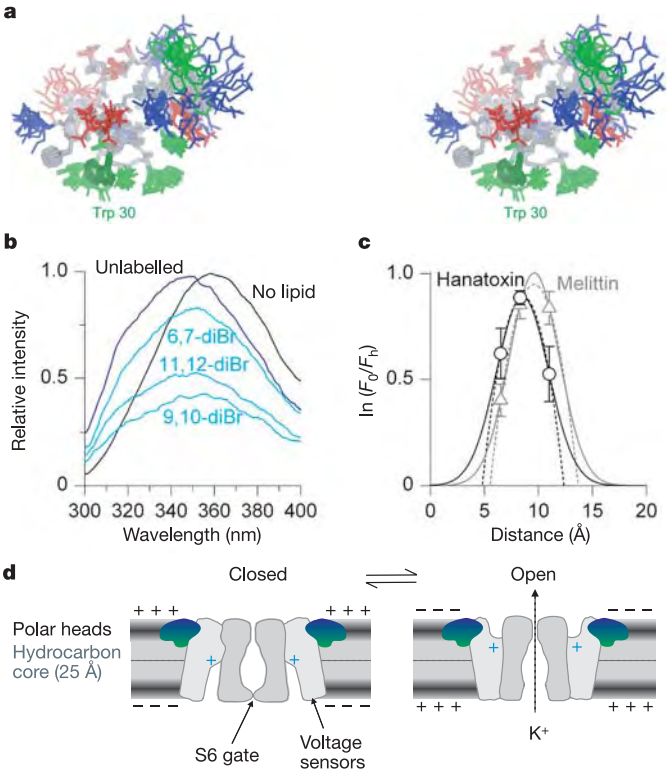


Figure 4 | Location of hanatoxin in lipid membranes. **a**, Stereo pairs of 21 converged structures of hanatoxin[7] (Protein Data Bank 1D1H). Side chain colours are: green, hydrophobic; blue, basic; red, acidic; pink, Ser/Thr; grey, other side chains and backbone atoms. Trp 30 is dark green. **b**, Fluorescence emission spectra for hanatoxin in the absence and presence of lipid vesicles (1.3 mM) containing unlabelled or brominated (diBr) lipids. **c**, Analysis of the depth-dependent quenching profiles for hanatoxin (circles) and melittin (triangles). Data show mean ± s.e.m. ($n = 3$–4). Fractional quenching (F_0/F_h) by brominated lipids plotted against average distance of bromine atoms from the centre of the bilayer[17]. Hanatoxin: $h_m = 8.5$ Å, $S = 5.3$ and $\sigma = 2.4$ Å using DA (solid line), and $h_m = 8.6$ Å, $R_c = 3.7$ Å and $f = 0.03$ using PM (dashed line). Melittin: $h_m = 9.6$ Å, $S = 5.8$ and $\sigma = 2.3$ Å using DA, and $h_m = 9.6$ Å, $R_c = 4$ Å and $f = 0.03$ using PM. **d**, Illustration of toxin interaction with voltage sensors in the outer half of the bilayer, near the interfacial region between polar head groups and the hydrophobic core.

At 2 μM [Tx_{aq}], rebinding occurs with $\tau = 4.5$ s (Fig. 3c, d), a value that is very similar to the onset of inhibition ($\tau = 5.1$ s) when the toxin is first added to the chamber for the same cell (Fig. 3a, b). The slow kinetics of rebinding in this experiment are incompatible with the kinetics of the binding step being fast.

The present results support the idea that the k_{off} for hanatoxin is low and that the binding affinity of the toxin is high. The long dwell time of the toxin on the resting voltage sensor, approaching 250 s for the wild-type channel (Fig. 1f)[2], is orders of magnitude slower than channel gating (on a millisecond timescale), strongly suggesting that the voltage sensors activate with the toxin continuously bound. These results rule out the possibility that membrane partitioning has an energetically dominant role[9], but leave open the possibility that partitioning is necessary for the toxin to reach the voltage sensor. Although we observe partitioning of hanatoxin into model membranes (see below), the strength of partitioning into physiological membranes remains unknown. If we assume that membrane partitioning is necessary for hanatoxin to reach the resting paddle, where in the membrane might toxin binding occur? To address this question we localized the toxin within phospholipid bilayers using a depth-dependent fluorescence-quenching approach[17–19]. Hanatoxin contains one tryptophan residue (Trp 30; Fig. 4a) that exhibits a blue-shift in fluorescence emission (towards shorter wavelengths) upon partitioning into phospholipid vesicles (Fig. 4b). In addition, this Trp fluorescence is quenched by bromine atoms covalently attached at different positions along the lipid hydrocarbon chain (Fig. 4b). Bromine atoms near the middle of the hydrocarbon tail (9,10-diBr) are more effective at quenching the toxin's fluorescence than bromine atoms closer to (6,7-diBr) or farther from (11,12-diBr) the polar head groups (Fig. 4b). Analysis of the depth-dependent quenching places Trp 30 from the toxin's hydrophobic protrusion at a depth of about 8.5 Å from the centre of the lipid bilayer (Fig. 4c), allowing the polar residues on the toxin to maintain favourable interactions with lipid polar head groups. This orientation and localization of the toxin is consistent with the observation that binding of the toxin to the voltage sensor does not change the quantity of gating charge[1], and with the kinetics of unbinding becoming voltage independent above +60 mV (Fig. 2c), both of which argue against penetration of charged toxin residues (blue and red residues; Fig. 4a) into the hydrophobic phase of the membrane.

Our results demonstrate that Kv channels can activate with a toxin bound to each of the four voltage sensors (Fig. 4d). The localization of the toxin within interfacial regions of the bilayer, together with the preference of the toxin for the resting voltage sensor, implies that the voltage-sensor paddle motif is positioned within the outer half of the bilayer in its resting conformation. This picture is compatible with a lipid-exposed paddle[20,21]. However, it rules out a membrane-translocating mechanism for paddle movement during gating[20], because it precludes the charges of the S4 helix from traversing the inner half of the bilayer and its associated electric field. Rather, these constraints imply that charge movement occurs across a focused membrane electric field[22–26], perhaps created by an intracellular water-filled crevice[27], allowing charges to traverse the full membrane electric field even though the paddle motif is restricted to the outer half of the bilayer.

METHODS

Channel constructs and electrophysiological recordings. The Kv2.1 Δ7 channel construct was expressed in *Xenopus* oocytes[1] and studied using a two-electrode voltage clamp (OC-725C; Warner Instruments) with a bath solution (200-μl

chamber) containing (in mM): 50 RbCl, 50 NaCl, 5–10 HEPES, 1 $MgCl_2$, 0.3 $CaCl_2$, pH 7.6 with NaOH. Data were filtered at 2 kHz and digitized at 10 kHz. Leak and background conductances, identified by blocking the channel with agitoxin-2, have been subtracted[1]. Hanatoxin was purified from *Grammostola spatulata* venom (Spider Pharm) using high-performance liquid chromatography[28].

Modelling of toxin partitioning and binding to Kv channels. The scheme in the main text (expanded below) incorporates membrane partitioning and a channel containing four independent and identical toxin-binding sites.

$$\mathrm{Tx_{aq}} \underset{k_{out}}{\overset{k_{in}}{\rightleftharpoons}} \mathrm{Tx_{mem}} + \mathrm{Ch} \underset{k_{off}}{\overset{4k_{on}}{\rightleftharpoons}} \mathrm{Tx_1{-}Ch}$$

$$\mathrm{Tx_{mem}} + \mathrm{Tx_1{-}Ch} \underset{2k_{off}}{\overset{3k_{on}}{\rightleftharpoons}} \mathrm{Tx_2{-}Ch}$$

$$\mathrm{Tx_{mem}} + \mathrm{Tx_2{-}Ch} \underset{3k_{off}}{\overset{2k_{on}}{\rightleftharpoons}} \mathrm{Tx_3{-}Ch}$$

$$\mathrm{Tx_{mem}} + \mathrm{Tx_3{-}Ch} \underset{4k_{off}}{\overset{k_{on}}{\rightleftharpoons}} \mathrm{Tx_4{-}Ch}$$

To model this scheme, seven rate equations were derived, forming an over-defined system of first-order nonlinear differential equations numerically solved while explicitly constraining the total concentration of toxin and channels (Stiffr, Mathcad 2000). The reaction between toxin and channel (the binding step) is a second-order reaction that depends on both the concentration of toxin in the membrane [Tx_{mem}] and the channel concentration [Ch]. For clarity, we modelled the surface reaction in a thin (50 Å) three-dimensional slab with a surface area of 0.2 cm^2 (a reasonable value based on oocyte structure and capacitance measurements). From our measured current density and the unitary properties of the Kv channel we estimate 10^6 to 10^7 channels per cell (10^{-8} to 10^{-7} M). We varied [Tx_{mem}] from 10^{-7} M to 10^{-3} M (the range from [Tx_{aq}] in our experiments to the concentration reached if the mole fraction partition coefficient for the toxin is around 10^5, as reported for VSTx and GsMTx[9,10,12]). To simulate the time course of recovery after removal of Tx_{aq}, the model was first equilibrated using assigned rate constants, k_{in} was set to 0 and the concentration of unbound channels calculated as the system relaxed. Figure 1d, e shows two extremes where k_{out} and k_{off} dominate the slow recovery time course and the results of large changes in K_d. A slow recovery time course can also be achieved in the context of $k_{out} = 0.04\ s^{-1}$ and fast unbinding (for example, $k_{off} = 10\ s^{-1}$), but only if k_{on} is exceedingly high ($> 10^9\ M^{-1}\ s^{-1}$), K_d is $< 10^{-8}$ M and the initial [Tx_{mem}] is $> 10K_d$. This scenario requires high initial toxin occupancy of the channel to achieve slow recovery, which is not observed in the case of hanatoxin, where slow recovery occurs at far lower occupancy. In addition, K_d values in this range are too high to explain the experimentally observed concentration dependence for occupancy (Fig. 1c).

Toxin–membrane interactions. Large unilamellar vesicles were formed by extrusion of lipid mixtures (1:1 molar ratios) of POPG (1-palmitoyl-2-oleoyl-*sn*-glycero-3-[phospho-*rac*-(1-glycerol)]) and either POPC (1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine), 6,7-diBr, 9,10-diBr or 11,12-diBr (1-palmitoyl-2-stearoyl(dibromo)-*sn*-glycero-3-phosphocholine) (Avanti Polar Lipids). The effects of lipid vesicles on the tryptophan fluorescence of hanatoxin or melittin (2 μM each) was examined using three lipid concentrations (0.3, 0.6 and 1.3 mM) in a solution containing HEPES (10 mM) and EDTA (1 mM), pH 7. Emission spectra were recorded between 300 and 400 nm (5-nm band pass, 0° polarizer) using an excitation wavelength of 270 nm (5-nm band pass, 90° polarizer) (SPEX FluoroMax 3 spectrofluorometer) and corrected for vesicle scattering[19]. Fluorescence-quenching profiles were analysed using both distribution analysis (DA)[29] and the parallax method (PM)[30]. The average insertion depth of the tryptophan residue (h_m) was calculated based on best fits of the following equations to the data: DA, $\ln(F_o/F_h) = S/(\sigma\sqrt{2\pi})\exp(-(h-h_m)^2/2\sigma^2)$; PM, $\ln(F_o/F_h) = \pi Cf[R_c^2 - (h-h_m)^2]$ when $h - h_m < R_c$, and $\ln(F_o/F_h) = 0$ when $h - h_m \geq R_c$. F and F_0 are the fluorescence intensities in the absence and presence of the diBr quenchers, h is the depth of the quencher based on X-ray diffraction measurements[17], C is the quencher concentration, σ is dispersion, S is area, R_c is radius of quenching, and f is the fraction of the quencher in the tryptophan vicinity. Indistinguishable results were obtained with all lipid concentrations using both DA and PM analyses. We also analysed the quenching data allowing for distribution of the toxin in both leaflets and obtained similar results.

Received 7 January; accepted 2 June 2005.

1. Lee, H. C., Wang, J. M. & Swartz, K. J. Interaction between extracellular Hanatoxin and the resting conformation of the voltage-sensor paddle in Kv channels. *Neuron* **40,** 527–536 (2003).
2. Swartz, K. J. & MacKinnon, R. Hanatoxin modifies the gating of a voltage-dependent K^+ channel through multiple binding sites. *Neuron* **18,** 665–673 (1997).
3. Swartz, K. J. & MacKinnon, R. Mapping the receptor site for hanatoxin, a gating modifier of voltage-dependent K^+ channels. *Neuron* **18,** 675–682 (1997).
4. Li-Smerin, Y. & Swartz, K. J. Localization and molecular determinants of the hanatoxin receptors on the voltage-sensing domain of a K^+ channel. *J. Gen. Physiol.* **115,** 673–684 (2000).
5. Li-Smerin, Y. & Swartz, K. J. Helical structure of the COOH terminus of S3 and its contribution to the gating modifier toxin receptor in voltage-gated ion channels. *J. Gen. Physiol.* **117,** 205–218 (2001).
6. Jiang, Y. *et al.* X-ray structure of a voltage-dependent K^+ channel. *Nature* **423,** 33–41 (2003).
7. Takahashi, H. *et al.* Solution structure of hanatoxin1, a gating modifier of voltage-dependent K^+ channels: common surface features of gating modifier toxins. *J. Mol. Biol.* **297,** 771–780 (2000).
8. Wang, J. M. *et al.* Molecular surface of tarantula toxins interacting with voltage sensors in Kv channels. *J. Gen. Physiol.* **123,** 455–467 (2004).
9. Lee, S. Y. & MacKinnon, R. A membrane-access mechanism of ion channel inhibition by voltage sensor toxins from spider venom. *Nature* **430,** 232–235 (2004).
10. Suchyna, T. M. *et al.* Bilayer-dependent inhibition of mechanosensitive channels by neuroactive peptide enantiomers. *Nature* **430,** 235–240 (2004).
11. Smith, J. J., Alphy, S., Seibert, A. L. & Blumenthal, K. M. Differential phospholipid binding by site 3 and site 4 toxins: Implications for structural variability between voltage-sensitive sodium channel domains. *J. Biol. Chem.* **280,** 11127–11133 (2005).
12. Jung, H. J. *et al.* Solution structure and lipid membrane partitioning of VSTx1, an inhibitor of the KvAP potassium channel. *Biochemistry* **44,** 6015–6023 (2005).
13. Rogers, J. C., Qu, Y., Tanada, T. N., Scheuer, T. & Catterall, W. A. Molecular determinants of high affinity binding of α-scorpion toxin and sea anemone toxin in the S3–S4 extracellular loop in domain IV of the Na^+ channel α-subunit. *J. Biol. Chem.* **271,** 15950–15962 (1996).
14. Catterall, W. A. Membrane potential-dependent binding of scorpion toxin to the action potential Na^+ ionophore. Studies with a toxin derivative prepared by lactoperoxidase-catalyzed iodination. *J. Biol. Chem.* **252,** 8660–8668 (1977).
15. Catterall, W. A. Binding of scorpion toxin to receptor sites associated with sodium channels in frog muscle. Correlation of voltage-dependent binding with activation. *J. Gen. Physiol.* **74,** 375–391 (1979).
16. McDonough, S. I., Lampe, R. A., Keith, R. A. & Bean, B. P. Voltage-dependent inhibition of N- and P-type calcium channels by the peptide toxin omega-grammotoxin-SIA. *Mol. Pharmacol.* **52,** 1095–1104 (1997).
17. McIntosh, T. J. & Holloway, P. W. Determination of the depth of bromine atoms in bilayers formed from bromolipid probes. *Biochemistry* **26,** 1783–1788 (1987).
18. Ladokhin, A. S. Evaluation of lipid exposure of tryptophan residues in membrane peptides and proteins. *Anal. Biochem.* **276,** 65–71 (1999).
19. Ladokhin, A. S., Jayasinghe, S. & White, S. H. How to measure and analyse tryptophan fluorescence in membranes properly, and why bother? *Anal. Biochem.* **285,** 235–245 (2000).
20. Jiang, Y., Ruta, V., Chen, J., Lee, A. & MacKinnon, R. The principle of gating charge movement in a voltage-dependent K^+ channel. *Nature* **423,** 42–48 (2003).
21. Cuello, L., Cortes, D. M. & Perozo, E. Molecular architecture of the KvAP voltage-dependent K^+ channel in a lipid bilayer. *Science* **306,** 491–495 (2004).
22. Yang, N., George, A. L. Jr & Horn, R. Molecular basis of charge movement in voltage-gated sodium channels. *Neuron* **16,** 113–122 (1996).
23. Larsson, H. P., Baker, O. S., Dhillon, D. S. & Isacoff, E. Y. Transmembrane movement of the shaker K^+ channel S4. *Neuron* **16,** 387–397 (1996).
24. Starace, D. M. & Bezanilla, F. A proton pore in a potassium channel voltage sensor reveals a focused electric field. *Nature* **427,** 548–553 (2004).
25. Ahern, C. A. & Horn, R. Specificity of charge-carrying residues in the voltage sensor of potassium channels. *J. Gen. Physiol.* **123,** 205–216 (2004).
26. Swartz, K. J. Towards a structural view of gating in potassium channels. *Nature Rev. Neurosci.* **5,** 905–916 (2004).
27. Islas, L. D. & Sigworth, F. J. Electrostatics and the gating pore of Shaker potassium channels. *J. Gen. Physiol.* **117,** 69–89 (2001).
28. Swartz, K. J. & MacKinnon, R. An inhibitor of the Kv2.1 potassium channel isolated from the venom of a Chilean tarantula. *Neuron* **15,** 941–949 (1995).
29. Ladokhin, A. S. Analysis of protein and peptide penetration into membranes by depth-dependent fluorescence quenching: theoretical considerations. *Biophys. J.* **76,** 946–955 (1999).
30. Abrams, F. S. & London, E. Calibration of the parallax fluorescence quenching method for determination of membrane penetration depth: refinement and comparison of quenching by spin-labelled and brominated lipids. *Biochemistry* **31,** 5312–5322 (1992).

Acknowledgements We thank J. Diamond, L. Milescu, S. Silberberg and members of the Swartz laboratory for discussions, and the NINDS DNA sequencing facility for DNA sequencing.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to K.J.S. (swartzk@ninds.nih.gov).

LETTERS

Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis

Kristin C. Gunsalus[1]*, Hui Ge[2]*, Aaron J. Schetter[1]*, Debra S. Goldberg[3]*, Jing-Dong J. Han[2], Tong Hao[2], Gabriel F. Berriz[3], Nicolas Bertin[2], Jerry Huang[1], Ling-Shiang Chuang[1], Ning Li[2], Ramamurthy Mani[3], Anthony A. Hyman[4], Birte Sönnichsen[5], Christophe J. Echeverri[5], Frederick P. Roth[3], Marc Vidal[2] & Fabio Piano[1]

Although numerous fundamental aspects of development have been uncovered through the study of individual genes and proteins, system-level models are still missing for most developmental processes. The first two cell divisions of *Caenorhabditis elegans* embryogenesis constitute an ideal test bed for a system-level approach. Early embryogenesis, including processes such as cell division and establishment of cellular polarity, is readily amenable to large-scale functional analysis. A first step toward a system-level understanding is to provide 'first-draft' models both of the molecular assemblies involved[1] and of the functional connections between them. Here we show that such models can be derived from an integrated gene/protein network generated from three different types of functional relationship[2]: protein interaction[3], expression profiling similarity[4] and phenotypic profiling similarity[5], as estimated from detailed early embryonic RNA interference phenotypes systematically recorded for hundreds of early embryogenesis genes[6]. The topology of the integrated network suggests that *C. elegans* early embryogenesis is achieved through coordination of a limited set of molecular machines. We assessed the overall predictive value of such molecular machine models by dynamic localization of ten previously uncharacterized proteins within the living embryo.

Global correlations between transcriptome profiling and interactome data sets have been used to derive network graphs that combine similarity relationships from transcription profiling with physical interactions between proteins[3,7–13]. Suggestive correlations between interactome or transcriptome data and phenotypic data sets[5,10,14,15] support the notion that these three types of data might complement one another in predicting functional relationships.

To model *C. elegans* early embryogenesis globally, we generated network graphs in which each node represents an early embryogenesis gene[6] and its product(s), and each edge represents a potential functional connection based on one of three data sets (Fig. 1a): (1) 6,572 binary physical interactions between 3,848 *C. elegans* proteins (WI7 data set; Supplementary Methods and Supplementary Table S1)[3]; (2) expression profiling similarity above a given threshold (transcriptional Pearson correlation coefficients (transcriptional PCCs) from a compendium of *C. elegans* microarray profiles[4]); and (3) phenotypic similarity above another threshold (described below).

For each of the 661 early embryogenesis genes identified[6], we used an RNA interference (RNAi) phenotypic signature[5] consisting of a vector describing specific cellular defects in early embryogenesis[6]. We defined a measure of phenotypic similarity between early embryogenesis genes as the uncentred Pearson correlation coefficient (phenotypic PCC) for each pair of signatures (Supplementary Methods). The level of phenotypic similarity correlates with similar functional attributes (Fig. 1b, top right), as measured by shared Gene Ontology (GO) terms (which provide a controlled vocabulary for gene function)[16]. We performed hierarchical clustering to group genes by phenotypic similarity, and observed that clusters tend to show significant enrichment for specific gene functions (Fig. 1b; see also Supplementary Methods, Supplementary Fig. S1 and Supplementary Table S2). These results suggest that phenotypic PCC derived from RNAi data represents a reasonable way to compare phenotypes quantitatively.

To evaluate whether all three functional relationships (physical interaction, expression similarity and phenotypic similarity) can be merged into predictive models, we asked whether they show correlations among early embryogenesis genes/proteins (Fig. 2). First, we found that the products of early embryogenesis genes are more interconnected by direct protein interactions than expected by chance (Fig. 2a; see also Supplementary Methods and Supplementary Table S3). Second, compared to genes partitioned randomly (Fig. 2b, right), genes clustered by phenotypic similarity encode proteins that are more likely to interact physically with one another, either directly (Fig. 2b, top left) or indirectly through a single shared interactor (Fig. 2b, bottom left). Third, expression correlations are significantly higher between protein interactors, early embryogenesis gene pairs and pairs in the same phenocluster (mean transcriptional PCC = 0.17, 0.16 and 0.19, respectively) relative to gene pairs selected at random from the genome (mean transcriptional PCC = 0.05) (Fig. 2c). An even more notable increase in expression correlation (ten times the level between random pairs) is seen among early embryogenesis or intra-phenocluster pairs that also interact directly (Fig. 2c).

Consistent with these global correlations, we found that the proportion of direct interactors rises as a function of both phenotype correlation (Fig. 2d) and expression correlation (Fig. 2e). Similarly, there is a positive relationship between phenotype and expression correlation values, particularly when only direct protein interactions are considered (Fig. 2f). Notably, interactors with strong expression correlation show strong phenotype correlation, suggesting membership in a common molecular assembly; interactors with low correlations of both types may represent either false positives or a different relationship (for example, a regulatory interaction). Overall, phenotypic and expression correlations both show a strong inverse relationship with distance in the interactome network (Fig. 2g). On the basis of the above correlations, we assigned edges between pairs of nodes

[1]Center for Comparative Functional Genomics, Department of Biology, New York University, New York, New York 10003, USA. [2]Center for Cancer Systems Biology and Department of Cancer Biology, Dana-Farber Cancer Institute and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Massachusetts 02115, USA. [4]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany. [5]Cenix BioScience GmbH, 01307 Dresden, Germany.
*These authors contributed equally to this work.

based on expression similarity (transcriptional PCC ≥ 0.7) or phenotypic similarity (phenotypic PCC ≥ 0.5).

The integrated early embryogenesis network—joining all 661 early embryogenesis genes/proteins by the union of all three types of relationship (Fig. 3a)—contains a main component with 31,173 edges characterized by an average of 0.9, 5.0 and 44 edges per node for protein interaction (Int), expression similarity (Tr) and phenotypic similarity (Ph), respectively. In this network, the number of gene/protein pairs with doubly supported edges is significantly higher than expected by chance ($P = 10^{-34}$, 10^{-61} and 10^{-109} for Int–Tr, Ph–Tr and Int–Ph associations, respectively). We examined the portion of this integrated early embryogenesis network containing only links with two or more types of functional support (Fig. 3b). In contrast to the full network, the topology of this 'multiple support network'—which contains about half (305) of the early embryogenesis genes/proteins—reveals distinct groups of highly interconnected genes/proteins and few or no links between the groups. To assess the predictive value of the early embryogenesis network on a global scale, we analysed the individual and combined networks for their ability to predict a specific shared function between two linked gene pairs using GO annotations (Supplementary Methods). Each individual network has significant power to detect shared function between linked gene pairs, and combining data types generally results in

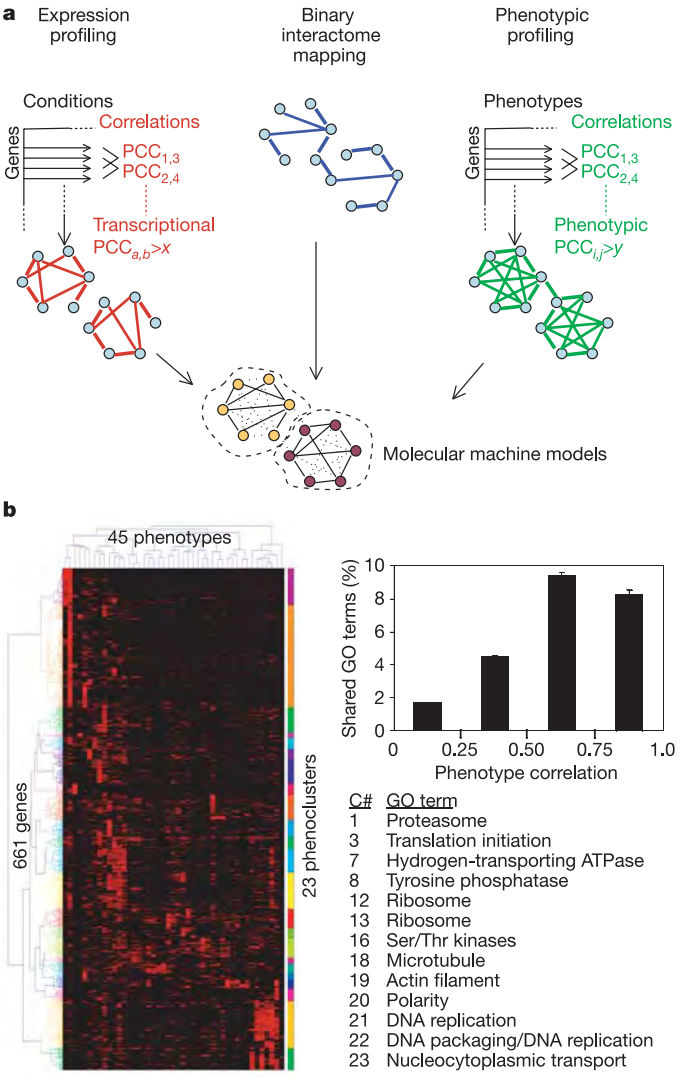


Figure 1 | Integrated networks and phenotypic profiling. **a**, Pairwise relationships between genes/proteins are determined from correlated transcript abundance (transcriptional PCC) (left), physical protein–protein interactions (centre) and phenotypic correlation (phenotypic PCC) (right). Graphs represent genes/proteins as nodes and relationships (transcriptional PCCs above threshold x, physical interactions, and phenotypic PCCs above threshold y) as edges. Highly interconnected regions represent models of molecular machines or processes (bottom). **b**, Six-hundred and sixty-one early embryogenesis genes clustered by phenotypic similarity using high-content early embryogenesis phenotypes (left; see also Supplementary Fig. S1). The fraction of shared functional annotations increases with phenotypic similarity (top right); most phenoclusters (indicated by cluster number, C#) are enriched for specific GO functional annotations (bottom right; see also Supplementary Table S1). Error bars represent standard errors of the mean.

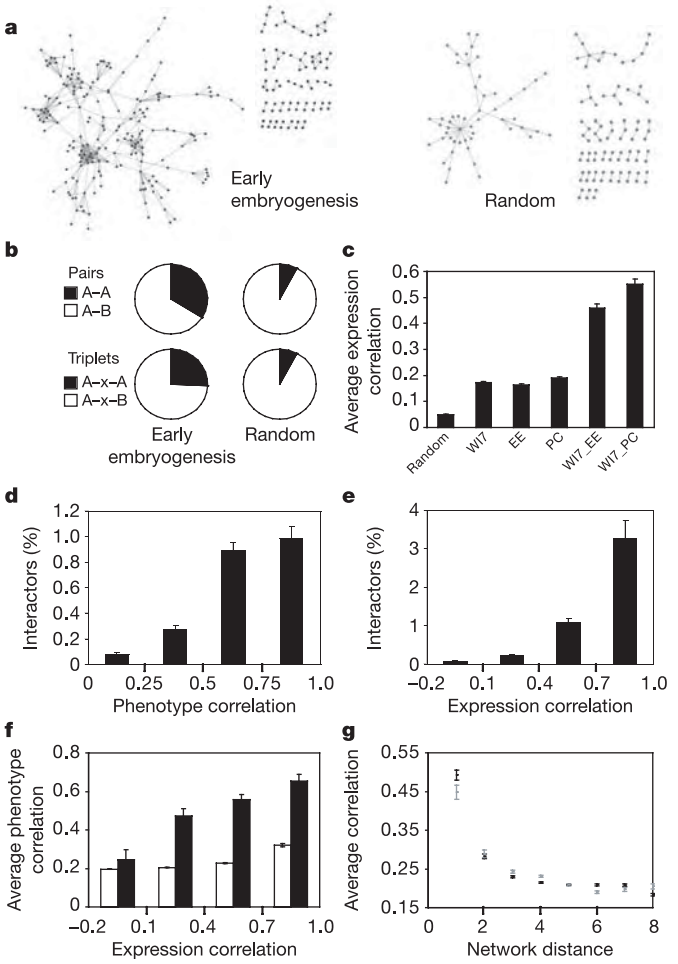


Figure 2 | Correlations between data sets (see also Supplementary Fig. S2). **a**, The early embryogenesis interactome subnetwork from WI7 (left) exhibits higher connectivity than networks of proteins chosen randomly (example at right; see also Supplementary Table S3). **b**, The early embryogenesis interactome subnetwork is enriched for interactions within the same phenocluster (A–A and A–x–A) relative to interactions between phenoclusters (A–B and A–x–B). **c**, Interacting proteins (WI7), random early embryogenesis (EE) pairs, intra-phenocluster early embryogenesis pairs (PC), pairs of interacting early embryogenesis proteins (WI7_EE) and interacting early embryogenesis proteins from common phenoclusters (WI7_PC) all show higher expression correlation than random pairs. **d**, **e**, The proportion of physical interactions increases with phenotypic (**d**) and expression correlation (**e**). **f**, Early embryogenesis genes with similar expression profiles are more likely to share similar RNAi phenotypes. All early embryogenesis gene pairs (open bars) and interacting early embryogenesis proteins (filled bars) were binned by expression correlation and plotted against average phenotypic correlation. **g**, Phenotype and expression correlation increase with interactome proximity. Average phenotype (black) and expression (grey) correlation decrease for early embryogenesis protein pairs as their distance (shortest path) increases. Error bars in **c**–**g** represent standard errors of the mean.

higher accuracy but lower sensitivity (Supplementary Tables S4 and S5). Considering only pairs for which both members have some GO annotation, the accuracy of the multiple support network is very high, with 88% of the links sharing a specific functional annotation.

To generate models representing the higher-level organization underlying early embryogenesis, we identified densely interconnected regions in the multiple support network using a graph theoretic clustering algorithm[17] followed by manual refinement based on functional annotations (Fig. 3b; see also Supplementary Table S6). We distinguished two types of highly interconnected regions among the resulting models. The first type of model contains a high density of links supported by both protein interactions and phenotypic correlations. These models represent known molecular complexes that constitute discrete molecular machines within the cell, such as the ribosome, proteasome, mitochondrial F_1F_0 ATPase, vacuolar H^+ ATPase, anaphase-promoting complex (APC) and COPI coatomer, as well as complexes involved in translation initiation, nucleocytoplasmic transport and cell polarity. Virtually all of the edges in the graph that are supported by all three types of evidence (41 out of 43 edges between 50 nodes) fall into such complexes. Proteins within such complexes function together as one physical unit, and depletion of any single member is likely to result in a very similar phenotypic profile.

The second type of model is dominated by edges supported by both phenotypic and expression correlations, containing few physical interactions. These models harbour genes that participate in distinct yet functionally interdependent cellular processes. Examples include messenger RNA/protein metabolism (mRNA transcription and processing, translational control, and protein modification and trafficking), chromosome maintenance/nucleocytoplasmic transport (DNA replication licensing and synthesis, chromosome segregation, nucleoporins and importins), and oocyte integrity/meiosis (oocyte development, extra-embryonic matrix and eggshell formation, and regulation of meiotic events). Within these models smaller molecular machines are found, supported by physical interactions and phenotypic similarity, such as the translation initiation, COPI coatomer, DNA replication licensing and importin complexes. Because current interactome maps have sampled only a small fraction of true interactions[3,18], such coordinated process models may serve to predict undiscovered protein interactions. Alternatively, these models may represent a qualitatively different type of functional unit, in which the phenotypic and expression profiling links reflect functional interdependencies dictated by the logical structure of the network, while the few protein interactions represent the physical path of information flow.

Putative functional interdependency, or cross-talk, between cellular events is evidenced in the multiple support network by links connecting distinct cellular processes. For instance, multiple edges

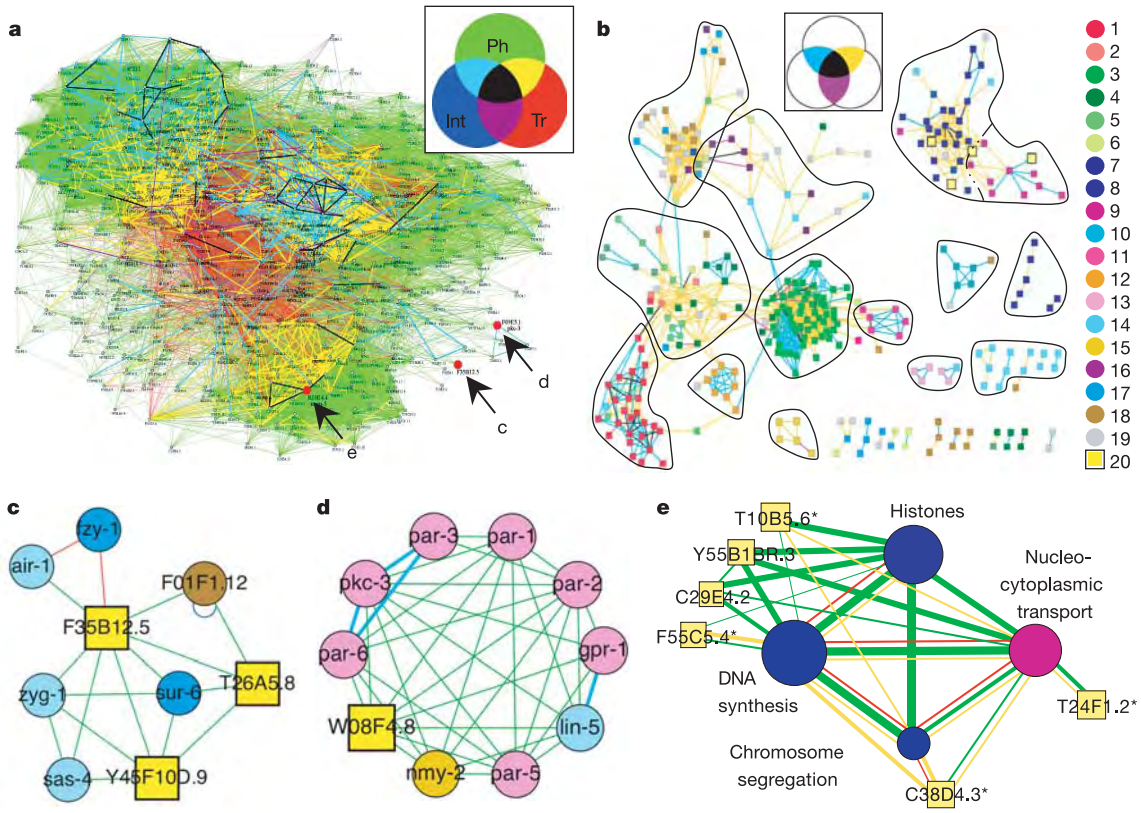


Figure 3 | Integrated network analysis. a, Entire early embryogenesis network graph (see also the Supplementary Data files). A Venn diagram (inset) shows the colour system for labelling edges based on available evidence: phenotypic profiling similarity (Phe; green), expression profiling similarity (Tr; red), physical interaction (Int; blue) and overlapping combinations of data types (intersecting regions). **b**, Multiple support network containing 305 nodes joined by 1,036 edges, each supported by two or three types of functional evidence. Predicted molecular machines are encircled (Supplementary Table S6 provides details). Nodes are colour-coded by function: 1, proteasome; 2, protein degradation; 3, ribosome/protein synthesis; 4, translational control; 5, protein/vesicular trafficking; 6, RNA synthesis/processing/binding; 7, histone; 8, DNA synthesis/replication and chromosome segregation; 9, nucleocytoplasmic transport; 10, APC; 11, mitochondrial F_1F_0 ATPase; 12, vacuolar H^+ ATPase; 13, cell polarity; 14, microtubule cytoskeleton; 15, actin cytoskeleton; 16, cell cycle; 17, signal transduction; 18, metabolism; 19, other/unknown; 20, analysed by protein localization. **c**–**e**, Subnetworks with proteins of unknown function (yellow nodes) analysed by localization (Fig. 4). **c**, Centrosome model. **d**, PAR cell polarity model. **e**, Nuclear function model. Fifty-five genes/proteins in four functional categories were grouped into 'metanodes' (with each metanode representing a collection of individual nodes). Line weights and metanode sizes approximate the number of underlying individual links and nodes, respectively. Four unknowns in this model are present in the multiple support network (asterisks in **e**).

are observed between DNA replication, DNA synthesis and nuclear membrane functions, and between regulation of protein synthesis, protein and vesicular trafficking, and protein degradation. We also observe connections and some overlap between genes required for oocyte integrity (oocyte development, extra-embryonic matrix and eggshell formation) and regulation of meiosis.

To investigate predictions from the early embryogenesis models, we selected ten genes of unknown function linked to groups of functionally characterized genes. To analyse potential participation in a molecular machine using an assay that was independent of the data used to generate the models, we used green fluorescent protein (GFP)-tagged proteins and visualized *in vivo* dynamic subcellular localization during early embryogenesis (Fig. 4; see also Supplementary Movies). We tested proteins with connections to three different early embryogenesis models: centrosomal function (Fig. 3c), cell polarity (Fig. 3d) and an elaborate network of molecular assemblies involved in DNA replication, chromatin architecture and nucleocytoplasmic transport (Fig. 3e).

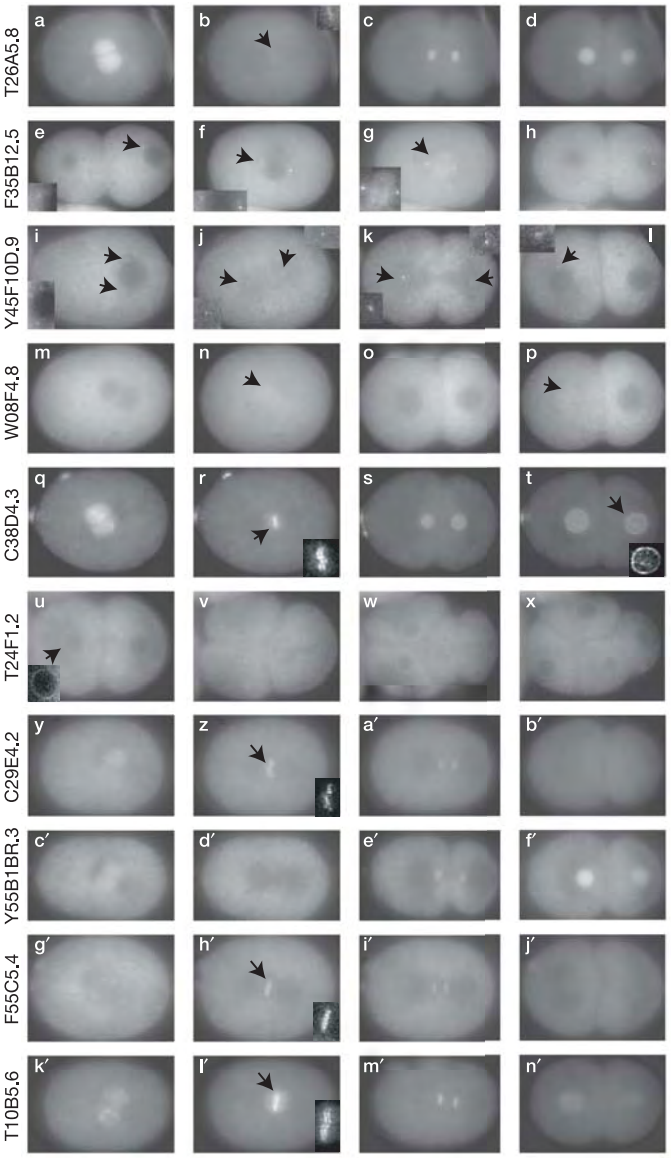


Figure 4 | Embryonic localization patterns of GFP-tagged fusion proteins analysed by time-lapse microscopy are largely consistent with model predictions. Arrows indicate areas expanded in insets (except in **n**, **p**). **a**–**d**, GFP–T26A5.8: nuclear during interphase and chromosomal during mitosis. **e**–**h**, GFP–F35B12.5 (SAS-5): appears on centriole and spindle. **i**–**l**, GFP–Y45F10D.9 (SAS-6): centriolar. **m**–**p**, GFP–W08F4.8: weakly nuclear (arrows in **n**, **p**) during mitosis but excluded from nucleus during interphase (compare two cells in **p**). **q**–**t**, GFP–C38D4.3: shuttles between nuclear periphery during interphase onto chromosomes during mitosis (double line at metaphase suggests kinetochore localization). **u**–**x**, GFP–T24F1.2: enriched on nuclear membrane. **y**–**b′**, GFP–C29E4.2: transient chromosomal localization (arrow in **z**). **c′**–**f′**, GFP–Y55B1BR.3: nuclear signal peaking during S phase (**f′**). **g′**–**n′**, GFP–F55C5.4 (**g′**–**j′**) and GFP–T10B5.6 (KNL-3) (**k′**–**n′**): transient chromosomal localization peaking at metaphase (double-line pattern suggests kinetochore localization). See Supplementary Movies for corresponding time-lapse recordings.

The centrosomal function model (Fig. 3c) connects several centriolar proteins (ZYG-1, AIR-1, SAS-4)[19–22] and FZY-1, a *C. elegans* homologue of Cdc20, which regulates the APC and is required for proper chromosome separation[23]. Remarkably, two of the three proteins we targeted (F35B12.5 and Y45F10D.9) localized to a pair of perinuclear puncta with dynamic behaviour matching that of centrioles (Fig. 4e–l; see also Supplementary Movies S2 and S3). These dynamics have subsequently been confirmed independently[24–26]. The third protein (T26A5.8) instead localized to the nucleus and faintly to metaphase chromosomes (Fig. 4a–d; see also Supplementary Movie S1); the nature of its apparent connection to centrosomal function remains to be elucidated.

The cell polarity model (Fig. 3d) contains all of the PAR proteins (partitioning-defective regulators of polarity) present in the early embryogenesis data set, as well as additional proteins known to be important for establishing polarity in *C. elegans*[27]. Interestingly, links in this subnetwork resemble mass spectrometry results for mammalian PAR homologues[28], and both implicate a homologue of *Saccharomyces cerevisiae* Cdc37 (W08F4.8 in Fig. 3d) in polarity. GFP–W08F4.8 did not reveal any asymmetric localization (a prominent feature of several PAR proteins), but appeared to fill the nucleus at prophase and be excluded from the nucleus at telophase (Fig. 4m–p; see also Supplementary Movie S4). The mammalian homologue of W08F4.8 (Cdc37) binds to the PAR-4 homologue (tumour suppressor LKB1), which shuttles in and out of the nucleus to regulate polarity[29]. The localization pattern of W08F4.8 suggests that it may participate in the nuclear shuttling of PAR-4 (PAR-4 is not shown because it is not in the early embryogenesis data set).

In each of six tests for the nuclear function model (Fig. 3e), the localization pattern of the targeted protein was consistent with its predicted role. The first, C38D4.3, is linked by expression and phenotypic similarity to members of the nuclear pore complex (*npp-2* and *npp-10*), a homologue of the centromere protein CENP-C (*hcp-4*) and DNA synthesis components (*mcm-2*, DNA polymerase B and a DNA topoisomerase II homologue). GFP–C38D4.3 localized in a remarkable dynamic pattern (Fig. 4q–t; see also Supplementary Movie S5), shuttling between the nuclear membrane (during interphase) and the chromosomes (during mitosis), appearing to coalesce at the nuclear envelope onto chromosomes (Fig. 4q), and then during metaphase forming a double-line pattern reminiscent of centromere/kinetochore proteins (Fig. 4r). The second protein, T24F1.2, is linked to several nuclear pore complex components by phenotypic similarity and to *ran-1*, a nuclear transport control factor, via expression correlation. Consistent with these connections, GFP–T24F1.2 localized to the nuclear envelope throughout the cell cycle (Fig. 4u–x; see also Supplementary Movie S6).

The last four genes/proteins tested have multiple links to the DNA synthesis and histone subnetworks (Fig. 3e). One of these, F55C5.4, is linked by both expression and phenotype to several DNA replication licensing factors (*mcm-3*, *mcm-5*, *mcm-6*, *mcm-7*) and DNA synthesis components (*lig-1*, *rnr-2*). GFP-tagged C29E4.2 and F55C5.4 both localized to condensing chromosomes exclusively around metaphase (Fig. 4g′–j′ and y–b′, respectively; see also Supplementary Movies S7 and S9). In contrast, GFP–Y55B1BR.3 showed a diffuse nuclear localization that peaked during S phase (Fig. 4c′–f′; see also Supplementary Movie S8). Finally, T10B5.6 is linked by phenotype and expression to both *hcp-4* (CENP-C) and *npp-19* (a nucleoporin),

and the GFP-tagged protein showed diffuse nuclear localization alternating with a double-line pattern typical of kinetochore/centromere proteins (Fig. 4k′–n′; see also Supplementary Movie S10). This protein's dynamic localization pattern and its function at the kinetochore was recently confirmed independently[30]. In total, eight out of ten experimental tests for potential new components of molecular machines gave rise to supportive evidence. Two localization experiments remain inconclusive, although for one (W08F4.8), proteomic experiments[28] support the hypothesis formulated here.

We have shown that phenotypic profiling data can be combined with interactome and expression profiling data to generate a network of functional relationships for *C. elegans* early embryogenesis. The combined evidence network suggests that the molecular machines acting in early embryogenesis are highly interconnected, and are likely to operate together through regulatory molecules that coordinate their activities. We have tested ten predictions from this integrated network by observing the dynamic *in vivo* localization pattern of GFP fusion proteins. The integrated network is a potential reservoir for hundreds of testable predictions about cellular processes in the early embryo. The approach presented here is scalable and can be extended to include additional data types. This general strategy is applicable to other biological processes and other organisms, including humans.

METHODS

Clustering and network analysis. Agglomerative hierarchical clustering of genes based on phenotypic similarity, construction of network graphs from protein–protein interaction data and similarity in expression and phenotypic profiles and corresponding statistical analyses are described in Supplementary Methods. GO annotations from WormBase version WS100 were used to test phenoclusters for functional enrichment and to test the integrated early embryogenesis network for its ability to predict a shared function between two linked gene/protein pairs, as detailed in Supplementary Methods.

Localization studies. For each gene selected, the sequence between the predicted initiation and termination codons was cloned into GFP vectors driving expression in the germ line and soma. Transgenic animals expressing extrachromosomal arrays were generated by injection, and animals from the F_2 generation were assayed for localization. Microscopy was carried out on Leica DMLA or DMRA microscopes using × 100 (1.3 N.A.) objectives and GFP filters. (See Supplementary Methods for details.)

Received 30 March; accepted 23 May 2005.

1. Alberts, B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. *Cell* **92,** 291–294 (1998).
2. Vidal, M. A biological atlas of functional maps. *Cell* **104,** 333–339 (2001).
3. Li, S. *et al.* A map of the interactome network of the metazoan *C. elegans*. *Science* **303,** 540–543 (2004).
4. Kim, S. K. *et al.* A gene expression map for *Caenorhabditis elegans*. *Science* **293,** 2087–2092 (2001).
5. Piano, F. *et al.* Gene clustering based on RNAi phenotypes of ovary-enriched genes in *C. elegans*. *Curr. Biol.* **12,** 1959–1964 (2002).
6. Sönnichsen, B. *et al.* Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* **434,** 462–469 (2005).
7. Marcotte, E. M., Pellegrini, M., Thompson, M. J., Yeates, T. O. & Eisenberg, D. A combined algorithm for genome-wide prediction of protein function. *Nature* **402,** 83–86 (1999).
8. Ge, H., Liu, Z., Church, G. M. & Vidal, M. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nature Genet.* **29,** 482–486 (2001).
9. Jansen, R., Greenbaum, D. & Gerstein, M. Relating whole-genome expression data with protein-protein interactions. *Genome Res.* **12,** 37–46 (2002).
10. Walhout, A. J. *et al.* Integrating interactome, phenome, and transcriptome mapping data for the *C. elegans* germline. *Curr. Biol.* **12,** 1952–1958 (2002).
11. Lee, I., Date, S. V., Adai, A. T. & Marcotte, E. M. A probabilistic functional network of yeast genes. *Science* **306,** 1555–1558 (2004).
12. Bader, J. S., Chaudhuri, A., Rothberg, J. M. & Chant, J. Gaining confidence in high-throughput protein interaction networks. *Nature Biotechnol.* **22,** 78–85 (2004).
13. Jansen, R. *et al.* A Bayesian networks approach for predicting protein-protein interactions from genomic data. *Science* **302,** 449–453 (2003).
14. Boulton, S. J. *et al.* Combined functional genomic maps of the *C. elegans* DNA damage response. *Science* **295,** 127–131 (2002).
15. Begley, T. J., Rosenbach, A. S., Ideker, T. & Samson, L. D. Damage recovery pathways in *Saccharomyces cerevisiae* revealed by genomic phenotyping and interactome mapping. *Mol. Cancer Res.* **1,** 103–112 (2002).
16. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genet.* **25,** 25–29 (2000).
17. Bader, G. D. & Hogue, C. W. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* **4,** 2 (2003).
18. Giot, L. *et al.* A protein interaction map of *Drosophila melanogaster*. *Science* **302,** 1727–1736 (2003).
19. O'Connell, K. F. *et al.* The *C. elegans zyg-1* gene encodes a regulator of centrosome duplication with distinct maternal and paternal roles in the embryo. *Cell* **105,** 547–558 (2001).
20. Schumacher, J. M., Ashcroft, N., Donovan, P. J. & Golden, A. A highly conserved centrosomal kinase, AIR-1, is required for accurate cell cycle progression and segregation of developmental factors in *Caenorhabditis elegans* embryos. *Development* **125,** 4391–4402 (1998).
21. Kirkham, M., Muller-Reichert, T., Oegema, K., Grill, S. & Hyman, A. A. SAS-4 is a *C. elegans* centriolar protein that controls centrosome size. *Cell* **112,** 575–587 (2003).
22. Leidel, S. & Gönczy, P. SAS-4 is essential for centrosome duplication in *C. elegans* and is recruited to daughter centrioles once per cell cycle. *Dev. Cell* **4,** 431–439 (2003).
23. Kitagawa, R., Law, E., Tang, L. & Rose, A. M. The Cdc20 homolog, FZY-1, and its interacting protein, IFY-1, are required for proper chromosome segregation in *Caenorhabditis elegans*. *Curr. Biol.* **12,** 2118–2123 (2002).
24. Delattre, M. *et al.* Centriolar SAS-5 is required for centrosome duplication in *C. elegans*. *Nature Cell Biol.* **6,** 656–664 (2004).
25. Dammermann, A. *et al.* Centriole assembly requires both centriolar and pericentriolar material proteins. *Dev. Cell* **7,** 815–829 (2004).
26. Leidel, S., Delattre, M., Cerutti, L., Baumer, K. & Gönczy, P. SAS-6 defines a protein family required for centrosome duplication in *C. elegans* and in human cells. *Nature Cell Biol.* **7,** 115–125 (2005).
27. Schneider, S. Q. & Bowerman, B. Cell polarity and the cytoskeleton in the *Caenorhabditis elegans* zygote. *Annu. Rev. Genet.* **37,** 221–249 (2003).
28. Brajenovic, M., Joberty, G., Kuster, B., Bouwmeester, T. & Drewes, G. Comprehensive proteomic analysis of human Par protein complexes reveals an interconnected protein network. *J. Biol. Chem.* **279,** 12804–12811 (2004).
29. Baas, A. F., Smit, L. & Clevers, H. LKB1 tumour suppressor protein: PARtaker in cell polarity. *Trends Cell Biol.* **14,** 312–319 (2004).
30. Cheeseman, I. M. *et al.* A conserved protein network controls assembly of the outer kinetochore and its ability to sustain tension. *Genes Dev.* **18,** 2255–2268 (2004).

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

Acknowledgements This work was supported by grants from NSF (to K.C.G. and D.S.G.), NIH/NHGRI (to M.V. and F.P.R.), NIH/NICHD (to F.P.) and Taplin Funds for Discovery (to F.P.R.). We thank members of our laboratories for discussions and comments on the manuscript.

Author Contributions K.C.G., H.G., D.S.G., J.-D.J.H., T.H., N.B., N.L., J.H., G.F.B. and R.M. performed informatic analyses; A.J.S. and L.-S.C. performed cloning and localization experiments; K.C.G., F.P.R., M.V. and F.P. wrote the manuscript; B.S., C.E. and A.A.H. shared critical data before publication; and F.P.R., M.V. and F.P. provided guidance for experimental and informatic analyses and interpretation of results.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to F.P. (fp1@nyu.edu), M.V. (marc_vidal@dfci.harvard.edu) or F.P.R. (fritz_roth@hms.harvard.edu).

LETTERS

The *ERECTA* gene regulates plant transpiration efficiency in *Arabidopsis*

Josette Masle[1], Scott R. Gilmore[1] & Graham D. Farquhar[1]

Assimilation of carbon by plants incurs water costs. In the many parts of the world where water is in short supply, plant transpiration efficiency, the ratio of carbon fixation to water loss, is critical to plant survival, crop yield and vegetation dynamics[1]. When challenged by variations in their environment, plants often seem to coordinate photosynthesis and transpiration[2], but significant genetic variation in transpiration efficiency has been identified both between and within species[3,4]. This has allowed plant breeders to develop effective selection programmes for the improved transpiration efficiency of crops[5], after it was demonstrated that carbon isotopic discrimination, Δ, of plant matter was a reliable and sensitive marker negatively related to variation in transpiration efficiency[3,4,6]. However, little is known of the genetic controls of transpiration efficiency. Here we report the isolation of a gene that regulates transpiration efficiency, *ERECTA*. We show that *ERECTA*, a putative leucine-rich repeat receptor-like kinase (LRR-RLK)[7,8] known for its effects on inflorescence development[7,9], is a major contributor to a locus for Δ on *Arabidopsis* chromosome 2. Mechanisms include, but are not limited to, effects on stomatal density, epidermal cell expansion, mesophyll cell proliferation and cell–cell contact.

We have previously demonstrated considerable natural variation in leaf carbon isotopic discrimination (Δ) among *Arabidopsis* accessions, and established that this variation fits the predicted pattern of a linear relationship with the ratio p_i/p_a of intercellular to ambient CO_2 partial pressure within the leaf[10], and hence with transpiration efficiency[6]. We therefore used Δ as a tool to assess the phenotype of a population of recombinant, inbred lines derived from a cross between Col-4 and Ler[11] ecotypes for variation in transpiration efficiency, with the aim of identifying genetic loci associated with this trait. Over the last 15 yr putative quantitative trait loci (QTL) for transpiration efficiency or carbon isotope discrimination have been identified in various species[12,13]; however, none of the contributing genes has been identified. Ribulose 1,5-bisphosphate carboxylase–oxygenase (Rubisco) activase has been suggested as a candidate gene controlling transpiration efficiency in tomato[14], but these results have not been confirmed. Our analysis revealed a significant QTL for rosette Δ on chromosome 2 (Fig. 1a), which, depending on growth conditions, explained 21–64% of the total phenotypic variation in Δ (see Supplementary Table 1).

That QTL, hereafter referred to as transpiration efficiency 1 (TE1), covered a small region of about 37 genes centred on the '*er*' marker (the *ERECTA* gene itself, At2g26330) at map position 50.64 cM. Furthermore, the distributions of Δ values for the subpopulations of recombinant lines carrying the *er1* and *ERECTA* alleles were always offset, with the mean values systematically lower for the latter population (Supplementary Table 1), indicative of significantly greater transpiration efficiency. This led us to test directly the effect of the *ERECTA* gene itself on Δ and transpiration efficiency.

As a first step, we compared Δ for three *erecta* mutants (Col*er2*, Col*er105* and Ler) to that of near-isogenic background lines homozygous for the *ERECTA* allele (Col-0, Col-0/Col-1 and N163/N3177, respectively). In each case, the *erecta* line displayed a higher Δ than the line homozygous for the *ERECTA* allele (Fig. 1b). This was especially so for Col*er105*, which on average differed from Col-0 by 1.1 per mil (that is, the $^{12}C/^{13}C$ ratio of the genotypes differed by

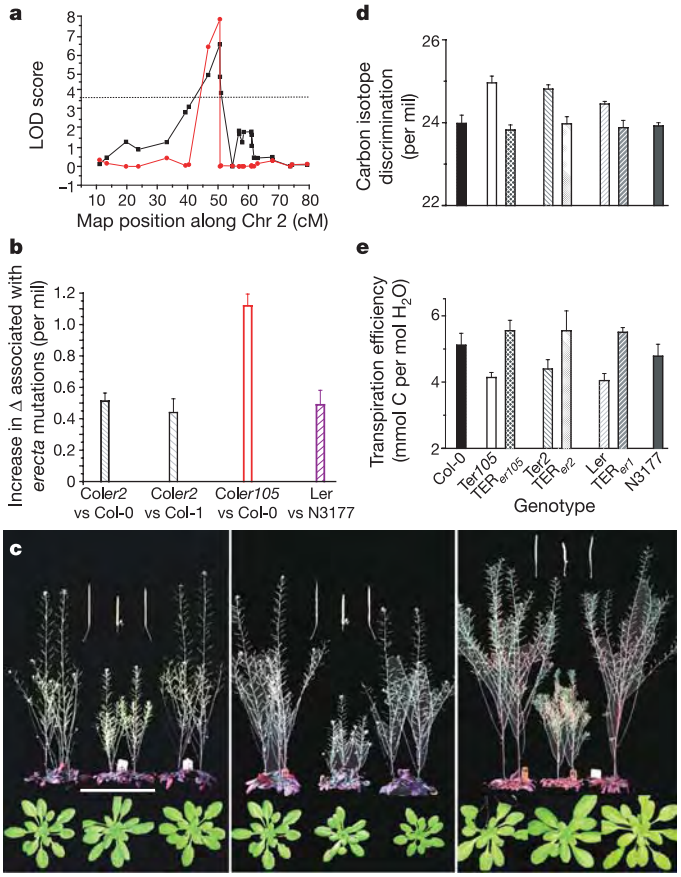


Figure 1 | *ERECTA*, a transpiration efficiency gene. a, A QTL for Δ on *Arabidopsis* chromosome 2 (48.96–51.02 cM). Simple (black symbols) and composite (red symbols) interval mapping. The dotted line indicates 1% LOD significance level. **b**, *Erecta* mutations cause increased Δ. Error bars denote s.e.m. **c**, Conversion of morphological phenotype of Col*er2*, Col*er105* and Ler (left, middle and right panels, respectively) by complementation with *ERECTA*. Left to right in each set of photographs: ER ecotype; *erecta* mutant; complemented TER_{er1} line. Scale bar, 75 mm (rosettes), 180 mm (mature plants), 30 mm (pods). **d**, **e**, Mutant complementation restores wild-type Δ and transpiration efficiency values (data are mean ± s.e.m.).

[1]Environmental Biology Group, Research School of Biological Sciences, The Australian National University, Canberra, ACT 2601, Australia.

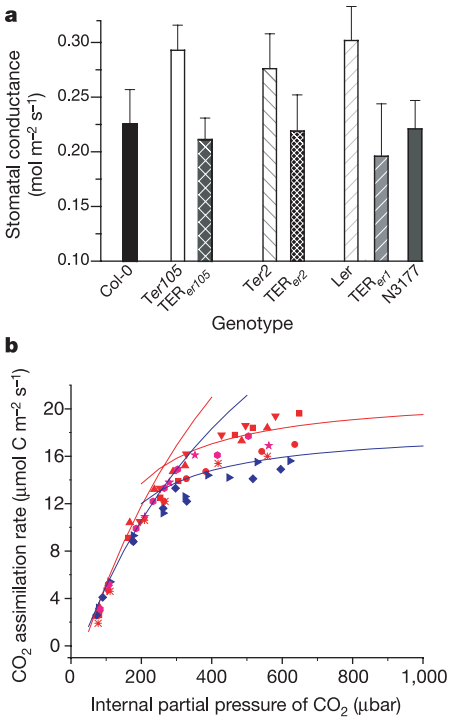
1.1×10^{-3} times the $^{12}C/^{13}C$ ratio of the source CO_2), implying an approximately 20% decrease in transpiration efficiency under our growth conditions. The Col*er105* mutant carries a null allele and is characterized by an absence of *ERECTA* transcripts[7,15]. The Col*er2* and Ler mutants, which both carry an I to K amino acid change in the protein kinase domain[8], showed an increase of 0.4–0.5 per mil in Δ compared to the isogenic ER (ERECTA) line.

To confirm the role of *ERECTA* in these differences, we transformed the Col*er2* and Col*er105* mutants with a pKUT196 construct carrying the Col-0 *ERECTA* gene (see Methods). All transgenic plants showed a typical *ERECTA* phenotype, indistinguishable from that of Col-0 or Col-1 (Fig. 1c). The abundance of *ERECTA* transcripts in the transgenic plants relative to that in Col-0 varied over roughly an order of magnitude depending on line (Supplementary Fig. 1). Ten independent T3 Columbia transgenic lines (denoted as TER_{er2} and TER_{er105} below) homozygous for the *ERECTA* transgene were selected for evaluation of Δ. All showed significantly reduced Δ values compared with the parental T*er2* and T*er105* mutant lines, resulting in values similar (Fig. 1d) or even lower (Supplementary Table 2) than those measured in Col ER ecotypes. Leaf transpiration efficiency was directly measured on the same plants by gas exchange techniques and, consistent with theory[6], showed variations opposite to those in Δ (Fig. 1e). These results were further confirmed by the analysis of a transgenic Landsberg ER line (TER_{er1}), a transformant of Ler also complemented with the pKUT196 construct[16] and showing typical *ERECTA* phenotype (Fig. 1c, right panel). The TER_{er1} line also displayed much lower Δ than Ler (Fig. 1d) and an improved transpiration efficiency (Fig. 1e), similar to or better than that measured in near-isogenic ERECTA lines N3177 or N163. These data demonstrate that *ERECTA* is directly involved in the control of leaf Δ and transpiration efficiency.

Leaf transpiration efficiency describes the ratio of photosynthesis to transpiration rates. Genetic variation in transpiration efficiency could arise from variation in either of these components. To determine the mode of action of *ERECTA* on transpiration efficiency and the underlying leaf properties, we first examined its effects on stomatal conductance, a driver of transpiration rate. This revealed that Col*er2*, Col*er105* and Ler mutants are all characterized by a much higher leaf stomatal conductance than the parental ER ecotype. Complementing these mutants with *ERECTA* markedly lowered stomatal conductance, restoring values similar to those observed in ER ecotypes (Fig. 2a). To examine whether *ERECTA* also modulates biochemical limitations to photosynthesis, we examined the response of the CO_2 assimilation rate to intercellular partial pressure of CO_2 (p_i) under saturating light, a diagnostic of variation in photosynthetic capacity. Data for all individual leaves could be well described by a biochemical model[17] in which, under low ambient CO_2 concentrations, CO_2 assimilation is limited by the ribulose-1,5-bisphosphate (RuBP)-saturated rate of Rubisco carboxylation, whereas under high concentrations, CO_2 assimilation is determined by the rate of RuBP regeneration (see example in Fig. 2b). Using this model, we derived estimates of maximum Rubisco carboxylation rate (V_{cmax}, related to the initial slope in Fig. 2b) and electron transport rate (J_{max}, related to the plateau in Fig. 2b). The *er2* mutation and to a lesser extent the *er105* mutation in the Columbia ecotype both depressed these rates, whereas the *er1* mutation in the Landsberg ecotype had no significant effect. These data on V_{cmax} and J_{max} demonstrate a role for the *ERECTA* gene in controlling leaf photosynthetic capacity and balancing biochemical and stomatal limitations on photosynthesis. They also indicate that this function of *ERECTA* is dependent on genetic background and is modulated by other genetic polymorphisms between the two ecotypes.

ERECTA has recently been implicated in the control of tissue patterning in stems and pedicels[18,19]. This led us to hypothesize that its effects on leaf gas exchange properties may partially reflect a role in the development of vegetative organs, leaves in particular, thereby impacting on stomatal conductance and CO_2 carboxylation capacity within the lamina. To test this hypothesis we examined the anatomy of the mature leaves used for gas exchange measurements, or of leaves of similar age. These observations revealed a greatly increased stomatal density in all three *erecta* mutants analysed (Fig. 3a), largely explaining their increased stomatal conductance (Fig. 3b). These leaves also displayed much smaller epidermal cells than leaves of ER lines (Fig. 3c), so that there was no significant change in stomatal index (Fig. 3d; see Methods for definition). Thus, *ERECTA* seems to modulate stomatal density mostly through a role in epidermal pavement cell expansion rather than in stomatal development and patterning *per se*[20].

There was no systematic difference in blade thickness between lines, but in T*er2* and T*er105* the anatomy of the mesophyll tissue was distinctly different compared with that in complemented TER lines or Col-0/Col-1 ecotypes, with fewer, looser packed mesophyll cells, especially in the spongy mesophyll (Fig. 3e). The measured reduction



c

Genotype	V_{cmax} (μmol C m^{-2} s^{-1})	J_{max} (μmol e$^-$ m^{-2} s^{-1})
T*er105*	47.2±3.4	71.2±3.6
TER_{er105}	53.2±1.5	75.9±2.2
T*er2*	48.9±1.1	71.3±3.9
TER_{er2}	58.0±2.9	87.8±3.0
Ler	51.8±2.04	81.5±2.5
TER_{er1}	51.8±3.6	85.3±0.7

Figure 2 | *ERECTA* regulates leaf gas exchange. a, Stomatal conductance to diffusion of water vapour calculated from gas exchange measurements on mature leaves (same plants as in Fig. 1d, e; error bars indicate s.e.m.). **b**, Relationships between CO_2 assimilation rate, measured under saturating light, and internal partial pressure of CO_2 in the substomatal cavity for Col*er2* (blue points) and complemented TER_{er2} plants (red points). Two-branch curves: theoretical relationships[17]. **c**, Electron transport capacity (J_{max}) and maximum Rubisco carboxylation rates (V_{cmax}) derived from the theoretical relationships such as those in **b**, for three *erecta* mutants and the complemented transgenic ERECTA lines.

of photosynthetic capacity (Fig. 2b, c) suggests that the reduction in the amount of photosynthetic machinery per unit leaf area more than offsets the benefits of greater porosity within the leaf[21]. Such anatomical features were hardly detectable and less systematic in the Ler mutant compared to its near-isogenic N163/N3177 or transgenic TER_{er1} lines, also consistent with the absence of a significant difference in V_{cmax} or J_{max} between those lines (Fig. 2c).

These data provide direct evidence that *ERECTA* is involved in leaf organogenesis and modulates cell expansion, cell division and cell–cell contact within leaves in a manner dependent on genetic background. Through these effects *ERECTA* modifies leaf diffusive properties and mesophyll capacity for photosynthesis, and thereby the trade-off between water loss through stomata and assimilated CO_2. A role of *ERECTA* in cell–cell and tissue–tissue signalling within leaves is consistent with its putative molecular identity as a LRR-RLK, and its broad expression in the shoot apical meristem and leaf primordia[22]. It has also recently been shown that *ERECTA* participates in the establishment of abaxial–adaxial polarity in leaf primordia[23] and in the control of cell proliferation and differentiation in the inflorescence[15,18]. Nevertheless, it is significant that even in the Landsberg background of Ler, where the effects of the mutation on mesophyll anatomy were small, the coordinated effect of the ER complementation on Δ, p_i/p_a and transpiration efficiency was still large, similar to that in the Columbia background (for example, $\Delta(\text{Ler}) - \Delta(TER_{er1}) = 0.83$ per mil, corresponding to a 16% difference in transpiration efficiency). It is therefore tempting to conclude that *ERECTA* does not affect transpiration efficiency solely through developmental mechanisms, but possibly through a separate pathway.

Plant transpiration efficiency, leaf development and gas exchange properties are strongly modulated by environmental factors, in particular those influencing plant water status[4]. That *ERECTA* improves transpiration efficiency under well-watered conditions does not necessarily imply that this should still be the case in dry environments where, from an ecological perspective, such a function would matter most. To test this proposition we challenged *erecta* mutants and near-isogenic ER lines with soil drought or low air humidity at constant leaf temperature. The *erecta* mutations conferred reduced leaf transpiration efficiency across the whole range of environments, and gene complementation led to restoration of higher transpiration efficiency, to levels similar to ER ecotypes (Fig. 4a). Similar observations held for long-term, whole-plant water use efficiency (Fig. 4b) measured over a drought period, which reduced plant growth by 20–30% depending on line.

We have identified a new function for the *ERECTA* gene in the control of transpiration efficiency. *ERECTA* is, to our knowledge, the first gene to be shown to act on the coordination between transpiration and photosynthesis, and therefore to be identified as

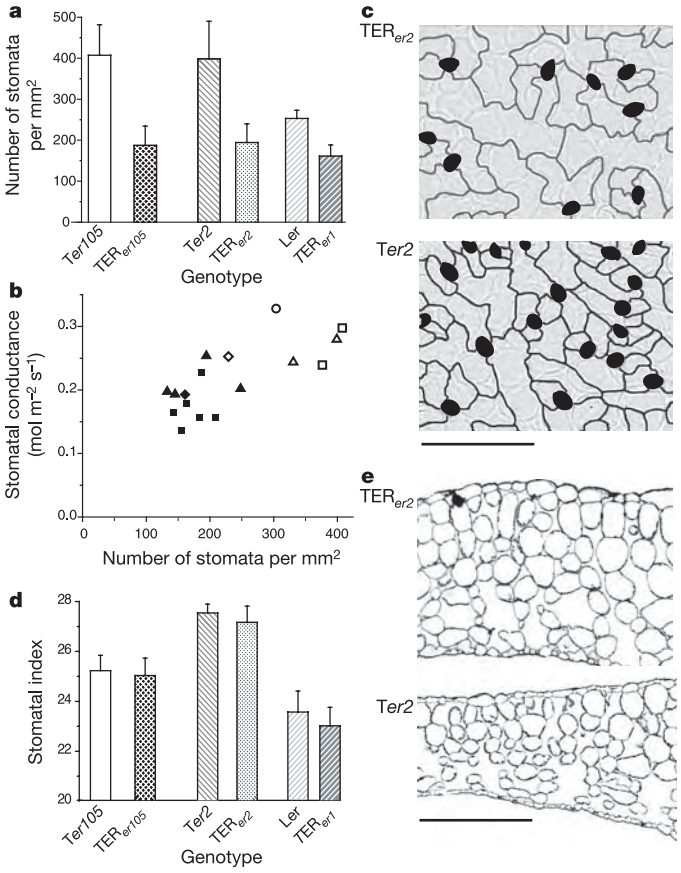


Figure 3 | Leaf anatomical features contributing to the effects of *ERECTA* on transpiration efficiency. **a**, Stomatal densities (means ± s.e.m.; adaxial epidermis). **b**, Relationship between stomatal conductance and stomatal density. Open symbols, T*er2* (squares), T*er105* (triangles) and Ler (diamonds); filled symbols, complemented lines. **c**, Coordinated effects of *ERECTA* on epidermal cell size and stomatal densities (data for adaxial epidermis, same pair of lines as for **a** with contours of cells outlined to facilitate comparison; scale bar, 155 μm). **d**, Stomatal index was unchanged. **e**, Cross-sections of leaves showing reduced mesophyll development and increased intercellular spaces in *erecta* leaves. Scale bar, 150 μm.

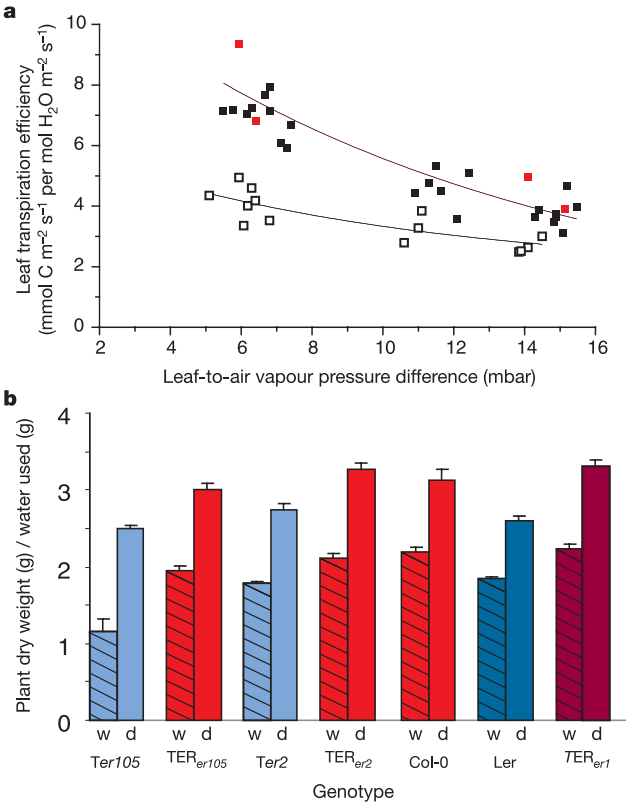


Figure 4 | *ERECTA* improved transpiration efficiency under both well-watered and drought conditions. **a**, Response of instantaneous leaf transpiration efficiency to leaf-to-air vapour pressure difference for T*er105* (open squares) and TER_{er105} (black squares) or Col-0 (red squares). Vapour pressure difference was varied by changing the humidity of air entering the chamber, keeping leaf temperature constant. Lines were fitted by nonlinear regression; ER plants had greater transpiration efficiency along the whole vapour pressure difference range. **b**, Whole-plant water use efficiency (days 16–30) for well-watered plants or those under drought conditions (labelled w or d, respectively on the *x* axis). Blue shades indicate *erecta* mutants (T*er105*, T*er2* and Ler); red shades indicate complemented TER lines and Col-0. Error bars indicate s.e.m.

a transpiration efficiency gene, as opposed to simply a stomatal or photosynthesis gene. We show that *ERECTA* can modify transpiration efficiency by acting on epidermal and mesophyll development, stomatal density and the porosity of leaves. Furthermore, even in backgrounds where changes in the mesophyll are small, the coordinated change in transpiration efficiency, as revealed by Δ or p_i/p_a, is still large. We have identified *ERECTA* homologues in diverse species (Supplementary Fig. 2 and our own unpublished observations). Phylogenetic analysis of these sequences shows that *ERECTA* has evolved during or before early angiosperm evolution. It is therefore not totally surprising that it should have a function in a trait like transpiration efficiency, which is of major importance for plant fitness and competitive ability under the selective pressure of one of the most variable and often limiting soil resources. The large magnitude of changes associated with complementation of null mutants suggests that *ERECTA* may be acting as a master gene. In several transgenic ER lines, Δ was decreased considerably below wild-type values, without detectable growth penalty, indicating the potential for manipulating *ERECTA* as a path towards improved crop performance. The ability to manipulate *ERECTA* in genetically and ecologically diverse species enhances prospects of deciphering its mode of action on the coordination between stomatal conductance and mesophyll capacity for photosynthesis at the molecular and cellular level, and of identifying the pathways involved. It will also assist in designing strategies for improved transpiration efficiency under dry conditions on the one hand, and removal of stomatal limitations and increase of yield potential in well-watered conditions on the other. Our data demonstrate that Δ constitutes a sensitive and integrative phenotypic marker of the effects of *ERECTA* on transpiration efficiency and underlying leaf gas exchange properties, opening great promise for evaluating the usefulness of this gene for breeding crop species[5].

METHODS

Plant material and growth conditions. Recombinant inbred lines were grown in controlled glasshouses or chambers under several combinations of irradiance, light spectrum, photoperiod, atmospheric humidity and temperature (see Supplementary Table 1). *Erecta* mutants, complemented transgenic lines (denoted as TER_{er1}, TER_{er2} and TER_{er105}) and near-isogenic Columbia and Landsberg ecotypes were grown in a controlled chamber (10 h photoperiod, 21 °C day/night, 400 μmol quanta $m^{-2} s^{-1}$ after acclimation at 100 μmol quanta $m^{-2} s^{-1}$). Water was continuously supplied to roots by capillarity, keeping plants well watered except for the 'drought' treatment (labelled 'd' in Fig. 4b), where from day 16 half of the plants were fed 60% only of the water used by control, well-watered plants of similar size. That regime led to a progressive water stress, causing an approximately 25% reduction in plant mass over 2.5 weeks compared to well-watered plants.

We generated the transgenic Columbia lines TER_{er2} and TER_{er105} by complementation of *erecta* mutants *er2* and *er105*, respectively, with pKUT196, a plasmid containing the entire *ERECTA* locus from Columbia[15] (see Supplementary Information for details of plant transformation and confirmation of *ERECTA* expression by quantitative PCR). The TER_{er1} Landsberg line (Ler mutant also complemented with the pKUT196 plasmid) was a gift from K. Torii. All other seeds were obtained from Nottingham Arabidopsis Stock Centre (stock N1899 (Recombinant Inbred Lines), N933 (Col-4), NW20 (Ler), N1093 (Col-0), N3176 (Col-1) and N163/N3177).

Carbon isotope discrimination. Analysis of carbon isotope composition was performed as described (see Supplementary Methods) on 4–5-week-old rosettes (50–70% of final size, vegetative stage or first flower buds just visible) from 4–6 plants per line. For the comparison of *erecta* mutants with near-isogenic ERECTA lines, an additional analysis was done at the time of gas exchange measurements and microscopy analysis, when 10–50% of flowers to be produced were opened. Carbon isotope discrimination was calculated as described (see Supplementary Methods).

QTL analysis of carbon isotopic discrimination. QTL analysis was done by simple and composite interval mapping[24,25] using QTL Cartographer[26] run as described for cofactor selection[27]. Walking speed and window size within model 6 (CIM) were left at default settings. Significant LOD (log likelihood ratios) threshold values were determined by running 1,000 permutations using the CIM model.

Physiological analyses. Physiological analyses were done on ten independent T3 transformants homozygous for the *ERECTA* transgene (lines denoted TER_{er1}, TER_{er2} and TER_{er105}, where the subscript denotes the *erecta* allele carried by the original mutant). These lines were compared to *erecta* mutants (Ler or negative T3 transformants T*er2* and T*er105*) as well as to the original mutants and to near-isogenic ER ecotypes Col-0 and Col-1 or N3177 and N163.

Rates of CO_2 assimilation (A) and transpiration (E) were measured in an open gas-exchange system, set up with a whole-leaf clamp-on chamber. Measurements were made on a young, fully expanded rosette leaf (on 4–8 plants per line and water supply treatment) at an irradiance of 500 μmol quanta $m^{-2} s^{-1}$, and leaf temperature of 22 °C. The first measurement was made at growth p_a (350 μbar) and a leaf-to-air vapour difference of 8–10 mbar; then either p_a or leaf-to-air vapour difference was varied to examine their effects on the rate of CO_2 assimilation. Calculations of gas exchange parameters were as described[28]. Instantaneous leaf transpiration efficiency was calculated as the ratio A/E.

From gas exchange measurements of CO_2 assimilation and p_i, *in vivo* estimates of the maximum rate of Rubisco carboxylation (V_{cmax}) and electron transport capacity (J_{max}) were obtained[29], assuming no limitation to the conductance for transfer of CO_2 between the substomatal cavities and the sites of carboxylation.

In the drought experiment, whole-plant water use efficiency was calculated from gravimetric measurements of increase in plant dry weight (day 16 to day 30) and water loss, providing an integrative measure over time and over the whole plant that also takes into account respiratory losses.

Leaf anatomy. Mature rosette leaves (leaves used for gas exchange or of similar age: three leaves per rosette and 4–6 replicated rosettes per line) were cleared and mounted on a light microscope fitted with a video camera for measurements of epidermal cell size (MTV program, Datacrunch), identification of stomatal complexes on both adaxial and abaxial epidermis, and determination of stomatal index (100 × stomatal density/(stomatal density + epidermal cell density)). These observations were performed on three small areas ($36–38 \times 10^3$ μm^2) selected on each side of the blade, symmetrical about the main vein. Before clearing, a small blade section was cut out mid-way along the blade, and fixed in 2.5% (v/v) glutaraldehyde, post-fixed in 1% (v/v) osmium tetroxide in 25 mM phosphate buffer, pH 7.0, slowly dehydrated in ethanol and embedded in Spurr's resin. Cross-sections were stained with toluidine blue and mounted on a light microscope fitted with a CCD camera for observation of mesophyll anatomy.

Data analysis. Statistical significance of results was tested by analysis of variance (ANOVA). Curves in Fig. 3 were fitted by nonlinear regression using ORIGIN7 (Microcal Software Inc.) statistical functions.

Received 13 January; accepted 19 May 2005.
Published online 10 July 2005.

1. Boyer, J. S. Plant productivity and environment. *Science* **218**, 443–448 (1982).
2. Wong, S. C., Cowan, I. R. & Farquhar, G. D. Stomatal conductance correlates with photosynthetic capacity. *Nature* **282**, 424–426 (1979).
3. Farquhar, G. D., Ball, M. C., von Caemmerer, S. & Roksandic, Z. Effect of salinity and humidity on $\delta^{13}C$ values of halophytes—evidence of diffusional isotope fractionation determined by the ratio of intercellular/atmospheric partial pressure of CO_2 under different environmental conditions. *Oecologia* **52**, 121–124 (1982).
4. Farquhar, G. D. & Richards, R. A. Isotopic composition of plant carbon correlates with water-use efficiency of wheat genotypes. *Aust. J. Plant Physiol.* **11**, 539–552 (1984).
5. Rebetzke, G. J., Condon, A. G., Richards, R. A. & Farquhar, G. D. Selection for reduced carbon isotope discrimination increases aerial biomass and grain yield of rain fed bread wheat. *Crop Sci.* **42**, 739–745 (2002).
6. Farquhar, G. D., O'Leary, M. H. & Berry, J. A. On the relationship between carbon isotopic discrimination and intercellular carbon dioxide concentration in leaves. *Aust. J. Plant Physiol.* **9**, 121–137 (1982).
7. Torii, K. U. *et al.* The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* **8**, 735–746 (1996).
8. Lease, K. A., Lau, N. Y., Schuster, R. A., Torii, K. U. & Walker, J. C. Receptor serine/threonine protein kinases in signalling: analysis of the *ERECTA* receptor-like kinase of *Arabidopsis thaliana*. *New Phytol.* **151**, 133–143 (2001).
9. Bowman, J. in *Arabidopsis: An atlas of morphology and development* (ed. Bowman, J.) (Springer, New York, 1994).
10. Masle, J., Shin, J. S. & Farquhar, G. D. in *Perspectives of Plant Carbon and Water Relations from Stable Isotopes* (eds Ehleringer, J., Hall, A. E. & Farquhar, G. D.) 371–386 (Academic, New York, 1993).
11. Lister, C. & Dean, C. Recombinant inbred lines for mapping RFLP and phenotypic markers in *Arabidopsis thaliana*. *Plant J.* **4**, 745–750 (1993).
12. Martin, B., Nienhuis, J., King, G. & Schaefer, A. Restriction fragment length

polymorphisms associated with water-use efficiency in tomato. *Science* **243,** 1725–1728 (1989).
13. Thumma, B. R. *et al.* Identification of causal relationships among traits related to drought resistance in *Stylosanthes scabra* using QTL analysis. *J. Exp. Bot.* **52,** 203–214 (2001).
14. Zhu, Y., Kuanhung, R. L., Huang, Y., Tauer, C. G. & Martin, B. A cDNA from tomato (*Lycopersicon pennellii*) encoding ribulose-1,5-bisphosphate carboxylase/oxygenase activase (accession No. AF037361) (PGR98–053). *Plant Gene Register* **116,** 1603 (1998).
15. Shpak, E. D., Lakeman, M. B. & Torii, K. U. Dominant-negative receptor uncovers redundancy in the *Arabidopsis ERECTA* leucine-rich repeat receptor-like kinase signalling pathway that regulates organ shape. *Plant Cell* **15,** 1095–1110 (2003).
16. Godiard, L. *et al. ERECTA*, an LRR receptor-like kinase protein controlling development pleiotropically affects resistance to bacterial wilt. *Plant J.* **36,** 353–365 (2003).
17. Farquhar, G. D., von Caemmerer, S. & Berry, J. A. A biochemical model of photosynthetic CO_2 assimilation in leaves of C_3 species. *Planta* **149,** 78–90 (1980).
18. Douglas, S. J., Chuck, G., Dengler, R. E., Pelecanda, L. & Riggs, C. D. *KNAT1* and *ERECTA* regulate inflorescence architecture in *Arabidopsis*. *Plant Cell* **14,** 547–558 (2002).
19. Shpak, E. D., Berthiaume, C. T., Hill, E. J. & Torii, K. U. Synergistic interaction of three *ERECTA*-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development* **131,** 1491–1501 (2004).
20. Nadeau, J. A. & Sachs, F. D. in *The Arabidopsis Book* (eds Somerville, C. & Meyerowitz, E.) 1–28 ⟨http://www.aspb.org/publications/arabidopsis/⟩ (American Society of Plant Biologists, Rockville, Maryland, 2002).
21. Farquhar, G. D. & Raschke, K. On the resistance to transpiration of the sites of evaporation within the leaf. *Plant Physiol.* **61,** 1000–1005 (1978).
22. Yokoyama, R., Takahashi, T., Kato, A., Torii, K. U. & Komeda, Y. The *Arabidopsis ERECTA* gene is expressed in the shoot apical meristem and organ primordia. *Plant J.* **15,** 301–310 (1998).
23. Xu, L. *et al.* Novel *as1* and *as2* defects in leaf adaxial-abaxial polarity reveal the requirement for *ASYMMETRIC LEAVES1* and *2* and *ERECTA* functions in specifying leaf adaxial identity. *Development* **130,** 4097–4107 (2003).
24. Lander, E. S. & Botstein, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121,** 185–199 (1989).
25. Zeng, Z.-B. Precision mapping of Quantitative Trait Loci. *Genetics* **136,** 1457–1468 (1994).
26. Basten, C., Weir, B. & Zeng, Z.-B. *QTL Cartographer* (North Carolina State Univ., Raleigh, North Carolina, 2000).
27. Loudet, O., Chaillou, S., Camilleri, C., Bouchez, D. & Daniel-Vedele, F. Bay-0 x Shahdara recombinant inbred line population: a powerful tool for the genetic dissection of complex traits in *Arabidopsis*. *Theor. Appl. Genet.* **104,** 1173–1184 (2002).
28. von Caemmerer, S. & Farquhar, G. D. Some relationships between the biochemistry of photosynthesis and the gas exchange of leaves. *Planta* **153,** 376–387 (1981).
29. Masle, J., Hudson, G. S. & Badger, M. R. Effects of ambient CO_2 concentration on growth and nitrogen use in tobacco (*Nicotiana tabacum*) plants transformed with an antisense gene to the small subunit of Ribulose-1,5-bisphosphate carboxylase/oxygenase. *Plant Physiol.* **103,** 1075–1088 (1993).

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

Acknowledgements We thank K. Torii for the pKUT196 plasmid, T. Baskin for Col*er105* seeds, and S. May, C. Somerville, S. C. Wong, R. Jost and O. Berkowitz for helpful discussions, encouragement and/or comments on the manuscript.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to J.M. (josette.masle@anu.edu.au).

LETTERS

Identification of JAK/STAT signalling components by genome-wide RNA interference

Patrick Müller[1], David Kuttenkeuler[2], Viola Gesellchen[2], Martin P. Zeidler[1] & Michael Boutros[2]

Signalling pathways mediating the transduction of information between cells are essential for development, cellular differentiation and homeostasis[1]. Their dysregulation is also frequently associated with human malignancies. The Janus tyrosine kinase/signal transducer and activator of transcription (JAK/STAT) pathway represents one such signalling cascade whose evolutionarily conserved roles include cell proliferation and haematopoiesis[2]. Here we describe a systematic genome-wide survey for genes required for JAK/STAT pathway activity. Analysis of 20,026 RNA interference (RNAi)-induced phenotypes in cultured *Drosophila melanogaster* haemocyte-like cells identified interacting genes encoding 4 known and 86 previously uncharacterized proteins. Subsequently, cell-based epistasis experiments were used to classify these proteins on the basis of their interaction with known components of the signalling cascade. In addition to multiple human disease gene homologues, we have found the tyrosine phosphatase Ptp61F and the *Drosophila* homologue of BRWD3, a bromo-domain-containing protein disrupted in leukaemia[3]. Moreover, *in vivo* analysis demonstrates that disrupted *dBRWD3* and overexpressed *Ptp61F* function as suppressors of leukaemia-like blood cell tumours. This screen represents a comprehensive identification of novel loci required for JAK/STAT signalling and provides molecular insights into an important pathway relevant for human cancer. Human homologues of identified pathway modifiers may constitute targets for therapeutic interventions.

Developmental genetic screens in *Drosophila* have identified multiple JAK/STAT pathway components on the basis of their segmentation phenotype[4–6], and subsequent analysis of the pathway has characterized evolutionarily conserved roles during immune responses, haematopoiesis and cellular proliferation[7–10]. The JAK/STAT signalling cascade in *Drosophila* comprises the extracellular ligand Unpaired (Upd)[5], a transmembrane receptor with homology to the interleukin 6 (IL-6) receptor family termed Domeless (Dome)[11], a single Janus tyrosine kinase (JAK) called Hopscotch (Hop)[4] and the Stat92E transcription factor[6,12] (Fig. 1a). Known regulators of JAK/STAT signalling, including a family of *SOCS*-like genes[13], *dPIAS/Su(var)2-10* (ref. 14) and *STAM* (ref. 15), are functionally conserved and were identified based on their homology to components originally characterized in mammalian cell culture studies[2]. Although successful in identifying the pathway members Upd, Dome, Hop and Stat92E, it is probable that forward genetic approaches have missed components, possibly due to non-saturating mutagenesis, genetic redundancy or phenotypic pleiotropy.

In order to identify novel pathway components and circumvent limitations of classical genetic screens, we have undertaken a genome-wide RNAi screen, a powerful technique for the identification of new components of diverse cellular pathways[16–19]. To this end, we devised a quantitative assay for JAK/STAT signalling activity in cultured *Drosophila* cells by multimerizing a STAT92E-binding site from the *Draf* promotor[20] to generate the *6 × 2 × DrafLuc* firefly luciferase reporter. Given the role for JAK/STAT signalling in haematopoiesis[9], we used *Drosophila* haemocyte-like Kc_{167} cells because of their endogenous ability to respond to pathway activation (Fig. 1b). On transfection of the *6 × 2 × DrafLuc* reporter and a plasmid to

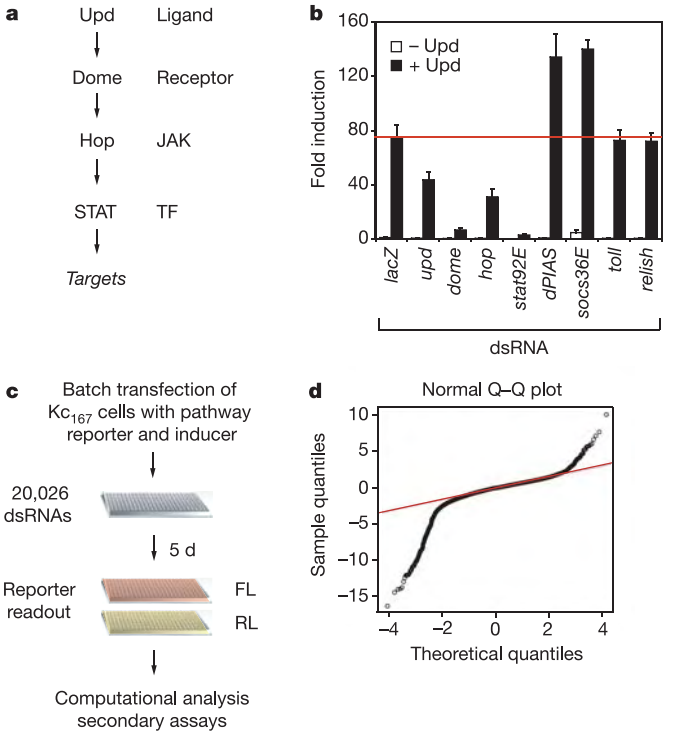


Figure 1 | Genome-wide RNAi screen for JAK/STAT signalling factors. **a**, Schematic representation of the *Drosophila* JAK/STAT signalling pathway. TF indicates transcription factor. **b**, Knock down of known JAK/STAT components leads to loss of pathway induction by Upd, whereas knock down of *lacZ*, *toll* and *relish* show no effect. The red line indicates an approximately 70-fold reporter induction relative to negative control dsRNA. Error bars represent standard deviations of six experiments. **c**, Schematic diagram of screening approach. A total of 20,026 dsRNAs were screened in duplicate in 384-well plates before computational analysis and re-testing. FL indicates firefly luciferase; RL indicates *Renilla* luciferase. **d**, Q–Q plot of normally distributed quantiles against actual pathway screening result quantiles. A fit to a normal distribution is represented by the red line. Tails of positively and negatively interacting dsRNAs at each extreme with a *z*-score threshold of >2 and <-2 represent RNAi experiments with significant phenotypes.

[1]Department of Molecular Developmental Biology, Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany. [2]Boveri-Group Signaling and Functional Genomics, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany.

constitutively express the ligand Upd, a robust induction of reporter gene activity was observed (Fig. 1b). We examined whether depletion of known pathway components by RNAi[21] modifies JAK/STAT signalling activity in Kc_{167} cells. We assessed the effect of double-stranded (ds)RNAs targeting the mRNA of the genes *dome*, *stat92E* and *hop*, as well as dsRNAs directed against the negative regulators *socs36E* and *dPIAS*. As shown in Fig. 1b, knock down of JAK/STAT components results in significant changes in reporter activity, whereas reporter activity in uninduced cells remains at low levels (Fig. 1b).

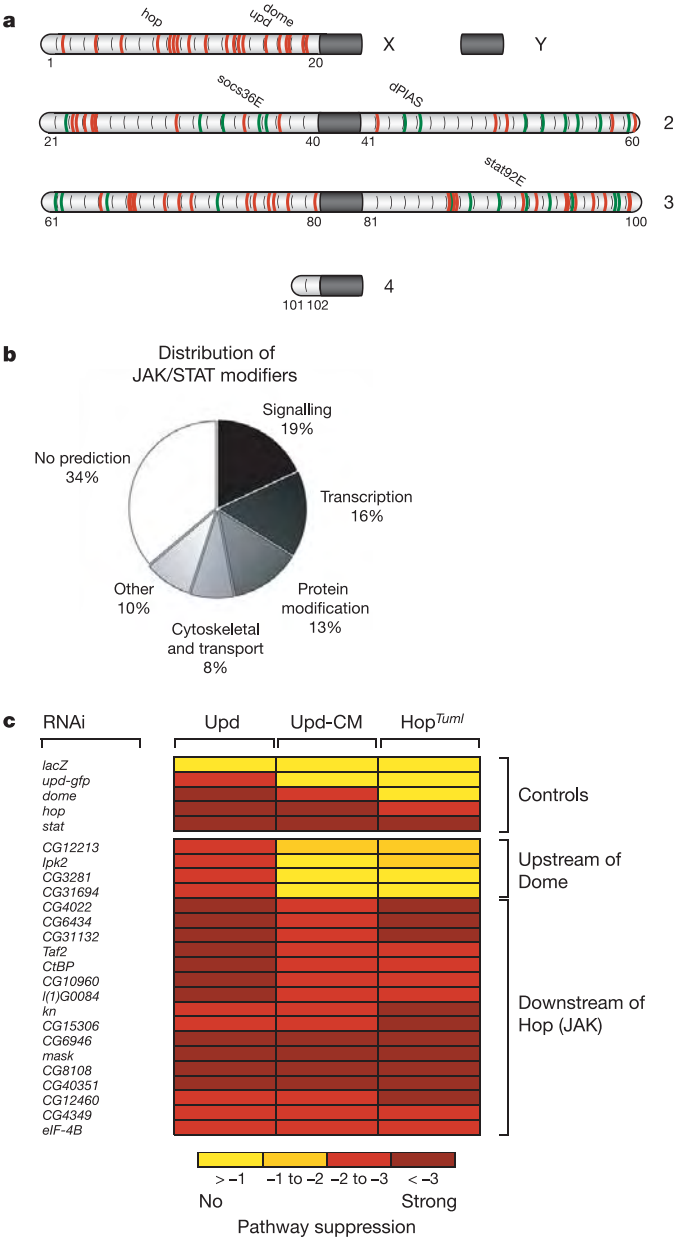


Figure 2 | **Analysis of JAK/STAT activity modulators.** **a**, Schematic representation of positive (red) and negative (green) regulator loci distributed within the *Drosophila* genome. An interactive version of this panel is available at http://www.dkfz.de/signalling/jak-pathway/. **b**, Distribution of predicted gene functions. **c**, Epistasis analysis of the indicated positive pathway regulators showing interactions graded from none (yellow squares) to strong (red squares). Results shown have been obtained in two independent replicate experiments. Abbreviations used are: ectopic expression of *upd* (Upd); Upd-conditioned medium (Upd-CM); and expression of a constitutively active JAK-allele (Hop^{Tuml}) (ref. 26). Colour coding of *z*-scores is shown.

We then set out to systematically identify genes required for JAK/STAT signalling by generating a library of 20,026 dsRNAs targeting ~91% of the predicted transcripts in the *Drosophila* genome (see Supplementary Information). Using this library we performed duplicate genome-wide screens as outlined in Fig. 1c and Supplementary Fig. S1. After computational analysis to identify candidate pathway modifiers (Fig. 1d; see also Supplementary Information), dsRNAs were resynthesized and assayed with an independent reporter, derived from the promoter of the pathway target gene *socs36E*[13], to exclude reporter-specific artefacts. These approaches confirmed the identification of 67 dsRNAs that decrease pathway activity (putative positive regulators) and 24 dsRNAs that

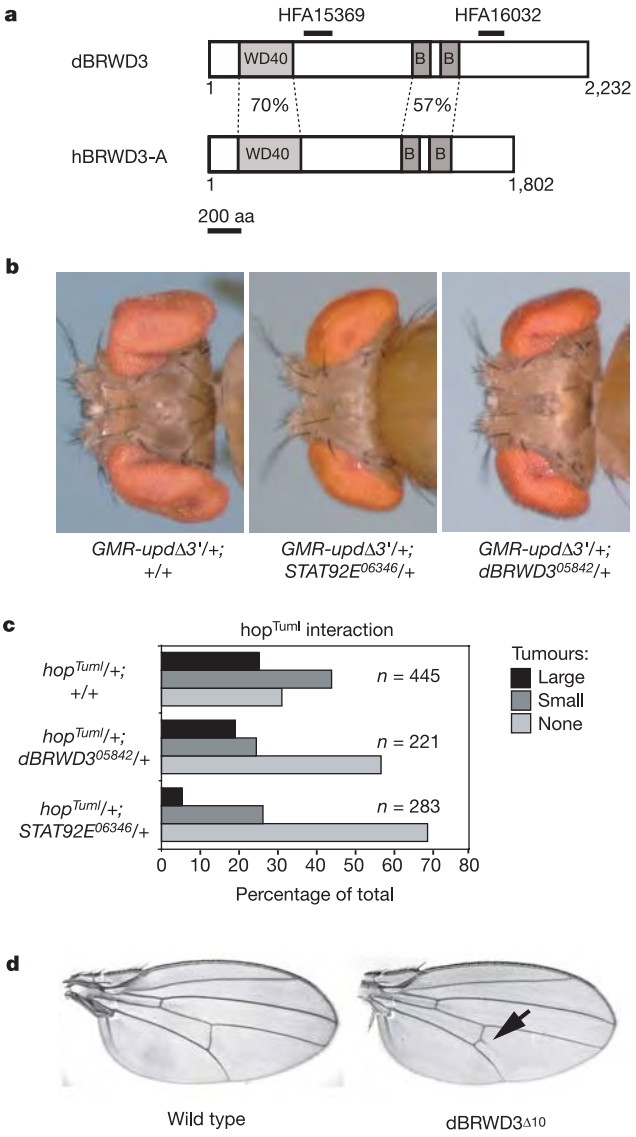


Figure 3 | **dBRWD3 functions as a JAK/STAT pathway component.** **a**, Domain structure and sequence similarity of *Drosophila* and human BRWD3 proteins. Percentages show the similarity in the amino acid sequence, and regions targeted by two independent dsRNAs independently recovered in the screen are shown. **b**, Adult *Drosophila* heads heterozygous for the *GMR-upd*$\Delta 3'$ transgene crossed to wild type (left panel), *stat92E* (middle panel) and *dBRWD3* (right panel) mutants. Note the reduction in eye overgrowth after removal of pathway components. **c**, The size and frequency of hop^{Tuml}-induced tumour formation is significantly decreased in *stat92E* and *dBRWD3* heterozygous backgrounds. **d**, Compared to adult wild-type wings (left), ectopic wing vein material (arrow) is present in homozygous $dBRWD3^{\Delta 10}$ mutant flies (a putative hypomorphic allele, right), a phenotype reminiscent of the $stat92E^{HJ}$ mutant[12].

increase pathway activity (putative negative regulators) (see also Supplementary Materials and Table 1), targeting a total of 90 predicted genes. Although most modifiers are distributed throughout the genome (Fig. 2a; see also Supplementary Table 7), the X chromosome appears to be devoid of negative regulators, a finding that may be linked to the role of the pathway in X:autosome ratio detection during *Drosophila* sex determination[22].

Based on Interpro and gene ontology (GO) annotations, pathway modifiers were classified according to their predicted functions (see Supplementary Information). Signalling factors, enzymes mediating post-translational protein modifications and transcription factors cumulatively represent 48% of the genes identified (Fig. 2b). Examples include *CG11501* encoding a putatively secreted, negative regulator of JAK/STAT signalling previously demonstrated to be a pathway target gene[8], *enok*/*CG11290* encoding an acetyl-transferase, and the *tumour suppressor protein 101*/*CG9712* gene, which encodes a ubiquitin conjugating enzyme. The molecular role of these genes in the regulation of JAK/STAT signalling remains to be determined. Seventy-four per cent of the identified loci possess human homologues (*E*-values $<10^{-10}$, compared to 62% for the whole proteome), of which 39% have been implicated in human disease (see Supplementary Tables 5 and 6).

A genetic technique to characterize signalling molecules is the determination of their epistatic relationship with respect to defined pathway components. We therefore performed cell-based epistatic assays to determine the pathway response to expression of *upd*, Upd conditioned medium or expression of the constitutively active JAK allele hop^{Tuml} (refs 5, 23) while simultaneously targeting a subset of positive regulators. In this way, dsRNA-inactivated genes required upstream in the pathway can be characterized on the basis of their rescue by pathway activation further downstream (Fig. 2c). For example, although depletion of the γ-interferon-related protein CG31694 results in downregulation of signalling stimulated by expression of *upd*, activation by Upd-conditioned medium or hop^{Tuml} is unaffected (Fig. 2c). This suggests that *CG31694* is required for the production and/or activity of the Upd ligand. Conversely, loss of pathway activity resulting from the knock down of *CG31132* cannot be rescued by any form of pathway stimulus, implying a function downstream of JAK (Fig. 2c). Although this analysis suggests a role for multiple genes upstream of Dome, this classification is based on the lack of interaction observed under differing experimental conditions and the molecular basis of these results remains to be established.

In order to confirm the function of candidate genes *in vivo*, we tested examples of both positive and negative regulators of the JAK/STAT signalling pathway. One positive regulator mentioned above is *CG31132*, which encodes a 2,232-amino-acid WD40- and bromodomain-containing protein homologous to human BRWD3 (Fig. 3a; see also Supplementary Information). In the screen, a strong reduction of pathway activity was observed for two independent dsRNAs that target different regions of the transcript (Fig. 3a). BRWD3 is a functionally uncharacterized locus recently identified at the break point of t(X;11)(q13;q23) translocations derived from multiple B cell chronic lymphocytic leukaemia (B-CLL) patients[3].

A previously uncharacterized mutagenic P element inserted in the fourth intron of *CG31132* (hereafter termed $dBRWD3^{05842}$) has been generated by the *Drosophila* genome project[24] and its remobilization indicates that the transposon insertion is responsible for late embryonic lethality (see Supplementary Information). We therefore tested for genetic interactions between *dBRWD3* and JAK/STAT signalling by crossing the $dBRWD3^{05842}$ allele to *GMR-updΔ3′* (ref. 25). The *GMR-updΔ3′* transgene ectopically misexpresses *upd* during eye development, resulting in cellular overproliferation and an enlarged adult eye (Fig. 3b, left panel). Furthermore, removal of one copy of *stat92E* significantly suppresses eye overgrowth (Fig. 3b, middle panel) due to a reduction in the potency of JAK/STAT signalling[25]. Removal of a single copy of *dBRWD3* was also able to suppress the *GMR-updΔ3′* phenotype (Fig. 3b, right panel) as expected for a

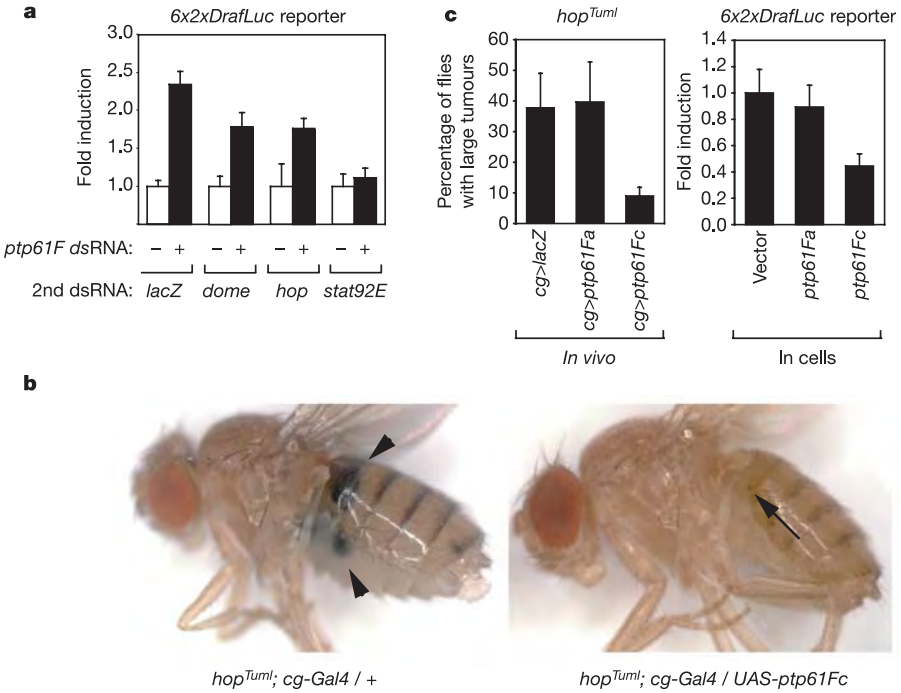


Figure 4 | Ptp61F is a tumour suppressor *in vivo*. **a**, Epistasis analysis of *ptp61F* dsRNA in cell culture indicates that it acts downstream of Hop and upstream or parallel to STAT92E. **b**, Haemocyte-specific misexpression of *ptp61F* can protect hop^{Tuml} mutants from melanotic tumour formation. Compare large black tumours in controls (arrowheads, left) with small tumours present in a *ptp61F*-expressing individual (right). **c**, Quantitative analysis of large tumour formation in hop^{Tuml} mutants expressing cytoplasmic Ptp61Fa and nuclear Ptp61Fc, showing specificity of rescue for the nuclear isoform (left), an effect that is mirrored by overexpression of the same isoforms in tissue-culture-based reporter assays (right). Error bars represent standard deviations of three or four independently tested transgenic lines or eight parallel cell culture experiments.

positive regulator of JAK/STAT signalling. A chromosomal deficiency removing the region has previously been independently identified as a suppressor of *GMR-updΔ3′* (ref. 25).

One phenotypic consequence of constitutive JAK/STAT activation caused by the gain-of-function JAK allele hop^{Tuml} is the overproliferation of haemocytes and the frequent formation of melanotic tumours, a phenotype described as a *Drosophila* model for leukaemia[23,26]. We found that removal of one copy of *dBRWD3* is sufficient to reduce the size and the frequency of hop^{Tuml}-induced melanotic tumours (Fig. 3c; and see also Supplementary Table 3). Moreover, homozygous escapers of a putative hypomorphic allele of *dBRWD3*, generated by excision of the original P-element, frequently develop ectopic wing vein material (Fig. 3d), a phenotype reminiscent of the weak loss-of-function $stat92E^{HJ}$ allele[27]. Taken together, these experiments suggest a role for *dBRWD3* in JAK/STAT signalling.

As a second example we analysed Ptp61F, a protein tyrosine phosphatase whose depletion led to an increase in JAK/STAT signalling activity. To perform an epistasis analysis we removed known pathway components and tested for the effect of simultaneously targeting *ptp61F*. Double RNAi against *ptp61F* together with *lacZ*, *dome* or *hop* results in pathway stimulation (Fig. 4a). However, simultaneous removal of *ptp61F* and *stat92E* is sufficient to prevent signalling (Fig. 4a). Loss of this phosphatase therefore results in the stimulation of Stat92E activity even in the absence of upstream components, indicating that Ptp61F negatively regulates the pathway downstream of JAK. We next asked whether Ptp61F also interferes with JAK/STAT signalling *in vivo* by using the *cg-Gal4* driver to misexpress *ptp61F* in blood cells of hop^{Tuml} mutants. Misexpression of *ptp61Fc* in a hop^{Tuml} mutant background resulted in suppression of melanotic tumour formation, with the average frequency of large tumours reduced by approximately fourfold, an effect also observed after the misexpression of *dPIAS*[14] (Fig. 4b; see also Supplementary Table 3). Alternative splicing of *ptp61F* gives nuclear and cytoplasmic protein forms that both contain the same phosphatase domain[28]. However, the tumour suppressor phenotype is only observed with nuclear Ptp61Fc (Fig. 4c), an effect that is reproduced by overexpression in cell culture (Fig. 4c). These results are consistent with our identification of Ptp61F as a negative regulator of pathway activity and suggest that it may function by targeting phosphorylated, nuclear-localized Stat92E for deactivation.

Aberrant JAK/STAT signalling has been implicated in multiple human malignancies and its components have been proposed as molecular targets for the development of therapeutic compounds[29,30]. The genome-wide screen presented here has identified known and previously unknown genes and we have characterized their probable level of interaction using cell-based epistasis analysis. Of the 90 JAK/STAT modifiers identified, many have human homologues that remain to be characterized. We have performed an analysis of two examples *in vivo* and demonstrate their roles in regulating the pathway in *Drosophila*. One of these is a homologue of human *BRWD3*, a gene recently identified at the break point of a translocation isolated from multiple B-CLL patients[3]. Given our functional analysis of *dBRWD3* and the known roles for JAK/STAT signalling during normal haematopoiesis, it is possible that a breakdown in *BRWD3*-mediated STAT regulation may represent a molecular mechanism leading to the development of B-CLL. Thus, comprehensive genetic surveys by RNAi using *Drosophila* as a model organism represent a powerful approach for identifying targets relevant to human diseases.

METHODS

Constructs. The JAK/STAT firefly luciferase reporter *6 × 2 × DrafLuc* was constructed by multimerization of a molecularly characterized Stat92E binding site present in the promoter of the endogenous target gene *Draf*[20], whereas the *4 × SocsLuc* reporter is based on a single 290-base-pair (bp) region containing four potential Stat92E binding sites present within the first intron of *socs36E* (ref. 13) (see Supplementary Information for details). A *Renilla* luciferase reporter gene under the control of the constitutively active *Actin5C* promoter was co-transfected and used to monitor cell number (see Supplementary Information).

Genome-wide RNAi screening. A genome-wide RNAi library based on polymerase chain reaction (PCR) templates with an average length of 408 bp (flanked by T7-promotor binding sites) was generated by *in vitro* transcription[17] (see Supplementary Information). Primer and amplicon sequence information can be found at http://rnai.dkfz.de. For screening experiments, *Drosophila* Kc_{167} cells were maintained in Schneider's medium (Invitrogen) supplemented with 10% fetal bovine serum (PAA) and 100 μg ml^{-1} penicillin-streptomycin (Invitrogen). A total of 57 384-well screening plates were loaded with approximately 75 nM dsRNA in 5 μl of 1 mM Tris at pH 7. Kc_{167} cells were batch transfected with the appropriate reporters, *pAct-Renilla* and *pAct-UpdGFP*, and transferred to dsRNA-containing screening plates in serum-free medium after 7 h. For 384-well plates, 15,000 cells in 20 μl were dispensed per well using an automated liquid dispenser (MultiDrop, Thermo Labsystems) and incubated for 60 min before addition of serum-containing medium. After 5 d, medium was removed, cells were lysed and both firefly and *Renilla* luciferase activities were determined (see also Supplementary Information for computational analysis).

Cell-based epistasis experiments. For epistasis experiments, cells were transfected with expression vectors to stimulate pathway activity (see below). After 7 h, 30,000 cells in 50 μl of serum-free medium were seeded into 96-well plates (Greiner) containing the dsRNAs to be tested (listed in Fig. 2c). One hour later, 75 μl medium supplemented with 10% fetal bovine serum was added to the cells. Cells were lysed after 5 d and luciferase activity measured.

Each dsRNA was tested for its ability to suppress pathway activity under three conditions: (1) in Upd-expressing cells (screening conditions); (2) in cells treated with Upd-conditioned medium (Upd-CM); and (3) in cells expressing hop^{Tuml}. Specifically, for Upd overexpression 5×10^6 Kc_{167} cells were transfected with 600 ng *pAct-UpdGFP*, 500 ng *6 × 2 × DrafLuc* reporter, 250 ng *pAc5.1-Sid-1*, 25 ng *pAct-RL* and *pAc5.1* to a total of 2 μg DNA. For Hop^{TumL} overexpression, 5×10^6 Kc_{167} cells were transfected with 200 ng *pAct-*hop^{TumL}, 500 ng *6 × 2 × DrafLuc* reporter, 250 ng *pAc5.1-Sid-1*, 25 ng *pAct-RL* and *pAc5.1* to a total of 2 μg DNA. To analyse processes upstream of Upd, two batches of cells were transfected separately to generate 'responder' and 'Upd-producer' cells. The responder cells were batch transfected with 500 ng *6 × 2 × DrafLuc* reporter, 250 ng *pAc5.1-Sid-1*, 25 ng *pAct-RL* and *pAc5.1* to a total of 2 μg plasmid DNA and subsequently seeded into 96-well plates containing the respective dsRNAs as described above. The Upd-producing cells were transfected with 2 μg *pAct-UpdGFP* and cultured in 10 cm dishes (Falcon). After 4 d, the cells were treated with 50 μg ml^{-1} heparin (Sigma) for 24 h and then the supernatant was collected, cleared by centrifugation and passed through a 0.2 μm filter (Millipore). 50 μl of this Upd-conditioned medium was then used to stimulate pathway activity in the responder cells for 24 h. Control medium from untransfected heparin-treated cells did not elicit pathway activity (data not shown). Epistasis analysis of *ptp61F* by double RNAi is described in Supplementary Information.

Genetics. For genetic interaction assays, females of the stock *y, w,* hop^{Tuml}*/FM7; P[*w^+*, cg-gal4.A]2* (ref. 26) were crossed to wild-type controls (*OreR* and w^{1118}) and mutations in *stat92E* and *dBRWD3*. The haemocyte specific Gal4 driver line *P[*w^+*, cg-Gal4.A]2* allowed specific misexpression constructs with upstream activating sequences (UAS) to be tested for their potential influence on tumour formation. Transgenic animals carrying *UAS-EGFP* or *UAS-β-galactosidase* were used as negative controls whereas *UAS-dPIASGFP* served as a positive control[14] (see Supplementary Table 3).

Crosses were incubated at 25 °C and adult females heterozygous for the hop^{Tuml} chromosome were scored within 24 h of eclosion for the presence of tumours, classified as small (one or two small melanotic spots as shown in Fig. 4b, right panel) or large (one or more large melanized growths or more than three small spots; Fig. 4b, left panel). Survival rates for hop^{Tuml} females seem to be independent of tumour frequency at the time point counted (data not shown). Assays were repeated at least twice for each genotype and a representative example from one experiment is shown (Fig. 4b).

Genetic interaction with *P[*w^+*, GMR-updΔ3′]′19* (termed *GMR-updΔ3′* in the text) was undertaken as described[25] using *OreR* and $stat92E^{06346}$ as negative and positive controls, respectively. Suppression of *GMR-updΔ3′*-induced eye overgrowth by $dBRWD3^{05842}$ was observed in multiple independent experiments in a majority of individuals of the appropriate genotype. *Drosophila* heads were photographed using a Zeiss STEMI 2000-C binocular microscope and Axiocam camera.

Received 9 March; accepted 26 May 2005.

1. Brivanlou, A. H. & Darnell, J. E. Jr Signal transduction and the control of gene expression. *Science* **295,** 813–818 (2002).

2. Hombria, J. C. & Brown, S. The fertile field of *Drosophila* Jak/STAT signalling. *Curr. Biol.* **12,** R569–R575 (2002).
3. Kalla, C. *et al.* Translocation t(X;11)(q13;q23) in B-cell chronic lymphocytic leukemia disrupts two novel genes. *Genes Chromosom. Cancer* **42,** 128–143 (2005).
4. Binari, R. & Perrimon, N. Stripe-specific regulation of pair-rule genes by *hopscotch*, a putative Jak family tyrosine kinase in *Drosophila*. *Genes Dev.* **8,** 300–312 (1994).
5. Harrison, D. A., McCoon, P. E., Binari, R., Gilman, M. & Perrimon, N. *Drosophila unpaired* encodes a secreted protein that activates the JAK signalling pathway. *Genes Dev.* **12,** 3252–3263 (1998).
6. Hou, X. S., Melnick, M. B. & Perrimon, N. Marelle acts downstream of the *Drosophila* HOP/JAK kinase and encodes a protein similar to the mammalian STATs. *Cell* **84,** 411–419 (1996).
7. Lagueux, M., Perrodou, E., Levashina, E. A., Capovilla, M. & Hoffmann, J. A. Constitutive expression of a complement-like protein in toll and JAK gain-of-function mutants of *Drosophila*. *Proc. Natl Acad. Sci. USA* **97,** 11427–11432 (2000).
8. Boutros, M., Agaisse, H. & Perrimon, N. Sequential activation of signalling pathways during innate immune responses in *Drosophila*. *Dev. Cell* **3,** 711–722 (2002).
9. Meister, M. & Lagueux, M. *Drosophila* blood cells. *Cell. Microbiol.* **5,** 573–580 (2003).
10. Mukherjee, T., Castelli-Gair Hombría, J. & Zeidler, M. P. Opposing roles for *Drosophila* JAK/STAT signalling during cellular proliferation. *Oncogene* **24,** 2503–2511 (2005).
11. Brown, S., Hu, N. & Castelli-Gair Hombria, J. Identification of the first invertebrate interleukin JAK/STAT receptor, the *Drosophila* gene *domeless*. *Curr. Biol.* **11,** 1700–1705 (2001).
12. Yan, R., Small, S., Desplan, C., Dearolf, C. R. & Darnell, J. E. Jr Identification of a *Stat* gene that functions in *Drosophila* development. *Cell* **84,** 421–430 (1996).
13. Karsten, P., Hader, S. & Zeidler, M. Cloning and expression of *Drosophila SOCS36E* and its potential regulation by the JAK/STAT pathway. *Mech. Dev.* **117,** 343–346 (2002).
14. Betz, A., Lampen, N., Martinek, S., Young, M. W. & Darnell, J. E. Jr A *Drosophila* PIAS homologue negatively regulates stat92E. *Proc. Natl Acad. Sci. USA* **98,** 9563–9568 (2001).
15. Mesilaty-Gross, S., Reich, A., Motro, B. & Wides, R. The *Drosophila* STAM gene homolog is in a tight gene cluster, and its expression correlates to that of the adjacent gene ial. *Gene* **231,** 173–186 (1999).
16. Kamath, R. S. *et al.* Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421,** 231–237 (2003).
17. Boutros, M. *et al.* Genome-wide RNAi analysis of growth and viability in *Drosophila* cells. *Science* **303,** 832–835 (2004).
18. Kittler, R. *et al.* An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* **432,** 1036–1040 (2004).
19. Hannon, G. J. & Rossi, J. J. Unlocking the potential of the human genome with RNA interference. *Nature* **431,** 371–378 (2004).
20. Kwon, E. J. *et al.* Transcriptional regulation of the *Drosophila raf* proto-oncogene by *Drosophila* STAT during development and in immune response. *J. Biol. Chem.* **275,** 19824–19830 (2000).
21. Clemens, J. C. *et al.* Use of double-stranded RNA interference in *Drosophila* cell lines to dissect signal transduction pathways. *Proc. Natl Acad. Sci. USA* **97,** 6499–6503 (2000).
22. Sefton, L., Timmer, J. R., Zhang, Y., Beranger, F. & Cline, T. W. An extracellular activator of the *Drosophila* JAK/STAT pathway is a sex-determination signal element. *Nature* **405,** 970–973 (2000).
23. Luo, H., Hanratty, W. P. & Dearolf, C. R. An amino acid substitution in the *Drosophila hop*$^{Tum\text{-}l}$ Jak kinase causes leukemia-like hematopoietic defects. *EMBO J.* **14,** 1412–1420 (1995).
24. Spradling, A. C. *et al.* The Berkeley *Drosophila* genome project gene disruption project: single P-element insertions mutating 25% of vital *Drosophila* genes. *Genetics* **153,** 135–177 (1999).
25. Bach, E. A., Vincent, S., Zeidler, M. P. & Perrimon, N. A sensitized genetic screen to identify novel regulators and components of the *Drosophila* janus kinase/signal transducer and activator of transcription pathway. *Genetics* **165,** 1149–1166 (2003).
26. Harrison, D. A., Binari, R., Nahreini, T. S., Gilman, M. & Perrimon, N. Activation of a *Drosophila* Janus kinase (JAK) causes hematopoietic neoplasia and developmental defects. *EMBO J.* **14,** 2857–2865 (1995).
27. Yan, R., Luo, H., Darnell, J. E. Jr & Dearolf, C. R. A JAK-STAT pathway regulates wing vein formation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **93,** 5842–5847 (1996).
28. McLaughlin, S. & Dixon, J. E. Alternative splicing gives rise to a nuclear protein tyrosine phosphatase in *Drosophila*. *J. Biol. Chem.* **268,** 6839–6842 (1993).
29. O'Shea, J. J., Pesu, M., Borie, D. C. & Changelian, P. S. A new modality for immunosuppression: targeting the JAK/STAT pathway. *Nature Rev. Drug Discov.* **3,** 555–564 (2004).
30. Levy, D. E. & Darnell, J. E. Jr Stats: transcriptional control and biological impact. *Nature Rev. Mol. Cell Biol.* **3,** 651–662 (2002).

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

Acknowledgements We are grateful to R. Paro for providing reagents to establish the genome-wide RNAi library. We wish to thank M. Yamayuchi and P. Karsten for the original *2 × DrafSTAT(wt)* and *Socs36E* promoter plasmids, C. Hunter for the *sid-1* plasmid, K. Bartscherer for advice and help with tissue culture experiments, T. Horn for bioinformatics support, and M. Stricker, B. Mosterman, I. Plischke and S. Häder for technical help. We thank H. Jäckle, S. Cohen, M. Osborn, R. Paro and N. Pelte for comments on the manuscript. P.M. was supported by a fellowship from the German National Academic Foundation (Studienstiftung). Research in M.B. and M.P.Z. laboratories was supported in part by Emmy-Noether grants from the Deutsche Forschungsgemeinschaft.

Author Contributions This work has been a collaborative effort between the groups of M.B. and M.P.Z.; the Boutros laboratory contributing the functional genomic expertise to dissect signalling pathways and the Zeidler laboratory expertise working with JAK/STAT signalling.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.B. (m.boutros@dkfz.de) or M.P.Z. (mzeidle@gwdg.de).

LETTERS

A high-resolution map of active promoters in the human genome

Tae Hoon Kim[1]*, Leah O. Barrera[1]*, Ming Zheng[3], Chunxu Qu[1], Michael A. Singer[4], Todd A. Richmond[4], Yingnian Wu[3], Roland D. Green[4] & Bing Ren[1,2]

In eukaryotic cells, transcription of every protein-coding gene begins with the assembly of an RNA polymerase II preinitiation complex (PIC) on the promoter[1]. The promoters, in conjunction with enhancers, silencers and insulators, define the combinatorial codes that specify gene expression patterns[2]. Our ability to analyse the control logic encoded in the human genome is currently limited by a lack of accurate information regarding the promoters for most genes[3]. Here we describe a genome-wide map of active promoters in human fibroblast cells, determined by experimentally locating the sites of PIC binding throughout the human genome. This map defines 10,567 active promoters corresponding to 6,763 known genes and at least 1,196 un-annotated transcriptional units. Features of the map suggest extensive use of multiple promoters by the human genes and widespread clustering of active promoters in the genome. In addition, examination of the genome-wide expression profile reveals four general classes of promoters that define the transcriptome of the cell. These results provide a global view of the functional relationships among transcriptional machinery, chromatin structure and gene expression in human cells.

The PIC consists of the RNA polymerase II (RNAP), the transcription factor IID (TFIID) and other general transcription factors[4]. Our strategy to map the PIC binding sites involves a chromatin immunoprecipitation-coupled DNA microarray analysis (ChIP-on-chip), which combines the immunoprecipitation of PIC-bound chromatin from formaldehyde crosslinked cells with parallel identification of the resulting bound DNA sequences using DNA microarrays[5,6]. We have previously demonstrated the feasibility of this strategy by successfully mapping active promoters in 1% of the human genome corresponding to the 44 genomic loci known as the ENCODE regions[6,7].

To apply this strategy to the entire human genome, we made a series of DNA microarrays[8] containing roughly 14.5 million 50-mer oligonucleotides, designed to represent all the non-repeat DNA throughout the human genome at 100-base pair (bp) resolution. We immunoprecipitated TFIID-bound DNA from primary fibroblast IMR90 cells using a monoclonal antibody that specifically recognizes the TAF1 subunit of this complex (TBP associated factor 1, formerly $TAF_{II}250$, ref. 9; Fig. 1a). We then amplified and fluorescently labelled the resulting DNA, and hybridized it to the above microarrays along with a differentially labelled control DNA (Fig. 1a). We determined 9,966 potential TFIID-binding regions using a simple algorithm that requires a stretch of four neighbouring probes to have a hybridization signal significantly above background. To independently verify these TFIID-binding sequences, we designed a condensed array that contained a total of 379,521 oligonucleotides to represent these sequences, and 29 control genomic loci selected from the 44 ENCODE regions[7] at 100-bp resolution. ChIP-on-chip analysis of two independent samples of IMR90 cells confirmed the binding of TFIID to a total of 8,597 regions, ranging in size from 400 bp to 9.8 kb (Fig. 1b). We further resolved a total of 12,150 TFIID-binding sites within the 8,597 fragments using a peak-finding algorithm that predicts the most likely TFIID-binding sites based on the hybridization intensity of consecutive probes with significant signals (Fig. 1b, see Supplementary Information for details).

Next, we matched these 12,150 TFIID-binding sites to the 5′ end of known transcripts in three public transcript databases (DBTSS[10], RefSeq[11] and GenBank human mRNA collection[12]) and the EnsEMBL gene catalogue[13]. To account for the uncertainty of our knowledge regarding the true 5′ end of transcripts and the uncertainty of predicted TFIID-binding positions due to noise within the microarray data, we chose an arbitrary distance of 2.5 kb as a measure of close proximity. We found that 10,553 (87%) TFIID-binding sites were within 2.5 kb of annotated 5′ ends of known messenger RNA. We resolved common TFIID-binding sites mapping to similar 5′ ends to define a non-redundant set of 9,328 5′-end-matched TFIID-binding sites. Of these TFIID-binding sequences 7,789 (83%) were found within 500 bp of the putative transcription start sites (TSS) (Fig. 1c). As these 9,328 DNA sequences were bound by TFIID *in vivo* and are within close proximity to the 5′ end of known transcripts, we defined them as promoters for the corresponding transcripts (Supplementary Table S1). Of these 9,328 promoters, 8,960 were mapped within 2.5 kb of the 5′ end or within annotated boundaries of 6,763 known genes in the EnsEMBL gene catalogue[13] (Fig. 1d and Supplementary Table S1). The remaining 368 promoters corresponded to transcripts not contained within these boundaries of EnsEMBL genes, and therefore provide support for inclusion of these transcripts to the current gene catalogues. The list of promoters also confirmed 5,118 previously annotated promoters[10], and defined 4,210 new promoters for at least 2,627 genes (Fig. 1e and Supplementary Table S1).

Four independent analyses validated the high specificity and accuracy of the active promoters detected in IMR90 cells. First, ChIP-on-chip analysis using an anti-RNAP antibody (8WG16) confirmed the binding of RNAP to at least 9,050 (97%) of the 9,328 promoters in IMR90 cells (Supplementary Fig. S1). Second, standard chromatin immunoprecipitation (ChIP) experiments performed on 28 promoters randomly selected from the above list confirmed the occupancy of RNAP on all but one promoter (Supplementary Fig. S2). Third, the 9,328 active promoters are enriched for known promoter-associated sequences such as CpG islands and the INR and DPE core promoter elements (Fig. 1f). The percentage

[1]Ludwig Institute for Cancer Research and [2]Department of Cellular and Molecular Medicine and Moores Cancer Center, UCSD School of Medicine, 9500 Gilman Drive, La Jolla, California 92093-0653, USA. [3]8125 Math Sciences Building, UCLA Department of Statistics, Los Angeles, California 90095-1554, USA. [4]NimbleGen Systems, Inc., 1 Science Court, Madison, Wisconsin 53711, USA.
*These authors contributed equally to this work.

of CpG-associated promoters (88%) was significantly higher than the previous estimate (56%, ref. 14), suggesting that CpG islands might play a more general role in gene expression than previously appreciated. Notably, we did not find the TATA box to be significantly enriched in these promoters (Fig. 1f). This might be due to a lack of conservation of the TATA box in human promoters, or it might alternatively indicate that the TATA box is not a general promoter motif for human genes. This observation is consistent with previous reports that the TATA box is only present in a small number of promoters in yeast and in *Drosophila*[15]. Fourth, ChIP-on-chip analysis using antibodies that recognize acetylated histone H3 (AcH3) or dimethylated lysine 4 on histone H3 (MeH3K4) showed that over 97% of the 9,328 promoters were associated with these known epigenetic markers of active genes[16] (Fig. 2a). The localization of MeH3K4 in these promoters was predominantly downstream of the TFIID-binding site (Fig. 2b), but the mechanisms for such chromatin organization at human promoters are currently unknown.

Among the 12,150 mapped TFIID-binding sites, 1,597 are found more than 2.5 kb away from previously defined 5′-ends of mRNA, and might represent promoters for new transcripts or genes (Supplementary Table S2). Of these, 607 non-redundant TFIID-binding sites were matched within 2.5 kb of the 5′ ends of the expressed-sequence-tag (EST)-based gene models, indicating that they may indeed produce mRNA (Supplementary Table S2). The remaining TFIID-binding sites were further filtered to a set of 632 putative promoters by requiring the occupancy of RNAP and presence of AcH3 and MeH3K4 within 1 kb of these sites (Supplementary Fig. S3). To verify that these promoters drive transcription, we analysed mRNA from the IMR90 cells using 50-mer oligonucleotide arrays that represent a 28 kb sequence surrounding 567 of the 632 unmatched putative promoters. At least 35 new transcription units were identified near the putative promoter regions, suggesting that these might represent new transcription units yet to be annotated in the human genome (Supplementary Table S3). The failure to detect mRNA from the other putative promoters might indicate that these transcripts are highly unstable. Indeed, at least one putative promoter is located in the 250 bp upstream of a predicted microRNA[17] (Supplementary Fig. S4), suggesting that some putative promoters could transcribe non-coding RNA that might have escaped detection by conventional mRNA-isolation techniques.

In total, we defined a set of 1,239 putative promoters that correspond to previously un-annotated transcription units (Supplementary Table S2). Evolutionarily conserved regions were found

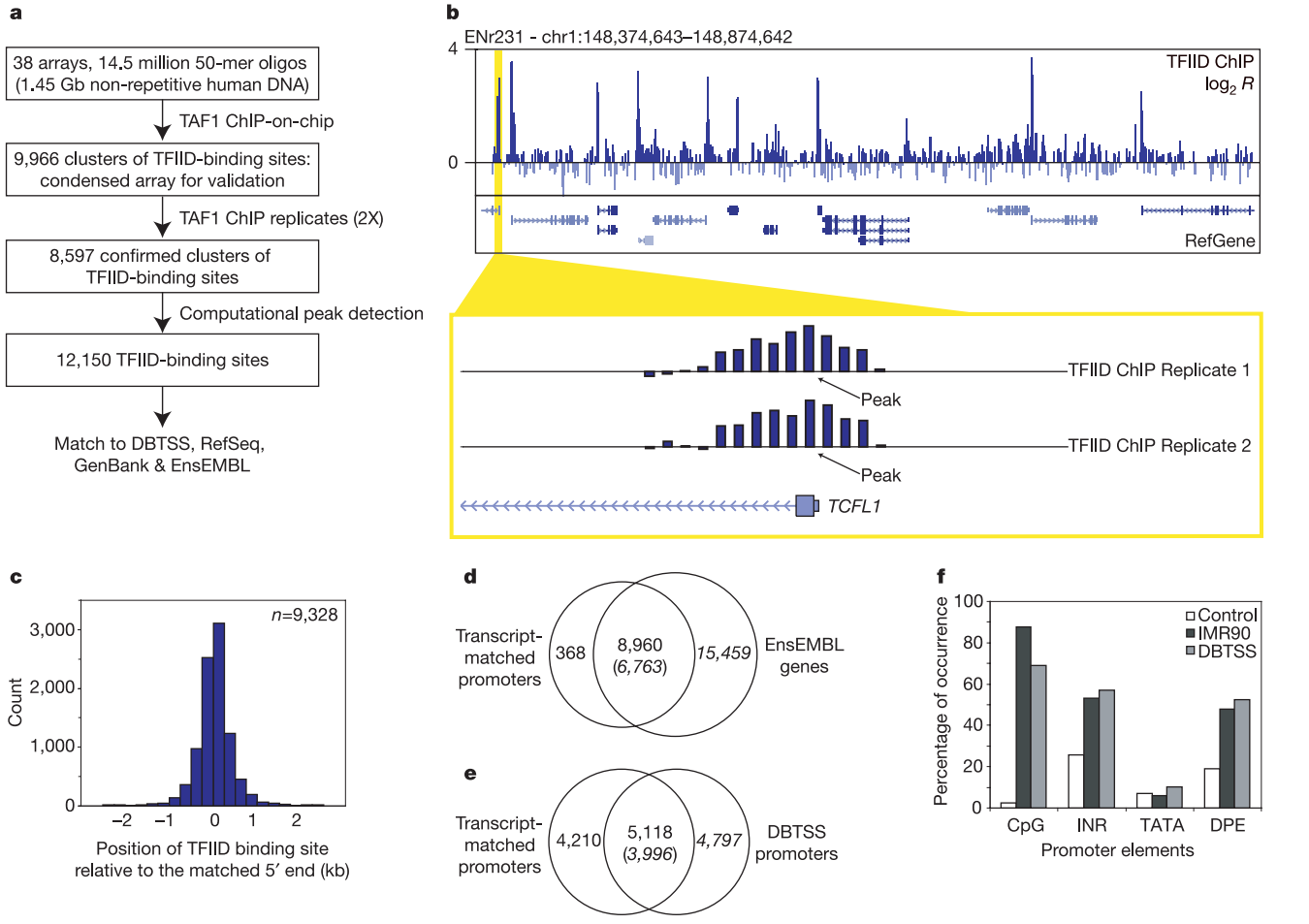


Figure 1 | Identification and characterization of active promoters in the human genome. **a**, Outline of the strategy used to map TFIID-binding sites in the genome. **b**, A representative view of the results from TFIID ChIP-on-chip analysis. Top panel, the logarithmic ratio ($\log_2 R$) of hybridization intensities between TFIID ChIP DNA and a control DNA. Middle panel, RefSeq gene annotation. Bottom panel, a close-up view of two replicate sets of TFIID ChIP-on-chip hybridization signals around the 5′ end of the *TCFL1* gene. Arrows indicate the position of the TFIID-binding site determined by a peak-finding algorithm. **c**, Distribution of TFIID-binding sites relative to the 5′ end of the matched transcripts. **d**, **e**, Venn diagrams showing the number of identified promoters that matched EnsEMBL genes (**d**) or promoters annotated in DBTSS (**e**). **f**, Chart showing the percentages of IMR90 or DBTSS promoters overlapping with CpG islands, or containing conserved TATA box, INR or DPE elements (see Supplementary Information for details).

in a majority of these putative promoters (Supplementary Fig. S5). In addition, they were significantly enriched for core promoter motifs including INR (46%) and DPE (40%), and overlapped with CpG islands (40%, Supplementary Fig. S6). These results indicate that many of the putative promoter sequences that we have defined by TFIID-binding sites may indeed be functional promoters. There are 828 putative promoters located in the intergenic regions. These promoters, together with the 368 promoters that matched to transcripts outside the EnsEMBL genes, suggest the existence of 1,196 new transcription units outside the current gene annotation[18]. This number corresponds to about 13% of the 8,960 promoters that were matched to known genes. We therefore estimate that there are probably an additional 13% of human genes that remain to be annotated in the genome. This number agrees well with a recent estimate of the total number of human genes[18], but is considerably lower than estimates based on the number of transcripts detected by microarrays, serial analysis of gene expression (SAGE) and other methods[19–22]. It is conceivable that promoters for many low-abundance transcripts may be infrequently occupied by TFIID and possibly escaped detection by our assays. Alternatively, it is possible that the new transcripts detected by the other studies are products of a different transcription machinery or process.

Two notable features were apparent in this map of active promoters. First, we observed that large domains of four or more consecutive genes were simultaneously bound by PIC and probably transcribed in the IMR90 cells. At least 256 clusters, consisting of 1,668 EnsEMBL genes, can be classified into such regions, and the number of clustered promoters is highly significant ($P < 0.001$, Supplementary Table S5). The clustering of active promoters is consistent with previous findings that co-regulated genes tend to be organized into coordinately regulated domains[23–26]. Second, a large number of genes contained two or more active promoters (Supplementary Table S4). In general, these multiple promoters correspond to transcripts with either different 5′ UTR sequences or distinct first exons (for example, *PTEN*) but do not affect the open reading frames. In some cases, however, distinct proteins were produced from multiple promoters (for example, *NR2F2* and *WEE1*). In other cases, transcripts undergo differential splicing and polyadenylation (for example, *NFKB2* and *STAT3*). The widespread use of multiple promoters in this single cell type indicates greater complexity of the cellular proteome than previously expected, and also reveals highly coordinated regulation of transcriptional initiation, splicing and polyadenylation throughout the genome[27].

To verify experimentally our observations regarding multiple promoter use in IMR90 cells, we selected the *WEE1* gene for further analysis. Two TFIID-binding sites were mapped within this gene, corresponding to the 5′ ends of two distinct mRNAs, NM_003390 and AK122837 (Fig. 3a). Each mRNA encodes a distinct protein: one encodes a well-characterized full-length version of WEE1 protein, and the other only the kinase domain. We detected both transcripts in a steady-state, asynchronous population of IMR90 cells (Fig. 3b). The shorter transcript appears to be most abundant in the G0 phase, and the longer transcript is highly transcribed in both G0 and S phase (Fig. 3c), suggesting that the two promoters in the *WEE1* gene might have distinct cell-cycle functions.

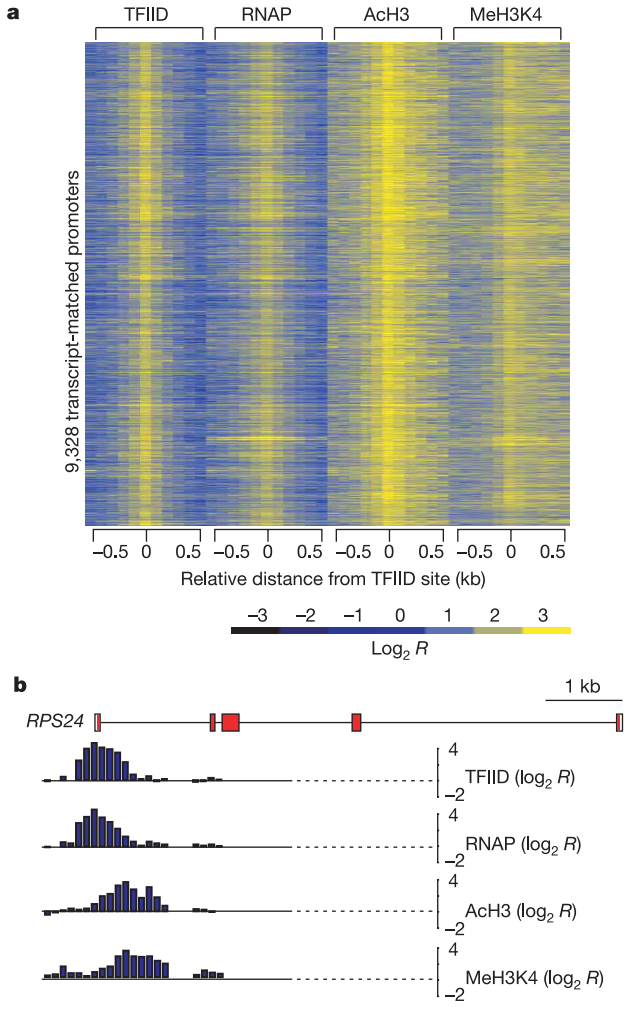


Figure 2 | The chromatin-modification features of the active promoters. **a**, Logarithmic ratios of the ChIP-on-chip hybridization intensities ($\log_2 R$) of probes from 0.5 kb upstream to 0.5 kb downstream of the identified TFIID-binding sites for TFIID, RNAP, AcH3 and MeH3K4 are plotted in a yellow–blue colour scale for 9,328 transcript-matched promoters. The bottom panel shows the colour scale with corresponding $\log_2 R$ values. **b**, A detailed view of TFIID, RNAP, AcH3 and MeH3K4 profiles on the promoter of *RPS24* gene.

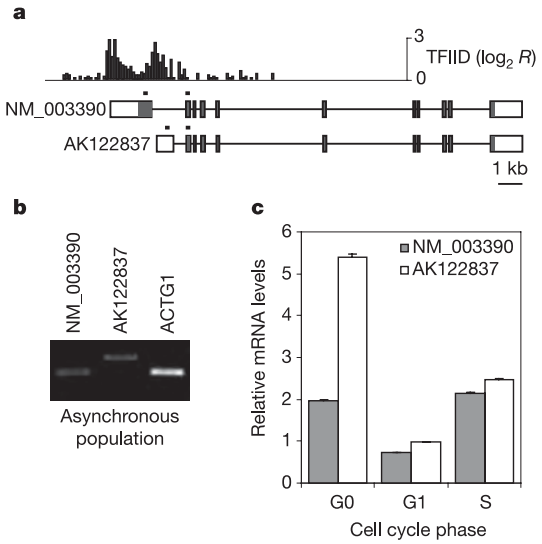


Figure 3 | Use of multiple promoters by human genes. **a**, Annotation of the *WEE1* gene locus and the corresponding TFIID-binding profile. Black bars over the first and second exons in transcripts indicate the positions of the primers used for analysis of each transcript, using real-time quantitative PCR with reverse transcription (RT–PCR). **b**, RT–PCR analysis of NM_003390 and AK122837 transcripts in an asynchronous population of IMR90 cells. **c**, Real-time quantitative RT–PCR analysis of NM_003390 and AK122837 transcripts in cell-cycle synchronized populations of IMR90 cells. Transcript levels observed for each cell-cycle phase were normalized to the level observed in the asynchronous population. Error bars represent standard deviation.

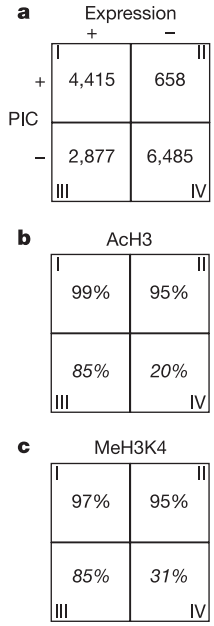


Figure 4 | Four distinct classes of promoters define the transcriptome of IMR90 cells. a, A 2 × 2 matrix describes the distribution of genes defined by expression and PIC occupancy on the promoter. **b**, **c**, Matrices showing the percentages of genes associated with AcH3 (**b**) or MeH3K4 (**c**) modification for each of the four classes of genes. Italicized numbers in some boxes represent extrapolation from the 29 ENCODE regions.

The active promoter map in IMR90 cells allowed us to systematically investigate the functional relationship between the transcription machinery and gene expression. We examined the genome-wide expression profiles of IMR90 cells and correlated the expression status of 14,437 EnsEMBL genes with promoter occupancy by the PIC. This comparison revealed four general classes of genes (Fig. 4 and Supplementary Table S6). Class I consists of 4,415 genes for which promoters were bound by the PIC and transcripts were detected. Class II includes 658 genes for which promoters were bound by the PIC but no transcript was detected. Class III contains 2,879 genes that were transcribed in IMR90 cells but for which the PIC was not detected on their promoters. Class IV contains the remaining 6,485 genes, for which the promoters were not bound by PIC and their corresponding transcripts were not detected.

The genes in class I and class IV, representing over 75% of the genes examined, support the general model that formation of the PIC on the promoters leads to transcription. The class II and III genes, on the other hand, are inconsistent with this model and may indicate that another mechanism is responsible for expression of these genes. We postulate that the discrepancy between PIC formation and transcription on the class II promoters can result from at least two possibilities. The first possibility is that the PIC assembles on these promoters, but that PIC formation is not sufficient to initiate transcription. Additional regulatory steps, such as promoter clearance or elongation, might be rate-limiting in the transcription of these genes[28]. Some notable examples in class II are the immediate early genes *FOS* and *FOSB*, the heat shock protein genes *HSPA6* and *HSPD1*, and the DNA damage repair genes *MSH5* and *ERCC4*. The second possibility is that transcription actually takes place at these promoters but that the resulting mRNAs are post-transcriptionally degraded, as in miRNA-mediated post-transcriptional silencing[29].

In contrast to class II, genes in class III appear to be transcribed, but the PIC binding on their promoters was not detected. This could simply be due to moderate sensitivity of our method[6]. To address this issue, we performed standard ChIP assays to detect binding of TFIID and RNAP on ten randomly selected class III gene promoters. Nearly 60% of the promoters were weakly associated with TFIID and RNAP in these cells, and were marked by enrichment ratios less than twofold but nonetheless above the observed background (Supplementary Fig. S2). Hence, the failure to detect TFIID and RNAP occupancy in roughly 60% of the class III promoters (~1,700) might be due to weak signals that fall below the detection sensitivity of our method. This result indicates that the promoters of a significant fraction of class III genes are open and accessible for transcription, but that PIC assembles on these promoters transiently, weakly or only during the early stage of fibroblast differentiation.

In order to understand the functional relationship between histone modification status and gene expression, we examined the AcH3 and MeH3K4 histone modifications in 29 ENCODE regions[7] (Supplementary Table S7), focusing specifically on the four classes of gene promoters. As expected, these epigenetic markers were associated with virtually all class I and class II genes, and the vast majority of class III genes. However, approximately 20% of the class IV genes were also associated with these markers (Fig. 4). This result indicates that a significant number of genes not actively transcribed are also associated with these epigenetic markers. We speculate that these histone modifications may serve to restrict genome expression potential and define the transcriptome capacity of the cell, and that transcription regulators and machinery collaborate with these epigenetic markers to further restrict the transcriptome to generate a unique pattern of genome expression.

Our results provide an initial framework for analysis of the *cis*-regulatory logic[30] in human cells. The high-resolution map of active promoters in IMR90 cells will enable detailed analysis of transcription factor binding sites within these regions. The promoter map described here can also serve as a reference for investigating gene expression in other cell types. We expect that a survey of additional cell types using the same approach will allow comprehensive mapping of all promoters in the human genome, and help elucidate the control logic that governs gene expression in different cell types in the body.

METHODS

Detailed descriptions of the experimental design and data analysis algorithms can be found in the Supplementary Information.

Briefly, IMR90 cells were obtained from the American Type Culture Collection and maintained under recommended conditions. ChIP-on-chip analysis was performed using commercial antibodies (anti-RNAP, MMS-126R, Covance; anti-TAF1, sc-735, Santa Cruz Biotechnology; anti-AcH3, 06-599, Upstate; anti-diMeH3K4, 07-030, Upstate) following the methods in ref. 6, with modifications. Microarray data from the initial 38 genome scan arrays were normalized, filtered and the TFIID-binding sites were identified as regions with a minimum of 4 probes separated by a maximum of 500 bp, with a logarithmic ratio of the ChIP-on-chip hybridization intensities ($\log_2 R$) greater than 2.5 standard deviations from the mean logarithmic ratio of the probes on each array. ChIP-on-chip hybridization intensities from the condensed arrays were normalized, averaged, and the TFIID-binding sites were identified using a computational peak-finding algorithm. The results were compared to annotated 5′-ends of transcripts from RefSeq, GenBank (downloaded from http://genome.cse.ucsc.edu; HG16, NCBI Build 34), DBTSS (http://dbtss.hgc.jp; Jan. 2004 version) and EnsEMBL (v26). Analysis of the promoter motifs was performed on a 400-bp sequence of each TFIID-binding site (from 200 bp upstream to 200 bp downstream) using matrices defined previously for the TATA box and the Inr and DPE elements. The analysis of CpG islands was carried out on a 1,200-bp sequence of each TFIID-binding site (from 1,000 bp upstream to 200 bp downstream). Standard ChIP assays were performed in duplicate with 0.5 ng of TFIID or RNAP ChIP DNA and the unenriched chromatin DNA from IMR90 cells using quantitative real-time polymerase chain reaction (PCR). Clusters of active promoters were defined by identifying runs of consecutive EnsEMBL genes with active promoters, and the significance of the number of genes found in the identified clusters was empirically determined by performing 1,000 times the same analysis on 6,763 randomly selected EnsEMBL genes. Gene expression analysis was performed in duplicate with total RNA extracted from the IMR90 cells using HU133 Plus 2.0 arrays (Affymetrix), according to the manufacturer's instructions.

Received 5 April; accepted 24 May 2005.
Published online 29 June 2005.

1. Smale, S. T. & Kadonaga, J. T. The RNA polymerase II core promoter. *Annu. Rev. Biochem.* **72,** 449–479 (2003).
2. Tjian, R. & Maniatis, T. Transcriptional activation: a complex puzzle with few easy pieces. *Cell* **77,** 5–8 (1994).
3. Trinklein, N. D., Aldred, S. J., Saldanha, A. J. & Myers, R. M. Identification and functional analysis of human transcriptional promoters. *Genome Res.* **13,** 308–312 (2003).
4. Reinberg, D. *et al.* The RNA polymerase II general transcription factors: past, present, and future. *Cold Spring Harb. Symp. Quant. Biol.* **63,** 83–103 (1998).
5. Ren, B. *et al.* Genome-wide location and function of DNA binding proteins. *Science* **290,** 2306–2309 (2000).
6. Kim, T. H. *et al.* Direct isolation and identification of promoters in the human genome. *Genome Res.* **15,** 830–839 (2005).
7. The ENCODE Project Consortium, The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science* **306,** 636–640 (2004).
8. Singh-Gasson, S. *et al.* Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nature Biotechnol.* **17,** 974–978 (1999).
9. Ruppert, S., Wang, E. H. & Tjian, R. Cloning and expression of human $TAF_{II}250$: a TBP-associated factor implicated in cell-cycle regulation. *Nature* **362,** 175–179 (1993).
10. Suzuki, Y., Yamashita, R., Sugano, S. & Nakai, K. DBTSS, DataBase of Transcriptional Start Sites: progress report 2004. *Nucleic Acids Res.* **32** (database issue), D78–81 (2004).
11. Pruitt, K. D., Tatusova, T. & Maglott, D. R. NCBI Reference Sequence project: update and current status. *Nucleic Acids Res.* **31,** 34–37 (2003).
12. Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J. & Wheeler, D. L. GenBank: update. *Nucleic Acids Res.* **32** (database issue)), D23–26 (2004).
13. Birney, E. *et al.* Ensembl 2004. *Nucleic Acids Res.* **32** (database issue), D468–470 (2004).
14. Antequera, F. & Bird, A. Number of CpG islands and genes in human and mouse. *Proc. Natl Acad. Sci. USA* **90,** 11995–11999 (1993).
15. Ohler, U., Liao, G. C., Niemann, H. & Rubin, G. M. Computational analysis of core promoters in the *Drosophila* genome. *Genome Biol.* **3,** RESEARCH0087 (2002).
16. Schubeler, D. *et al.* The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes Dev.* **18,** 1263–1271 (2004).
17. Griffiths-Jones, S. The microRNA Registry. *Nucleic Acids Res.* **32** (database issue), D109–111 (2004).
18. International Human Genome Sequencing Consortium, Finishing the euchromatic sequence of the human genome. *Nature* **431,** 931–945 (2004).
19. Bertone, P. *et al.* Global identification of human transcribed sequences with genome tiling arrays. *Science* **306,** 2242–2246 (2004).
20. Kampa, D. *et al.* Novel RNAs identified from an in-depth analysis of the transcriptome of human chromosomes 21 and 22. *Genome Res.* **14,** 331–342 (2004).
21. Saha, S. *et al.* Using the transcriptome to annotate the genome. *Nature Biotechnol.* **20,** 508–512 (2002).
22. Rinn, J. L. *et al.* The transcriptional activity of human chromosome 22. *Genes Dev.* **17,** 529–540 (2003).
23. Su, A. I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl Acad. Sci. USA* **99,** 4465–4470 (2002).
24. Spellman, P. T. & Rubin, G. M. Evidence for large domains of similarly expressed genes in the *Drosophila* genome. *J. Biol.* **1,** 5 (2002).
25. Roy, P. J., Stuart, J. M., Lund, J. & Kim, S. K. Chromosomal clustering of muscle-expressed genes in *Caenorhabditis elegans*. *Nature* **418,** 975–979 (2002).
26. Caron, H. *et al.* The human transcriptome map: clustering of highly expressed genes in chromosomal domains. *Science* **291,** 1289–1292 (2001).
27. Maniatis, T. & Reed, R. An extensive network of coupling among gene expression machines. *Nature* **416,** 499–506 (2002).
28. Krumm, A., Hickey, L. B. & Groudine, M. Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation. *Genes Dev.* **9,** 559–572 (1995).
29. Ambros, V. The functions of animal microRNAs. *Nature* **431,** 350–355 (2004).
30. Yuh, C. H., Bolouri, H. & Davidson, E. H. Genomic *cis*-regulatory logic: experimental and computational analysis of a sea urchin gene. *Science* **279,** 1896–1902 (1998).

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

Acknowledgements We thank J. Kadonaga, R. A. Young, R. Kolodner, W. K. Cavenee, S. Van Calcar and C. K. Glass for discussion and comments on the manuscript. This research was supported by a Ruth L. Kirschstein National Research Service Award (T.H.K.) a Ford Foundation Predoctoral Fellowship (L.O.B.); the Ludwig Institute for Cancer Research (B.R.); NIH grants (B.R.) and the NSF (Y.W.).

Author Contributions B.R. and T.H.K. conceived the experimental design; T.H.K. performed the experiments; data analysis was by L.O.B. and C.Q.; microarray fabrication, hybridization and data acquisition were by M.A.S., T.A.R. and R.D.G.; M.Z. and Y.W. worked on the computational peak detection program; writing of the manuscript was primarily by T.H.K. and B.R.

Author Information The microarray data sets are available from GEO (Gene Expression Omnibus) under accession number GSE2672, and from http://licr-renlab.ucsd.edu/download.html. Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare competing financial interests: details accompany the paper on www.nature.com/nature. Correspondence and requests for materials should be addressed to B.R. (biren@ucsd.edu).

CORRIGENDUM

doi:10.1038/nature03941

EphB receptor activity suppresses colorectal cancer progression

Eduard Batlle, Julinor Bacani, Harry Begthel, Suzanne Jonkheer, Alexander Gregorieff, Maaike van de Born, Núria Malats, Elena Sancho, Elles Boon, Tony Pawson, Steven Gallinger, Steven Pals & Hans Clevers

Nature 435, 1126–1130 (2005)

In the author list of this paper, Suzanne Jonkheer's surname was incorrectly spelled 'Jonkeer'.

CORRIGENDUM

doi:10.1038/nature04044

Similar response of labile and resistant soil organic matter pools to changes in temperature

Changming Fang, Pete Smith, John B. Moncrieff & Jo U. Smith

Nature 433, 57–59 (2005)

The *x* axis of Fig. 1 of this Letter was mislabelled. The correct tick labels (from left to right) should read: 'Root-free soil 0–10 cm'; 'Intact soil 0–10 cm'; 'Root-free soil 20–30 cm'; and 'Intact soil 20–30 cm'. These errors do not affect any of our conclusions.

ERRATUM

doi:10.1038/nature04042

An integrated view of the chemistry and mineralogy of martian soils

Albert S. Yen, Ralf Gellert, Christian Schröder, Richard V. Morris, James F. Bell III, Amy T. Knudson, Benton C. Clark, Douglas W. Ming, Joy A. Crisp, Raymond E. Arvidson, Diana Blaney, Johannes Brückner, Philip R. Christensen, David J. DesMarais, Paulo A. de Souza Jr, Thanasis E. Economou, Amitabha Ghosh, Brian C. Hahn, Kenneth E. Herkenhoff, Larry A. Haskin, Joel A. Hurowitz, Bradley L. Joliff, Jeffrey R. Johnson, Göstar Klingelhöfer, Morten Bo Madsen, Scott M. McLennan, Harry Y. McSween, Lutz Richter, Rudi Rieder, Daniel Rodionov, Larry Soderblom, Steven W. Squyres, Nicholas J. Tosca, Alian Wang, Michael Wyatt & Jutta Zipfel

Nature 436, 49–54 (2005)

Reference 34 should have read: "34. Goetz, W. *et al.* Indication of drier periods on Mars from the chemistry and mineralogy of atmospheric dust. *Nature* doi:10.1038/nature03807 (this issue)." The HTML version has been corrected online.

ADDENDUM

doi:10.1038/nature04075

Evidence for magmatic evolution and diversity on Mars from infrared observations

P. R. Christensen, H. Y. McSween Jr, J. L. Bandfield, S. W. Ruff, A. D. Rogers, V. E. Hamilton, N. Gorelick, M. B. Wyatt, B. M. Jakosky, H. H. Kieffer, M. C. Malin & J. E. Moersch

Nature **436**, 504–509 (2005); advance online publication, 6 July 2005 (doi: 10.1038/nature03639)

We omitted to cite as a 'Note added in proof' two recently published papers[50,51] that independently identified pyroxene- and olivine-bearing rocks, making use of data from the Mars Express spacecraft.

50. Bibring, J. P. *et al.* Mars surface diversity as revealed by the OMEGA/Mars Express observations. *Science* **307**, 1576–1581 (2005).
51. Mustard, J. F. *et al.* Olivine and pyroxene diversity in the crust of Mars. *Science* **307**, 1594–1597 (2005).

THE CAREERS MAGAZINE FOR SCIENTISTS Vol 436 | Issue no. 7052 | 11 August 2005

FOCUS
SPOTLIGHT
RECRUITMENT
ANNOUNCEMENTS
EVENTS

naturejobs

Searching questions

During the Internet boom of the mid-1990s, information technology (IT) researchers could practically write their own ticket. Then the bubble burst, and many IT experts found themselves scanning the 'help wanted' adverts. Although the halcyon days are history, demand for IT experts may once more be on the rise, this time from search-engine companies such as Yahoo and Google.

Last month, both firms made aggressive moves to attract top talent from computer giants IBM, Apple and Microsoft, and there are signs of more to come. Yahoo recruited Prabhakar Raghavan, who helped IBM develop a search technology called 'Clever', and another senior IBM technology researcher Andrew Tomkins. It also took on Apple's Larry Tessler in a bid to help the company become more user friendly. Google, meanwhile, attracted Kai-Fu Lee from Microsoft to help it build a laboratory in China.

Such moves don't come without some retribution. Microsoft warned Lee that he may be in violation of a 'no-compete' clause in his contract, because he has key knowledge about Microsoft technology that is still in the works. And the former IBM employees now at Yahoo have said they need to tread carefully around intellectual-property laws.

Although these moves are notable for their high-profile nature, questions remain over whether IT jobs as a whole will recover, or whether this is just a quick flurry of activity limited to a few senior positions at a few companies. Will Apple, Microsoft and IBM fight back by hiring more search-engine scientists? Will Google and Yahoo expand their teams further, creating even more opportunities for IT experts? And will smaller search companies enter into the hiring fray?

Whatever the outcome, researchers looking for employment in the field should temper their optimism with caution, lest their bubble burst again.

Paul Smaglik, *Naturejobs* editor

CONTACTS
Publisher: Ben Crowe
Editor: Paul Smaglik
Marketing Manager: David Bowen

US Head Office, New York
345 Park Avenue South, 10th Floor,
New York, NY 10010-1707
Tel: +1 800 989 7718
Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager/Corporations:
Peter Bless
Classified Sales Representatives
Tel: +1 800 989 7718

New York/Pennsylvania/
Latin America: Kelly Roman
Midwest USA/Maryland/
NIH: Wade Tucker
East USA/Canada:
Janine Taormina

San Francisco Office
Classified Sales Representative:
Michaela Bjorkman
West USA/West Corp. Canada
225 Bush Street, Suite 1453
San Francisco, CA 94104
Tel: +1 415 781 3803
Fax: +1 415 781 3805
e-mail: m.bjorkman@naturesf.com

European Head Office, London
The Macmillan Building,
4 Crinan Street,
London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

***Naturejobs* Sales Director:** Nevin Bayoumi (4978)
European Sales Manager: Andy Douglas (4975)

Advertising Production Manager: Billie Franklin
To send materials use London address above.
Tel: +44 (0) 20 7843 4814
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

***Naturejobs* web development:** Tom Hancock
***Naturejobs* online production:** Niamh Shields

European Satellite Office
Patrick Phelan
e-mail: p.phelan@nature.com

Japan Head Office, Tokyo
Chiyoda Building,
2-37 Ichigayatamachi,
Shinjuku-ku,
Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746
Asia-Pacific Sales Director: Rinoko Asami
e-mail: r.asami@naturejpn.com

NATURE|Vol 436|11 August 2005

Breaking open a closed system

Malaysia's research system is closed and isolated. What are scientists with a yen for rigorous research to do? **David Cyranoski** finds out.

In March 2005, neuroscientist Ishwar Parhar had coffee with the king of Malaysia at a Tokyo hotel. "He asked me why I don't come back to Malaysia," Parhar remembers. A desire to do more research than the infrastructure would allow had driven him away years before, and he had had no indication that the government was serious about plans to strengthen science. Months later, officials from Malaysia's science and health ministries negotiated to let Parhar, currently at Nippon Medical School in Tokyo, build a new brain-science research institute in the capital, Kuala Lumpur. "The meeting with the king was good," he says. "I think it got the ministries to talk to me."

But not everyone can meet the king, or have a new institute set up to back a recruitment drive. Outsiders in Malaysia have difficulty finding attractive research positions. Cronyism is rife. Foreigners aren't the only ones to miss out: policies designed to promote development of the ethnic Malay majority put the affluent, educated Chinese and Indian minorities at a disadvantage in education and employment — and in the forefront of the brain drain from Malaysia.

Pay is unattractive, and a dearth of colleagues and postgraduates to collaborate or discuss research with makes this an isolated world. There is little pressure on researchers to publish in international journals. Doing little to attract researchers and much to drive them away, Malaysia has a serious human-resource problem. Efforts to create a bioindustrial hub flopped mainly for that reason (see *Nature* **436,** 620–621; 2005).

But the government is trying to encourage change. In 2003, for example, it partially reversed a 30-year-old policy, created to strengthen the ethnic Malay majority, of having all courses taught in Malay; mathematics and science are now taught in English. This has removed an extra impediment to Chinese and Indian scholars and is helping to promote Malaysian scientists' ability to interact internationally.

Top heavy? Kuala Lumpur's high-tech city centre belies a weak science base.

Y. LIU/CORBIS

Private enterprises

The government's decision in 1996 to allow private universities is shaking things up even more. The institutions that are now emerging could form a much-needed bridge to the international research community.

"It was a bold initiative to make Malaysia a centre of excellence," says inorganic chemist Kumar Das, vice-chancellor of the science department at the private Asian Institute of Medicine, Science and Technology (AIMST).

The new private institutions — including AIMST, Universiti Tunku Abdul Rahman (UTAR), Monash University Malaysia, Penang Medical College and the International Medical University — are expanding. UTAR's science and technology faculty, established only a year and a half ago, has 65 staff and 1,500 students; plans are to boost these numbers to 400 and 7,000, respectively, over the next four years. Mostly with private support, all are constructing large campuses.

The private universities and colleges were formed to fill a gap in the Malaysian education system created by a 'Malays first' policy. The policy guarantees places for ethnic Malays, reducing the options for the Chinese and Indian minorities. To offer more opportunities to home-grown talent, Chinese Malaysians with political power established UTAR and their Indian counterparts set up AIMST. Admission is open to students of all ethnic groups through a purely merit-based selection process. All classes are taught in English.

But the effect of these institutions is likely to reach well beyond offering places to minorities. They have scientific ambitions and internationalism on the brain. A quarter of the staff at Monash are foreign, mostly recruited from the region: these universities actively

seek staff from abroad or with training overseas. And the administration is creating collaborations with overseas universities.

Furthermore, the new institutes put an emphasis on cutting-edge research, such as brain science at Monash and organometallic chemistry at UTAR. They send staff to international academic conferences and encourage publication in international journals.

"Right now the emphasis on research is not high," says Tham Choy-Yoong, dean of the faculty of science and engineering at UTAR. An astrophysicist formerly at the University of Cambridge's Cavendish Laboratory, UK, Tham says he wants to introduce what he learned there. "In Britain, research is everything. It's more important than stressing how many students you can train."

A Malaysian developmental biologist who recently moved from a national university, and asked not to be named, says the independent institutes have a different philosophy. "The problem with the national universities is that there is too much central control, so the area of innovation is very limited. Universities should be a place to experiment."

Cash flow

That road will not be easy, as funding and human resources will continue to be a problem. The private universities only achieved the right to get grants from the main funding source, Intensification of Research in Priority Areas grants, two months ago. But even these have their problems, says Merilyn Liddell, vice-chancellor of Monash Malaysia. She is particularly concerned that grants do not provide for major infrastructure or salary. "From the university's perspective, it costs money to get the grant," she says.

Moreover, the grants specify priority areas, focusing on applied science over basic research, says Liddell. Many researchers in Malaysia echo her concern, saying that R&D investment seems focused on schemes to turn a quick profit rather than on long-term development. Much of the country's national research is on rubber, forestry and palm oil.

The grant system still causes concern, because the peer-review ideal is often replaced by patronage. "In Malaysia, the pool of people reading the grants is so small, they will clearly know who has written them," says Parhar. In a place where cronyism from the top down is rife, and ethnic lines clearly demarcated, the toll on merit-based assessment can be high.

This February, Monash Malaysia provided a report for the government advising it to put millions of dollars into basic-research activities at foreign-branch universities. "By training people in research, we could contribute to the human capital of the country," says Liddell.

The new universities are struggling to provide competitive salaries so that established scientists can be recruited to form a research core. One foreign scientist working in Malaysia says she couldn't even pay rent on her salary.

A shortage of colleagues and staff is also a lingering concern. There's a dire need for postgraduate students, says Das, "but Malaysian postdocs get only a small stipend, and foreigners get none". Parhar is bringing six of his own postdocs — four Japanese and two Chinese — to help set his institute up. But wholesale importing won't be an option for many.

Brain work: Ishwar Parhar intends to set up a centre of excellence in neuroscience.

D. CYRANOSKI

Taking the long view: the Asian Institute of Medicine, Science and Technology has ambitious expansion plans.

Malaysia's famed biodiversity could be a great draw for researchers. But even that has been tainted by bad policies, say researchers. One foreign ecologist says that environmental-protection laws are not enforced, and considerable red tape must be hacked through before researchers can get out to the rainforests.

Meanwhile, the pool of scientists who could collaborate on biodiversity projects has shrunk. "Taxonomists are an endangered species," says Abdul Razak Mohd Ali, head of the Forest Research Institute of Malaysia, which is heading a national project on a biodiversity inventory. Many blame late-1990s policies designed to cultivate biotechnologists while supplanting taxonomists, but attitudes seem to be changing (see *Nature* **436,** 313; 2005).

If the new institutes manage to attract top researchers, the Malaysian research community could develop basic and applied research capacity in line with international trends, as neighbouring Singapore has done (see *Nature* **425,** 746–747; 2003).

Parhar is hopeful. He plans, among other things, to establish an active neuroscience and endocrinology society that could host the International Brain Research Organization's annual meeting.

"I want this to become a regional centre of excellence in neuroscience," says Parhar. "Once we show good-quality research, we will become attractive to scientists in the region and even to some in the United States and Europe."

David Cyranoski is *Nature*'s Asia-Pacific correspondent.

WEB LINKS

Universiti Tunku Abdul Rahman
www.utar.edu.my
Monash University, Malaysia
www.monash.edu.my
Asian Institute of Medicine, Science and Technology
www.aimst.edu.my
Penang Medical College
www.pmc.edu.my
International Medical University
www.imu.edu.my

Correction

A story in *Naturejobs* about England's Golden Triangle (*Nature* **436,** 144–147; 2005) stated incorrectly that Millennium Pharmaceuticals left Cambridge after six weeks. Millennium had a presence in Cambridge during 2000–03, but decided to leave the country six weeks after moving into a new facility at Cambridge's Granta Park. The story also incorrectly said that no Cambridge area firm floated on the stock market in 2004. Millennium spin-out Sareum went public in October 2004, floating on the AIM market on the London Stock Exchange. *Naturejobs* regrets the errors.

NATURE|Vol 436|11 August 2005

MOVERS

Eric Staeva-Vieira, business analyst, Rodman and Renshaw, New York

2003–05: Programme manager, Science Alliance, New York Academy of Sciences
1998–2003: Researcher, Skirball Institute, New York University Medical Center, New York
1996–98: Senior research technician, Memorial Sloan-Kettering Cancer Center, New York

Most scientists approach their craft with a tight focus on a particular area of science. But a few, such as Eric Staeva-Vieira, take a broader view, trawling the vast sea of science to understand its role in society. In plotting his own course, Staeva-Vieira has helped make the less-travelled route a little easier for like-minded scientists.

When Staeva-Vieira was an undergraduate studying molecular biology, an adviser handed him a copy of *The Coming Plague* by Laurie Garrett, which sparked his interest in public health and emerging diseases. This led him to start a doctorate in parasitology at New York University's School of Medicine. While there, he became so intrigued with the impact that the Human Genome Project could have on drug and therapy development that he completed a PhD in developmental genetics.

To understand not only the scientific but also the ethical and social implications of the project, he sought the counsel of the university's famed science-and-society expert Dorothy Nelkin — a move that confirmed his broad view of science. "I think scientists should be more involved in politics, business and legal discussions," says Staeva-Vieira.

Eager to show the impact that non-bench scientists can have — and help graduate students and postdocs grapple with similar career issues — he took a position as manager of the Science Alliance career-development programme at the New York Academy of Sciences. There, he helped to create a successful business training course, and sought to dismiss the myth that career transitions require multiple degrees. Courses in finance and economics helped to prepare him for his role at Rodman and Renshaw, an early-stage technology-investment firm in New York.

"My goal is to get an inside view of industry," he says, adding that gaining insight into early-stage technologies, such as therapies using RNA interference, is exciting.

"I look for opportunities," Staeva-Vieira says. "I'm constantly trying to figure out what skill set I should get next." His ultimate goal, he says — combining all the skills he's honed throughout his career — will probably be a position in the nexus between the non-profit and for-profit world, seeking solutions to global health problems.

A firm believer that scientists can contribute more than R&D, he says that the decision to pursue a non-traditional track is a difficult one, fraught with fear of the unknown. "The challenge," he says, "is staying true to your gut." And recognizing when, if you stayed on the traditional path, you'd be a fish out of water. ■

RECRUITERS & ACADEMIA

A class act

When the women's studies department at Duke University asked us — two graduate students in the environmental sciences programme — if we could design and teach a class about gender and environmental science, we jumped at the chance. Given how much coverage there has been about women in science, we were curious to explore how gender might affect environmental problems and solutions.

There were three other reasons we pursued this opportunity. First, we wanted to experience the entire teaching process from designing to teaching to evaluation. We had worked as teaching assistants, but were involved mainly in administrative tasks. We wanted to discover how to balance research needs with preparing and teaching a class. We shared ideas and split up tasks, began preparations a year ahead, and used flexible periods, including the summer session and December.

Second, we wanted to experiment with the latest teaching techniques, including active learning, where the focus is on engaging students in the process. In our class we explored several methods such as class discussions, role-plays and small-group activities. A scenario about a community exposed to high levels of pesticides emphasized the complexity of environmental problems. In small groups, students prepared a budget to balance the community's medical needs with the longer-term benefits of toxicology research. This realistic environmental problem showed students how difficult it can be to meet both scientific and societal goals.

Third, we welcomed the challenge of teaching an interdisciplinary class. Environmental problems are complex because they involve the interaction between the environment and people. By the end of the semester the entire class agreed that gender plays an important role. A gender perspective can inform our analysis of the problems and help us carry out solutions.

From this teaching experience, we realized that perspectives from a variety of disciplines can provide context for the issues and should be included in environmental-science curricula.

Mid-way through our PhD programme, we had lost the larger context of our research. Exploring an interdisciplinary topic with enthusiastic students revitalized us. Teaching this class took a great deal of time and effort, but it was also one of our most rewarding and inspiring experiences. ■

Ariana Sutton-Grier and Melissa Kenney are graduate students at the Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina.

GRADUATE JOURNAL

A tale of a whale

It's been two months since I defended my PhD and I have temporarily escaped to a place where people don't ask where you're doing your postdoc: Alaska. With the hope of clearing our minds and catching some fish, another recent graduate and I headed up north. If you love the outdoors, there's just no better place than here.

So far, the trip is not pushing the reset button for my stresses in the way I thought it would, but perhaps that was too much to expect. I still wake up with mini-panics about this past year and the one to come. All neuroses aside, it has been an amazing journey and I'm feeding upon the wonders of nature everywhere we go. I've experienced a few salient moments of clarity too.

One day, I awoke stressed about the hiking boots I had left on a river bank and cursed my forgetfulness. Then, as I was kayaking in Tutka Bay near Homer, I watched a humpback whale leap from the brine in a full breach. It was an astounding sight and I was instantly healed from that little stress attack about some dumb pair of boots. I laughed as I realized that one can always buy new boots, but experiences like that are priceless. Hopefully I can re-enter my career remembering not to sweat about the small stuff, as there's big stuff like whales that deserve far more attention. ■

Jason Underwood completed his PhD in molecular biology at the University of California, Los Angeles, in June.

NATURE|Vol 436|11 August 2005

Prometheus unbound, at last

And not a moment too soon.

Kim Stanley Robinson

Please append your report here

This novel postulates that science is an ongoing utopian proto-political experiment poorly theorized as such and lacking a paradigm within which to exert power in human affairs commensurate with its actual productive capacity and life-maintenance criticality. Scientists are first seen marginalized from macro-decision-making in a backstory (written in the style of a Cold War thriller) in which agents sequester science by convincing Truman *et al.* that science's metastasizing wartime ability to create new technologies crucial to victory (radar, penicillin, atom bomb, etc.) might constitute a threat to postwar civilian-corporate control of society.

Scientists, subsequently inoperative in surplus value investment and allocation decisions, produce goods and services unconscious of themselves as a group and individually willing to work within the existing hierarchical extractive non-sustainable system for $100,000 ± 50,000 annually plus pension, stock options and a light teaching load. (This chapter is in the form of a zombie novel, highly amusing.)

Then the scientifically augmented human population catastrophically overshoots the long-term carrying capacity of the planet. Scientists in their various toothless non-decision-making organizations conclude that the anthropogenically initiated climate change, and mass extinction event associated with it, probably threatens their descendants' welfare, and thus scientists' own evolutionary fitness. The sleepers awake.

Meanwhile a certain proportion of humanity makes a cost–benefit analysis comparing 15 years' work learning a science with saying "I believe" and through group political action controlling more calories per capita than scientists do, also more power over funding and rather more offspring. Many conclude faith-based parasitism on science less costly to the individual, so more adaptive. (Vampires living off zombies, guns brandished, chases by night: the novel gets pretty lurid at this point.)

Then at a modelling conference a discussion springs up concerning Hamilton's rule, which states that altruism should evolve whenever the cost to the giver, C, is less than the fitness benefits, B, obtained by helping another individual who is related by r, with r being calculated as the proportion of genes these two individuals share by common descent (as in Hrdy, 1999): $C \leqslant Br$.

A geneticist at the conference points out that as humans share 60% of their genes with fruitflies, and all eukaryotes share 938 core genes, r is probably always higher than heretofore calculated. An ecologist mentions the famous *Nature* article in which the benefits provided by the biosphere to humans were estimated at $33 trillion a year (R. Costanza *et al. Nature* **387,** 253–260; 1997). An economist suggests that the cost for individual scientists wanting to maintain these benefits could be conceptualized in the form of a mutual hedge fund, with initial investment set for the sake of discussion at $1,000 per scientist.

JACEY

Comic scene here as modellers debate the numbers, with a biologist pointing out that the benefit of life to every living organism could justifiably be defined as infinity, considerably altering equation's results. (Shouting, fights, saloon demolished in Wild West manner.)

Conference attendees conclude altruism is probably warranted, and hedge fund is established. (Readers of novel wishing to pre-invest are directed to a website www.sciencemutual.net.) Participating scientists then vote to establish a board; a model constitution for all governments to adopt; a policy-research institute tasked with forming a political platform; and a lobbying firm. All scientific organizations are urged to join the fund. Fund's legal team goes to World Court to claim compensation for all future biospheric damage, to be paid into the fund by those wreaking the damage and the governments allowing it.

Many meetings follow, no doubt explaining the presence in this chapter of most of the novel's sex scenes. Author seemingly familiar with and perhaps overfond of the bonobo literature. Strenuous attempts to maximize reproductive success in Davos, Santa Fe, Las Vegas, etc.

Novel's style shifts to amalgam of legal thriller and tolkienesque high fantasy as scientists take power from corporate military-industrial global élite. A spinradian strategic opacity here obscures the actual mechanism that would allow this to work in the real world, said opacity created by deployment of complicated syntax, phrases low in semantic content ("information cascade"), especially active stage business (man runs through with hair on fire), explosions, car chases, and reinvocation of Very Big Numbers — in this case Science Mutual's potential assets if World Court returns positive judgment, after which subsequent chapter (with toll-free number as epigraph!) emerges in newly utopian space, looking plausible to those still suspended in coleridgean willed non-disbelief.

Speed of narration accelerates. Science Mutual arranges winners in all elections everywhere. Hedge fund continues to grow. Scientific organizations form international supra-organization. Black helicopters proliferate. Entire population decides to follow new scientific guidelines indicating that reproductive fitness is maximal the closer behaviour conforms to palaeolithic norms, this being the lifestyle that tripled brain size in only 1.2 million years. Widespread uptake of this behavioural set augmented by appropriate technology (especially dentistry) reduces global resource demand by an order of magnitude despite demographic surge to UN-predicted mid-range peak of ten billion humans. A rationally balanced positive feedback loop into maximized universal fitness obtains. (Novel ends with standard finale, singing, dancing, reproducing. All Terran organisms live optimally ever after.)

Please give your recommendation

Reader recommends acceptance for publication, but suggests that the apparent size of the text's strategic opacity be reduced to three seconds of arc or less. Publisher should take steps to secure domain name sciencemutual.com. (Also, more car chases.) ■

Kim Stanley Robinson's next novel *Fifty Degrees Below*, sequel to *Forty Signs of Rain*, will be published by HarperCollins in September.

FUTURES