Science

MAAAS

Epithelial cells defy fate

Committed cells can regain stem cell properties in regeneration and cancer *p. 1243*

CONTENTS

1234
An evo-devo framework
for innovation

CONTENTS

1229 & 1289
Epidemiological dynamics in a fungal-plant system

RESEARCH

1224 & 1262

1250

ON THE COVER

Immunofluorescence microscopy image of squamous cell carcinoma (SCC). SCCs contain tumor-initiating stem cells and their early progeny, which, like normal cells, reside along the tumor-stroma interface and express high levels of integrins (white) and progenitor keratins (green). Stem cells with high tumor-initiating capacity are shown in red; nuclear DNA is in blue. Average cell size within the tumor, 10 micrometers. See page 1243 and http://dx.doi.org/10.1126/science.1242281. *Image: Courtesy of Markus Schober and Elaine Fuchs [see Proc. Natl. Acad. Sci. U.S.A.* **108**, *10544–10549 (2011)]*

Science
Epithelial cells defy fate

1235

Published by AAAS

Time to speak up for research

As U.S. National Institutes of Health (NIH) Director Francis Collins stated in recent testimony to Congress, "Our nation has never witnessed a time of greater promise for advances in medicine." Researchers are moving closer to developing a universal flu vaccine, cancer immunotherapy (harnessing the immune system to attack tumor cells) is on the horizon, and the public/private BRAIN initiative will provide deeper knowledge of billions of nerve cells to advance research on Alzheimer's disease, autism, and other brain disorders. Yet the general public, and in particular elected officials, have failed to embrace the promise of cutting-edge science as a means to improve health and the economy.

Federal funding for research and innovation in the United States is on the decline; NIH's budget for fiscal year 2014 (FY14) is 11.7% below the FY04 peak.* The pain of budget cuts extends beyond the science community to the patients and families who await cures. With sequestration-level spending still the law of the land, new therapies that could save

"Scientists must take off their lab coats and engage the people of their communities and states."

lives, improve the quality of life, and reduce health care costs will remain out of reach. If a treatment became available in 2015 that delayed the onset of Alzheimer's disease by 5 years—similar to anticholesterol drugs preventing heart disease—annual Medicare and Medicaid spending would be $42 billion less by 2020.† Policy-makers who are determined to shrink the size of government should also remember that they have a responsibility to stabilize the economy and address costs that have nothing to do with the size of federal agencies. Cutting research is not a pathway to deficit reduction; it is a pathway to increased health threats, lost lives, and economic insecurity. The real money needed

to control the nation's long-term deficit can be found in the entitlement reform and tax reform issues that policy-makers are reluctant to tackle, even more so during an election year.

Inspiration is needed to bolster public appreciation for science and support for making federal funding for research a very high national priority. We must seize the opportunity to cultivate more champions for research. Yet there has been little outreach by scientists to the public to help them understand how science contributes to better health, job creation, and global competitiveness. Few Americans can name a living scientist or an institution where research takes place, according to polling commissioned by Research!America.‡ To put it bluntly, scientists remain largely invisible to the public. Yes, this says a great deal about the nation's people, but it says even more about scientists and their lack of engagement with the public.

The midterm elections this year provide an opportunity for scientists to become stronger advocates for research and converse with citizens to a much greater degree. Attend campaign debates and speak up about the health and economic benefits of research. Offer to be a scientific advisor for candidates or help them create a science advisory committee that can play a critical role in shaping policy. Write op-eds and letters to the editor about the latest medical breakthroughs and their implications for treating and preventing disease. Volunteer to speak at local organizations and chambers of commerce.

Scientists must take off their lab coats and engage the people of their communities and states. They must be willing to defend and spread the good news about science. If scientists themselves are unwilling to defend science, how can we expect others to do so?

– **John Edward Porter**

John Edward Porter is a former U.S. congressman, a partner in the law firm of Hogan Lovells, and chair of Research!America.
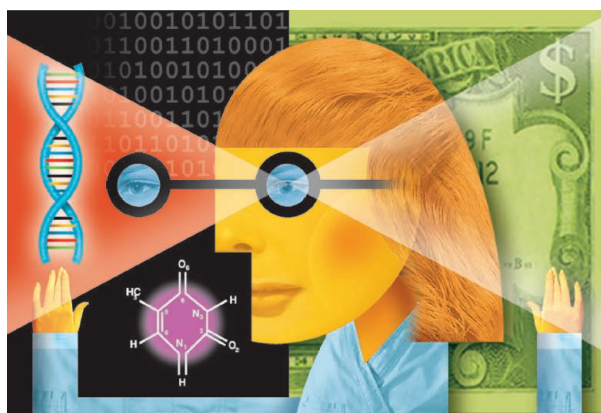
IMAGES: (RIGHT) MICHAEL MOENNING; (ILLUSTRATION) GORDON STUDER

*www.aaas.org/sites/default/files/NIHBud_1.jpg. †www.alz.org/documents_custom/trajectory.pdf. ‡www.researchamerica.org/uploads/MostAmericansCantNameaLivingScientist.pdf.

10.1126/science.1256929

IN BRIEF

Price hike for brain initiative
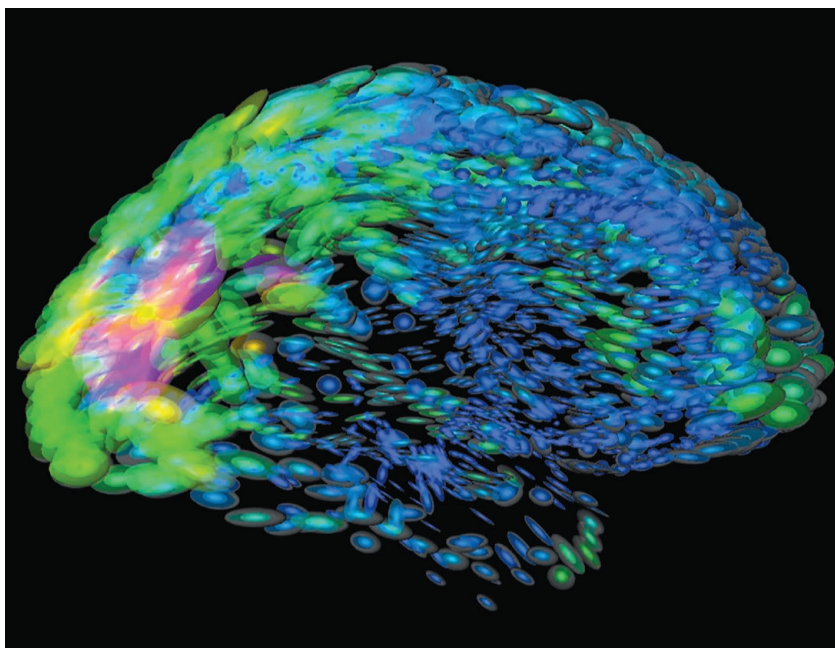
Similarities (blue ovals) and differences (other colors) in brain structure and activity between two people.

The price tag for the White House's project to map the human brain in action, the Brain Research through Advancing Innovative Neurotechnologies (BRAIN) Initiative, just skyrocketed. The president allocated $100 million this year, spread over the National Institutes of Health (NIH), the National Science Foundation, and the Defense Advanced Research Projects Agency, to kick-start BRAIN. Last week, an NIH-convened working group offered a more realistic appraisal of the funding needed: $4.5 billion over a decade for the agency's portion of the project. The report lays out a plan for NIH to invest $300 million to $500 million each year, beginning in fiscal year 2016, to develop new tools to monitor and map brain activity and structure. Getting the extra money "won't be fast, it won't be easy, and it won't be cheap," NIH Director Francis Collins told reporters on 5 June. The first round of requests for NIH grant applications went out last fall; awardees will be announced in September. http://scim.ag/BRAINprice

Stem cell patent case thrown out

WASHINGTON, D.C. | A consumer advocacy group's attempt to invalidate a patent on embryonic stem cells held by the Wisconsin Alumni Research Foundation failed last week when an appeals court dismissed the case. Santa Monica, California–based Consumer Watchdog (CW) hoped to invalidate the patent, awarded in 2006, which it claims puts a burden on California's taxpayer-funded research. When the patent office rejected its petition last year, CW filed a case with the U.S. Court of Appeals for the Federal Circuit, arguing that the cells are unpatentable "products of nature" and that their isolation was an "obvious" step (*Science*, 24 January, p. 359). But on 4 June, the court found that CW lacks legal standing to bring the case because it wasn't harmed by the patent. CW could appeal to the Supreme Court, but the patent is set to expire in 2015. http://scim.ag/WARFout

Female scientists are awesome

BILLUND, DENMARK | Your child could soon clamor for a LEGO play set featuring a trio of female scientists, thanks to geochemist Ellen Kooijman, whose "Research Institute" design—featuring a paleontologist, a chemist, and an astronomer—was approved by The Lego Group last week. A longtime LEGO enthusiast, Kooijman, of the Swedish Museum of Natural History in Stockholm, says she "noticed two things about the available LEGO sets: a skewed male/female ratio and a stereotypical

Ellen Kooijman's LEGO paleontologist.

representation of the available female figures." She designed a set intended to "make LEGO communities more diverse" and submitted it to the Denmark-based toy company's LEGO Ideas blog last year; the collection received 10,000 votes from the fan-based online community, earning it a review from LEGO. "Diversity is beneficial for science and technology"—and diversifying the roles of women in the toy market can help with that, Kooijman says. The completed set will be available in stores in August.

Venus Express taps the brakes

PARIS | A European spacecraft is about to take a dip into the sulfurous atmosphere of Venus, 8 years after it began orbiting the planet. Between 18 June and 11 July, the Paris-based European Space Agency's Venus Express will conduct an aerobraking campaign that brings the spacecraft into the upper reaches of the atmosphere. During the descent, the closest approach of the satellite's elliptical polar orbit has been shrunk from 250 kilometers above the planet's north pole to 130 kilometers—where the probe will feel the friction of the atmosphere. Along with scientific observations, the campaign is a way to learn about the effects of aerobraking, which is one way for a spacecraft to conserve fuel when entering orbit around a planet. If the probe
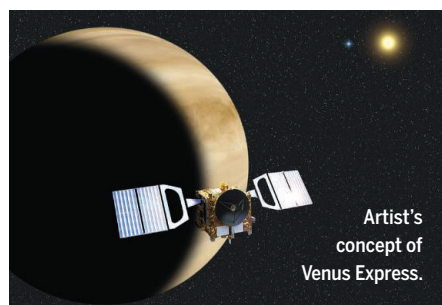
Artist's concept of Venus Express.

survives the campaign, mission managers will use any remaining fuel to boost it to a higher orbit; it is then expected that the spacecraft will plunge to its doom by the end of the year.

Germany bows out of giant array

BERLIN | Shocking its partners, Germany has withdrawn from an international collaboration to build the €2 billion Square Kilometre Array (SKA), the world's biggest radio telescope, stating that it will end its participation in June 2015. The announcement "came out of the blue," says SKA Director General Philip Diamond. The 20-country project will create a single huge telescope from thousands of individual dishes and antennas across southern Africa and Australia, to test relativity and study galactic evolution. Germany's contribution has been

small so far, at €3.8 million, but financial priorities appear to be the reason for its withdrawal. The country is in the process of building two large international facilities on its territory: the European XFEL x-ray laser facility and FAIR, an accelerator center for nuclear physics. http://scim.ag/GermanySKA

NEWSMAKERS

Three Q's

In 2001, physicist **Neil Gershenfeld**, who heads the Massachusetts Institute of Technology's (MIT's) Center for Bits and Atoms in Cambridge, created a few digital fabrication workshops. That spawned an international network of "fab labs," helping inspire the thriving maker movement—including the White House's first Maker Faire on 18 June. Gershenfeld spoke to *Science* in an interview edited for clarity.

Q: **What is digital fabrication?**
A: Digitizing materials rather than just designs. We're developing tabletop [micro]chip fab by assembling bricks of electronic materials, and a year ago in
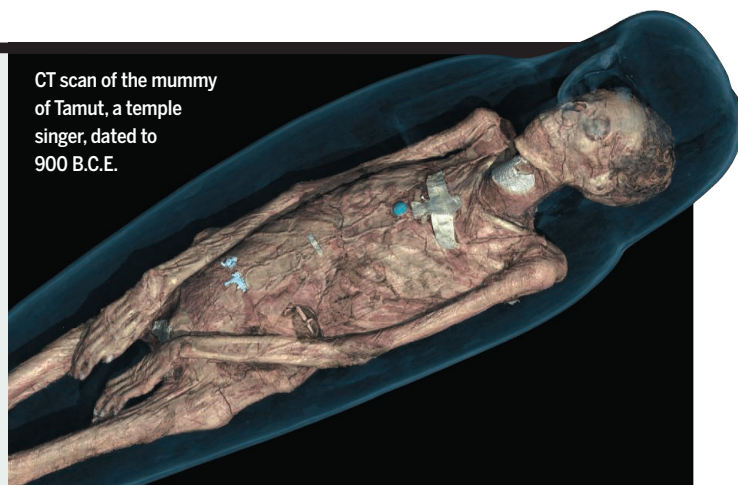
CT scan of the mummy of Tamut, a temple singer, dated to 900 B.C.E.

RANDOM SAMPLE

Mummies, (virtually) unwrapped

Think if you've seen one mummy, you've seen them all? An exhibit at London's British Museum challenges that assumption by delving into the lives of eight mummified individuals—without removing a single bandage. In 200 years of studying the 120 mummies in its collection, museum scientists have never unwrapped one, for fear of damaging fragile human remains. Instead, the new exhibit, called *Ancient Lives*, uses videos from a CT scanner as it virtually slices through the bodies. The ancient eight, who lived along the Nile River in Egypt and Sudan, hailed from time periods ranging between 3500 B.C.E. and 700 C.E. and from many different walks of life. They include an ordinary man preserved in the desert sands; the daughter of a high-ranking priest from Thebes; a temple doorkeeper and a temple singer, both also from Thebes; and a medieval Christian woman with a tattoo of the Archangel Michael on her thigh. The exhibit, which opened last month, continues until 30 November.

Skeleton of juvenile *Platypterygius hauthali*, an early Cretaceous ichthyosaur.

An ichthyosaur graveyard

Ten years ago, in the wake of a receding glacier in Torres del Paine National Park near the southern tip of Chile, scientists discovered a giant cache of fossils of ichthyosaurs—Mesozoic-era marine reptiles that were contemporaries with the dinosaurs. A German-Chilean research team led by geoscientist Wolfgang Stinnesbeck of Heidelberg University in Germany found 46 articulated, nearly complete specimens (including some soft tissues and embryos). The remains belong to four different species and date to the early Cretaceous, about 125 million years ago, the researchers reported last month in the *Geological Society of America Bulletin*. The scientists paint a rich picture of how the cache came to be: The ichthyosaurs, they suggest, lived at the edge of a deep sea that once separated Antarctica from Patagonia, hunting squid and fish in an underwater canyon. As the continent broke apart, underwater earthquakes and mudflows sucked the air-breathing ichthyosaurs deep into the ocean, burying them in sediment.

Science we showed how to make the world's highest performance, ultralight composites by reversibly linking loops of carbon fiber.

Q: How might digital fabrication impact science beyond manufacturing?
A: Bringing the campus to the student rather than the student to the campus. A fab lab can make other kinds of labs. Students in fab labs working with peers and mentors locally are connected globally. Rather than science limited by scarcity, it's a lot more accessible.

Q: How does your MIT class "How to Make (Almost) Anything" illustrate that?
A: We've expanded from on-campus teaching to hands-on teaching globally in the Fab Academy. Students create tools for molecular biology and electrical engineering. We're finding the kind of bright, inventive people who get attracted to a place like MIT in Arctic villages and shantytowns.

Schmidt to head climate lab

NASA has named climate modeler **Gavin Schmidt**, 46, to lead the Goddard Institute for Space Studies, an influential climate lab in New York City. "Schmidt has readily embraced the role of brash culture warrior … mixing it up with 'deniers' in the blogosphere," says communications expert Matthew Nisbet of American University in Washington, D.C. Will Schmidt, who co-founded climate blog RealClimate.com in 2004, follow in the footsteps of his predecessor, outspoken climatologist James Hansen? Schmidt says he'll steer clear of policy prescriptions—though, he says, scientists should differentiate between what science says "is" versus decisions they think society "ought" to make.
http://scim.ag/SchmidtGISS
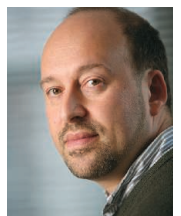
PHOTOS: (TOP TO BOTTOM) WOLFGANG STINNESBECK; BRUCE GILBERT
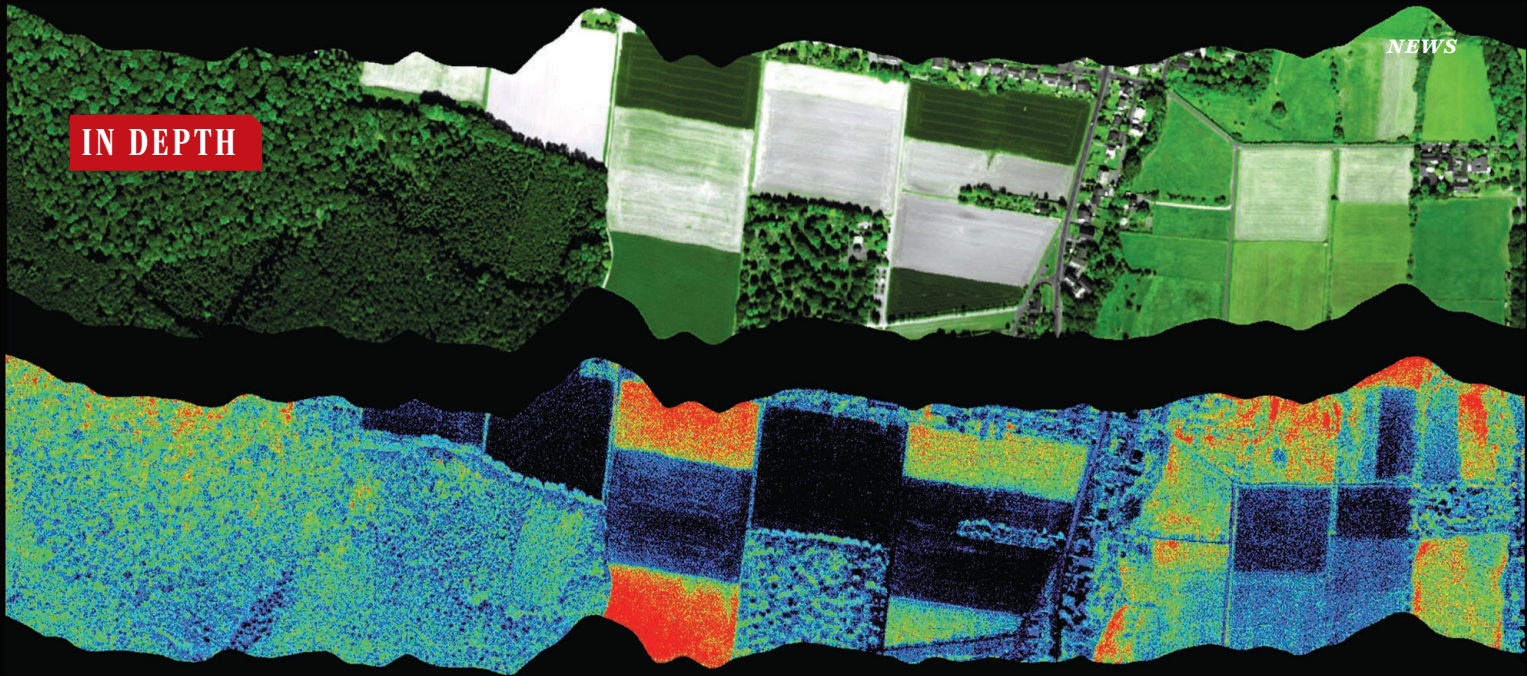
Chlorophyll fluorescence from crops can be detected, here by an aerial sensor called HyPlant.

REMOTE SENSING

Carbon-mapping satellite will monitor plants' faint glow

Launch of Orbiting Carbon Observatory-2 promises precise mapping of chlorophyll fluorescence from space

By **Eric Hand**

It's a lesson learned in grade school: Sunlight falls on the leaf, and, catalyzed by chlorophyll, atmospheric carbon dioxide is absorbed and fixed inside a sugar molecule. But there's also a startling extra credit factoid: The leaf re-emits 1% of that sunlight as a faint red glow.

Plant physiologists have known about chlorophyll fluorescence for decades. However, the ability to map this weak signal from space, through the confounding murk of Earth's atmosphere, has arrived only in the past few years. Now, scientists are about to get their clearest look yet at chlorophyll fluorescence—and with it, their best accounting of how vegetation worldwide soaks up carbon dioxide—with the launch of NASA's Orbiting Carbon Observatory-2 (OCO-2) on 1 July.

Fluorescence mapping was never part of the original rationale for the $468 million OCO-2, a replacement for a carbon-monitoring satellite that crashed into the ocean 5 years ago. But many on its science team say fluorescence mapping is now the mission's raison d'être. "In my opinion, this is the most innovative and revolutionary observation that satellites will make," says Mike Gunson, the mission's project scientist at the Jet Propulsion Laboratory (JPL) in Pasadena, California.

Other popular proxies for photosynthesis, such as greenness and leaf area, come with problems: An evergreen forest, for example, remains green year-round even though it takes up little carbon in winter. By contrast, a plant fluoresces only when photosynthesis is happening, so the glow directly reflects the amount of carbon that plants take up. Accurate maps of it could correct gaping inaccuracies in global carbon budgets and provide a new tool for assessing how ecosystems will behave under the drought and heat stresses of climate change. The ability to make such maps by satellite is one silver lining in the long wait for an OCO replacement, says JPL's David Crisp, the science lead for OCO-2 who was principal investigator for OCO. "We didn't even know we had a measure," he says. "This is going to tell us where the action is."

Minutes after launch in 2009, a protective clamshell failed to open on the Taurus XL rocket in which OCO rode, and the mission plunked down in the Pacific Ocean. The loss was devastating for climate scientists, who were counting on the satellite to make up for

spotty ground measurements of CO_2. OCO was designed to monitor a narrow column of air for absorption lines associated with CO_2, yielding a global map of its sources and sinks. Some scientists pointed out that the satellite could also help monitor whether countries were complying with emissions-reductions requirements in possible future international climate treaties.

Researchers rounded up political support for a replacement. By 2010, JPL had authorization and money to build OCO-2, with as few changes as possible. Then, in 2011, a similar failure of the Taurus XL—built by Orbital Sciences, of Dulles, Virginia—sent another NASA earth science mission, Glory, to a watery grave. NASA had had enough: It decided to move OCO-2 onto a Delta II rocket. That forced another delay on the OCO-2 team—but one that wasn't entirely unhelpful. In that time, the team learned about problems with a brand of spinning inertial wheels, used for pointing satellites, that crippled the planet-hunting telescope Kepler in 2013. The team had time to swap them out.

Meanwhile, chlorophyll fluorescence science came of age. In 2009, the Japanese space agency launched the Greenhouse Gases Observing Satellite (GOSAT). GOSAT can't map CO_2 in as much detail as OCO-2 will, but it has similar spectral resolution: the ability to tease apart reflected sunlight. Working with GOSAT data, several groups realized that to measure CO_2 amid the interference from clouds and aerosols, they would have to identify—and subtract—the fluorescence signal anyway. "As we say in our business, one person's noise is another's signal," says Joanna Joiner, a remote sensing scientist at NASA's Goddard Space Flight

Center in Greenbelt, Maryland, and one of the pioneers of the technique.

The CO_2-mapping ability of GOSAT and OCO-2 gets at the net exchange of CO_2 above a particular region. With the fluorescence signal, researchers can drill down farther and get the two components that make up the net exchange: carbon uptake through photosynthesis and carbon losses through respiration. And they can watch how these factors evolve with time under different climate conditions. Already, GOSAT data has provided clues: a study published in March in the *Proceedings of the National Academy of Sciences* found that the U.S. corn belt has a peak fluorescence brighter than anywhere else in the world—and that climate models may be underestimating the region's carbon uptake by 50% to 75%.

Fluorescence measurements could also help resolve a long-standing debate about how the Amazon rainforest responds to droughts—which are expected to occur more often as the world warms. Some scientists have argued that photosynthesis in the Amazon's dense canopy is limited not by water but by light, so increased sunshine during droughts can lead to a "greening up." A 2013 paper in the *Proceedings of the Royal Society B* based on GOSAT data suggests otherwise: During the dry season in some parts of the Amazon, fluorescence—and photosynthesis—went down, even as another proxy for production, the leaf area index, peaked.

The applications could extend well beyond climate science. OCO-2's resolution is about 2 kilometers. With subkilometer spatial resolution, a fluorescence mapper could assess the productivity of different crops in patchwork fields. That could lead to a more accurate picture of global crop yields and how they respond to drought or heat waves—valuable information for adapting to climate change. "Are we planting the right crops in the right place in a changing climate?" asks Matthias Drusch, the European Space Agency (ESA) project scientist for the Fluorescence Explorer (FLEX), a proposed mapper with 300-meter resolution. The €100 million FLEX is one of two finalists competing for an ESA mission selection in 2015.

Crisp knows firsthand how hard it can be to see these missions to fruition. He says he has been working 80-hour weeks ever since first proposing OCO to NASA in 2001. Though he feels some trepidation about the upcoming launch, he plans on watching it from the launch site at Vandenberg Air Force Base in California. "We've done everything we can," he says. "God, I hope we get it right. I really want to get it up there. We've just got to do this." ∎

INFECTIOUS DISEASES

Polio eradicators struggle to prevent the next outbreak

A new "Red List" identifies the countries where the virus may strike next

By **Leslie Roberts**

They should have seen it coming. Instead, the Global Polio Eradication Initiative (GPEI) was caught flat-footed by the massive outbreak last year in Somalia, Ethiopia, and Kenya, countries where the disease had previously been eliminated. At least that is the opinion of an influential board of outside experts tasked with keeping the eradication effort on track.

that's a major reason, along with killings of polio workers and turmoil in Pakistan, that the goal of stopping all polio transmission in 2014 is at "extreme risk."

"We got hammered," says Bruce Aylward, who runs GPEI as an assistant director-general at the World Health Organization (WHO) in Geneva, Switzerland. But he admits that the May report is "100% fair."

Under the harsh language lies a very real question: Just how good are the prognostications, and even if they are right, is it re-

The Red List

Polio outbreaks waiting to happen

- Red List
- Infected
- At risk*

*Vulnerable, but coordinated response already under way

Source: The Global Polio Eradication Initiative
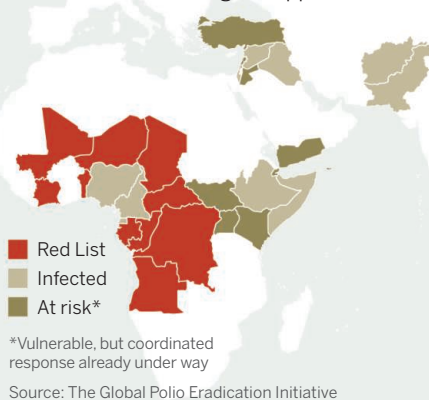
RED LIST COUNTRY	PROBABILITY OF IMPORTATION	POTENTIAL SEVERITY
Angola	Medium	High
Benin	Medium	High
CAR	High	High
Chad	High	High
Congo	High	Medium
DR Congo	Medium	High
Gabon	High	High
Ivory Coast	Medium	High
Mali	High	High
Niger	High	High

In its characteristic blunt style, the Independent Monitoring Board (IMB) had blasted the eradication program: "The outbreak in the Horn of Africa shows up the program's surprising disregard for the value of preventing outbreaks," it wrote in an October 2013 report. "In failing to address a plethora of red flags, it is almost as if the program has operated in the belief that it would 'remain lucky.'"

In response, GPEI last month unveiled a carefully honed "Red List"—a global top 10 list of countries at greatest risk of a devastating polio outbreak—which is supposed to help the eradication campaign focus its finite resources. But IMB still isn't satisfied.

In a report released in late May, it pushes GPEI to go further, calling the program's approach to preventing outbreaks "rudimentary," "slow," and "unsophisticated." And

ally possible to prevent new outbreaks in every polio-free country?

Much of GPEI's effort goes into chasing the polio virus from the only three countries where viral transmission has never stopped: Afghanistan, Pakistan, and Nigeria. But surprise outbreaks of the virus happen every so often in places where the disease was already wiped out and vaccination rates have fallen, costing the program precious time, money, and credibility. In addition to the three countries in the Horn of Africa, five others were hit last year, bringing the total case count to 407, up from 2012's historic low of 223.

For the past few years, WHO, the U.S. Centers for Disease Control and Prevention (CDC), and the Bill & Melinda Gates Foundation, which supports the Seattle, Washington–based modeling group Global Good,

have been trying to predict such emergencies. (The three are members of GPEI, along with Rotary International and the United Nations Children's Fund.) "The challenge is that outbreaks are actually extremely rare events, but with terrible consequences," says Nicholas Grassly, an epidemiologist at Imperial College London who has done his own outbreak modeling.

All three groups have developed models that score each country on two basic questions: How likely is it to be reinfected? And just how bad would an outbreak be? Lots of variables go into the equation, explains Gregory Armstrong, the lead person for polio eradication at CDC, whose team developed the agency's model. Is a country close to one of the three main holdouts or to a country where an outbreak is ongoing? Has imported virus sparked an outbreak there previously—and if so, from where? How much travel is there between it and an infected country? How many kids are un-

Tallying vaccinations in Somalia, where polio is back.

or underimmunized—a surprisingly tough number to get? Does conflict prevent vaccination teams from doing their work? Is the surveillance system good enough to detect a poliovirus importation right away, or only after it has taken off, as happened in Cameroon last year?

Until now, the three groups had come up with different but overlapping lists. Pushed by IMB, they sat down together to reconcile the differences in April. "There wasn't too much bloodshed," says Hamid Jafari, WHO's director of global polio eradication and research, who led the effort. "People are passionate about this."

At the top of the new Red List is the conflict-torn Central African Republic (CAR), which has low vaccination rates and is a neighbor to Cameroon, where the outbreak remains out of control. Other calls

were tougher. Ukraine did not make the list, despite IMB pushing for it in its October report. An outbreak there would definitely be severe because vaccination rates have fallen, but the models suggest that the virus is unlikely to land in Ukraine. Angola did not make the cut initially because the risk of an importation was deemed low; that changed when the modelers realized that oil workers travel frequently between Angola and Equatorial Guinea, where polio was detected in March. (The team excluded at-risk countries in the Horn and the Middle East where massive vaccination campaigns are already under way.)

GPEI knows what needs to be done in countries at risk: Start tightly supervised vaccination campaigns that aim to reach every child under age 5, for instance, make sure any sign of the virus is detected as soon as it gets in, and mount a speedy response if that happens. The program has already scheduled extra campaigns in all of the Red List countries.

But Aylward and others note that such campaigns can be challenging in Red List countries, which tend to be poor, corrupt, and violent. It is also hard to convince a country to invest money and time in a vaccination campaign when it hasn't seen polio in years, Jafari says.

The risk team also admits that outbreaks can occur in countries not on the list. "Absolutely, there will be surprises," Jafari says. "There is an element of randomness in where the virus goes." And sadly, some outbreaks simply can't be prevented, Aylward says. When polio resurfaced in Somalia, for instance, 1 million kids were at risk because of a ban on vaccination by the extremist group al-Shabab—and there was very little anyone could do about it.

IMB, however, thinks GPEI can do more: It calls on the program to set up a new team dedicated exclusively to outbreak prevention, and another devoted to outbreak response, where it says the program is also falling seriously short.

Formed in 2010, IMB has had harsh words for the initiative before (*Science*, 3 August 2012, p. 514). Although the reports can be difficult to swallow, people within GPEI say they take IMB's tough love to heart. "I don't think they are expecting us to change red to green in all these places," Aylward says. But the program must try harder. "This is the absolute clarity of expectations you need for an eradication program." ■

Racing for disaster?
Breeding thoroughbreds for speed may harm their health

By **Ann Gibbons**

When California Chrome narrowly missed becoming the first horse in 36 years to win racing's most prestigious prize, the Triple Crown, much was made of this thoroughbred's meteoric rise from humble origins. *The New York Times* declared the colt, whose winning streak ended last weekend at the Belmont Stakes, "common stock," born "on the wrong side of the tracks." His mother was an "unspectacular mare," and his father was "an equally unheralded stallion," according to *The Washington Post*.

But run a background check on California Chrome and it quickly becomes apparent that he comes from the same gene pool as all the other elite thoroughbreds running alongside him, with great-great-grandsires that included Kentucky Derby winner Northern Dancer and Triple Crowner Seattle Slew, and further back, the legendary Secretariat. "They're so inbred they're like purebred dogs," says Carrie Finno, a veterinarian at the University of California (UC), Davis. "Thoroughbreds are almost like clones compared to other breeds," adds Doug Antczak, an immunologist and vet who specializes in horses at Cornell University.

Recent genetic studies are finding that these animals are getting even more alike—and creating a potentially dangerous situation for the breed. With a smaller and smaller number of the fastest stallions siring more of the foals, some veterinarians think inbreeding is hurting thoroughbred stock. Meanwhile, hard times in racing have reduced the total number of new thoroughbred foals registered annually from a high of 51,000 in 1986 to 23,000 in 2013, further shrinking the gene pool. Finno calls the combination of these trends "alarming."

The resulting genetic bottleneck, she says, could make thoroughbreds defenseless against emerging infections and more apt to retain genes that cause disease, infertility, or malformed legs, for example. Other researchers dismiss Finno's concern, saying that breeding for performance has prevented the inheritance of many devastating diseases in thoroughbreds, because sick ani-

Thoroughbreds such as California Chrome (above in Belmont Stakes) are becoming increasingly inbred.

mals can't run as fast. "One of the breeds of dogs with the fewest diseases is the racing greyhound—they're bred in the same way as thoroughbreds," Antczak says. "This is not a fragile breed," agrees geneticist Ernest Bailey of the University of Kentucky in Lexington.

Thoroughbreds started out inbred: All trace their paternal origins to one of three stallions in England in the early 1700s. These days, stallions command stud fees that depend on their speed, earnings on the track, and how well they produce other winners. This system balanced durability with speed quite well until the mid-1980s, when stud fees soared to $1 million for Northern Dancer and year-old foals sold for as much as $13 million.

This kind of money led to a new kind of air traveler—"shuttle stallions" flown in for spring breeding season in different hemispheres. Some jet setter studs mate with 300 to 400 mares a year compared with up to 40 a year 40 years ago. This creates a sort of Genghis Khan effect, where some stallions dominate the gene pool, says Bailey, who with his colleagues reported the increase in inbreeding in 2012 in *Animal Genetics*.

Foals sired by top stallions command the highest prices at auction, so there's plenty of incentive for owners to keep paying high stud fees. But breeding for speed is not the same as selection for health or durability. Today's fastest thoroughbreds have larger muscles balanced on slimmer legs and smaller hooves, resulting in top-heavy animals whose limbs are more liable to break at high speed—as when Barbaro broke his hind limb in the 2006 Preakness Stakes and the filly Eight Belles broke both front ankles and

had to be euthanized after she crossed the finish line second in the Kentucky Derby in 2008. As inbreeding has increased, individual thoroughbreds start in fewer races and retire earlier, prompting widespread speculation that the horses aren't as durable.

With 700 thoroughbreds to minister to, Jeanne Bowers, the resident veterinarian at Harris Farms in Coalinga, California, where California Chrome was bred and raised, says she has seen it all—thoroughbreds whose leg bones chip at the joint, causing arthritis prematurely; whose lungs bleed at high speeds; who "roar" or wheeze when they run because their throat muscles freeze, making it hard to get enough air; and foals that are particularly susceptible to respiratory ailments. Infertility and miscarriages due to inbreeding have become a "huge" problem in thoroughbreds, she says.

Some breeders say there are fewer recessive diseases in thoroughbreds than in other breeds, but Finno challenges that, suggesting that research funding hasn't been readily available to find the relevant genes. "Everyone knows they're inbred," she says. "The question is, what are they going to do about it?"

Genetic technology could help. Since the sequencing of the horse genome in 2007, Finno says, "[t]here has been an explosion of research," leading to the discovery of the mutations responsible for 35 diseases and traits caused by single genes in horses. She is searching for the genes responsible for respiratory diseases in young thoroughbreds, as well as for neuroaxonal dystrophy, a devastating inherited neurodegenerative disease in which horses are uncoordinated

and have trouble balancing.

Her colleague at UC Davis, vet and cell biologist Fern Tablin, is working on the genetics and mechanisms underlying the bleeding from the nostrils and lungs that often afflicts horses when they run. Other groups are looking for the genetic cause of about a dozen different thoroughbred diseases or disorders that affect behavior, bone formation, fractures, and athletic ability.

Researchers have also developed a horse gene chip with 70,000 DNA sequences that can scan for genes associated with diseases, coat color, or performance; a chip with 670,000 probes is expected by the end of the year.

Breeders around the world have begun to use genetics to test yearlings for the so-called speed gene, discovered several years ago by a team led by Emmeline Hill, a geneticist at University College Dublin and chair of Equinome, a thoroughbred genetics company. The gene controls variation in muscle development in horses—and predicts whether a horse is better at short sprints or longer races.

Racehorse trainer David Hayes, the owner of Lindsay Park Racing in Creightons Creek, Australia, says testing for that gene prevented him from retiring a 2-year-old filly who wasn't winning any 1200-meter races. The genetic test predicted she would perform better at longer distances. Sure enough, with training, she has won several races including a stakes race longer than 2100 meters. The key question, though, is whether breeders will begin using such genetic information to breed healthier horses—or simply ones that will cross the finish line first. ∎

CELL BIOLOGY

STAP cells succumb to pressure

Retraction plans for easy stem cell recipe leave scientists wondering how the papers came to be published

By **Dennis Normile** and **Gretchen Vogel**

Two papers that electrified—and confused—the stem cell field just 6 months ago appear to have lost their last defenders. Last week, lead author Haruko Obokata of the RIKEN Center for Developmental Biology (CDB) in Kobe, Japan, agreed to retract a *Nature* paper that described a startlingly easy way to produce stem cells; she had agreed to the retraction of a related paper a week earlier. Obokata's Japanese co-authors had previously called for the moves, and it has been widely reported that another key author, Charles Vacanti, a tissue engineering specialist at Brigham and Women's Hospital in Boston, has agreed to the retractions as well. (He declined to speak with *Science*.)

Last week's developments, which also included new technical challenges to the claims, only add to the puzzlement of many in the field about how the papers were published in the first place. They described a method, called stimulus-triggered acquisition of pluripotency, or STAP, which involves briefly bathing blood cells from newborn mice in a mildly acidic solution, or applying external pressure to the cells, and then carefully culturing them. The authors claimed such stimuli generated pluripotent stem cells that could produce mouse embryos and placentas.

But soon after the papers were published, questions emerged about manipulated images. RIKEN, the home institution of several authors, conducted an investigation that found Obokata guilty of research misconduct. Since then, no independent lab has been able to replicate the production of STAP cells. The saga has raised anew perennial questions about the responsibility of senior authors on a paper and the efficacy of the peer-review process—although *Science* has learned that an initial STAP cell paper was rejected by three journals, including *Nature*. Still, the work was ultimately published, leaving open a major question: How did leading stem cell scientists who were co-authors fail to detect the myriad apparent problems with the work?

Hans Schöler, a stem cell scientist at the Max Planck Institute for Molecular Biomedicine in Münster, Germany, offers one lesson from the affair: "Repeat [the experiment] right away. Other people in the lab have to do it before it goes out."

Obokata still says, through her lawyer, that despite the problems in the papers, the procedure works as she claimed. A group at RIKEN, led by co-author Hitoshi Niwa, is trying to determine if she is right. But more problems with the papers have come to light. Last week, Japanese media reported that genetic analysis of cell lines provided by co-author Teruhiko Wakayama of the University of Yamanashi to a third party for testing suggests that what were thought to be STAP cell lines were not derived from the mouse strains supposedly used to create the cells.

Japanese media also reported that Takaho Endo, a researcher at the RIKEN Center for Integrative Medical Sciences in Yokohama, looked at RNA sequences the authors had deposited in databases for a particular type of stem cells derived from STAP cells. He, too, concluded that the sequences came from mouse strains different from those identified in the paper. Moreover, Endo said, the RNAs, which reflect a cell's gene activity, suggested the cells were derived from a mix of embryonic stem cells and trophoblast stem cells—precursors of the placenta.

Many are wondering how the papers got through the peer-review process to begin with. On 4 February, just 6 days after the papers appeared online, an anonymous contributor to PubPeer, a website where published papers are discussed, pointed out that in one article, an image of an electrophoresis gel showing a genetic analysis, Figure 1i, appeared to have been spliced together contrary to normal practice. At about the same time, allegations of plagiarism in the methods section of the article were circulating on Twitter. A RIKEN investigating committee in a 13 March report confirmed those and several other problems with the papers. The committee later concluded that the gel image manipulation constituted research misconduct.

Sources in the scientific community confirm that early versions of the STAP work were rejected by *Science*, *Cell*, and *Nature*. For the *Cell* submission, there were concerns about methodology and the lack of supporting evidence for the extraordinary claims, says Schöler, who reviewed the paper and, as is standard practice at *Cell*, saw the comments of other reviewers for the journal. At *Science*, according to the 8 May RIKEN investigative committee's report, one reviewer spotted the problem with lanes being improperly spliced into gel images. "This figure has been reconstructed," the RIKEN report quotes from the feedback provided by a *Science* reviewer. The committee writes that the "lane 3" mentioned by the *Science* reviewer is probably the lane 3 shown in Figure 1i in the *Nature* article. The investigative committee report says Obokata told the committee that she did not carefully consider the comments of the *Science* reviewer.

Schöler says it is possible that the *Nature* reviewers placed too much trust in some of the authors on the papers. Asked if other reviewers should

---

Co-authors take STAP back

Some in the stem cell community say they believed Haruko Obokata's discovery of STAP cells because of the reputations of several authors on the papers.

TERUHIKO WAKAYAMA

University of Yamanashi in Kofu; well-known mouse cloning pioneer; formerly at RIKEN

Called for retraction: 10 March

HITOSHI NIWA

Internationally respected stem cell researcher at RIKEN CDB

Agreed to consider retraction: 14 March

YOSHIKI SASAI

Director of Neurogenesis and Organogenesis Group at RIKEN CDB

Agreed to consider retraction: 14 March

CHARLES VACANTI

Tissue engineer at Brigham and Women's Hospital. Initial work by Obokata was done in Vacanti's lab, following up on controversial stem cell research by his team.

Agreed to retraction: 30 May, reportedly, in a letter to *Nature*.

have noticed the spliced gel image, he said, "I wouldn't have even thought to look for it because Hitoshi [Niwa], Teruhiko [Wakayama], and Yoshiki [Sasai] were on the paper; there are some people you 100% trust." Niwa and Sasai are respected stem cell researchers at RIKEN CDB. All three are co-authors of the *Nature* papers, but it is not clear at what point they joined the STAP cell team or which ones were co-authors on the previous submissions.

Both *Science* and *Nature* declined to comment on any particular submission, review, or retraction. "The science of the two papers was rigorously, robustly peer-reviewed as part of our usual editorial procedures. Any inaccuracies in the presentation of data that may have come to light since the peer review are being investigated," a *Nature* representative wrote in an e-mail. "We are currently conducting our own evaluation and we hope that we are close to reaching a conclusion and taking action."

It remains to be seen if the STAP cell debacle touches off the kind of broader soul-searching that followed the unraveling of the Woo Suk Hwang scandal. Hwang, then of Seoul National University, and colleagues made several breakthrough announcements in cloning and stem cell work in *Science* and *Nature* in 2004 and 2005, only for much of it to be revealed as fraudulent in spring 2006.

"Everyone's awareness of the potential for image manipulation has increased significantly since 2006," writes *Science* Executive Editor Monica Bradford in an e-mail. And partly in response to the Hwang scandal, "*Science* and *Nature* and their associated journals are now collaborating to develop standards that promote the reproducibility of research publications," another *Nature* representative wrote in an e-mail.

Many in the field are still waiting expectantly for the results of Niwa's attempts to reproduce the results, hoping they will shed some light on how so many top labs were apparently misled. Schöler says he won't pass final judgment until those results are made public. "I'm going to wait for Hitoshi's final word." ∎

Dinosaur metabolism neither hot nor cold, but just right

Growth rates suggest answer to long debate: Dinosaurs were "mesothermic," like today's tuna and echidna

By **Michael Balter**

Call it the Goldilocks solution. Paleontologists have struggled for 50 years to determine whether dinosaurs were cold-blooded ectotherms like today's reptiles, making little effort to control their body temperatures, or endotherms, like most modern mammals and birds, which keep their body temperatures at a constant, relatively high set point. The answer greatly influences our view of dinosaurs, as endotherms tend to be more active and faster growing.

A study on page 1268 concludes that dinosaur blood ran neither cold nor hot but something in between. Examining growth and metabolic rates of nearly 400 living and extinct animals, the researchers conclude that dinosaurs, like a handful of modern creatures including tuna and the echidna, belonged to an intermediate group that can raise their body temperature but don't keep it at a specific level. The researchers christen these creatures mesotherms.

Establishing a new metabolic category is "audacious," admits lead author John Grady, an evolutionary biologist at the University of New Mexico, Albuquerque. And some still think dinosaurs were "just fast-growing ectotherms," as vertebrate physiologist Frank Paladino of Indiana University–Purdue University Fort Wayne insists. But paleobiologist Gregory Erickson of Florida State University in Tallahassee calls the paper "a remarkably integrative, landmark study" that transforms our view of the great beasts.

For the first 150 years after their discovery, dinosaurs were considered ectotherms like today's reptiles. Ectothermy makes some sense: "It requires much less energy from the environment," explains Roger Seymour, a zoologist at the University of Adelaide in Australia. But it has drawbacks, too: "The animal cannot feed in cold conditions and has a much more limited capacity for sustained, powerful activity, even if warmed by the sun," he says.

Beginning in the late 1960s, researchers put forward the then-heretical idea of dinosaurs as endotherms, and evidence for this has accumulated. Annual growth rings in dinosaur bones suggest fast, energy-hungry developmental rates. Birdlike air sacs may have boosted their respiratory efficiency, suggesting rapid movements. And isotopic data from fossils suggest higher body temperatures (*Science*, 22 July 2011, p. 443).

Giant endotherms pose their own puzzles, however, such as the huge quantities of food

Metabolic middle ground
Dinosaurs and some living animals are mesotherms, raising but not maintaining their body temperature.
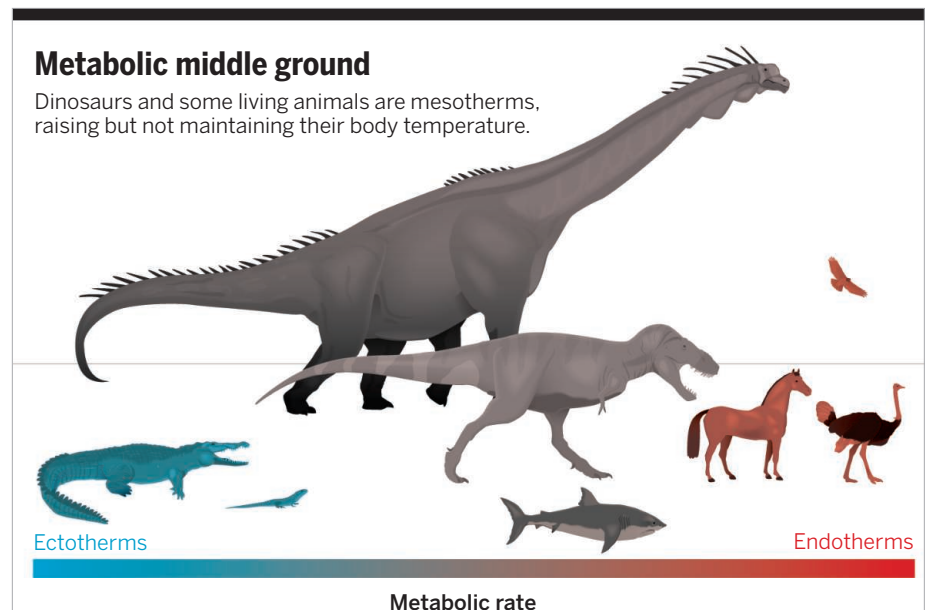
Ectotherms   Endotherms

Metabolic rate

needed to sustain them. An endothermic *Tyrannosaurus rex* "would probably have starved to death," Grady says.

He and his colleagues tackled the problem by examining the relationship between an animal's growth rate—how fast it becomes a full-sized adult—and its resting metabolic rate (RMR), a measure of energy expenditure. Earlier studies, based on limited data, had suggested that growth rates scale with metabolic rates. That is, the more energy an animal can expend, the faster it can grow and the bigger it can get. The team pulled together updated data on 381 living and extinct vertebrates, including 21 species of dinosaurs, and developed mathematical equations that predict the relationship between metabolic rate, growth rate, and body size in living animals.

These equations show that ectotherms and endotherms fall into distinct clusters when growth rate is plotted against metabolic rate. High-energy endotherms grow fast and have high metabolic rates, whereas ectotherms have low values of both. Those two categories include most living species, but the team found that a handful, such as fast-swimming sharks, tuna, reptiles such as large sea turtles, and a few odd mammals like the echidna, fall into an in-between state: mesothermy. These animals use their metabolism to raise their body temperatures, but do not "defend" a set temperature.

Using their equations, the team calculated dinosaur RMRs, plugging in reliable published data on these extinct animals. Dino growth rates can be estimated because rings of bone, which give a measure of age, were laid down annually, and body size can be estimated from bone size. The results placed dinosaurs squarely among the mesotherms. The earliest birds—direct descendants of dinosaurs—plotted as mesotherms, too.

Grady and colleagues think mesothermy may have allowed dinosaurs to grow large and active with lower energy costs. Geochemist Robert Eagle of the California Institute of Technology in Pasadena agrees: "In a world that was generally hotter than today, it wasn't really necessary to be a full endotherm." Previous studies have suggested that during the Mesozoic, even mammalian endotherms kept their bodies at a lower set point than they do today, he says.

Grady suggests that mesothermy might even help explain why dinosaurs ruled the Earth: They could easily outcompete other reptiles, which were lethargic ectotherms. And by getting big quickly, they occupied the large-animal niches, and prevented the small, energy-hungry endothermic mammals from getting bigger themselves. Until, of course, the fateful asteroid struck, and dinosaurs vanished. ∎

SOUTH AMERICA

Legal highs make Uruguay a beacon for marijuana research

Government funds new lab to study the biochemistry and medical uses of the drug

By **Lizzie Wade**

A year ago, biophysicist Nelson Bracesco was studying yerba mate, a popular caffeine-laden infusion brewed in South America. But a legal about-face in his home country has given his research an edgier focus. Last December, attempting to head off violent crime associated with drug trafficking, Uruguay became the first country to legalize the production, sale, and use of marijuana for both recreational and medical purposes; the government finalized key regulations last month. Legally unshackled, Bracesco and his colleagues at the medical school of Uruguay's University of the Republic (Udelar) in Montevideo are launching a lab dedicated to marijuana's chemistry, genetics, medical uses, and side effects.

Scientists in countries still hamstrung by marijuana's illicit status say the Udelar lab is poised to fill in glaring research gaps. "It's absolutely precedent-setting in the world," says Alan Shackelford, a physician who recommends marijuana as part of his practice in Denver. Even though the drug is legal in Colorado and Washington, he says, federal red tape continues to be an impediment to good science in the United States.

First on the to-do list for the Uruguayan lab, funded at the outset primarily by the government's new Institute for Regulation and Control of Cannabis (IRCCA), is a genetic database of marijuana strains in the country. IRCCA will import and sell seeds to licensed growers, and Uruguayans can grow their own plants—up to six per household, so long as they register strains they are cultivating. All will be studied by the Udelar lab.

That will allow researchers to begin characterizing each strain's chemical constituents, says Udelar neuroscientist Atilio Falconi, another founding member of the new lab, which now counts about 20 collaborators. Strains can vary in content of tetrahydrocannabinol (THC), the drug's main active component, for instance, and of cannabidiol. Variations may affect therapeutic efficacy; marijuana plants chock-full of THC are rumored to be better at easing posttraumatic stress disorder, while cannabidiol-rich strains may be better at curbing seizures. "We're interested in finding out if this popular knowledge has a real scientific basis," Falconi says.

The lab plans to test efficacy claims. Few are well grounded; after reviewing studies conducted since 1948, Barbara Koppel, a neurologist at Metropolitan Hospital in New York City, concluded that there is sufficient evidence of marijuana's effectiveness only

Marijuana activists light up in Montevideo.

for multiple sclerosis. Promising preliminary data for Parkinson's disease did not stand up in follow-on studies, and other conditions are open questions. Also understudied are the effects of recreational use, says Marilyn Huestis, chief of chemistry and drug metabolism at the U.S. National Institute on Drug Abuse. "We have lots, lots more to learn about how THC works in the brain … especially [how] it changes the way the brain develops."

Now that marijuana is legal, Uruguayans may be more willing to accurately report use to researchers. Adults registered with IRCCA can buy up to 40 grams per month through licensed pharmacies. They also have the option of joining a cannabis club that grows and distributes marijuana to members. Although information about personal use and the availability of strains could be useful for helping the government keep tabs on the marijuana market, Bracesco pledges that the lab will operate with "scientific rigor and absolute academic independence." ∎

Camel racing—in which lightweight robots have replaced human jockeys since 2005—has become a multimillion-dollar business in Qatar.

Mission to MERS

A massive study in Qatar is yielding the most detailed data yet on how a new virus spreads between camels—and how it jumps to humans

By **Martin Enserink**
in Qatar

"Now comes the difficult part." Elmoubasher Farag, a public health epidemiologist at the Qatari health ministry, grins as he walks to the backside of a camel and pulls out a long, sturdy-looking cotton swab. While a caretaker holds up the animal's tail, Farag briefly sticks the swab into its rectum. He comes back and shows off the brownish tip to a group of Dutch visitors before sliding it into a plastic tube with a screw cap.

Doha's animal market, the scene of this encounter, is the camel equivalent of an international airport. In one 25-square-meter pen stands a group of about 20 camels imported from Sudan, a mark on one of their thighs betraying their origins. Right next to it is a pen with nearly

20 camels from Saudi Arabia. A bit farther are camels from Oman. Animals can spend months here waiting for buyers.

Farag is part of an unusual collaboration between scientists from Qatar and the Netherlands aiming to better understand the baffling puzzle of Middle East respiratory syndrome (MERS), the viral disease that emerged 2 years ago and that has infected more than 800 people to date, killing more than 300. A lot is still unknown about how people become infected, but dromedary camels—of which Qatar has about 65,000—are believed to play a major role.

The Dutch team, led by virologist Marion Koopmans of Erasmus MC in Rotterdam, was the first to show that camels might be involved in MERS, when they found antibodies in every single one of 50 camels

tested in Oman in August 2013. That fall, Mohammed Al Hajri, a high-ranking official at the Qatari health ministry, invited the team to help investigate a small farm outbreak of MERS. (So far, Qatar has seen a total of only nine human cases, the last one in November, including three deaths.) The scientists found viral RNA in a camel at the farm; since then, they and others have also discovered widespread MERS infection in camels from Saudi Arabia, Egypt, Tunisia, Nigeria, and Kenya.

To be sure, contact with camels is not the only way people become infected. The virus occasionally spreads between humans, and it can be transmitted easily in hospitals if hygiene measures are inadequate. (This appears to account for many of the more than 500 cases that Saudi Arabia has reported in

the past 3 months.) For some patients, the route of infection is completely unclear. Still, most scientists now believe that camels are the key to controlling MERS.

But how, when, and where do camels become infected? And what type of contact with camels puts humans at risk—taking care of them, slaughtering them, or drinking their raw, warm milk? Is the local practice of drinking the highly concentrated urine from virgin female camels dangerous as well?

The answers are likely to be found in the Arab world's intricate ecosystem of people and camels. Qatari members of the team—which also includes Qatar's Ministry of Environment and the Dutch National Institute for Public Health and the Environment—have taken more than 14,000 samples from a wide range of animals, people, and the environment, sending many to Rotterdam to be tested for evidence of infection. Now, Koopmans and her colleagues Chantal Reusken and Bart Haagmans are in Qatar to discuss the results, which link camels to human infections more directly than ever.

They also want to see exactly how and where the samples were taken and get a fine-grained picture of how camels and people interact. "To interpret the results, you have to see everything yourself," Reusken says. "You have to ask a ton of questions."

CAMEL CARETAKERS At the market, it's 41° outside, and Farag, a big, bearded physician from Sudan who leads the Qatari ministry's MERS studies, has finished his demonstration. He takes us to a small hut with a corrugated metal roof that offers a reprieve from the blistering sun. There are rugs on the dirt floor; three camel caretakers, clad in traditional Sudanese jalabiyas, sit on old sofas, sipping tea. Sticky flytraps are dangling from the ceiling; the flies glued to them will be tested for possible presence of the virus.

The Dutch researchers fire off a series of questions for the men, which Farag translates. Are they only in contact with live animals, or also with meat? Well, after camels are taken to the nearby slaughterhouse, the owners sometimes give them the liver as a gift, they say, which they eat raw. "It is so delicious," Farag translates.

"Do they eat the lungs as well?" Koopmans asks.

"No, never the lungs. Only the liver, the kidneys, and the heart."

Do the men ever see sick camels, or animals with runny noses? Sometimes, the caretakers say; mostly in animals up to about 10 months old. If all the MERS virus does in camels is cause the sniffles, Reusken explains later, their infection could be easily overlooked.

After we step outside the hut, a middle-aged Qatari man jumps from a white pickup truck and asks what's going on. He has a camel farm in Al Shahaniya, the town where the 2013 outbreak occurred, but he's skeptical

Camel country

Numbers of camels* (black) and human MERS cases** (red) in selected Middle East countries

Jordan
13, 18

Kuwait
11.5, 3

United Arab Emirates
364, 70

Qatar
65, 9

Egypt
137, 1

Saudi Arabia
260, 689

Oman
133.5, 2

Yemen
407, 1

* In thousands
** As of 5 June

Source: FAO/ECDC

that camels are the culprit; it's more likely that humans infect them rather than vice versa, he says. "We have lived with camels for thousands of years, and we've always been healthy," he says. "Why would they start making us sick now?"

Driving into the desert on a deserted eight-lane highway, Farag tries to answer that question. It's true that people in the Middle East have depended on camels for meat, milk, hides, and transportation for millennia. "They're almost members of the family," Farag says. He believes MERS has only recently become a problem because the ecology in the Gulf states is changing; he calls it the "new camel lifestyle." As Qatar's population and wealth have exploded, for instance, so has demand for camels and their meat—which is why you see animals from many countries so close together at the market. Farag suspects that this may have given

the virus many more opportunities to travel between camel populations—and occasionally to humans.

CAMEL FARMS The roughly 700 camel farms in the town of Al Shahaniya, some 30 kilometers west of Doha, consist of walled compounds the size of one or two football fields, often with a nice villa; besides camels, there are often sheep, goats, and pigeons or chickens. These farms are really second homes; wealthy Qataris take their families there on weekends to escape the traffic and noise in Doha. Owning a farm in Al Shahaniya can easily cost $5000 per month, Farag says.

At one farm, helpers demonstrate how they milk a camel; first, a young animal is allowed to suck the udder to get the milk flowing; then it is shunted aside and human hands take over. Soon, the men have several liters of very thick, bright white milk in an aluminum pan. The owner takes a gulp, cracks a joke about the coronavirus, and offers the milk to his guests. Nobody is tempted. Farag had warned that declining would be rude—sharing camel milk is an almost sacred tradition—but he saves the situation by saying we would very much like to try it later. The helpers fill up eight bottles and load them into the trunk of his car.

In one of its most important findings to date, just submitted to *Eurosurveillance*, the Dutch-Qatari team has found the MERS virus in the milk of more than half of the animals with an active infection. It's still unclear, however, whether it is secreted directly from the mammary glands or if it enters the milk via other routes—for instance, through newborn camels' saliva or traces of fecal matter. In April, U.S. researchers also showed that virus added to unpasteurized camel milk can survive for at least 3 days. It's unclear what the infection risk is, but boiling is advisable, Koopmans says.

As we head out of the village, Farag makes another stop and knocks on the red metal gate of the farm where the 2013 outbreak took place. A young man with a shy smile opens the gate and invites us in.

Just what happened to trigger the outbreak, which the group documented in *The Lancet Infectious Diseases*, may never be entirely clear. The owner declared that his camels never left the farm, and the research team doesn't know how they

became infected. It's possible that people acquired the virus elsewhere and infected the camels, or that both were infected from some other source. The two human cases were the 61-year-old owner and the young man, a 23-year-old immigrant worker from South Asia whose first name is Mohammed, who was kept in isolation in the hospital for a month. His main symptom was fatigue, but he's on the official list of MERS cases and has entered the medical literature as "Case 2." Case 2 waves as we leave the farm.

AL SHAHANIYA RACETRACK A centuries-old sport and a favorite pastime for the wealthy, camel racing has recently become an international, multimillion-dollar industry, in which a single camel can be worth $5 million or even $10 million. New racetracks have gone up around the Gulf; owners ship camels, along with their caretakers, to Saudi Arabia, Oman, and the United Arab Emirates to race, where they mingle—and swap viruses with—the local population. That, too, is part of the "new lifestyle," Farag says.

> *"To interpret the results, you have to see everything yourself. You have to ask a ton of questions."*
>
> **Chantal Reusken,**
> Erasmus MC

going here all through the night, Farag says. He wonders whether this new biorhythm stresses the animals and weakens their immunity.

DOHA SLAUGHTERHOUSE Between 10 and 20 camels meet their end here daily, as do some 500 sheep. We watch as a hydraulic sling slowly lowers a young, kneeling camel to the floor at the center of a large, clean-looking hall while a butcher sharpens two knives by rapidly scraping them against each other. Two men pull back the camel's head with a rope, exposing its neck to the butcher. A fast cut, a pulsating red fountain,

to help interpret the data. Contamination happens easily: If you cut through one tissue to get a sample of another, your knife could transport the virus, for instance. But they're happy with what they see.

The camel carcass is carved up and the meat is picked up by one of the 12 Doha butchers who sell camel meat. All have agreed to be tested in the months ahead, along with their employees, Farag says.

A week after the team finished its tour of Qatar's camel ecosystem, the health ministry announced results from their months of work at a press conference in Doha. Among the team's key findings: 8.7% of 109 people working with camels have antibodies against MERS, a sign that they have been infected at some point. (No antibodies were found in people in Qatar and in Europe who had no animal contact or contact only with sheep.) None of the camel workers who were infected—most of them young, healthy men—had been severely ill. Some 20% of camels infected with the virus shed it in their feces.

The findings suggest that MERS is both more widespread and less lethal than believed. "It's the first time we get a better idea of how much MERS is out there in the community," says Peter Ben Embarek of the World Health Organization (WHO) in Geneva, Switzerland, who has followed the research closely and calls the Qatari-Dutch collaboration "a showcase of what should be done" in the region. Whether people with a mild or asymptomatic infection can transmit the virus to others—which might account for some of the unexplained cases—is an important next question, he adds.

Based on the findings, WHO will soon update its recommendations on how to prevent MERS, Ben Embarek says. The Qatari government has already issued new guidelines

Mohammed Al Hajri of Qatar's health ministry (*left*) and Marion Koopmans (*right*) of Erasmus MC discuss results of their joint investigation into the epidemiology of MERS in Qatar.

Al Shahaniya's public racetrack is ultramodern; at night, colossal masts spaced 50 meters apart bathe its 7-kilometer circuit in bright white light, like some intergalactic runway. Qatar banned the use of young boys as camel jockeys in 2005 after an international outcry over child labor; now specially developed lightweight robots take their places during races.

A few groups of camels are training tonight, while the owners follow their prized animals in air-conditioned white Land Cruisers. Camels will be coming and

and, briefly, a panic-stricken look on the animal's face—then it's over. The head slumps into a rapidly expanding lake of blood on the floor.

The Qatari team has already collected hundreds of samples from the workers here to see if they have antibodies; they've also taken camel samples covering "everything from head to butt," Koopmans says, including the nose, mouth, trachea, intestines, urine, lymph nodes, kidney, and liver. The Dutch want to see exactly how these samples were taken, and how the abattoir operates,

to protect camel workers, such as frequent hand washing, wearing facemasks "as far as possible," given the scorching heat, and wearing protective clothing, which should be washed daily. The government also advised that camel milk should be boiled before drinking.

Whether people follow that advice remains to be seen, Ben Embarek says. For people in the Arabian Desert, it will be hard to accept that the animals that allowed their survival for millennia have become a 21st century threat. ■

Tropical storm systems confound
conventional climate models but
might yield to stochastic simulations.

A touch of the random

As researchers seek ever-larger supercomputers to crunch climate models of baffling complexity, some are calling for a fresh, statistics-based approach

By **Colin Macilwain**

Three statisticians go hunting and flush a duck. The first shoots high, over the bird's head. The second aims too low and sends a bullet whistling meters below the duck's belly. So the third statistician jumps up and down yelling, "We got him! We done got him!"

Not the best way to bag waterfowl, perhaps—but in their own habitat, researchers have found similar scattershot methods very effective for predicting the behavior of complex systems. "Stochastic" techniques, in which computer simulations spit out clouds of possible outcomes, have been widely applied in economics, physics, engineering, and weather forecasting. They have not found a home, however, in one of the highest profile and most contentious areas of forecasting: climate modeling.

There, researchers have usually aimed for a deterministic solution: a single scenario for how climate will respond to inputs such as greenhouse gases, obtained through increasingly detailed and sophisticated numerical simulations. The results have been scientifically informative—but critics charge that the models have become unwieldy, hobbled by their own complexity. And no matter how complex they become, they struggle to forecast the future.

"The house used to have two floors. Now it has eight. It is bearing all this weight, and cracks are appearing in the walls," says Christian Jakob, a climate modeler based at Monash University, Clayton, in Australia. "That gives you two choices," Jakob says. "You can go in and strengthen the foundations. Or maybe, it's time to build a new house."

Now, some researchers are calling for a major overhaul: The models, they say, should be remodeled along stochastic lines. Later this month, for example, a special issue of the *Philosophical Transactions of the Royal Society A* will publish 14 papers setting out a framework for stochastic climate modeling.

One key reason climate simulations are bad at forecasting is that it's not what they were designed to do. Researchers devised them, in the main, for another purpose: exploring how different components of the

system interact on a global scale. The models start by dividing the atmosphere into a huge 3D grid of boxlike elements, with horizontal edges typically 100 kilometers long and up to 1 kilometer high. Equations based on physical laws describe how variables in each box—mainly pressure, temperature, humidity, and wind speed—influence matching variables in adjacent ones. For processes that operate at scales much smaller than the grid, such as cloud formation, scientists represent typical behavior across the grid element with deterministic formulas that they have refined over many years. The equations are then solved by crunching the whole grid in a supercomputer.

The approach has proven itself very useful for probing the workings of Earth's climate system—how fossil fuel emissions, atmospheric carbon dioxide, and global temperatures all interact, for example. But it falls short when asked to predict where, when, and how severely future climate changes will unfold.

Last year, for example, scientists on the Intergovernmental Panel on Climate Change (IPCC) systematically compared the predic-

> ## "You can go in and strengthen the foundations. Or maybe, it's time to build a new house."
>
> **Christian Jakob,**
> Monash University, Clayton

tions of 20 major climate models against the past 6 decades of climate data. The results were disappointing, says Ben Kirtman, a climate scientist at the University of Miami in Florida and coordinating author of the near-term predictability chapter of last year's fifth IPCC assessment report. The models performed well in predicting the global mean surface temperature and had some predictive value in the Atlantic Ocean, but they were virtually useless at forecasting conditions over the vast Pacific Ocean.

The most comprehensive models aren't necessarily the most useful ones, Kirtman says. As climatologists keep adding components to simulate processes as detailed as leaf growth and termite distribution, the models have become bloated with features and run sluggishly without making better predictions. "It's a fundamental problem," he says. "They keep trying to add in everything." And parts of the models don't simulate nature well at all, Jakob adds: "Some of the old problems have not been solved. On things like simulating rainfall and cloud formation,

progress has been painfully slow."

Much of the problem boils down to grid resolution. "The truth is that the level of detail in the models isn't really determined by scientific constraints," says Tim Palmer, a physicist at the University of Oxford in the United Kingdom who advocates stochastic approaches to climate modeling. "It is determined entirely by the size of the computers." Roughly speaking, an order-of-magnitude increase in computer power is needed to halve the grid size. Typical horizontal grid size has fallen from 500 km in the 1970s to 100 km today and could fall to 10 km in 10 years' time. But even that won't be much help in modeling vitally important small-scale phenomena such as cloud formation, Palmer points out. And before they achieve that kind of detail, computers may run up against a physical barrier: power consumption. "Machines that run exaflops [10^{18} floating point operations per second] are on the horizon," Palmer says. "The problem is, you'll need 100 MW to run one." That's enough electricity to power a town of 100,000 people.

Faced with such obstacles, Palmer and others advocate a fresh start. Climate modelers, they say, need to step backward and draw some inspiration from weather forecasting—specifically techniques developed at the European Centre for Medium-Range Weather Forecasts (ECMWF) in Reading, U.K.

Starting in the 1990s, researchers at ECMWF shook up weather forecasting worldwide by introducing stochastic approaches into their models. Also known as Monte Carlo methods, these techniques were first developed by physicists in the World War II Manhattan Project to model how neutrons diffuse through materials, bouncing off atomic nuclei as they go—a process they had struggled to model deterministically. The idea is analogous to repeatedly rolling dice or spinning a roulette wheel: Run the calculations many times to produce a range of different outcomes, then "tune" the model by aggregating the results and comparing them with empirical observations. Researchers can extend the "hindcasts" into the future to make predictions expressed as probabilities, with uncertainties plainly evident in the scatter of results. Such techniques are widely used today in many branches of physics and engineering, by insurers calculating risk, and even by computational biologists to model cell membranes or proteins.

To apply the method to weather forecasting, the ECMWF modelers introduced small, random perturbations in the initial weather conditions. By running its model many times with slightly different initial conditions—an approach known as ensemble modeling—the center produced

Computational challenge

Models simulate climate by crunching equations for a wide range of interacting processes, parceled out among the compartments in a huge, global 3D grid.

superior weather forecasts. The approach has been adopted by most major weather forecasters worldwide.

Palmer, who led the development of the ensemble approach at ECMWF, says closely related approaches could transform climate modeling. One major difference is in timescale: Unlike weather modelers, who quickly find out whether their predictions were correct, climate researchers think decades ahead. To tune a model, they must feed it climate records up to a certain year—say, 1990—and see how well it would have predicted climate patterns for the decade that followed.

Stochastic models, Palmer says, could get a grip on components of the climate system that are too slippery for traditional deterministic models to handle. For example, tropical thunderstorm systems—each of which can release the energy of a hydrogen bomb—are "crucially important" in the global climate system, he says. But because their cores are only a few kilometers wide, ordinary model grids can't capture them, and the fixed mathematical descriptions on which they rely contain "significant errors" that a probabilistic approach would avoid, Palmer says.

In the special issue of *Philosophical Transactions A*, researchers propose several complementary approaches to incorporate such approaches into climate models. The ideas include building supercomputers to work faster by allowing some stochastic errors at the transistor level, and modeling different processes on different scales to ease the computational burden.

A web of subroutines
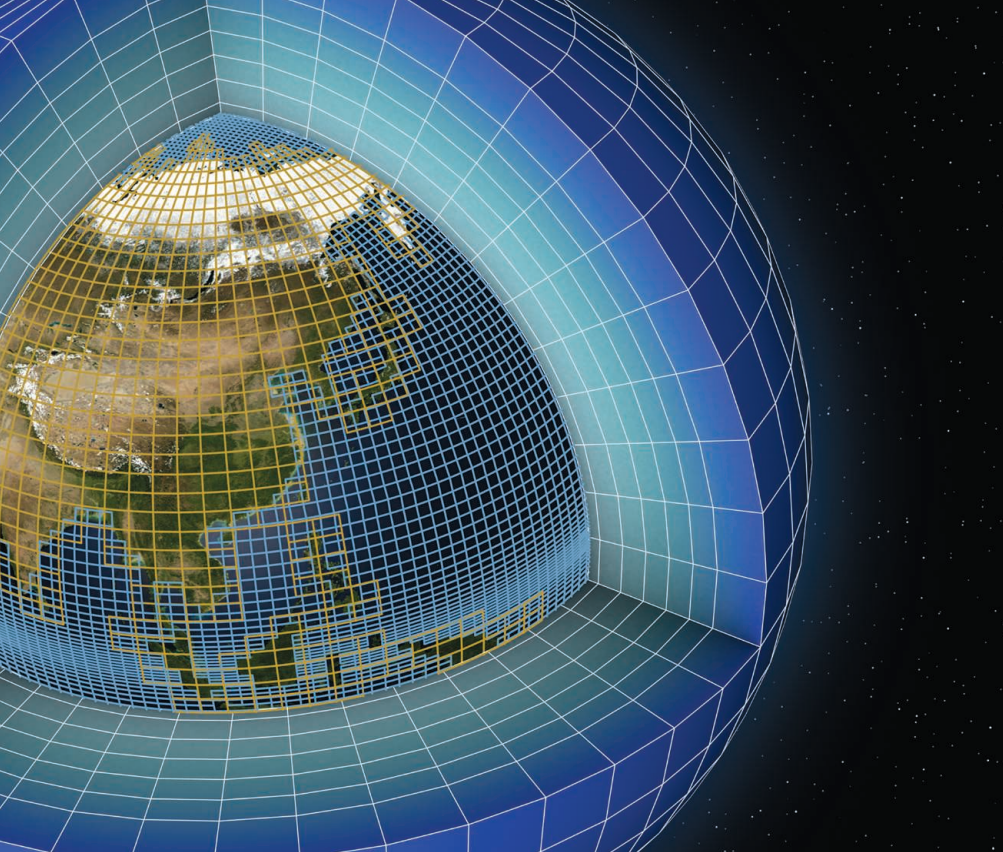
build the stochasticity in at a fundamental level, to make it more consistent with the underlying laws of physics."

Others would go still further and predict climate change purely on the basis of statistical data from the climate record. Last year in the *Journal of Climate*, Leonard Smith of the London School of Economics and Political Science and his colleague Emma Suckling reported the results of an experiment in which they took global climate data for various locations over the past half-century, fed it into the most prominent climate models, and compared the predictions with those of a very simple statistical model that extrapolated the future climate from that of the recent past. The simple model—which worked by taking mean-temperature changes for periods of 1 to 10 years over the past century and using these to extrapolate the future— fared best, Smith says.

Supporters of deterministic modeling say they are not about to abandon an approach they consider tried and trusted. Existing climate models are "continuously scrutinized in the scientific literature" and pass muster, says Erland Källén, a veteran climate modeler and director of research at ECMWF. And Chris Bretherton of the University of Washington, Seattle, who chaired a 2012 study of climate models by the U.S. National Research Council, says critics of existing models have yet to present convincing evidence that stochastic modeling could do a better job. "There's a feeling that it is a valid approach. But I don't think that there's a consensus that we'll need to have stochastical parameterization," he says.

Advocates of stochastic approaches, however, say only a drastic change of course can jolt predictive climate modeling out of its current rut. With policymakers clamoring for robust forecasts of how temperature and precipitation will change region by region in coming decades, Smith says, time is running out: "The question is, when will we have significantly better quality information than we have today? I think we may have our answer from the climate before we get it from the physics."

With current physics-based models struggling to predict the climate a decade or two out, modelers may be inclined to give stochastic methods a roll of the dice—especially as better tools emerge for testing models against the climate record. "We need to be unforgiving," says Miami's Kirtman, and "make hard-nosed comparisons" of models' predictive performance. "I think we still have an enormous amount to learn." ■

In another approach, modelers would take a particular component of existing climate models, such as a subroutine that simulates cloud formation, and replace it with a stochastic equivalent. In an unpublished study submitted to the *Journal of the Atmospheric Sciences*, Andrew Majda, a mathematician at New York University's Courant Institute of Mathematical Sciences in New York City, and colleagues did just that. He showed that slotting a stochastic cloud simulation into a National Center for Atmospheric Research (NCAR) climate model could closely reproduce an important tropical weather pattern that numerical models have struggled to capture.

In the pattern, known as the Madden-Julian Oscillation (MJO), rainfall sweeps eastward across the Indian and Pacific

oceans every 30 to 60 days. The researchers used a stochastic process called a Markov chain to capture the behavior of various types of clouds. Plugged into NCAR's High-Order Methods Modeling Environment (HOMME), the probabilistic simulation "drastically improves the results of the deterministic model," the authors say. Among other real-world characteristics that conventional models fail to capture, the souped-up HOMME model correctly forecasts the MJO's propagation speed and its tendency to spawn "trains" of two or even three such weather patterns in quick succession.

Some researchers believe such approaches can be fully tested only by building entirely new climate models. "Ideally, I think we will need to go back and design the models from scratch," Palmer says. "You really want to

Colin Macilwain is a science writer based in Edinburgh, U.K.

INSIGHTS

Measuring malaria
elimination *p. 1230*

Promising perspective from
focus on novelty *p. 1234*

HOMOLOGY,
GENES, AND
EVOLUTIONARY
INNOVATION

GÜNTER P. WAGNER
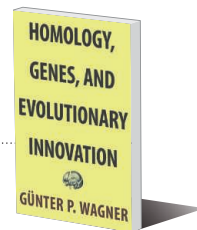
The arm of an interferometer used for
detecting gravity waves, artificially
backlit at night.

QUANTUM MECHANICS

The quantum nondemolition derby

Monitored resonators can evade back-action from shot noise in the detection laser

By **Dirk Bouwmeester**

Heisenberg's uncertainty principle demands that a very precise measurement of a particle's position x not only will increase the uncertainty in a measurement of its momentum p, but also affects any additional measurement of x because the high momentum uncertainty affects the spreading of the particle's wave function. However, it is possible to avoid this problem by cleverly choosing what to measure. A measurement variable can be constructed so that

the inevitable increases in uncertainties in other quantities will have no influence on the outcome of its repeated measurement. Such "quantum nondemolition" (QND) measurements can be used to detect the turning on of an arbitrarily small force. A specific QND method, referred to as back-action evading (BAE), was conceived in the context of detecting gravitational waves by their effect on large test masses or on the freely suspended mirrors of interferometers (*1–3*). This technically demanding method was originally based on optomechanical systems, and on page 1262 of this

issue, Suh *et al*. (*4*) now report an implementation of BAE using microwaves in a superconducting resonator circuit with a movable capacitor plate.

The first QND measurement, proposed by Braginsky and Vorontsov (*5*) in order to measure the tiny vibrational excitations (phonons) of a gravity wave detector, is a precise measurement of the total number of discrete phonon excitations N. This measurement inevitably results in a large uncertainty in the phase of the oscillator. However, the phase uncertainty does not influence future measurements of N. This

"quantum-counting" QND measurement proved difficult [but eventually was demonstrated (6)], and an alternative QND measurement was introduced that measured either the real or imaginary part of the oscillator's complex amplitude (X_1 or X_2 respectively) (1–3, 7, 8). X_1 and X_2 satisfy the uncertainty relation $\Delta X_1 \Delta X_2 \geq \hbar/2m\omega$ (where $\hbar$ is Planck's constant divided by 2π, ω is the resonant frequency, and m is the mass of the oscillator) and also satisfy the condition for a QND measurement: An arbitrarily precise measurement of X_1 will increase the uncertainty in X_2 without having an impact on future measurements of X_1.

To appreciate why this scheme is a BAE measurement for X_1 (or X_2), it is helpful to understand the quantum back-action that is present in "standard" optical measurements of the dynamics of a harmonic oscillator. Consider a mirror attached to the harmonic oscillator, and let the mirror's position x change the length of an optical cavity or of one arm of an interferometer. Measuring the optical signal at an output port of the interferometer or measuring near-resonant light leaving the cavity provides information on the oscillator's motion.

To determine the optimal sensitivity at which the dynamics can be determined, several quantum properties must be taken into account. First, the mechanical harmonic oscillator has a minimum spread in x and p, referred to as zero-point fluctuations, such that the ground-state energy is $\hbar\omega_m/2$ and is equally distributed over potential and kinetic energy. Second, typically a coherent state produced by a laser is used for the phase-sensitive optical measurement. The coherent state can be considered as a stream of photons that arrive at random and follow a Poisson distribution with an average photon number N (the signal) during a given measurement time interval t and fluctuations of $\sqrt{N}$ (shot noise). Because the signal-to-noise ratio scales as $\sqrt{N}$, it might seem best to boost the laser power and measurement time interval as much as possible. However, the shot noise acts back on the motion of the mechanical system by radiation pressure (9). When the $\sqrt{N}$ back-action leads to fluctuations equal to the zero-point fluctuations of the mechanical oscillator, the optimal sensitivity, or standard quantum limit (SQL), has been achieved.

Third, for interferometers, the quantum properties following from the central beam splitter are also essential. The laser field enters through one input port of the beam splitter while vacuum fluctuations
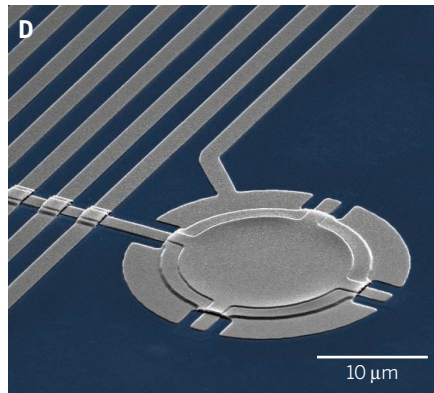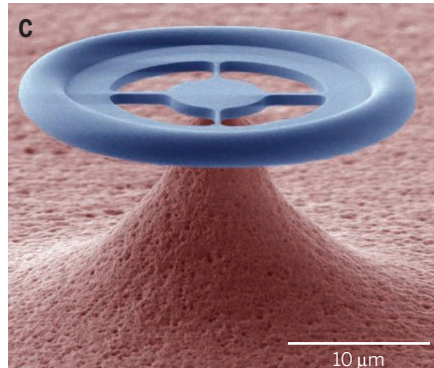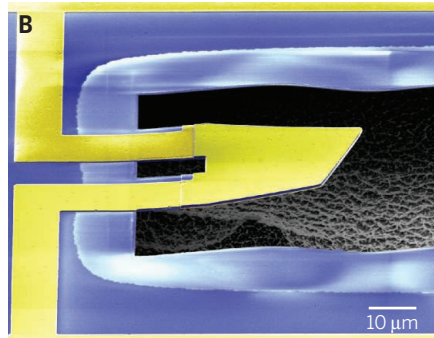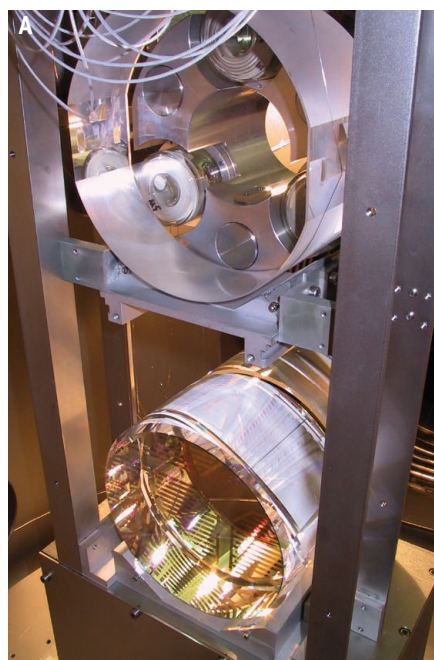
Oscillators hall of fame. Examples of some of the most spectacular optomechanical, piezoelectric-mechanical, and electromechanical systems for which the theoretical description are similar but the implementations are quite different. (**A**) Kilogram-size mirrors used for interferometric gravitational wave detectors operating beyond the quantum shot noise limit (12). (**B**) High-frequency (6 GHz) piezo-electromechanical (dilatational) resonator that has been cooled to ground state and excited into a superposition of the ground and first excited state (13). (**C**) Microtoroid for which quantum-coherent coupling of a mechanical oscillator to an optical cavity mode has been demonstrated (14). (**D**) Superconducting LC circuit with drum-mode capacitor that was used to demonstrate side-band cooling to the mechanical ground state (15).

enter through the other. The interference of the two will cause fluctuations of $\sqrt{N}$ on the number of photons reflecting from the mirror on the mechanical resonator, with N the total number of photons in the laser field entering at the input port during the measurement interval. Whether a cavity, an interferometer, or a combination of both is used, the SQL firmly applies.

The various ways in which to bypass the SQL all monitor only a certain aspect of the harmonic oscillator motion rather than the full dynamics. A QND measurement can in principle lead to a maximally squeezed state (a state in which the Heisenberg uncertainty principle is satisfied by unevenly distributed uncertainties for the noncommuting observables). For an optical QND measurement, the laser power can be increased well above the optimum set by the SQL. In the specific case of the BAE scheme, the back-action can be evaded in the variable X_1; all of the radiation back-action ends up in X_2.

Several experimental implementations of the BAE scheme have been proposed, and it is remarkable that the general system of electromagnetic radiation coupled to a mechanical oscillator has been considered for oscillators with masses from 10^3 g to 10^{-22} g (see the figure). Suh *et al.* now report on the first realization of the "continuous two-tone" BAE scheme. The experiment is based on a superconducting circuit that represents an electromagnetic resonator at microwave frequency ω_c, with one of the capacitor plates suspended so that it can oscillate at a mechanical frequency ω_m and couples to the inductor-capacitor (LC) resonance photons. These superconducting electromechanical devices can now be fabricated and controlled with high quality and operate at temperatures down to a few millikelvin, so that the mechanical sideband frequencies can be resolved, shot noise–limited microwaves can be applied, and microwave photons can be de-

Department of Physics, University of California, Santa Barbara, CA 93106, USA, and Huygens Kamerlingh Onnes Laboratory, Leiden University, 2300 RA Leiden, Netherlands. E-mail: bouwmeester@physics.ucsb.edu

tected with near–single-photon sensitivity (*4, 10, 11*).

Among the most recent results with such devices are the entangling of mechanical motion with microwave fields (*11*) and the implementation by Suh *et al.* of the BAE technique (*4*). Suh *et al.* applied two input or pump microwave tones to the LC circuit. One, at frequency $\omega_c - \omega_m$ (where ω_m is the resonant frequency of the movable capacitor plate), is called the red sideband, and the other at frequency $\omega_c + \omega_m$ is the blue sideband, relative to ω_c. The resulting interaction between the total number of microwave photons N_c in the LC resonator and the motion of the capacitor plate is proportional to $N_c\{X_1[1 + \cos(2\omega_m t)] + X_2 \sin(2\omega_m t)\}$. A time-averaged monitoring of the system will wash out the oscillating terms at $2\omega_m$ and only detects the slowly varying interaction with X_1. There is in principle no limit to the accuracy with which X_1 can be determined, provided that $\Delta X_1 \Delta X_2$ satisfies the Heisenberg uncertainty relation.

To experimentally analyze the effect of the two-tone BAE method, Suh *et al.* applied two additional but much weaker probe microwave tones. From the beating between those tones and the two pump tones, the spectral densities of the red- and blue-sideband tones are determined. Those spectral densities contain all of the information needed to extract the fluctuations in X_1 and X_2. In this way, Suh *et al.* convincingly demonstrated the BAE scheme. This QND measurement method, enabling squeezed states of the mechanical motion, is likely to lead to important applications in weak force detection and quantum information science. ∎

REFERENCES

1. V. B. Braginsky, Y. I. Vorontsov, *Sov. Phys. Usp.* **17**, 644 (1975).
2. K. S. Thorne, R. W. P. Drever, C. M. Caves, M. Zimmermann, V. D. Sandberg, *Phys. Rev. Lett.* **40**, 667 (1978).
3. V. B. Braginsky, Y. I. Vorontsov, K. S. Thorne, *Science* **209**, 547 (1980).
4. J. Suh *et al.*, *Science* **344**, 1262 (2014).
5. V. B. Braginsky, in *Gravitational Radiation and Gravitational Collapse* (Reidel, Dordrecht, Netherlands, 1974), p. 28.
6. C. Guerlin *et al.*, *Nature* **448**, 889 (2007).
7. C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sanberg, M. Zimmermann, *Rev. Mod. Phys.* **52**, 341 (1980).
8. A. A. Clerk, F. Marquardt, K. Jacobs, *New J. Phys.* **10**, 095010 (2008).
9. C. M. Caves, *Phys. Rev. Lett.* **45**, 75 (1980).
10. J. D. Teufel, T. Donner, M. A. Castellanos-Beltran, J. W. Harlow, K. W. Lehnert, *Nat. Nanotechnol.* **4**, 820 (2009).
11. T. A. Palomaki, J. D. Teufel, R. W. Simmonds, K. W. Lehnert, *Science* **342**, 710 (2013).
12. LIGO Scientific Collaboration, *Nat. Phys.* **7**, 962 (2011).
13. A. D. O'Connell *et al.*, *Nature* **464**, 697 (2010).
14. E. Verhagen *et al.*, *Nature* **482**, 63 (2012).
15. J. D. Teufel *et al.*, *Nature* **475**, 359 (2011).

MOLECULAR BIOLOGY

Molecular basis of transcription pausing

RNA polymerase progression is regulated by a ubiquitous DNA consensus sequence to mediate RNA utilization

By **Jeffrey W. Roberts**

During RNA synthesis, RNA polymerase moves erratically along DNA, frequently resting as it produces an RNA copy of the DNA sequence. Such pausing helps coordinate the appearance of a transcript with its utilization by cellular processes; to this end, the movement of RNA polymerase is modulated by mechanisms that determine its rate. For example, pausing is critical to regulatory activities of the enzyme such as the termination of transcription. It is also essential during early modifications of eukaryotic RNA polymerase II that activate the enzyme for elongation. Two reports analyzing transcription pausing on a global scale in *Escherichia coli*, by Larson *et al.* (*1*) and by Vvedenskaya *et al.* (*2*) on page 1285 of this issue, suggest new functions of pausing and reveal important aspects of its molecular basis.

The studies of Larson *et al.* and Vvedenskaya *et al.* follow decades of analysis of bacterial transcription that has illuminated the molecular basis of polymerase pausing events that serve critical regulatory functions. A transcription pause specified by the DNA sequence synchronizes the translation of RNA into protein with the transcription of leader regions of operons (groups of genes transcribed together) for amino acid biosynthesis; this coordination controls amino acid synthesis in response to amino acid availability (*3*). A protein-induced pause occurs when the *E. coli* initiation factor σ70 restrains RNA polymerase by binding a second occurrence of the "–10" promoter element. This paused polymerase provides a structure for engaging a transcription antiterminator (the bacteriophage λ Q protein) (*4*) that, in turn, inhibits transcription pauses, including those essential for transcription termination.

Knowledge about the interactions between nucleic acids and RNA polymerase that induce pausing comes partly from studies on the *E. coli* histidine biosynthesis operon. RNA polymerase pauses at the leader region of this cluster of genes (the "*his* pause"), allowing an essential RNA hairpin structure to form just upstream of
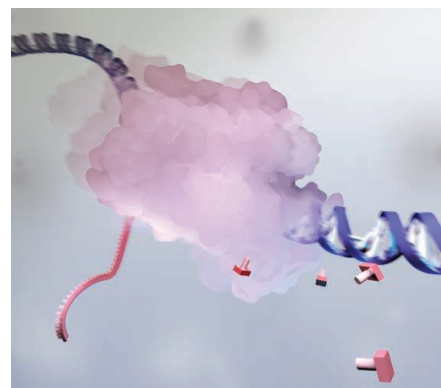
the RNA-DNA hybrid where RNA synthesis is templated in the polymerase's catalytic cleft. Importantly, however, other sequence elements are required to induce and stabilize the *his* pause—particularly the nucleotide at the newly formed, growing end of the RNA (pausing is favored by pyrimidines rather than purines) (*5*), and at the incoming nucleotide position [pausing is favored particularly by guanine (G)] (*6*), as well as surrounding elements. Biochemical and structural analyses have identified an endpoint of the pausing process called the "elemental pause" in which the catalytic structure in the active site is distorted, preventing further nucleotide addition (*7*). The elemental paused state also involves distinct conformational changes in the polymerase that may favor transcription termination and allow the *his* and related pauses to be stabilized by RNA hairpins (*8*).

Single-molecule analysis of transcribing RNA polymerase, at nearly single-nucleotide resolution, identified many specific pause sites in the *E. coli* genome (*9*). Pausing occurs on essentially any DNA, and very frequently—every 100 nucleotides or so. These "ubiquitous" pauses are only partly efficient (i.e., not always recognized as the enzyme transits), and mostly have not been associated with specific functions. However, their existence is consistent with biochemical experiments showing that the progress of RNA polymerase is generally erratic. A consensus sequence for ubiquitous pauses was identified, with two important elements: a preference for pyrimidine [mostly cytosine (C)] at the newly formed RNA end,
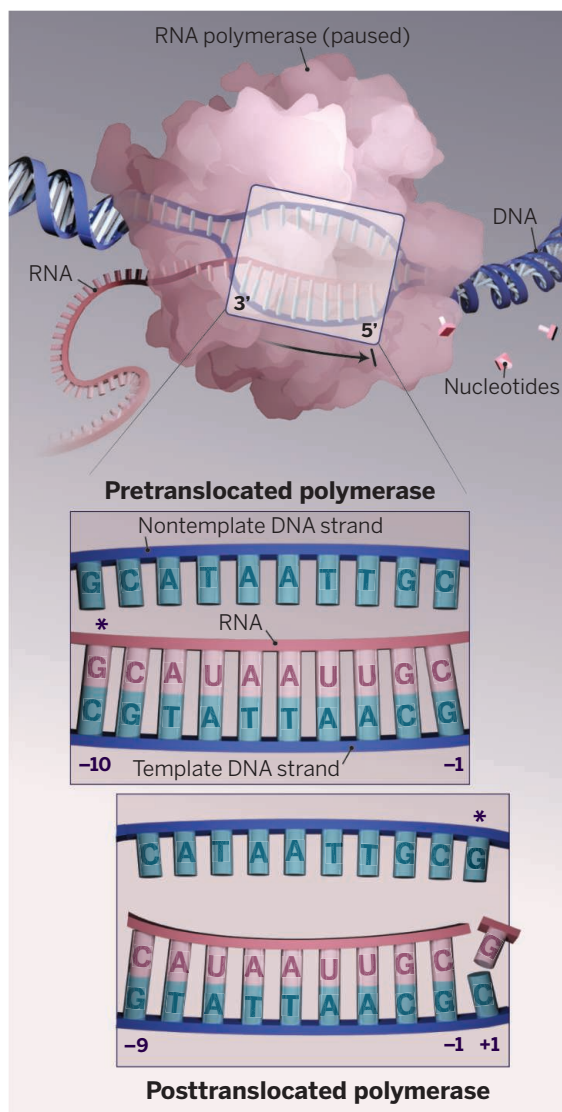
Polymerase, paused. During transcription, RNA exists in two states as RNA polymerase progresses: pretranslocated, just after the addition of the last nucleotide [here, cytosine (C)]; and posttranslocated, after all nucleic acids have shifted in register by one nucleotide relative to the enzyme, exposing the active site for binding of the next substrate molecule [here, guanine (G)]. The pretranslocated state is dominant in the pause. The critical G-C base (RNA-DNA) pair at position −10 in the pretranslocated state and the nontemplate DNA strand G bound in the polymerase in the posttranslocated state are marked with an asterisk.

followed by G to be incorporated next—just as found for the *his* pause; and a preference for G at position −10 of the RNA (10 nucleotides before the 3′ end), which is at the upstream boundary of the RNA-DNA templating hybrid. Remarkably, the tendency of a G in this position to induce pausing was recognized earlier, when DNA could be sequenced only through its transcript (*10*); it was thought that inhibited unwinding of the RNA-DNA hybrid underlies the pause.

Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14853, USA. E-mail: jwr7@cornell.edu

This ubiquitous pausing consensus sequence now has been refined and mapped exhaustively in the *E. coli* genome by Larson *et al.* and Vvedenskaya *et al.* (see the figure). In an analysis called native elongating transcript sequencing (NET-Seq) (*11*), transcripts associated with the whole cellular population of RNA polymerase are isolated from abruptly frozen cells and their growing ends are sequenced, giving a snapshot at nucleotide resolution of global transcription activity; DNA sites that are highly populated by RNA polymerase represent pauses. Larson *et al.* identified ~20,000 transcription pause sites in the *E. coli* genome, including those expected from previous analysis of known sites like the *his* pause. Their analysis raises interesting questions about the role of such abundant pausing sequences.

Primarily, Larson *et al.* note that pauses frequently occur exactly at the site of translation initiation, suggesting an important role in gene expression. This coincidence of events is understandable when you examine the sequences. The consensus sequence in RNA for RNA polymerase pausing is $G_{-10}Y_{-1}G_{+1}$ [G at position −10 and at the site after the pause; Y denotes either C or uracil (U) at the RNA end] according to Larson *et al.* and Vvedenskaya *et al.* The Shine-Dalgarno consensus sequence in RNA that the small-subunit ribosome recognizes is AGGAGG [adenine (A)] providing the G at the −10 position; the downstream initiation codon for RNA translation is AUG, providing (for *E. coli*) the U at the pause end at position −1, with a following G at position +1. A slightly modified pausing consensus sequence in the bacterium *Bacillus subtilis* accommodates the difference in spacing between the Shine-Dalgarno sequence and the initiation codon. What might be the role of a pause exactly at the translation initiation site? Because the ribosome binding site is physically concealed by RNA at the pause, pausing may enable some process that prepares the RNA for translation once RNA polymerase transits the pause site. Larson *et al.* suggest that the pause allows upstream

RNA secondary structure to resolve in order to present the initiation region properly to the ribosome.

A particularly informative application of NET-Seq that provides new mechanistic information about pausing is based on the discovery of a specific binding site in RNA polymerase [the core recognition element (CRE)] for G in the nontemplate DNA strand (the strand not transcribed), at position +1 in the "posttranslocated" structure (*12*). It could be that specific binding of a nucleotide to the enzyme in this position enhances pausing by slowing translocation; surprisingly, however, Vvedenskaya *et al.* find the opposite. Cells altered to destroy the G binding site have up to twice as many sites of pausing as in wild-type cells, with a greater preference for G as the incoming nucleotide. However, this result is understandable in terms of the translocation cycle of RNA polymerase and the ubiquitous pausing sequence that has G at position +1. Binding of G at position +1 to CRE only occurs in the posttranslocated state, which would thus be favored over the pretranslocated state. Hence, if G binding inhibits pausing, then the rate-limiting paused structure must be in the pretranslocated state (a conclusion also made by Larson *et al.* from biochemical experiments). This is an important insight into the sequence of protein–nucleic acid interactions that occur in pausing. Vvedenskaya *et al.* suggest that the actual role of the G binding site is to promote translocation and thus inhibit pausing, to smooth out adventitious pauses in genomic DNA.

The studies by Larson *et al.* and Vvedenskaya *et al.* provide a refined and detailed analysis of DNA sequence–induced transcription pausing. As a core process in gene expression, this understanding is relevant not only for the basic biology of transcription, but also has applications in synthetic biology and the design of genetic circuits. ∎

REFERENCES

1. M. H. Larson *et al.*, *Science* **344**, 1042 (2014).
2. I. O. Vvedenskaya *et al.*, *Science* **344**, 1285 (2014).
3. R. Landick, J. Turnbough, C., C. Yanofsky, in *Escherichia coli and Salmonella*, F. Neidhardt, Ed. (American Society for Microbiology, Washington, DC, 1996), vol. 1, pp. 1263–1286.
4. S. A. Perdue, J. W. Roberts, *J. Mol. Biol.* **412**, 782 (2011).
5. C. L. Chan, R. Landick, *J. Mol. Biol.* **233**, 25 (1993).
6. D. N. Lee, L. Phung, J. Stewart, R. Landick, *J. Biol. Chem.* **265**, 15145 (1990).
7. I. Toulokhonov, J. Zhang, M. Palangat, R. Landick, *Mol. Cell* **27**, 406 (2007).
8. A. Weixlbaumer, K. Leon, R. Landick, S. A. Darst, *Cell* **152**, 431 (2013).
9. K. M. Herbert *et al.*, *Cell* **125**, 1083 (2006).
10. W. Gilbert, in *RNA Polymerase*, R. L. a. M. J. Chamberlin, Ed. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1976), pp. 193–205.
11. L. S. Churchman, J. S. Weissman, *Nature* **469**, 368 (2011).
12. Y. Zhang *et al.*, *Science* **338**, 1076 (2012).

OCEANS

A salty start to modern ocean circulation

Water flow out of the Mediterranean is linked to large-scale ocean circulation patterns

By **Gabriel Filippelli**

In a small area of the high North Atlantic, dense water sinks to the deep ocean and begins a ~2000-year journey through the Atlantic, Southern, Indian, and Pacific oceans until it emerges again off the west coasts of North and South America (1). Yet 6 million years ago, ocean circulation patterns were quite different, with gateways through the Central American Seaway (2) and the Indonesian Seaway (3) allowing deep ocean flow at mid-latitudes. On page 1244 of this issue, Hernández-Molina et al. (4) document the history and impact of ocean flow out of the Mediterranean Sea at another such gateway: the Strait of Gibraltar. The results highlight the importance of this gateway in modulating global ocean circulation and climate patterns.

Numerous studies have found evidence for links between tectonic events, such as the opening and closure of ocean gateways, and ocean circulation patterns. Yet one key question has remained: What is the history of the warm salty ocean flow out of the Mediterranean Sea? Water flowing out of the Mediterranean is far from the high North Atlantic but is nonetheless critical for deep water production in that region, contributing salty and warm water to the northward flow at intermediate depth and providing one of the dense ingredients of North Atlantic Deep Water (NADW).

Hernández-Molina et al. now report the first results from the recently completed Integrated Ocean Drilling Program (IODP) Expedition 339 in the Gulf of Cadiz, just west of the Strait of Gibraltar. The data provide valuable clues to the complicated links among tectonic events in the narrow Strait of Gibraltar, dense water flow out of the restricted Mediterranean, and thick continental margin–hugging sediment deposits (contourites) produced when the Mediterranean Outflow Water (MOW) enters the Atlantic Ocean in a powerful jet (see the figure).

Department of Earth Sciences, Indiana University–Purdue University Indianapolis, Indianapolis, IN 46202, USA. E-mail: gfilippe@iupui.edu

5.6 to 5.2 Ma: No flow from closed Mediterranean

4.5 to 3.5 Ma: Weak flow through restricted gateway
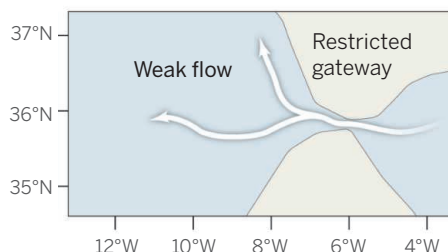
3.2 Ma to present: Moderate to strong flow through open gateway
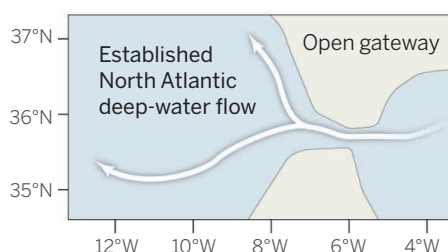
How tectonics and ocean circulation are coupled. Hernández-Molina et al. show that tectonic changes have led to substantial variations in the flow of water out of the Strait of Gibraltar over the past ~6 million years, with strong impacts on global ocean heat transport.

The typical approach to developing high-resolution paleoceanographic records is to drill into marine sediments in places where accumulation rates are high and erosion is very low, typically far away from ocean gateways and tectonically active regions. In contrast, Expedition 339 intentionally targeted the Gulf of Cadiz, near the Strait of Gibraltar and the entry point of MOW into the Atlantic. The choice of expedition drilling sites was guided by a detailed seismic survey of the region, which provided information about the geometry and extent, but not the ages, of sedimentary layers along the continental margin. During the expedition, five sites were drilled in the Gulf of Cadiz and two sites off the West Iberian margin. The goal was to obtain information on the composition and ages of the sedimentary layers. The timing of the sedimentary layers and of the gaps (hiatuses) between them would be indicative of the flow of MOW through the Strait of Gibraltar, which in turn was largely controlled by the tectonics of this gateway.

The general tectonic evolution of the Strait of Gibraltar has been studied before (5, 6). The results from Expedition 339 now reveal the detailed history of MOW flow through this gateway. The MOW is restricted to a very narrow flow path out of the Strait of Gibraltar that amplifies and directs the flow in much the same way that a river speeds up through narrows. Sedimentation in the targeted drilling region thus provides a sensitive record of MOW flow history. Interpreting the sedimentary history recorded in cores and seismic data is by no means straightforward, because material eroded from one place on the continental margin is delivered down-slope or farther away and the MOW flow direction varies as a result of changes in seafloor topography in this very tectonically active region. But these issues were largely addressed by the drilling strategy and the metadata provided by seismic surveys of sedimentary layers and their three-dimensional geometry, resulting in an extremely successful expedition.

Results from Expedition 339 revealed several pulses of MOW production. These pulses coincided with major events in ocean circulation. A relatively weak early MOW began ~4.5 million years ago (Ma), after the Strait of Gibraltar opened at ~5.3 Ma. This initial MOW component probably had only a minor influence on North Atlantic circulation (7), although the picture is complicated somewhat by a reorganization in low-latitude flow patterns from the closure of the Central American Seaway, which began at this time (8). During the mid-Pliocene warm period (3.2 to 3.0 Ma), enhanced production of warm, salty MOW may have initiated the general North Atlantic circulation pattern seen today. This finding provides an important time constraint on MOW influence in North Atlantic circulation, given the final closure of the Central American Seaway at the same time.

MOW intensified again in the early Quaternary (2.4 to 2.1 Ma), likely contributing to the near-modern interplay among the MOW, Atlantic Meridional Overturning Circulation (AMOC), and NADW, and ultimately the heat balance dynamics that initiated a major glaciation event at 2.15 Ma. An additional intensification of MOW between 0.9 and 0.7 Ma coincides with the Mid-Pleistocene Transition that marks the onset of the current 100,000-year glacial/interglacial cycles.

More information about the global impacts of MOW and related ocean circulation patterns over the Pleistocene may come from the Shackleton Site, one of the two drilling sites off the West Iberian margin. Drilling at this site is an important component of IODP Expedition 339 aimed at recovering substantial quantities of high-resolution core material. Initial results from these materials (9, 10) promise the potential for a marine-sediment analog to the polar ice cores over the past ~1 million years from this important mixing point of ocean circulation systems.

The near-source drilling approach used in Expedition 339 has some limitations. The sedimentary record is lost or at best altered substantially during the deposition hiatuses that mark intervals of enhanced MOW production. More traditional data from distant sites are thus required to understand the true impact of MOW production. The expedition data also cannot resolve what thresholds in the MOW density or production rate must be met to alter North Atlantic circulation patterns. Nonetheless, the bold move to examine integrated tectonic, gateway, and ocean circulation dynamics by combining ocean drilling and detailed seismic interpretations has paid off. The initial successes provide support for other proposed efforts, such as drilling downstream of the Drake Passage gateway to further understand how the Antarctic Circumpolar Current developed. ∎

REFERENCES

1. C. Wunsch, *Science* **298**, 1179 (2002).
2. G. Bartoli *et al.*, *Earth Planet. Sci. Lett.* **237**, 33 (2005).
3. M. A. Cane, P. Molnar, *Nature* **411**, 157 (2001).
4. F. J. Hernández-Molina *et al.*, *Science* **344**, 1244 (2014).
5. C. Roque *et al.*, *Mar. Geol.* **303–306**, 42 (2012).
6. P. Martinez-Garcia, M. Comas, J. I. Soto, L. Lonergan, A. B. Watts, *Basin Res.* **25**, 361 (2013).
7. M. Rogerson, E. J. Rohling, G. R. Bigg, J. Ramirez, *Rev. Geophys.* **50**, RG2003 (2012).
8. M. A. Maslin, X. S. Li, M.-F. Loutre, A. Berger, *Quat. Sci. Rev.* **17**, 411 (1998).
9. F. Abrantes, D. Hodell, G. Carrara, L. Batista, H. Duarte, *Sci. Drill.* **9**, 50 (2010).
10. D. A. Hodell, L. Lourens, D. A. V. Stow, J. Hernández-Molina, C. A. Alvarez Zarikian, *Sci. Drill.* **16**, 13 (2013).

EPIDEMIOLOGY

It helps to be well connected

Highly connected populations of the weedy plant *Plantago* are less likely to be colonized by a fungal pathogen

By **Meghan A. Duffy**

Parasites are everywhere. Yet despite this ubiquity, they are not in all places at all times. For a particular host-parasite pairing, infections vary greatly over space and time, with some but not all host populations suffering infections in a given year. What drives this variation? This question is important not just because parasites can have major ecological and evolutionary impacts, but also because disease outbreaks can have devastating impacts in agricultural systems and on species of conservation concern (1, 2). On page 1289 of this issue, Jousimo *et al.* (3) investigate how being highly connected to other populations affects the likelihood of a *Plantago* population being colonized by a fungal pathogen.

The results presented by the authors come from a long-term, large-scale study on a common weedy plant, *Plantago lanceolata*, and its fungal pathogen, powdery mildew. The study involved a small army of about 40 field assistants who conducted annual censuses of ~4000 populations over a 12-year period. The pattern revealed by this remarkable effort was surprising: The more connected a *Plantago* population was to other populations, the less likely it was to be colonized by the fungal pathogen (see the figure).

This finding seems counterintuitive. If all else were equal, traditional theory for a set of populations linked by migration (a metapopulation) would predict that being closely connected to other populations should make colonization more likely, not less (4). But of course, all else is rarely equal. The apparent discrepancy between theory and Jousimo *et al.*'s findings can be resolved by considering differences in the average resistance of plants in different populations. Jousimo *et al.* show experimentally that plants from highly connected populations are more resistant to the pathogen than those from less connected populations. This higher resistance makes it harder for the pathogen to establish itself in highly connected populations. This key finding fits well within a growing body of literature showing that evolution can be rapid and has the potential to profoundly affect ecological dynamics (5).

Why are more connected populations more resistant? One possible reason is that more connected populations are exposed to the pathogen more often, selecting for higher resistance. This would maintain high resistance in the highly connected populations, whereas resistance would be lost from the less connected populations because of fitness costs associated with resistance. There is some evidence for this in the data: The only low-resistance populations are ones with low connectivity. However, this explanation would also require highly connected populations to be routinely colonized by the pathogen. Jousimo *et al.*'s data

Plantago lanceolata
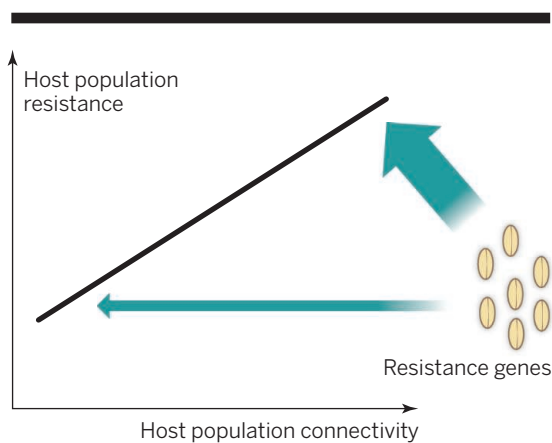
Why highly connected populations fare better. Jousimo *et al.* show that highly connected populations of the weedy plant *Plantago lanceolata* are more resistant to a fungal pathogen and thus less likely to be colonized by it. One possible driver of this pattern is increased gene flow to highly connected populations, leading to a greater influx of resistant alleles that selection can then favor. The authors estimated connectivity by combining the size of a population (which influences how much space it occupies) and the distance between populations.

suggest that this is not the case, although it seems possible that highly susceptible genotypes might become infected and die at a very early stage, before those infections can be detected by the annual survey.

Further work is thus needed to elucidate the eco-evolutionary dynamics in this system and to determine the drivers of the observed pattern. Knowledge of the mechanism underlying the results will help to predict when similar patterns of low resistance are likely to be seen in populations that are less connected to others in the metapopulation. This is particularly pressing given widespread habitat fragmentation; the results of the current study suggest that, in some cases, fragmentation might increase the likelihood of a population suffering disease outbreaks.

It also remains to be shown whether the observed pattern holds for less fragmented metapopulations. The *Plantago* metapopulation studied by Jousimo *et al.* on the Åland archipelago in Finland is highly fragmented; do metapopulations with higher rates of migration between populations show similar patterns?

Another notable aspect of this *Plantago*-mildew system is the very high turnover in the pathogen population. In any given year, only a small fraction (generally 1 to 7%) of the *Plantago* populations were infected by the pathogen. Moreover, there was sub-

stantial turnover between years; only a single population was infected in all 12 years of the study. When populations were reinfected, the pathogen usually came from an overwintering population located within a few kilometers. However, occasional long-distance leaps allowed the pathogen to reach new regions, possibly facilitated by populations along roadsides. Mostly local transmission with occasional large-scale leaps facilitated by modes of transportation has also been found for other plant pathogens (*6*) as well as for diseases of humans (*7*) and other animals (*8*), which suggests that this might be a common phenomenon.

Few studies to date have analyzed host-parasite interactions in a spatially explicit manner (*9*), but it is already clear that there is much to be gained from this approach. Perhaps most important, studying infectious diseases in a metapopulation framework allows us to understand not just what permits outbreaks of infectious diseases to occur, but also the factors that determine the persistence of parasites between outbreaks, as shown previously for human infectious diseases (*10–12*). The main factor preventing more widespread application of the metapopulation approach to nonhuman infectious diseases is a lack of suitable data sets. Jousimo *et al.*'s results demonstrate that studies of this kind can provide important and unexpected insights into the drivers of disease dynamics in nature. ■

REFERENCES

1. C. Combes, *Parasitism: The Ecology and Evolution of Intimate Interactions* (Univ. of Chicago Press, Chicago, 2001).
2. R. S. Ostfeld, F. Keesing, V. T. Eviner, Eds., *Infectious Disease Ecology: Effects of Ecosystems on Disease and of Disease on Ecosystems* (Princeton Univ. Press, Princeton, NJ, 2008).
3. J. Jousimo et al., *Science* **344**, 1289 (2014).
4. I. Hanski, *Oikos* **87**, 209 (1999).
5. T. W. Schoener, *Science* **331**, 426 (2011).
6. E. S. Jules, M. J. Kauffman, W. D. Ritts, A. L. Carroll, *Ecology* **83**, 3167 (2002).
7. C. Viboud et al., *Science* **312**, 447 (2006).
8. C. A. Russell, D. L. Smith, J. E. Childs, L. A. Real, *PLOS Biol.* **3**, 382 (2005).
9. M. J. Keeling, P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton Univ. Press, Princeton, NJ, 2008).
10. M. J. Keeling, C. A. Gilligan, *Nature* **407**, 903 (2000).
11. C. J. E. Metcalf, K. Hampson, A. J. Tatem, B. T. Grenfell, O. N. Bjornstad, *PLOS ONE* **8**, e74696 (2013).
12. Y. Xia, O. N. Bjørnstad, B. T. Grenfell, *Am. Nat.* **164**, 267 (2004).

Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48104, USA. E-mail: duffymeg@umich.edu

10.1126/science.1255350

Measuring the path toward malaria elimination

Developing new targets and milestones from standard surveillance data

By **Thomas S. Churcher,**[1] **Justin M. Cohen,**[2] **Joseph Novotny,**[2,3] **Nyasatu Ntshalintshali,**[2,3] **Simon Kunene,**[4] **Simon Cauchemez**[1,5]*

In many parts of the world, malaria elimination—defined by the World Health Organization (WHO) as the absence of locally acquired malaria cases in the country—is being considered as a target because of recent successes in reducing disease burden (*1*, *2*). Rigorous evaluation of malaria elimination programs is essential for financial and political support to be maintained. Yet such evaluation remains challenging, and appropriate metrics to ascertain "success" are needed.

POLICY

Although the long-term focus on elimination is commendable, evaluation of programs cannot rely on a dichotomous approach where "success" would correspond to no locally acquired malaria cases, and anything else would be seen as failure. Evaluation needs to take into account the local and regional epidemiological circumstances, because countries that manage to control local transmission to relatively low levels but receive large numbers of imported cases are likely to see locally acquired cases (*3*). Effective programs may therefore be wrongly perceived as unsuccessful, which would jeopardize their long-term viability. In addition to the long-term objective of elimination, we must develop intermediate milestones that better capture and acknowledge these scenarios.

Such a milestone should be the interruption of endemic transmission, meaning that the country would eventually see malaria go away if importation ceased. The country could then aim to reduce chains of transmission generated from imported cases down to zero. Although this sounds sensible, there are no operational measures and tools for evaluating local transmission

Malaria is transmitted by female anopheline mosquitoes.

as it approaches elimination, because incidence estimates require either surveillance data with perfect case detection or surveys with impractically large sample sizes (*4*). In addition, incidence does not measure local transmission in the context of importation.

Here, we present simple, theoretically sound, and robust metrics to ascertain progress toward elimination. Aside from their importance for program evaluation, these metrics also allow more-refined assessment of the epidemiological situation, which could lead to more accurate evaluation of different elimination strategies (for example, differentiating between persistent low-level endemic transmission or repeated importation and extinction events that may require different public health approaches).

HAS ENDEMIC TRANSMISSION BEEN HALTED? The term "controlled nonendemic malaria" describes areas where interventions have interrupted endemic transmission but where some local transmission from imported cases remains (*3*). The reproduction number *R* (average number of new persons infected by a person

with malaria given current control interventions) would be a theoretically sound measure to ascertain controlled nonendemic malaria (as *R* < 1 indicates that transmission cannot be self-sustaining locally) (*3*). However, these considerations have had limited practical impact because (i) it is not possible to directly measure *R* in the field, because chains of transmission are typically not observed; and (ii) estimates of *R* derived from mathematical models require local information, which is seldom collected.

To overcome these issues, we propose a method to determine the status of controlled non-endemic malaria that only requires standard data (the numbers of local and imported malaria cases detected by surveillance). The approach uses a simple criterion—Is the proportion of imported cases among detected cases above a certain threshold?—to test the hypothesis *R* ≥ 1. The approach accounts for factors such as incomplete case detection and considers a worst-case scenario to ensure that the test for *R* ≥ 1 is conservative. Last, it can easily be tuned to consider more stringent milestones, for example, *R* < 0.5, in which

imported cases cause shorter chains of transmission and the disease is closer to local elimination. A description of the method and a user-friendly tool can be found in the supplementary materials (SM).

Given the number (*n*) of malaria cases detected by surveillance, the proportion of cases that must be imported to provide statistical evidence that controlled non-endemic malaria has been reached (i.e., reject the hypothesis that *R* ≥ 1) is shown in part (A) of the chart. This proportion is 48% when *n* = 50 but drops to 32% when *n* reaches 500 cases (see also table S1).

We apply this approach to Swaziland, a country that embarked on an elimination campaign in 2008 (*5*). Malaria is a notifiable disease (required by law to be reported) in Swaziland. Since 2009, all cases confirmed by a rapid diagnostic test and/or microscopy are reported to a central database. An investigation of 64% of confirmed cases was carried out, and travel history was ascertained, with those reporting travel to endemic regions within the previous 2 to 8 weeks classified as imported cases. Cases identified by screening people associated with cases identified by routine surveillance were omitted from the analysis to avoid overestimating *R* (see SM) (*6*).

The results from weekly epidemiological reports are shown in part (B) of the chart. Thirty-six percent (52 out of 143), 45% (170 out of 377), and 67% (153 out of 229) of investigated cases were imported in 2010, 2011, and 2012, respectively. Thus, the status of controlled non-endemic malaria was reached in 2011 and 2012 but not in 2010 (borderline). This provides evidence that, since 2011, Swaziland has halted endemic transmission at the national level and that malaria would be eliminated if the current level of control was continued and importations ceased. Swaziland has a very effective malaria surveillance network (*7*), but the approach would remain valid if the case-detection rate (the proportion of malaria cases identified) or the investigation rate (the proportion of confirmed cases where travel history is ascertained) were lower (see SM).

ESTIMATING *R* BY SEASON. Beyond determining whether *R* is beneath a certain threshold, it may be necessary to estimate *R* and its associated uncertainty directly. For example, it may be important to separate *R* estimates for high and low seasons to evaluate the effectiveness of local interventions and to predict their impact in countries with different seasonal patterns. For instances in which there is a good understanding of local surveillance and health systems (e.g., the proportion of malaria cases de-
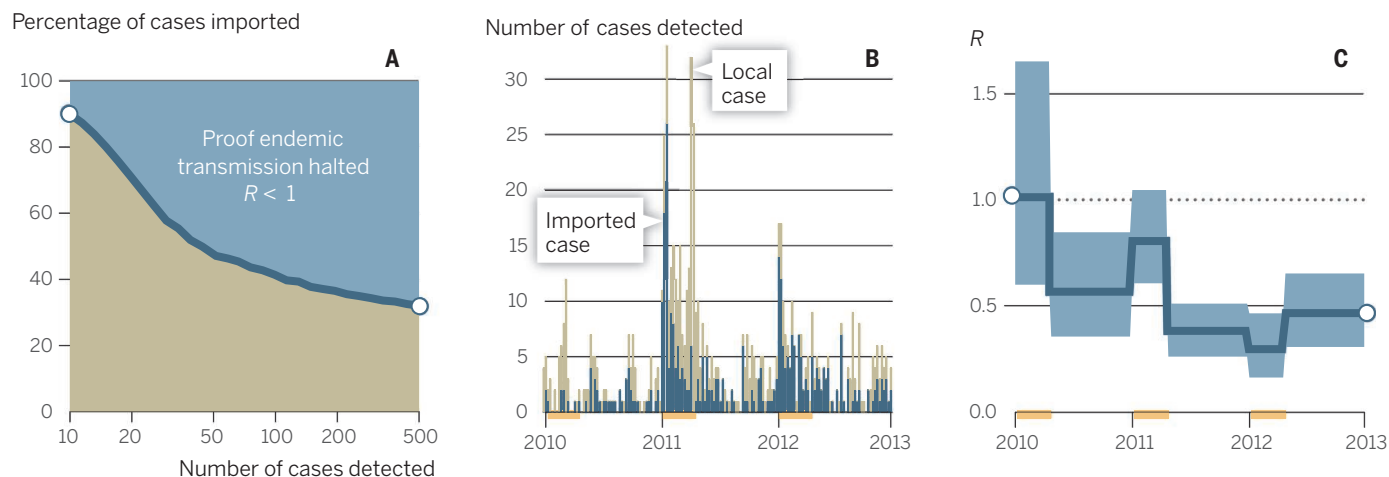
Tracking malaria transmission. (A) The percentage of imported cases required to confirm that endemic malaria transmission has been halted. If the percentage of imported cases is greater than the solid dark blue line (blue area), there is statistical evidence that malaria is no longer endemic. The tan areas show where the hypothesis that $R \geq 1$ cannot be rejected, either because there is insufficient evidence or because endemic transmission is ongoing. **(B)** The weekly incidence of malaria cases investigated in Swaziland. All cases are likely to be falciparum malaria (9). Orange lines on the x axis indicate the high-transmission season. **(C)** Estimates of the reproduction number R for each season in Swaziland, with shaded area indicating 95% credibility intervals. See SM for details.

tected and treated), we built upon recent advances in statistics (8) to estimate temporal trends in R from the epidemic curves of imported and local cases [see chart (B); see SM]. These trends are shown in part (C) of the chart for Swaziland. The reproduction number R was not significantly different from 1 during the high season for 2010 and 2011; but was significantly smaller for 2012. A sensitivity analysis (SM) shows that estimates were robust to changes in the assumed case-detection rates. This method is more precise than the threshold framework presented on the left in the figure, although the model-fitting process makes it less user-friendly.

LIMITATIONS AND FURTHER DEVELOPMENTS. Here, we generated countrywide estimates of R. Given the important spatial heterogeneities in malaria transmission, national statistics may mask the presence of more localized pockets of endemic transmission. Further work is needed to determine at which optimal spatial resolution analyses should be carried out (see SM). Our approach was developed for the monitoring of falciparum malaria elimination. However, a majority of countries affected by malaria are likely to have cases of *Plasmodium vivax* (2). In this species, new cases

can be caused by a relapse of a previous infection as well as by local transmission. Our test for $R \geq 1$ remains valid in areas affected by *P. vivax* although its power is diminished (see SM).

DO WE NEED TO CHANGE STRATEGY? Our results show, for the first time, that a country in mainland sub-Saharan Africa has halted endemic transmission. The method can also help predict the likely impact of interventions. For $R = 0.5$, 500 imported cases are expected to result in about 500 subsequent local cases before the disease is eliminated. If further investment in vector control could reduce R to 0.3, the number of local cases would drop to about 200. In contrast, relaxing control measures, raising R to 0.6, would generate longer chains of transmission and a total of 750 local cases. Understanding the current level of transmission might have programmatic implications. For example, when R is very low it might be feasible and economically advantageous to reduce routine disease control and concentrate on outbreak investigation and response.

In endemic countries, estimating R may help determine whether malaria elimination is feasible in a particular location by quantifying the current level of local transmission and then using mathematical models to predict what additional control is required to push R below 1. For example, if R is estimated to be 3 in a particular location and surveys indicate current effective bed-net coverage is 40%, increasing effective bed-net coverage to 60% in that loca-

tion should be sufficient to reach controlled non-endemic malaria (9).

There is little doubt that financial and political support for the malaria elimination agenda will fade out if programs are systematically failing to reach their targets. Instead of just presenting numbers of cases, control programs should report estimates of R to better reflect the current level of transmission given their level of disease importation. This will allow a more rigorous evaluation of different elimination strategies to enable the successes of countries like Swaziland to be maintained and replicated elsewhere. ■

REFERENCES AND NOTES

1. WHO, "Disease surveillance for malaria elimination: An operational manual" (WHO, Geneva, 2012).
2. WHO, "World malaria report: 2013" (WHO, Geneva, 2013).
3. J. M. Cohen, B. Moonen, R. W. Snow, D. L. Smith, *Malar. J.* **9**, 213 (2010).
4. S. I. Hay, D. L. Smith, R. W. Snow, *Lancet Infect. Dis.* **8**, 369 (2008).
5. S. Kunene, A. A. Phillips, R. D. Gosling, D. Kandula, J. M. Novotny, *Malar. J.* **10**, 313 (2011).
6. S. Cauchemez *et al.*, *PLOS Med.* **10**, e1001399 (2013).
7. M. S. Hsiang *et al.*, *PLOS ONE* **7**, e29550 (2012).
8. J. Dureau, K. Kalogeropoulos, M. Baguelin, *Biostatistics* **14**, 541 (2013).
9. D. L. Smith, S. I. Hay, A. M. Noor, R. W. Snow, *Trends Parasitol.* **25**, 511 (2009).

¹Department of Infectious Disease Epidemiology, Imperial College London, London, UK. ²Clinton Health Access Initiative, Boston, MA 02127, USA. ³Global Health Group, University of California, San Francisco, CA 94143, USA. ⁴National Malaria Control Program, Manzini, Swaziland. ⁵Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, Paris, France.
*Corresponding author. E-mail: simon.cauchemez@pasteur.fr

10.1126/science.1251449

Gary Becker (1930-2014)

An economist used sociology and economics to examine life and explain human behavior

By **Edward L. Glaeser** and **Andrei Shleifer**

Gary Becker, who died on 3 May 2014 at the age of 83, redefined economics both in its methodology and scope. He radically expanded the sphere of economic analysis. As the range of issues and especially data in economics increased over the last half century, Becker's approach became more and more relevant and modern. He was awarded the 1992 Nobel Prize in Economics for "having extended the domain of microeconomic analysis to a wide range of human behavior and interaction, including nonmarket behavior."

Becker grew up in Brooklyn, New York, and earned a B.A. at Princeton University in 1951 and a Ph.D. at the University of Chicago in 1955, where he was a student of the Nobel Prize–winning economist Milton Friedman. Becker grasped that Friedman's methodological definition of economics—as a discipline based on formal models that yielded refutable hypotheses—meant that the field was no longer tied to any traditional topic, such as monetary policy or international trade. Even as a graduate student, Becker used the methods of economics to investigate political behavior and racial disparities. His doctoral dissertation on "The Economics of Discrimination" showed the full power of economics. The study showed how racism could be a money-losing proposition for businesses. Becker argued that market pressure forces individuals and firms to bear the costs of racism, and would thus discourage its practice. Although he was clear that competitive pressures do not always destroy discrimination, he showed how competition can be a force for tolerance.

Two broad themes reappear throughout Becker's work. One is the centrality of rational individual choice, even in domains widely regarded as emotional and immune to economic analysis, such as racism. Another is the power of competition in shaping broad social and political outcomes, and not just prices of traditional commodities.

Becker left Chicago for Columbia University, New York, in 1957 where he taught until his return to Chicago in 1969. His

Harvard University, Boston, MA, USA. E-mail:shleifer@fas.harvard.edu

best-known book from this period, *Human Capital* (1964), is a masterly analysis of investment in education, combining formal theory, statistics, and descriptive work to solve old riddles and generate new questions. Becker wondered whether credit constraints were the reason why so many people failed to take advantage of the large financial returns
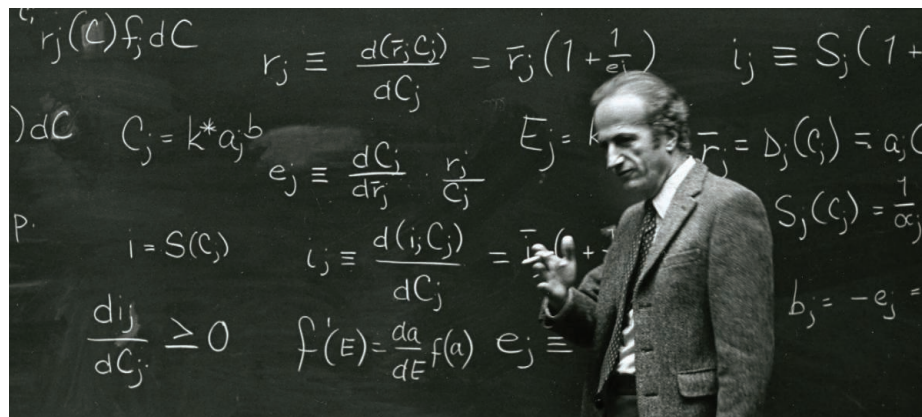
from education, thus explaining the need for student loans. He asked why firms should invest in training if workers could then turn around and use those skills to bargain for higher wages or leave for other firms.

Becker's period at Columbia University was marked by a surge of creativity. He wrote about time allocation within the household, explaining why the value of time for non-working spouses could be just as high as that of their highly paid companions. He wrote on fertility, explaining why women's education reduced child-bearing by increasing the value of their time. He described the trade-off between the quantity and quality of children, suggesting that greater educational attainment should lead to smaller households. The analysis of fertility led Becker to predict that world population growth would decline as countries grew richer and better educated, a prediction that proved true and is self-evident today, but was highly controversial in the 1960s, when most pundits forecast exponentially growing world population.

In Becker's view, economics was defined by more than just formal models and testable hypotheses. Economics must begin with decision-making agents who rationally maximize "utility" (the hedonic value of something). Contrary to the caricature of

economics in which all people care about is money and goods, Becker embraced an extremely broad conception of human nature, in which love, hatred, and altruism are all part of utility.

In his brilliant 1962 essay "Irrational Behavior and Economic Theory," Becker pointed out that economics could still yield meaningful predictions if humans acted in a random fashion but were subjected to resource constraints. Still, Becker insisted on the assumption of human rationality, despite his extremely broad view of what enters human utility. Perhaps he feared that opening the doors to irrationality would create a discipline without order, nor the ability to generate falsifiable predictions. Only

in recent years, with tremendous growth of experimental and survey evidence, have economists been able to move away from rationality and still generate testable and falsifiable predictions.

Becker would continue writing papers and books until his death. He worked on the family, fads and fashions, advertising, addiction, health, and politics. As he approached each new topic, even late in life, his work would often be the lodestone that defined the issues for years to come. But his influence is only partially reflected in his writings. He followed Friedman in teaching an introductory microeconomics course (called "Price Theory") to Ph.D. students at the University of Chicago. This course was the central intellectual experience in the life of many of them [including one of the authors (E.L.G.) of this reflection]. Becker was astonishing in seminars, with an ability to cut immediately to the heart of any paper.

Gary Becker saw scholarship as a high and moral calling and expected a devotion to scholarship from his colleagues and his students. He was enormously generous to all around him and continues to inspire thousands of social scientists. Economics has lost one of its greatest leaders. ∎

EVOLUTION AND DEVELOPMENT

Enriching macroevolution

By **Carl Simpson** *and*
Douglas H. Erwin

Homology, Genes, and Evolutionary Innovation
Günter P. Wagner
Princeton University Press, 2014. 494 pp.

What drives evolutionary innovation? This question has become a central focus of evolutionary biologists over the past decade as comparative evolutionary developmental biology ("evo-devo") has established a mechanistic basis for understanding the evolution of form and, increasingly, the evolution of developmental processes themselves. The challenge has been to build a conceptual framework that makes sense of the facts of development and explains how development can evolve while producing viable offspring. Prior to the codification of evo-devo, heterochrony was invoked for this purpose (*1*). Heterochrony produces new phenotypes by varying the rates of somatic and reproductive maturation, and for a time it was a promising conceptual organizer of developmental evolution. But heterochrony can only modify existing allometric variation and thus cannot explain the origin of novel traits—and the history of life is strung with novel traits.

This issue of the developmental origin of novel traits is the central focus of Günter Wagner's rich and insightful *Homology, Genes, and Evolutionary Innovation*. Wagner, an evolutionary biologist at Yale, has been studying evolutionary novelty empirically and conceptually for over two decades. This deep engagement is evident in the rigor of his arguments for a new approach to the study of developmental evolution. In the book, Wagner addresses the mechanisms that allow the generation of new morphological novelties, in contrast with others who have focused on innovation in networks of DNA, RNA, or proteins (*2*). Wagner emphasizes the primacy of traits that require a substantial but specific genetic change: the origin of novel character identity networks (ChINs). ChINs are self-regulatory networks of genes

The reviewers are in the Department of Paleobiology, MRC-121, National Museum of Natural History, Washington, DC 20013-7012, USA. E-mail: simpsoncg@si.edu and erwind@si.edu

State for a character identity. The elytra (protective cover) of beetles, such as this glorious scarab (*Chrysina gloriosa*), is a character state of insect forewings.

associated with a morphologic character and occurring in an intermediate position within gene regulatory networks. Serving to produce characters that share identity even if they are dissimilar or in another place on the body, ChINs thus generate homologous traits. [ChINs are conceptually very similar to the kernels discussed by Eric Davidson and one of us (*3*, *4*), indeed more similar than Wagner acknowledges.]

Homology is conceptually messy because traits change both as morphological features—a bat's wings are homologous to a frog's arms—and at the developmental level, where different processes can generate a homologous character. The challenge is to understand how traits can share identity while similarities at the phenotypic and genetic levels dissipate over evolutionary time.

Wagner argues that homologous characters need share only a common ChIN: neither a character's embryological origin nor its outward form are relevant. This powerful idea focuses our attention on one important type of organismal characteristic out of the many in the developmental fray. When homologs first evolve, they and their associated ChINs are novelties: new and unique characters and gene networks. The structure produced by a ChIN can be of any complexity, and, importantly, the parts of a homologous structure are not automatically homologous
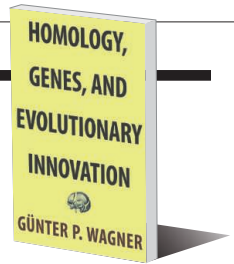
themselves. Previous notions of homology are hierarchical: limbs and the digits within are each homologous among tetrapods. Wagner's notion does not recognize digits as being homologous unless they are generated by a separate ChIN from the one that initiates the limb. ChINs allow a precise operational distinction between character identity and character state. Homologous traits have a shared identity that evolves by the origin of a novel ChIN, and they indicate specific ways in which development evolves. In contrast, character states are variations of a homologous trait. As Wagner distinguishes between the two kinds of entities: "[T]here are character identities—for example, forewings and hind wings of insects. [And] there are various character states that insect wings can assume; the forewing can be a wing blade or an elytra or even a haltere … and the hind wing can be a wing blade or a haltere."

Missing from the modern synthesis, development has had an uneasy relationship with the rest of evolutionary biology. It has never been clear how evolutionary changes in development can be encompassed within a population genetic framework. Ron Amundson (*5*) noted that population genetics focuses on vertical transmission of genetic information, while evo-devo is more concerned with the generation of form within an individual and how developmental mechanisms evolve over time. By using ChIN-based homology, Wagner takes a different approach, establishing the developmental circumstances where population genetics directly applies and what limits development places on the remit of population genetics. This boundary characterizes systems where little heritable variation exists until new genetic elements are introduced into a gene network (most dramatically with the origin of a new ChIN). Given a specific and stable gene network population, genetic methods can measure a population's ability to evolve and adapt to environmental changes by microevolution. The origin of novel ChINs alters the structure of gene networks substantially and thus imposes nonuniformitarian evolutionary changes. Consequently, this limits our ability to extrapolate population genetic methods to macroevolutionary timescales. Another issue that will challenge some readers is Wagner's explicit and welcome defense of developmental types, long the bête noir of right-thinking

microevolutionists. Wagner argues that these represent "suites of characters actively maintained through developmental or functional constraints."

Wagner's articulation of the ChIN framework and its application to homology and evolutionary novelty occupies the volume's first half. After an interregnum that addresses some of the philosophical issues raised by this conceptual framework, the final five chapters provide in-depth case studies of its application to some well-studied examples: the origin of cell types; skin and derivatives such as feathers; limbs; digits; and flowers.

Among the book's many great strengths is the virtual absence of a disquisition on the term homology. Wagner clearly defines homology in terms of the development of ChINs, but he disposes of the debate over the term with a few sentences. Although we suspect that those who enjoy wallowing in arguments over homology will fault Wagner, we view this as an inspired decision as it frees him to focus on the power of his conceptual framework.

Readers will find obvious extensions to the author's framework; Wagner identifies some, such as arthropod appendages, himself. Major evolutionary transitions provide another opportunity for extending the framework. There, Wagner's focus on the individuation of characters through the structure of regulatory networks could be applied to the study of new evolutionary individuals.

Homology, Genes, and Evolutionary Innovation makes a seminal contribution to evolutionary biology. As Wagner argues, his view provides an opportunity for a major research program on the study of novelty as distinct from adaptation. This means focusing research on the various mechanisms that generate variation in character identity rather than the consequences of changes in character state. The benefits of such research will prove as substantial as the study of speciation has been to macroevolution and microevolution. There is new biology to be discovered in macroevolution now that we know where to look. ∎

REFERENCES AND NOTES

1. S. J. Gould, *Ontogeny and Phylogeny* (Harvard Univ. Press, Cambridge, MA, 1977).
2. A. Wagner, *The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems* (Oxford Univ. Press, Oxford, 2011).
3. E. H. Davidson, *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution* (Academic Press, San Diego, CA, 2006).
4. E. H. Davidson, D. H. Erwin, *Science* **311**, 796–800 (2006).
5. R. Amundson, *The Changing Role of the Embryo in Evolutionary Thought: Roots of Evo-Devo* (Cambridge Univ. Press, Cambridge, 2005).

EXERCISE PHYSIOLOGY

See how we run

By **Stephen D. R. Harridge**

Exercise is an essential component of the study of human physiology. The world faces enormous healthcare challenges from diseases (e.g., cardiovascular disease, obesity, and type 2 diabetes) inextricably linked to a lack of physical activity. Thus the need to understand how our bodies move and adapt to exercise has never been more relevant.

Insufficient exercise is clearly not a personal issue for Christopher Gillen—runner, physiologist (Kenyon College), and author of *The Hidden Mechanics of Exercise*. Despite its title, the book doesn't focus on whole-body joints and moment arms. Instead, Gillen's enjoyable account emphasizes the molecules and protein structures that allow us to move, run, jump, control fuel use, and regulate adaptations to exercise training.

Using analogies to make the science accessible, Gillen takes readers through critical systems of the human body, from the molecular protein structures that allow us to store and reuse elastic energy in our tendons to the processes by which hemoglobin molecules bind and release oxygen. He effectively integrates traditional physiology with more recent developments in cell and molecular biology and genetics. He links together topics ranging from the molecular chain of events in the brain that cause us to initiate movement to the behavior of myosin (the molecular motor protein in skeletal muscle that produces force and ultimately generates movement). Refreshingly, Gillen approaches the subject from the system down rather than the molecule up. Throughout the book, he emphasizes how tiny changes in protein structures scale up to produce whole-body movements.

Gillen comes to the topic with the perspective of a long-distance runner (he has

The reviewer is at the Centre of Human and Aerospace Physiological Sciences, School of Biomedical Sciences, King's College, London SE1 1UL, UK. E-mail: s.harridge@kcl.ac.uk

The Hidden Mechanics of Exercise: Molecules That Move Us
Christopher M. Gillen
Harvard University Press, 2014. 352 pp.
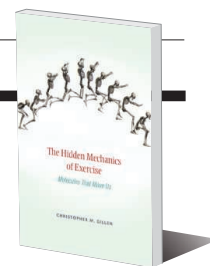
completed 10 ultramarathons) and not that of a power athlete. Thus he can be forgiven for some confusion regarding hyperplasia in the hypertrophic response of whole muscle to high-resistance strength training. Satellite cells may well proliferate, but that does not mean that the number of functioning muscle fibers increases. His description of proteins "morphing" also becomes a little irksome after a while, but that should not detract from the enjoyable way he takes us through the different biological systems, linking each with various aspects of physical and sports performance.
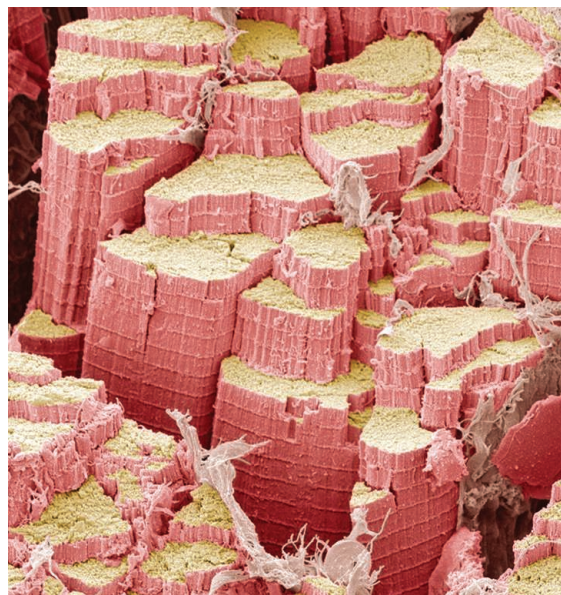
Force producer. Colored scanning electron micrograph of a freeze-fractured skeletal (striated) muscle.

Overall, Gillen offers exercise enthusiasts wishing to understand the science behind their training an interesting read. The book also serves as an engaging primer for exercise-science students who want to begin to understand some of the underlying molecular mechanisms. *The Hidden Mechanics of Exercise* introduces concepts that make the step to the specialized textbook or research article easier. At the same time, researchers studying the behavior of the individual molecules may find Gillen's account enlightening in regard to the functional implications of their work at the whole-body level. ∎

LETTERS

Edited by **Jennifer Sills**

Making waves about spreading weeds

THE LIFE IN *SCIENCE* "Weeds making waves" (D. S. Johnson, 18 April, p. 255) told the story of a middle school project that made a boat out of the highly invasive reed *Phragmites australis*. The teacher lashed together the reed stems to make a boat that was "twice as long as the truck bed" and floated for at least 2 hours. Although finding creative ways to use invasive species is admirable, this particular use is ill advised because *Phragmites* is one of the world's most invasive wetland plants.

Boats or natural debris containing stems of this plant can travel 6.5 km on a single tide (*1*). They serve as efficient dispersal vessels because *Phragmites* can regenerate from green stem nodes and fragments, and bundles of reed stems can also contain viable seed and rhizomes. Water dispersal occurs naturally, but increasing opportunities via reed rafts can expedite colonization of new sites, aggregate introduced genotypes from distant locales, and promote novel genetic admixtures. This increased genetic diversity facilitates hybridization and heterosis—factors that contribute to invasions (*2*). This is especially relevant for *Phragmites*, as North America has become

a genetic melting pot for this genus (*3*). Non-native genotypes and hybrids have been identified from Quebec to Florida and across the United States to California. Millions of dollars are spent annually to manage and remove *Phragmites* because it reduces plant and animal diversity and changes ecosystem functions (*4*).

Education about the negative consequences of invasive species is especially laudable in K–12 schools, and finding uses for plant invaders is a potential management strategy. However, promoting the inadvertent spread of invasive species under the guise of humor should be avoided from both educational and ecological standpoints.

Laura A. Meyerson,[1]* Jan Pergl,[2] Petr Pyšek[2,3]

[1]Natural Resources Science, Coastal Institute 111, University of Rhode Island, Kingston, RI 02881, USA. [2]Department of Invasion Ecology, Institute of Botany, Průhonice, 252 43, Czech Republic. [3]Department of Ecology, Faculty of Science, Charles University, Viničná 7, Prague, 128 44, Czech Republic.

*Corresponding author. E-mail: lameyerson@mail.uri.edu

REFERENCES

1. T. E. Minchinton, *Aquat. Bot.* **84**, 372 (2006).
2. N. C. Ellstrand, K. A. Schierenbeck, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 7043 (2000).
3. L. A. Meyerson, C. Lambertini, M. K. McCormick, D. F. Whigham, *AoB Plants* **2012**, pls022 (2012).
4. L. J. Martin, B. Blossey, *Estuaries Coasts* **36**, 626 (2013).

Response

I RECOGNIZE MEYERSON *et al.*'s scientific arguments, but argue that the increased risk of putting a boat made of senesced aboveground shoots of *Phragmites* without

Phragmites australis.

seed heads into the February waves, is small, especially given that it was launched near already towering *Phragmites* stands. Furthermore, I do not recommend nor foresee a flotilla of *Phragmites* boats hugging our coastlines.

Humor may have clouded my scientific points. To clarify, my piece highlights the need for creativity in controlling invasive species, while conceding that *Phragmites* boats may not be an effective management strategy. Meyerson *et al.* correctly highlight that we spend $4 million annually on *Phragmites* management, citing Martin and Blossey (*1*). Those same authors question the efficacy of current *Phragmites*

The eureka story

I N HIS 11 APRIL LETTER "Immunotherapy: It takes a village" (p. 149), D. Pardoll recognized some of the scientists and research that led to the discovery highlighted as *Science*'s 2013 Breakthrough of the Year: cancer immunotherapy. Online commenters pointed out additional factors worthy of recognition. An excerpt from those comments is below. Read the full comments at http://comments.sciencemag.org/content/10.1126/science.344.6180.149-a.

A selection of your thoughts:

THE LETTER BY D. PARDOLL pays tribute to the basic science and its primary contributors behind breakthroughs in cancer immunotherapy. A major aspect of the breakthrough is the approach of blocking co-inhibitory receptors.... Pardoll's Letter leaves one asking, when was the "eureka" moment when immunologists understood that peripheral tolerance was not simply a consequence of AgR signals, the previous paradigm, but instead due to such co-inhibitory signals? ...Nicholas R.

StC. Sinclair did experiments that led him to publish in 1971 the model that B cells are made tolerant not by AgR engagement but instead co-engagement of AgR with a hypothesized receptor for the Fc portion of IgG on the B cell. The B cell FcR receptor was later discovered by others and was shown in tremendous cellular and molecular detail to be a co-inhibitory receptor.... [B]etween 1990 and 1993, Sinclair proposed that these co-inhibitory receptors underlie negative feedback regulation of T cells and not just B cells, a conceptualization driven by his quite radical view at the time that the AgR transmits positive rather than negative signals; hence the need for other signals to shut down lymphocytes. One wonders how such breakthroughs in understanding could have received so little attention. It raises broader questions of what we value in basic science. If we are to do our best to justify funding of basic science, we must ensure that the stories of such eureka moments illuminating a general principle in biology are told.

C. Anderson, U Alberta;
A. Panoskaltsis-Mortari, U Minnesota

management, which yields little ecological benefit (*1*). This is precisely my point in writing that "We have tried to cut it, poison it, burn it, bury it, till it, and drown it, and yet its tufted heads still sway in the wind." Our current management strategies for *Phragmites* are simply ineffective.

Phragmites invasions can certainly have devastating consequences, but this is not universally the case. For instance, *Phragmites* in certain instances can positively influence invertebrate populations (*2*). *Phragmites* is effective in phytoremediation of metal, hydrocarbon, and nutrient pollution (*3*, *4*) and can allow for faster marsh accretion (*5*), an important factor in maintaining elevation with an ever-rising sea.

Given the lack of management efficacy and the possible positive effects of *Phragmites* in certain instances, should we be spending $4 million annually on *Phragmites* management? I do not advocate abandoning management efforts—the conversion of brackish tidelands into *Phragmites* is a serious ecological concern—nor do I advocate the spread of any invasives. I do suggest that we triage our efforts and seriously re-evaluate current approaches (*1*). Finally, I applaud any teacher who uses creativity to engage students.

David Samuel Johnson

Marine Biological Laboratory, Woods Hole, MA 02543, USA. E-mail: manayunkia@gmail.com

REFERENCES

1. L. J. Martin, B. Blossey, *Estuaries Coasts* **36**, 626 (2013).
2. R. S. Warren *et al.*, *Estuaries* **24**, 90 (2001).
3. M. Borin, M. Salvato, *Ecol. Eng.* **46**, 34 (2012).
4. A. Adams, A. Raman, D. Hodgkins, *Water Env. J.* **27**, 373 (2013).
5. J. E. Rooth, J. C. Stevenson, J. C. Cornwell, *Estuaries* **26**, 475 (2003).

TECHNICAL COMMENT ABSTRACTS

Comment on "Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA"

Francis V. Chisari, William S. Mason, Christoph Seeger

■ Lucifora *et al.* (Research Articles, 14 March 2014, p. 1221) report that the hepatitis B virus (HBV) transcriptional template, a long-lived covalently closed circular DNA (cccDNA) molecule, is degraded noncytolytically by agents that up-regulate APOBEC3A and 3B. If these results can be independently confirmed, they would represent a critical first step toward development of a cure for the 400 million patients who are chronically infected by HBV.

Full text at http://dx.doi.org/10.1126/science.1254082

Response to Comment on "Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA"

Yuchen Xia, Julie Lucifora, Florian Reisinger, Mathias Heikenwalder, Ulrike Protzer

■ Chisari *et al.* challenge our central conclusion that the hepatitis B virus (HBV) persistent form, the covalently closed circular DNA (cccDNA), is degraded in a noncytotoxic and specific fashion in the nucleus of infected hepatocytes. Specificity of the assays used, exclusion of cell division or death, and activity of APOBEC3 deaminases in the nucleus, however, were addressed in the paper.

Full text at http://dx.doi.org/10.1126/science.1254083

TECHNICAL COMMENT

VIROLOGY

Comment on "Specific and nonhepatotoxic degradation of nuclear hepatitis B virus cccDNA"

Francis V. Chisari,[1]* William S. Mason,[2] Christoph Seeger[2]

Lucifora et al. (Research Articles, 14 March 2014, p. 1221) report that the hepatitis B virus (HBV) transcriptional template, a long-lived covalently closed circular DNA (cccDNA) molecule, is degraded noncytolytically by agents that up-regulate APOBEC3A and 3B. If these results can be independently confirmed, they would represent a critical first step toward development of a cure for the 400 million patients who are chronically infected by HBV.

Viral clearance during acute hepatitis B virus (HBV) infection is mediated by CD8-positive cytotoxic T lymphocytes (CTL) that kill infected cells and secrete antiviral cytokines that noncytolytically inhibit HBV gene expression and replication. In contrast, the CTL-induced mechanisms that mediate clearance of the nuclear pool of viral covalently closed circular DNA (cccDNA), the long-lived transcriptional template of HBV, have been difficult to resolve, largely because of the absence of a cell-based HBV-infection system. Based on in vivo experiments in duck, woodchuck, mouse, and chimpanzee animal models, two competing, but not mutually exclusive, hypotheses of cccDNA clearance have emerged.

On one hand, there is ample evidence that cccDNA elimination requires the turnover of infected hepatocytes in ducks infected with duck hepatitis B virus (DHBV) and in woodchucks infected with woodchuck hepatitis virus (WHV), in which a large fraction of the infected hepatocyte population is killed and replaced by hepatocyte cell division (1). On the other hand, CTL-induced, interferon gamma (IFN-γ)–mediated clearance of HBV replicative DNA intermediates is well documented in HBV transgenic mice (2), and clearance of HBV in acutely infected chimpanzees is initiated by the influx of IFN-γ–producing CD8 T cells and characterized by a disproportionate loss of both cytoplasmic HBV replicative DNA intermediates and nuclear cccDNA relative to multiple simultaneous markers of cell death (3, 4).

These latter results suggest that cccDNA may be susceptible to noncytolytic clearance by antiviral cytokines; that is, clearance not requiring the death or proliferation of infected cells. Confirmation of noncytolytic cccDNA clearance in experimentally tractable systems would be a major step forward in finding antiviral therapies for chronic hepatitis B. We were therefore very encouraged by the paper by Lucifora et al. (5) reporting that high-dose interferon alpha (IFN-α) and lymphotoxin β receptor (LTβR)–mediated stimulation of APOBEC3A or 3B expression noncytopathically reduces cccDNA levels in differentiated HepaRG cells and in primary human hepatocytes. We were also encouraged by evidence suggesting that this loss might include, as an intermediate step, a detectable level of cccDNA depurination. However, after a careful reading of this Report, which entails a wealth of technically difficult experiments, we believe several critical points warrant further investigation before these conclusions can be exploited.

1) Was the quantitative polymerase chain reaction (qPCR) assay used in this study cccDNA-specific? Quantifying HBV cccDNA in infected cells is extremely challenging because of the vast excess of non-cccDNA replicative forms in total DNA extracts like those that were used for qPCR analysis in this study. However, cccDNA can be physically separated from other viral DNA forms by Hirt extraction, allowing specific qPCR amplification. Also, it coisolates with mitochondrial DNA that can be used to normalize results. Although Hirt extraction was used to prepare cccDNA before Southern blot analysis shown in figure 2C and S1D of (5), the results were not normalized against mitochondrial DNA, so the small lane-to-lane differences observed in those figures could reflect differences in extraction efficiency or loading rather than true differences in cccDNA. Moreover, qPCR assays for cccDNA did not employ Hirt extraction, raising concerns about their specificity. Hopefully, future studies employing both Hirt extraction and mitochondrial DNA normalization for qPCR analysis will confirm that cccDNA can be noncytopathically degraded by IFN-α and LTβR-mediated APOBEC editing, as reported by Lucifora et al.

2) Was IFN-α– or LTβR-mediated cccDNA loss noncytopathic, or due to cell death or regeneration? This study is based on specialized cell culture systems where, typically, only a fraction of cells are productively infected. Hence, these cells likely differ from the bulk of cells in the culture and may be more sensitive to high-dose IFN-α– or LTβR-mediated toxicity. Because infection levels in this study were not reported, it is a formal possibility that some or all of the cccDNA decline is due to cell death. This might not have been apparent if very few of the cells are infected. cccDNA loss at

division could also be a factor if infected cells divided. LTβR agonists have been shown to trigger tumor cell apoptosis (6) and proliferation (7), both of which might explain the observed loss of cccDNA. LTβR-induced hepatocellular proliferation almost certainly occurred in some of their experiments, because it is the only known mechanism that can explain the disappearance of nuclear HBcAg in the HBV transgenic mouse hepatocytes shown in figure S6E of Lucifora et al. (1, 8). Thus, induced cell death and proliferation must be ruled out before the noncytopathic nature of the cccDNA loss can be accepted.

3) Did APOBEC3s edit cccDNA (nuclear) or just replicative intermediate (cytoplasmic) HBV DNAs? Members of the APOBEC3 family can deaminate the nascent, reverse transcribed (minus) DNA strand of HBV, leading to G to A hypermutation in plus strands (9). In this Report, hypermutation of the plus strand of cccDNA involved almost exclusively G to A mutations, suggesting that mutation preceded cccDNA formation from replication intermediates or was minus-strand specific on a cccDNA target. Distinguishing these possibilities, or even showing that the minus strand of cccDNA was the APOBEC target is difficult, because cccDNA accumulates at much lower levels (~0.1 to 1.0%) than the relaxed circular minus- and plus-strand DNA replicative forms in the cytoplasm of infected hepatocytes. The authors stated that their qPCR assay amplified cccDNA with 1000-fold higher specificity than the cytoplasmic DNA forms, but they did not offer evidence of that specificity. Moreover, the subsequent three-dimensional (3D) PCR, used to selectively amplify edited DNA for sequencing, did not have cccDNA specificity. Rather, this method would amplify any low–melting point, AT-enriched DNA present in the products of the first PCR. Hence, the possibility that the 3D PCR products actually originated from cytoplasmic DNA, not from cccDNA, has not been excluded. Quantitative differences between treated and untreated cells were also unclear.

4) Is therapeutic LTβR activation tenable in the context of chronic HBV infection? Because LTβR agonists have been shown to trigger apoptosis (6), hepatocellular proliferation (7), inflammation, and hepatocellular carcinoma (10, 11), safety considerations are likely to preclude regulatory approval of LTβR activation for the "development of new therapeutics" as suggested by Lucifora et al.

REFERENCES AND NOTES

1. J. Summers et al., Proc. Natl. Acad. Sci. U.S.A. 100, 11652–11659 (2003).
2. L. G. Guidotti et al., Immunity 4, 25–36 (1996).
3. L. G. Guidotti et al., Science 284, 825–829 (1999).
4. S. F. Wieland, H. C. Spangenberg, R. Thimme, R. H. Purcell, F. V. Chisari, Proc. Natl. Acad. Sci. U.S.A. 101, 2129–2134 (2004).
5. J. Lucifora et al., Science 343, 1221–1228 (2014).
6. X. Hu et al., Carcinogenesis 34, 1105–1114 (2013).
7. A. V. Tumanov et al., Gastroenterology 136, 694–704, e4 (2009).
8. L. G. Guidotti, V. Martinez, Y. T. Loh, C. E. Rogler, F. V. Chisari, J. Virol. 68, 5469–5475 (1994).
9. T. F. Baumert, C. Rösler, M. H. Malim, F. von Weizsäcker, Hepatology 46, 682–689 (2007).
10. J. Haybaeck et al., Cancer Cell 16, 295–308 (2009).
11. M. J. Wolf, G. M. Seleznik, N. Zeller, M. Heikenwalder, Oncogene 29, 5006–5018 (2010).

[1]The Scripps Research Institute, La Jolla, CA 92037, USA. [2]Fox Chase Cancer Center, Philadelphia, PA 19111, USA. *Corresponding author. E-mail: fchisari@scripps.edu

RESEARCH

Crayfish can become nervous too in stressful situations *Fossat* et al. *p.* 1293

IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

Integrated Ocean Drilling Program expedition in the Gibraltar gateway.

PALEOCEANOGRAPHY

The when of Mediterranean water outflow

The trickle of water that began to flow from the Mediterranean Sea into the Atlantic Ocean after the opening of the Strait of Gibraltar turned into a veritable flood by the end of the Pliocene 2 to 3 million years ago. It then began to influence large-scale ocean circulation in earnest. Hernández-Molina *et al.* describe marine sediment cores collected by an ocean drilling expedition (see the Perspective by Filippelli). The results reveal a detailed history of the timing of Mediterranean outflow water activity and show how the addition of that warm saline water to the cooler less-salty waters of the Atlantic was related to climate changes, deep ocean circulation, and plate tectonics. — HJS

Science, this issue p. 1244; see also p. 1228

NEURONAL REPAIR

Improving stroke recovery by timing treatment

Patients recovering from strokes often fight a long uphill battle, with mixed results. Studying the effect of physical training on regeneration from damaged nerves in a model of stroke in rats, Wahl *et al.* show that timing matters. First, the researchers gave the rats a stroke, which damaged their ability to reach for food pellets with their forelimbs. The researchers then gave them physical training and treated them with an antibody to encourage neural regeneration. The rats improved more when the researchers waited until after the antibody treatment to start the training. Damaged circuits, it seems, need a little time to regrow before being called into action. — PJH

Science, this issue p. 1250

DINOSAUR PHYSIOLOGY

Not too fast, not too slow, somewhere in between

In early depictions, dinosaurs lumbered slowly, dragging their tails. More recently, we have imagined them lifting their tails and running. The question boils down to whether dinosaurs had energetic systems closer to those of rapidly metabolizing mammals and birds, or to those of slower reptiles that do not internally regulate their body temperature. However, determining the metabolic rate of extinct organisms is no easy task. Grady *et al.* analyzed a huge data set on growth rate in both extinct and living species, using a method that considers body temperature and body size. Dinosaur metabolism seems to have been neither fast nor slow, but somewhere in the middle—so, dinosaurs did not fully regulate their internal temperature but they were also not entirely at the whim of the environment; neither slow goliaths nor supercharged reptiles. — SNV

Science, this issue p. 1268

NONHUMAN GENETICS

Male chimps evolve faster with age

Chimpanzees are evolving faster than humans. Venn *et al.* examined the genetics of three generations of western chimpanzees and found that overall the mutation rate is similar between humans and chimpanzees. However, while male humans had three to four times the mutation rate of females, in chimpanzees the sex difference was even higher, with a male mutation rate five to six times that of females. Blame aging dads. For every extra year of the father's age, baby

A family of western chimpanzees.

chimpanzees exhibited approximately one extra mutation. This finding will inform future studies of primate evolution. — LMZ

Science, this issue p. 1272

QUANTUM GASES

Tilting just right makes atoms tunnel

One of the most fascinating phenomena in the quantum world is the ability of particles to go through an energy barrier — a process called quantum tunneling. Meinert *et al.* studied the dynamics of quantum tunneling in an optical lattice of strongly interacting atoms. When the lattice was suddenly tilted, the atoms, originally each in their own lattice site, tunneled to non-neighboring sites. — JS

Science, this issue p. 1259

CANCER IMAGING

Taking a broader view of cancer imaging

Many people think the best way to visualize tumors is to target imaging agents to specific cancers at the molecular level. Kuo *et al.* feel differently: They developed a new class of small molecules, called alkylphosphocholine (APC) analogs, which are broadly taken up by nearly all cancers, without such molecular specificity. Compared to normal cells, cancer cells have a strong taste for APC analogs. By attaching fluorescent

labels or radiolabels to the APC analogs, the researchers could image more than 50 different human cancers in animal models, as well as brain, lung, and liver tumors in human patients. These broadly applicable APC-based agents for imaging—and possibly for treatment—are now poised for further translation to clinical trials. — MLF

Sci. Transl. Med. **6**, 240ra75 (2014).

EARTH'S INTERIOR

Cycling water through the transition zone

The water cycle involves more than just the water that circulates between the atmosphere, oceans, and surface waters. It extends deep into Earth's interior as the oceanic crust subducts, or slides, under adjoining plates of crust and sinks into the mantle, carrying water with it. Schmandt *et al.* combined seismological observations beneath North America with geodynamical modeling and high-pressure and -temperature melting experiments. They conclude that the mantle transition zone—410 to 660 km below Earth's surface—acts as a large reservoir of water. — NW

Science, this issue p. 1265

CANCER METASTASIS

Copper for breast cancer metastasis
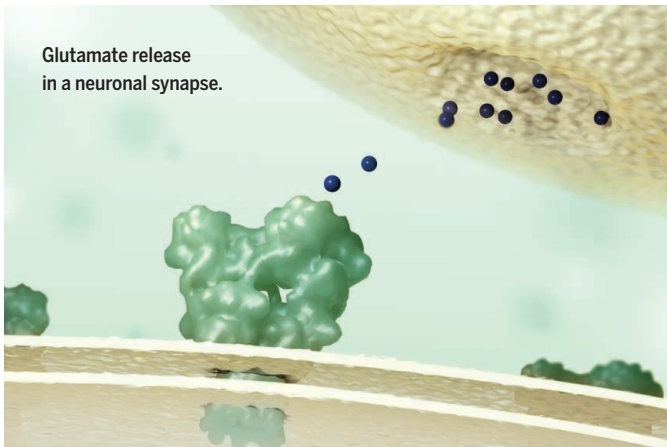
Many patients with breast cancer die from metastases, when cancer cells spread from the primary tumor to other sites. Some of the intracellular proteins that help cells move from one location to another can be activated by a chemical modification called oxidation. MacDonald *et al.* found that the enzyme Memo bound copper, enhancing the oxidation of proteins involved in cell movement. Mice with tumors formed from breast cancer cells that lacked Memo had fewer lung metastases, and human patients with breast cancers that had high levels of Memo were more likely to develop metastases. — WW

Sci. Signal. **7**, ra56 (2014).

Edited by **Kristen Mueller** and **Jesse Smith**

Glutamate release in a neuronal synapse.

NEUROSCIENCE

Making excitatory synapses slow (or fast)

Glutamate is the most widely used neurotransmitter in the brain. Glutamate receptors in the mammalian brain are macromolecular complexes assembled from a pool of different proteins. Several proteins in these complexes, including cornichon homolog 2 (CNIH2), are known to influence the workings of certain glutamate receptors of the AMPA type, but their precise roles are unknown. Boudkkazi *et al.* recorded synaptic transmission between hippocampal neurons and inhibited CNIH2 expression with small interfering RNAs to investigate the role CNIH2 plays at different synapses. They found that CNIH2 profoundly affected the timing of synaptic transmission—CNIH2 slowed things down, whereas its absence sped things up. Such careful kinetic regulation is necessary because speed is more important for some tasks, whereas reliability, the successful transmission of a signal across the synapse, is more important for others. — PRS

Neuron **82**, 848 (2014).

NEUROLOGICAL DISEASE

Skin may hold the key for Parkinson's

In Parkinson's disease, a degenerative movement disorder of the central nervous system, a protein called phosphorylated alpha-synuclein builds up in neurons, damaging the brain. The disease is hard to diagnose early or monitor over time because the protein builds up so slowly and so deep inside the brain. Doppler *et al.* now report that patient skin samples hold key insights. The authors detected phosphorylated alpha-synuclein in autonomic and sensory nerves found in the skin samples in 16 out of 31 people diagnosed with Parkinson's disease and in 0 out of 35 healthy volunteers. Because skin is far more accessible than brain tissue, these observations could lead to diagnostic tests to identify and follow the progression of Parkinson's disease. — PJH

Acta Neuropathol. 10.1007/s00401-014-1284-0 (2014).

EDUCATION

Active learning: The twilight of Chem 101?

Should professors continue to use traditional lectures in

ALSO IN *SCIENCE*

Edited by **Stella Hurtley**

STEM CELL PLASTICITY

The versatility of epithelial stem cells

Stem cells are very important in the maintenance of our bodies' tissues and organs. Blanpain and Fuchs review how different populations of naturally lineage-restricted epithelial stem cells and committed progenitors can also display remarkable plasticity. These cells can re-acquire long-term self-renewing capacities and multilineage differentiation potential during physiological and regenerative conditions. These abilities depend on whether the stem cell remains within its resident niche or has been mobilized to repair a wound. Such cellular plasticity has implications for regenerative medicine and for cancer. — BAP

Science, this issue p. 1243

QUANTUM NONLOCALITY

Testing nonlocality for many particles

Distant parts of a quantum-mechanical system can be correlated in ways that cannot be described classically—a concept known as nonlocality. Tura *et al.* propose a simple test for nonlocality in systems with multiple particles. The test involves quantities that should readily be measurable in, for example, cold atom experiments. This is an improvement over currently available tests, which are difficult to implement experimentally. — JS

Science, this issue p. 1256

QUANTUM MECHANICS

Avoiding back-action in quantum measurements

The very process of measuring a quantum system has an influence on the system through the process of back-action. Suh *et al.* used a back-action evasion scheme to monitor the motion of a miniature oscillator without influencing its motion (see the

Perspective by Bouwmeester). The scheme should help in the understanding of the fundamental limits associated with measurement and will have practical implications in providing a low-temperature thermometer and a probe of extremely weak forces. — ISO

Science, this issue p. 1262

NEURAL MIGRATION

Dissecting how signaling directs axon growth

During development of the nervous system, nerve cells send out projections called axons that must be guided to their proper targets. Netrins are secreted proteins that bind to receptors to either attract or repel the growing axons. Xu *et al.* present x-ray structures that show that complexes of netrin with two different receptors, neogenin and DCC, have different architectures. How netrin signals remains to be understood in detail, but netrin's ability to create different assemblies probably plays a role in the diverse signaling outcomes it mediates. — VV

Science, this issue p. 1275

HUMAN GENETICS

The population structure of Native Mexicans

The genetics of indigenous Mexicans exhibit substantial geographical structure, some as divergent from each other as are existing populations of Europeans and Asians. By performing genome-wide analyses on Native Mexicans from differing populations, Moreno-Estrada *et al.* successfully recapitulated the pre-Columbian substructure of Mexico. This ancestral structure is evident among cosmopolitan Mexicans and is correlated with subcontinental origins and medically relevant aspects of lung function. These findings exemplify

the importance of understanding the genetic contributions of admixed individuals. — LMZ

Science, this issue p. 1280

TRANSCRIPTION

Pausing for control of gene expression

Pausing during gene transcription can play a critical role in gene regulation. Vvedenskaya *et al.* mapped pause sites across the whole genome in actively growing *Escherichia coli* (see the Perspective by Roberts). Thousands of undocumented pause sites were identified across well-transcribed genes, allowing the definition of a consensus pause sequence that is dependent on specific interactions of RNA polymerase with the DNA template and nascent RNA transcript. — GR

Science, this issue p. 1285; see also p. 1226

DISEASE ECOLOGY

Many connections are not always bad for health

Contrary to expectations, highly connected populations can experience less impact from infectious disease than isolated groups. What happens to pathogens in natural populations has been poorly studied, because they rarely cause devastating disease outbreaks. Thanks to a long-term study of an inconspicuous fungal-plant disease system, we have now gained some surprising insights. During a 12-year study, Jousimo *et al.* discovered that clustered and linked host-plant patches showed lower levels of fungal damage and higher fungal extinction rates than more distant patches (see the Perspective by Duffy). This phenomenon is explained by high gene flow and rapid evolution of host resistance within the connected patches. Populations of the modest weed *Plantago,*

growing on the Åland Islands in the Baltic, were less than 10% infected by the *Podosphaera* mildew fungus in any given year, but infection turnover was high. These findings have broad implications for ecology, disease biology, conservation, and agriculture. — CA

Science, this issue p. 1289; see also p. 1229

COMPARATIVE BEHAVIOR

The crayfish that was afraid of the dark

We tend to assume that complex emotions, such as anxiety, only occur in mammals or other cognitively complex vertebrates. But a heightened sense of awareness and the avoidance of novel or dangerous environments could be helpful for any animal species. Fossat *et al.* show that crayfish exposed to a stressful electric field refuse to enter dark arms in a light/dark maze, even after the electric field has been removed. The animals calmed down when they were injected with an anxiolytic drug used to treat anxiety in humans, and they entered the dark as normal. The stressed animals had increased levels of the neurotransmitter serotonin in the brain, and injections of serotonin induced anxiety-like behavior in control animals. Thus, these invertebrates display a primitive form of anxiety that shares a mechanism with the more complex emotions displayed by vertebrates. — SNV

Science, this issue p. 1293

REVIEW SUMMARY

Plasticity of epithelial stem cells in tissue regeneration

Cédric Blanpain[1,2]* and **Elaine Fuchs**[3]*
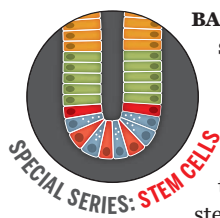
BACKGROUND: At the surface of body organs, epithelial tissues must withstand harsh external environments. To do so, they rely heavily upon stem cells to replenish and repair wounds and replace the many cells that die from this wear and tear. To maintain tissue size, the number of cells lost must be compensated by cell divisions. Tissue homeostasis and wound repair are ensured by stem cells, located within specialized microenvironments, referred to as niches. Each niche is tailored to accommodate the regenerative needs of its tissue. Some tissues—for instance, skin epithelium—harbor multiple stem cell niches, each with their own responsibility for maintaining cellular balance within their particular domain. Governance of discrete tissue units has ancient origins and is also seen in *Drosophila* gut epithelium.

Identifying stem cells and tracking their progeny is accelerated by lineage tracing, a technique in which a stem cell is genetically marked in its niche and in a way such that their subsequent progeny retain marker expression. Although interpretation of these experiments has been complicated by the lack of specificity of most stem cell markers, this method can be helpful in evaluating the contribution of stem cells to tissue homeostasis and wound repair. Additional tools include live imaging of marked stem cells and ablating stem cells in situ either by laser or by targeted expression of diphtheria toxin/receptor in stem cells.
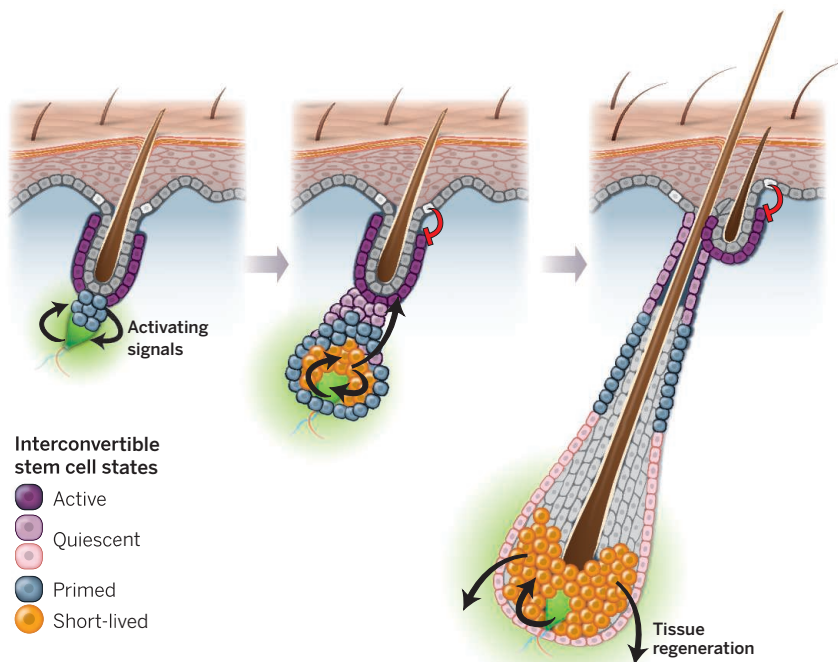
ADVANCES: Accumulating evidence on bone marrow, intestinal stem cell crypts, and hair follicles suggests that stem cells often exist in two distinct states based upon their relative activity and/or their ease of activation during homeostasis and/or wound-induced regeneration. Recent studies on the hair follicle reveal that signals emanating from both heterologous niche cells and from lineage progeny influence the timing and length of stem cell activity. This in turn can profoundly affect the amount of tissue regenerated. Stem cell ablation studies on both intestinal and hair follicle stem cell niches further show that the two states are interconvertible, perhaps best exemplified by the ability of a single intestinal stem cell to eventually outcompete its siblings during rounds of turnover within an intestinal villus.

Additional new findings suggest that fates and multilineage potentials of epithelial stem cells can change, depending upon whether a stem cell exists within its resident niche and responds to normal tissue homeostasis, whether it is mobilized to repair a wound, or whether it is taken from its niche and challenged to de novo tissue morphogenesis after transplantation. In this Review, we discuss how naturally lineage-restricted populations of stem cells and committed progenitors can display such remarkable plasticity under these different conditions.

OUTLOOK: Although the molecular mechanisms underlying cellular plasticity, fate conversion, and reacquisition of stem cell properties in committed and/or differentiated cells still remain poorly understood, this cellular plasticity and lineage reversibility may represent adaptive mechanisms for the self-preservation of epithelia to repair body surfaces and linings in whatever ways possible following injuries. When gone awry, these repertoires become the curse of epithelial stem cells, contributing in major ways to human cancers.

Interconvertible stem cell states
- Active
- Quiescent
- Primed
- Short-lived

Coordinating stem cell activity to match tissue output. Stem cells (purple) often exist in two states, one more quiescent than the other. Primed stem cells are closer to activating niche signals (green). They typically respond faster and generate shorter-lived progenitors (orange), which also signal, fueling tissue production. Each stem cell niche must be responsive to the regenerative demands of tissue homeostasis and wound repair and adjust niche activating and inhibitory signals as necessary.

ILLUSTRATION: K. SUTLIFF/*SCIENCE*

STEM CELL PLASTICITY

Plasticity of epithelial stem cells in tissue regeneration

Cédric Blanpain[1,2]* and Elaine Fuchs[3]*

Tissues rely upon stem cells for homeostasis and repair. Recent studies show that the fate and multilineage potential of epithelial stem cells can change depending on whether a stem cell exists within its resident niche and responds to normal tissue homeostasis, whether it is mobilized to repair a wound, or whether it is taken from its niche and challenged to de novo tissue morphogenesis after transplantation. In this Review, we discuss how different populations of naturally lineage-restricted stem cells and committed progenitors can display remarkable plasticity and reversibility and reacquire long-term self-renewing capacities and multilineage differentiation potential during physiological and regenerative conditions. We also discuss the implications of cellular plasticity for regenerative medicine and for cancer.

Epithelia are cellular sheets often residing at the interface between the external environment and body organs, including skin, gut, airway tracts, kidney, liver, mammary glands, and prostate. They perform a diverse array of physiological functions, including the ability to retain body fluids, absorb nutrients, filter and eliminate toxic by-products of metabolism, and regulate body temperature. Each epithelium is morphologically and molecularly suited to its particular task, a feature that necessitates specialized cell lineages.

Most epithelia replenish themselves through a process called tissue homeostasis, in which the number of cell divisions within a tissue compensates for the number of cells lost (*1*). Tissue homeostasis is ensured by the existence of stem cells (SCs) located within specialized microenvironments, referred to as niches. Each niche is tailored to accommodate the regeneration needs of the tissue (*2*).

The skin epidermis and its appendages (hair follicles, sebaceous glands, and sweat glands) harbor spatially distinct SC niches. The innermost (basal) layer of interfollicular epidermis (IFE) harbors proliferative progenitors, which generate the stratified layers of the skin barrier. Every few weeks, the IFE renews itself almost entirely, placing a constant demand on its SCs. Sebaceous glands (SGs) also turnover continuously during adult homeostasis. By contrast, hair follicles (HFs) cycle through bouts of hair growth

and degeneration, necessitating only periodic use of SCs, whereas sweat gland (SwG) cells are mostly quiescent (Fig. 1A).

Other epithelia also have distinct requirements for tissue homeostasis, which must be met by their resident SCs. In the small intestine, the epithelium is organized into a crypto-villus unit (Fig. 1B). The crypt is composed of columnar basal cells (CBCs) intermingled with Paneth cells at the crypt base; an overlying compartment of transit-amplifying (TA) cells divides several times and then terminally differentiates to generate the absorptive and secretory cells

of the villus. Villus cells are subsequently shed into the lumen (*3*), which results in continual turnover of the entire crypt every 3 to 5 days. CBCs, now known to be SCs, fuel the process.

Functionally validating stemness of epithelial cells in vitro

Different methods have been elaborated throughout the years to study the fate, renewal, and differentiation potential of epithelial SCs. The first functional demonstration of an epithelial SC was made when methods were identified to culture human epidermal keratinocytes under conditions where they could be maintained and propagated for hundreds of generations without losing stemness (*4*). When grown from an unaffected region of a burn patient, expanded epidermal cultures could be stably engrafted onto the damaged skin (*5*). Engrafted epidermis did not develop cancer or other abnormalities, which indicated that, under the right conditions—in this case, coculture with irradiated dermal fibroblasts—in vitro SC expansion and differentiation can be achieved without deleterious consequence.

The requirement of dermal neighbors for successful culturing of epidermal SCs highlights the reliance of SCs on cross-talk with their niche microenvironment. Indeed, by elucidating key heterologous niche components and/or the cross-talk involved, SCs from many different epithelia have since been successfully cultured. For intestinal stem cells (ISCs), it took BMP and Notch inhibition together with Wnt activation to recapitulate in vitro the long-term proliferative capacity and multipotency normally conferred to ISCs by their niche (*6*). These studies underscore the complexities of signaling circuitry governing SC behavior and the need to
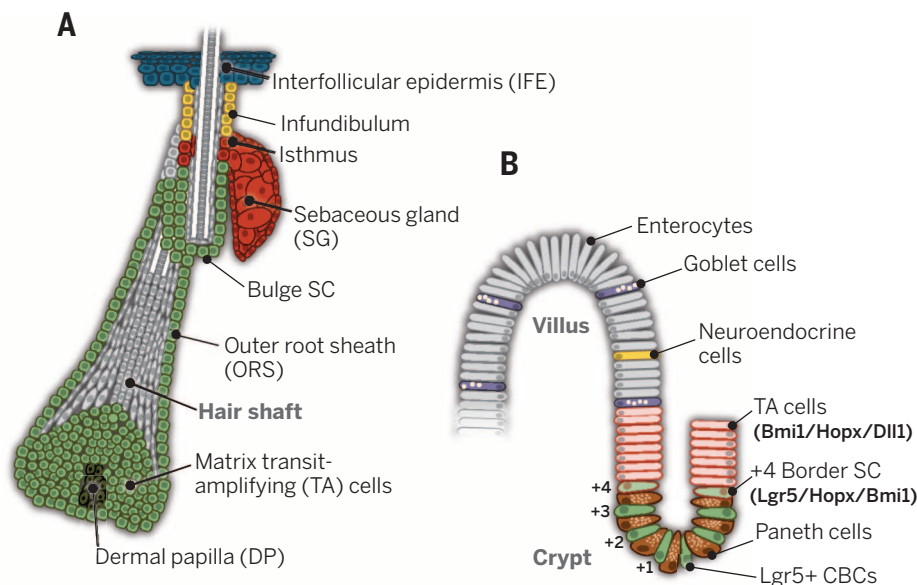
Fig. 1. Skin and intestinal epithelia: paradigms for epithelial stem cell biology. (**A**) Schematic illustrating the epithelial lineages of hairy skin, color-coded here, which derive from at least four distinct stem cell populations. (**B**) Schematic illustrating the location of intestinal crypt stem cells (green), giving rise to TA cells and, in turn, four distinct cell types, three in the villus and one in the crypt.

[1]Institut de Recherche Interdisciplinaire en Biologie Humaine et Moléculaire (IRIBHM), Université Libre de Bruxelles, Brussels B-1070, Belgium. [2]Walloon Excellence in Life Sciences and Biotechnology (WELBIO), Université Libre de Bruxelles (ULB), Brussels B-1070, Belgium. [3]Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10065, USA.
*Corresponding author. E-mail: fuchslb@rockefeller.edu (E.F.); cedric.blanpain@ulb.ac.be (C.B.)

understand this to maintain SCs in the absence of other heterologous cell types in vitro.

Identifying epithelial SCs in vivo and probing their roles in tissue homeostasis
HF homeostasis

Lineage tracing entails the genetic marking of one or a group of cells in their normal physiological context in a way that their subsequent progeny retain marker expression. This method is powerful in evaluating the contribution of SCs to tissue homeostasis (1). The fluctuations of HFs through synchronized bouts of hair growth and inactivity present an interesting variation on this theme (Fig. 2A). Before modern-day genetics, cells with proliferative potential that spent extended periods in quiescence were marked and monitored by nucleotide analog pulse-chase experiments. Such label-retaining cells (LRCs) reside at the base of the resting HF, a region now referred to as the bulge and its associated hair germ (HG) (7). LRCs are SCs, as demonstrated by using a regulatable fluorescent histone to label LRCs and monitor their cell divisions, as well as lineage tracing to follow their fate (8–12) (Fig. 2B).

Both bulge and HG share many molecular features of stemness, including expression of *Lgr5* and *Sox9* (12, 13). However, HG cells are always the first to be activated at the start of each new hair cycle, and they undergo more divisions than bulge cells (13). Their close proximity to the underlying mesenchymal signaling center, the dermal papillae (DP), functions in dictating this early response.

Activated HG cells do not maintain stemness in vitro (13), and in vivo, they generate the TA cells that produce the hair and its channel (14, 15). By contrast, once the new hair cycle initiates, some bulge cells leave their niche and form an inverse proliferative gradient along the emerging outer root sheath (ORS). Early in the hair-growth phase, TA cells stimulate remaining bulge cells to proliferate and replenish the niche (15). ORS cells closest to the bulge return to quiescence soon thereafter and form a new bulge and HG for the next cycle (12, 16). The ability of bulge and HG SCs to generate the seven different HF lineages underscores their multilineage potency. Additionally, even though bulge normally gives rise to HG, HG can replenish an empty bulge niche, as shown by laser ablation and live imaging (16), which underscores their close relation and capacity to interconvert when necessary (see below).

Although the above studies disclose insights into the behavior and maintenance of cycling HFs, lineage tracings reveal the existence of at least two additional SC populations—SG and infundibulum—within the noncycling HF segment. SGs are maintained by unipotent $Lgr6^{+}Lrig1^{+}$ SCs that arise from *Blimp1*-expressing progenitors (17). In adults, Lgr6-expressing cells mark and sustain SGs (18, 19), whereas Lrig1 expression extends to SCs fueling infundibulum homeostasis (19) (Fig. 2A). One other SC population in the upper bulge region has been suggested on
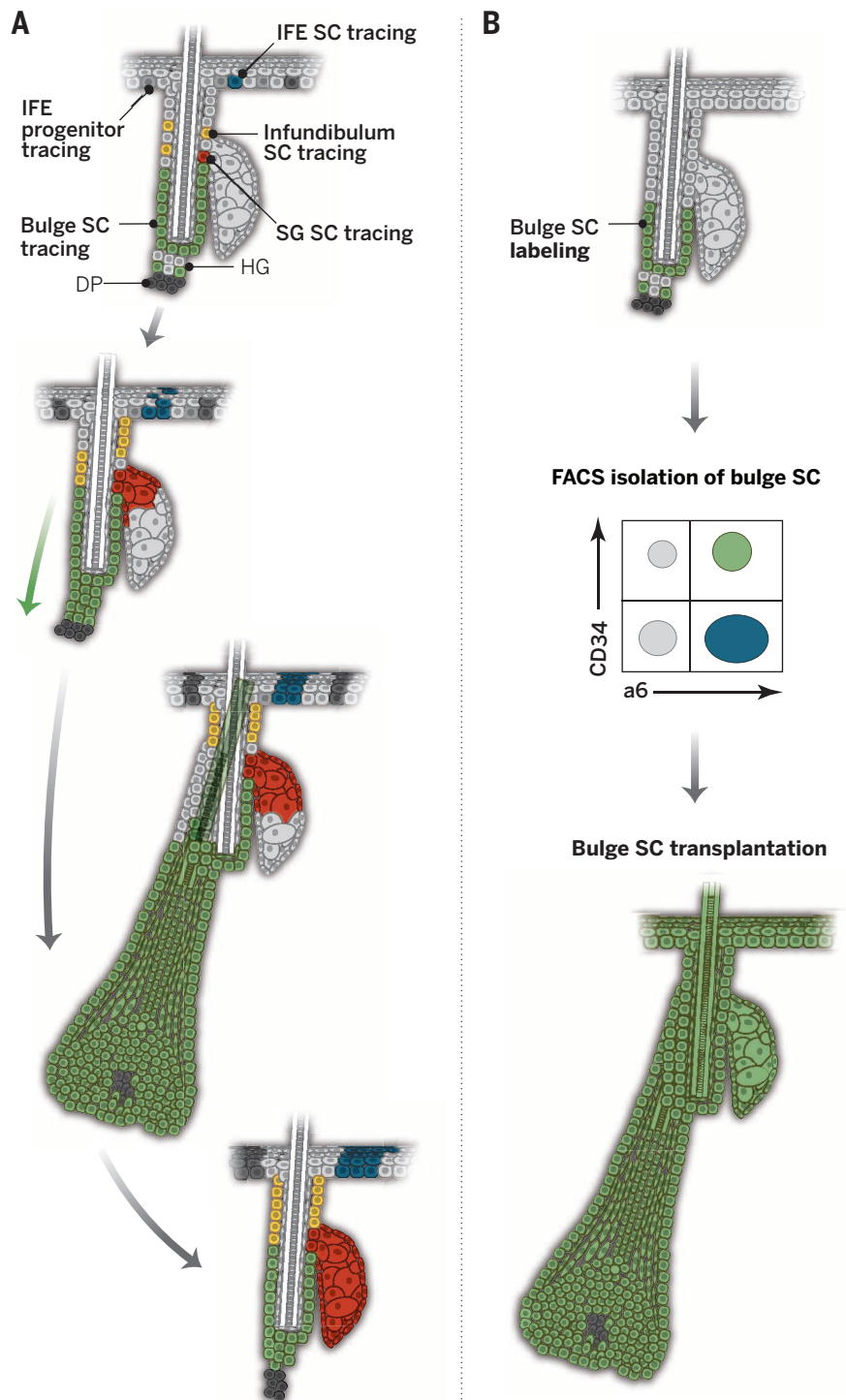
Fig. 2. Epidermal homeostasis is achieved through distinct pools of stem cells. (**A**) Schematic illustrating the outcome of five separate lineage tracings of Rosa26-floxed-stop-floxed–reporter mice. In each experiment, a different inducible Cre recombinase was expressed in the desired SC or progenitor compartment. Because the Rosa26 promoter is generic, once Cre is activated and the stop codon is excised, the marked cells and all their downstream progeny express the reporter. The results shown here illustrate that each SC compartment is responsible for sustaining tissue homeostasis within a discrete skin domain. (**B**) We purified fluorescently marked bulge SCs (green) by fluorescence activated cell sorting (FACS) and cultured them as individual colonies of cells before transplanting the cells to a hairless mouse. The experiment illustrated that a clone from a single bulge SC can regenerate the entire skin epithelium, which documents the stemness and multipotency of the cells (9, 69, 70). We now know that when taken out of their native niche and engrafted, epithelial SCs are often less restricted in their fates.

the basis of its encasement by sensory nerve sheaths (20). Whether these cells represent an independent pool of functional SCs remains unresolved.

A sharp boundary exists between infundibulum-derived *Lrig1*+ cells and IFE (19), and little if any contribution to the IFE has been observed by the various adult SCs thus far identified in the HF (9, 10, 19, 21–23) (Fig. 2A). This argues against the prior view that a single "master" SC population presides over all skin lineages, as initially postulated based upon embryonic *Lgr6*-Cre lineage tracing (18). Indeed, the paradigm for segmental-tissue governance by SC units has ancient origins, as, like the HF, *Drosophila* intestinal epithelium is also compartmentalized into discrete units maintained by separate SC populations (24).

IFE homeostasis

The IFE is maintained by juxtaposition of small units of proliferation containing stem and/or progenitor cells (1). During embryogenesis, the single layer of K14+ epidermal basal progenitors undergoes a spindle orientation shift from >90% symmetric to ~70% asymmetric cell divisions, which leaves one daughter in the basal layer and one suprabasal differentiating daughter cell (25). Postnatally, SCs and transient progenitors coexist within the IFE basal layer, and both express K14 but can be distinguished by their survival rate, mode of division, gene expression, and ability to respond to tissue damage

A Lgr5+ CBC lineage tracing during homeostasis

Labeling   Weeks   Months

B Dynamic of Lgr5+ bottom/centre cells

+4
+3
+2
+1

Bottom/center cells

C Dynamic of Lgr5+ border cells

+4
+3
+2
+1

Border cells

Short-lived clones

Crypt colonization by border cells

D Intestinal homeostasis following Lgr5 cell ablation

Bmi1 tracing

Lgr5+ ablation

E Monoclonal drift of Lgr5+ CBCs

+4
+3
+2
+1
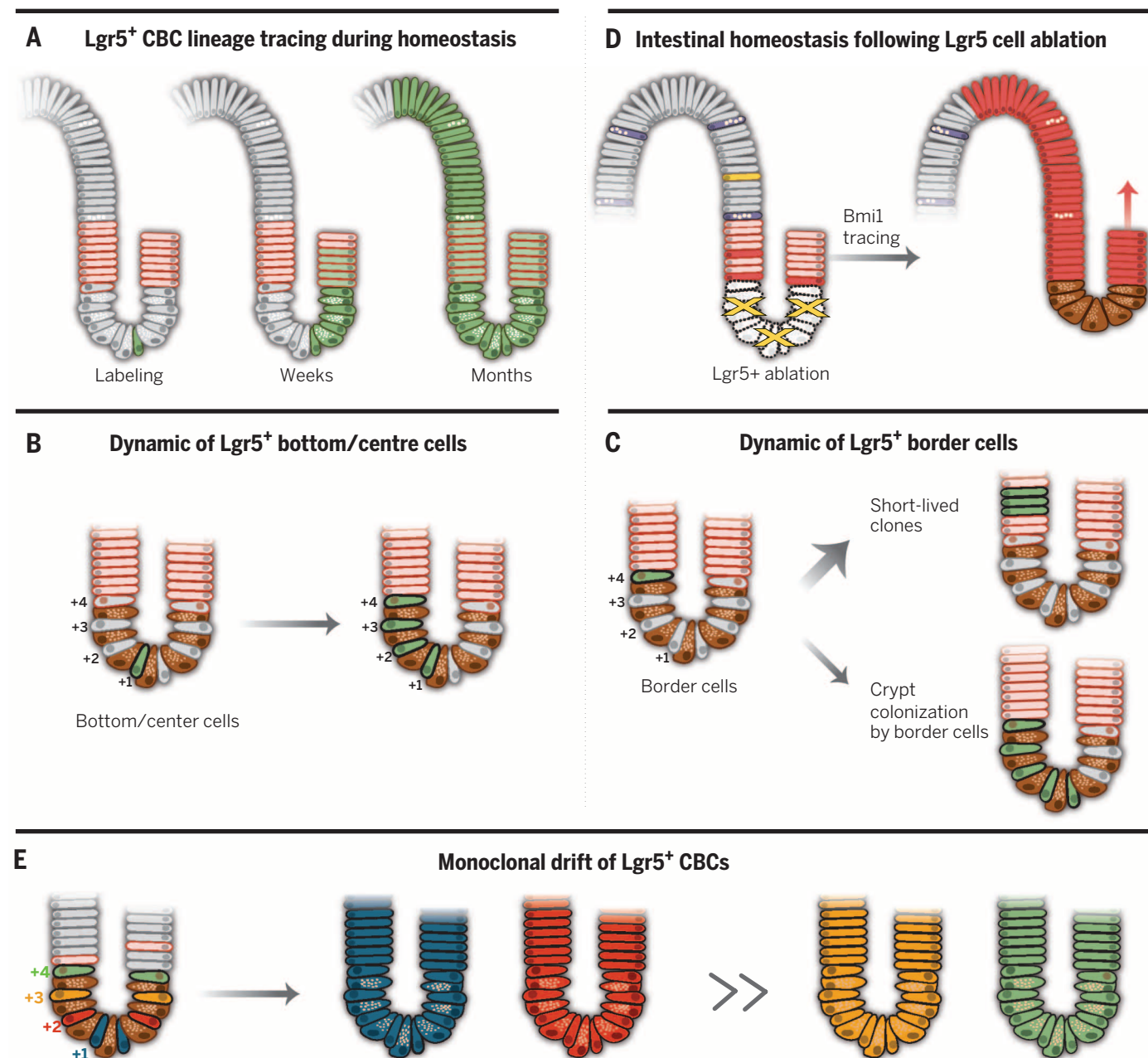
Fig. 3. Interconversion and monoclonal drift of intestinal stem cells. (**A**) Lineage tracing of *Lgr5*+ cells (green) showing that these crypt cells give rise to all intestinal lineages during homeostasis (38). (**B** and **C**) Intravital microscopy showing the colonization of the crypt from *Lgr5* cells at bottom center. *Bmi1*+ border (+4) cells either colonize the bottom of the crypt or give rise to TA cells (red) (42). (**D**) Lineage ablation of *Lgr5*+ (yellow X's) prompts *Bmi1*+ cells (red) to convert into *Lgr5*+ crypt cells, and thus gut homeostasis is not impaired (43). (**E**) Multicolor lineage tracing rapidly leads to unicolor crypts, which demonstrate the monoclonal drift of ISCs (49).

(26) (Fig. 2A). Basal progenitors targeted by *Ah*-Cre^{ER} (27, 28), *Inv*-Cre^{ER} (26), and possibly *Axin2*-Cre^{ER} (29), divide mostly asymmetrically, whereas K14$^+$ basal SCs are integrin-rich and divide mostly symmetrically to generate two long-lived daughters (26, 30). Although the exact nature of the imbalance between SC and progenitor division is not yet clear, each SC-progenitor division must also be accompanied by some differentiation, driven in part by Notch signaling (31–34).

Crypt homeostasis: A one-cell–winner competition

Ultrastructural analyses and proliferative capacity of the intestinal crypt led to the initial hypothesis that CBCs are ISCs (35) (Fig. 1B). Subsequent assignment of stemness favored cells at the +4 position, given their mode of chromosome segregation (36) and higher resistance to DNA damage–induced cell death (37). Lineage tracings of +4 CBCs with *Bmi1*, *mTER*, and *Hopx*-Cre^{ER} and 0→+3 CBCs with *Lgr5*-Cre^{ER} revealed that all crypt CBCs behave as interconvertible multipotent ISCs (38–42) (Fig. 3, A to C). This is further exemplified by diphtheria toxin (DT)–targeted ablation of *Lgr5*-expressing cells, which does not impair intestinal homeostasis (43) (Fig. 3D). Thus, despite their markedly different regenerative demands, both HF and intestine have spatially discrete interconvertible SCs existing in quiescent and primed and/or activated states (bulge and HG versus +4 and 0→+3 crypt cells).

Although it was initially proposed that all *Lgr5*$^+$ ISCs cycle rapidly (38), a recent study using yellow fluorescent protein and histone H2B label–retention assays reveals that ~20% of *Lgr5*-expressing cells cycle less frequently, exhibit a mixed ISC–Paneth cell transcriptional profile, and differentiate into Paneth and neuroendocrine cells (44). Although these slow-cycling cells do not contribute to crypt homeostasis during physiological conditions, they can form organoids in vitro with comparable efficiencies as rapidly cycling *Lgr5*$^+$ ISCs and can mediate crypt regeneration after injuries (44).

Despite these behavioral distinctions among ISCs, their cellular dynamics within the crypt systematically drift toward monoclonality (45–49). Thus, over time, each crypt-villus unit derives from a single ISC (Fig. 3E). The mechanism leading to crypt monoclonality is thought to derive from neutral competition between an equipotent pool of ISCs that includes both *Lgr5* and *Bmi1-Hopx* ISCs (49). In contrast to epidermis (in which progenitors divide asymmetrically), ISCs are thought to divide symmetrically and compete for niche space (48, 49). Based initially on *Lgr5* expression and mathematical modeling (48, 49) and subsequently on a novel method of continuous labeling (50), it is estimated that between 5 and 16 *Lgr5*$^+$ ISCs compete with each other for niche space in a neutral drift manner.

Live imaging of *Lgr5*-Cre^{ER} lineage tracing has recently enabled the visualization of these displacements during ISC divisions. Ironically, with each division, ISCs reorganize their position within the crypt, which underscores their interconvertibility (42) (Fig. 3, B and C). In the end, one ISC outcompetes the others. It will be interesting to see in the future whether such competition happens in other SC niches and how the competition unfolds at a molecular level.

Crypt monoclonality underscores the multilineage potential of ISCs. Increasing evidence suggests that their fate choices are rooted at the transcriptional level. Thus, equipotent progenitors undergoing Notch-mediated lateral inhibition quickly enable distinct—in this case, reversible—cell fates to establish progenitor cell lineages as either absorptive or secretory. Moreover, *Atoh1*, a secretory-specific transcription factor expressed by ISCs, controls lateral inhibition through *Dll* genes and also drives expression of secretory lineage genes, which suggests that intestinal crypt lineage plasticity involves a lineage-restricted transcription factor expressed by multipotent ISCs (51).

Switch from multipotency to unipotency in glandular epithelia

Mammary glands (MGs), SwGs, and prostate glands are composed of an inner luminal layer, surrounded by an outer layer of myoepithelial and/or basal cells. Their morphogenesis begins late in embryogenesis and is completed postnatally.

As judged by lineage tracing, both MGs and SwGs and their associated ducts originate from K14-expressing multipotent embryonic epidermal progenitors (52–54). Although it was recently suggested that some bipotent SCs persist within the myoepithelial layer (55), myoepithelial and luminal lineages of MGs, SwGs, and prostate are largely maintained postnatally by distinct pools of unipotent SCs (52–54, 56–60) (Fig. 4A).
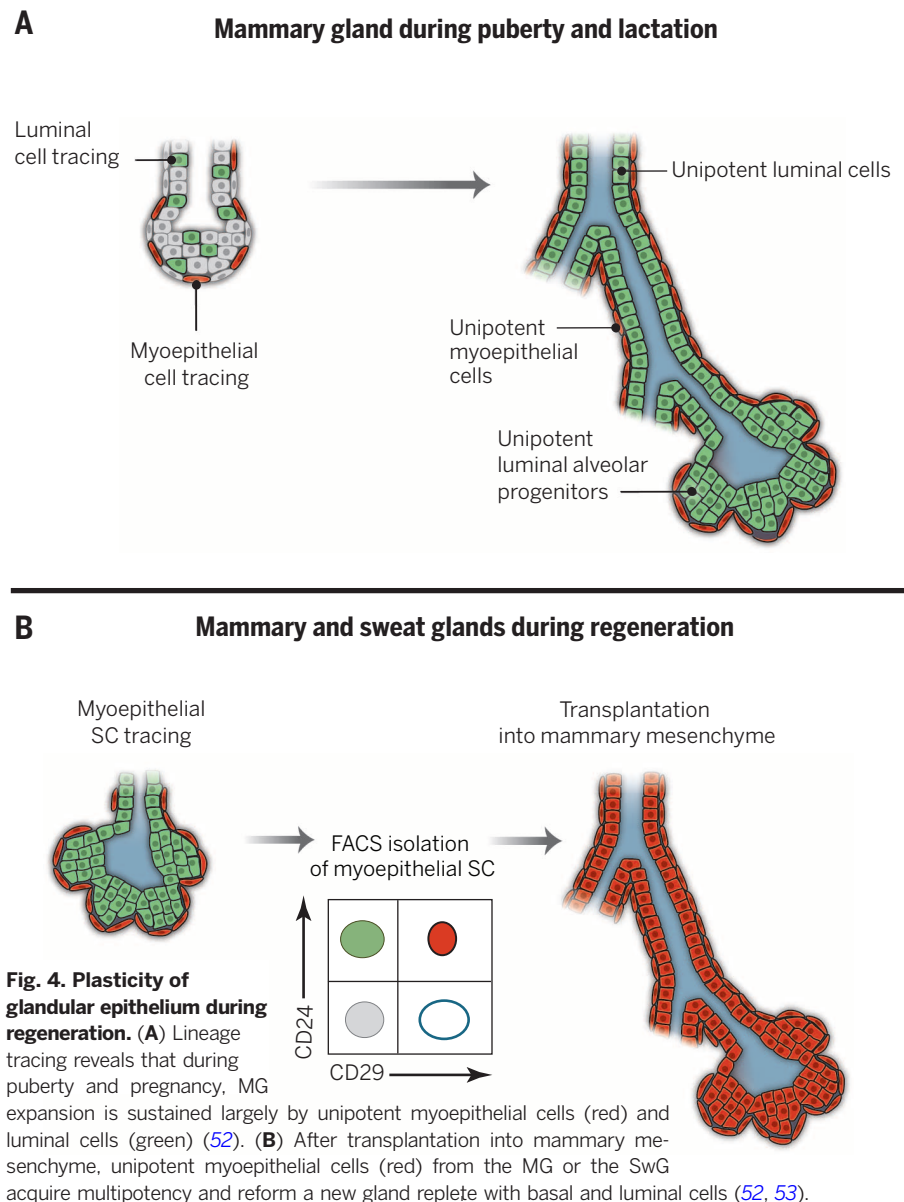
Mammary gland during puberty and lactation

Luminal cell tracing

Myoepithelial cell tracing

Unipotent luminal cells

Unipotent myoepithelial cells

Unipotent luminal alveolar progenitors

Mammary and sweat glands during regeneration

Myoepithelial SC tracing

FACS isolation of myoepithelial SC

Transplantation into mammary mesenchyme

CD24

CD29

Fig. 4. Plasticity of glandular epithelium during regeneration. (**A**) Lineage tracing reveals that during puberty and pregnancy, MG expansion is sustained largely by unipotent myoepithelial cells (red) and luminal cells (green) (52). (**B**) After transplantation into mammary mesenchyme, unipotent myoepithelial cells (red) from the MG or the SwG acquire multipotency and reform a new gland replete with basal and luminal cells (52, 53).

In the adult, both myoepithelial and luminal epithelial SwG SCs display very little turnover during homeostasis (*53*). By contrast, MG's SCs exert tremendous tissue-generating potential during puberty and pregnancy, making them especially well suited for studying glandular SC biology (*52, 54*). Heterogeneity within luminal and alveolar compartments has been seen with *Notch2*-Cre[ER] and *Notch3*-Cre[ER] lineage tracing (*59, 60*). Whether
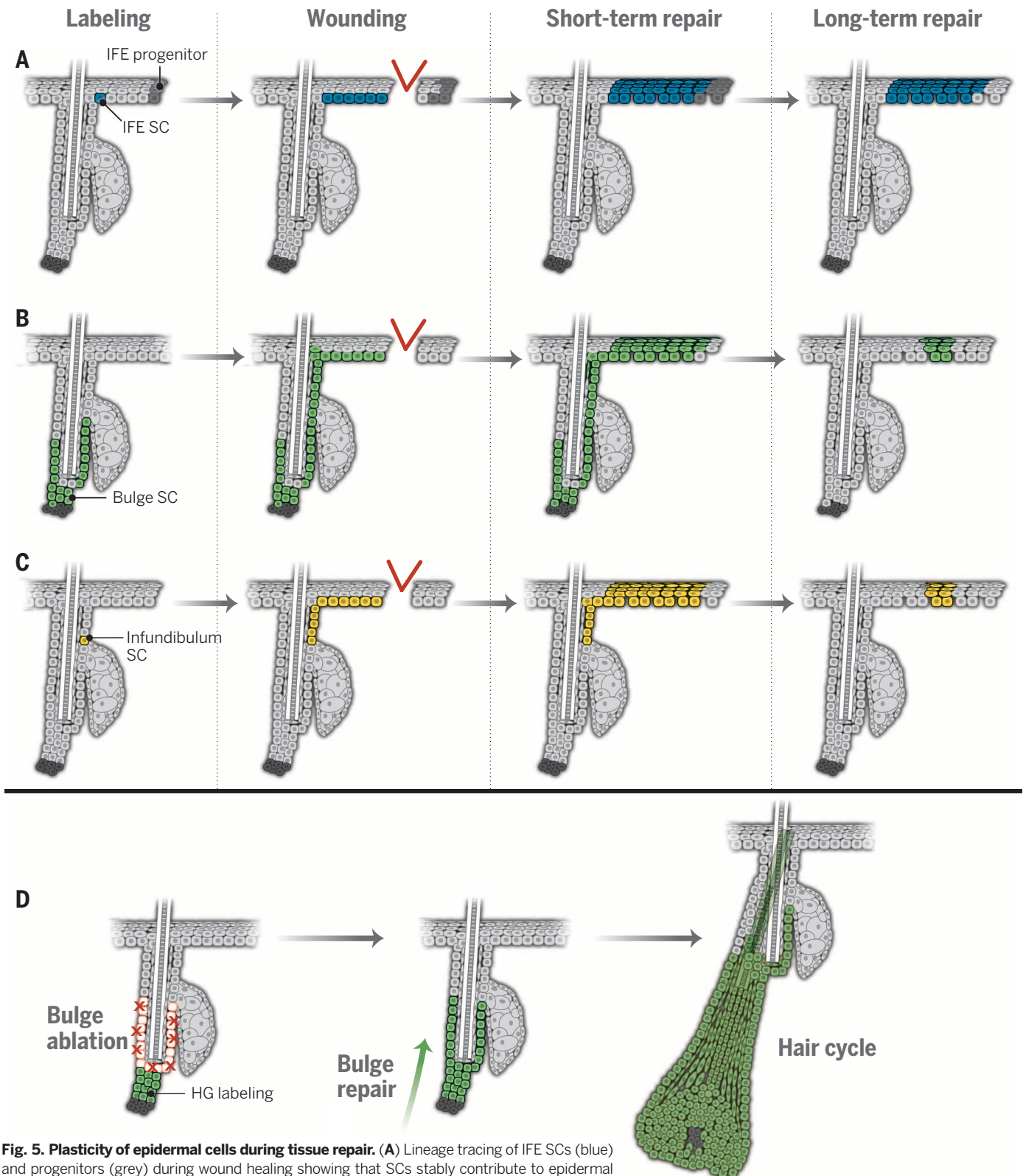

Fig. 5. Plasticity of epidermal cells during tissue repair. (**A**) Lineage tracing of IFE SCs (blue) and progenitors (grey) during wound healing showing that SCs stably contribute to epidermal repair while progenitor contribution is only transient (*26*). (**B** and **C**) Lineage tracing of bulge (B) and infundibulum SCs (C) demonstrate that adult HFSCs are rapidly recruited to IFE during wounding, but very few cells survive and contribute to IFE homeostasis after wound repair (*19, 21*). (**D**) After ablation of bulge cells (red X's), hair germ (HG) cells (green) recolonize the bulge niche and mediate hair regeneration (*16*).

these two luminal populations can interconvert remains unknown.

During prostate development, clonal analyses also suggest heterogeneity, this time in the basal compartment. Bipotent and unipotent basal progenitors have been identified, as well as basal cells already committed to the luminal lineage (*61*). Whether this apparent cellular heterogeneity reflects the existence of distinct progenitors or, alternatively, stochastic fate decisions of a single multipotent progenitor remains to be determined (*61*).

Altogether, lineage-tracing experiments performed in different glandular epithelia show that they initially develop from multipotent progenitors which are progressively replaced by unipotent SCs for adult tissue homeostasis and repair (*52*, *53*, *56*, *61*). However, despite similar histologies and SC behaviors, their multipotency → unipotency switch occurs at different times during development (*52*, *53*, *56–58*, *61*).

Transient plasticity of epithelial SC during tissue repair

Over evolution, homeostasis has been optimized for different SC compartments to replace local cells that die. However, if one SC compartment is damaged, other SCs must be recruited to repair the injury. A series of recent studies reveals that the fate and differentiation potential of epithelial cells can broaden during tissue regeneration after wounding. In some cases, unipotent progenitors acquire multipotency, whereas, in others, normally committed cells revert back to a SC-like state to ensure tissue regeneration. The cellular plasticity and reversibility observed in adult epithelial tissues have not been associated with "transdifferentiation" into completely unrelated fates but rather with contribution to the repair of the tissue from which the cells originated. In this regard, the plasticity seems to arise through a process of dedifferentiation and/or redifferentiation.

How SCs respond to injuries and repair tissue wounds varies dramatically depending not only on the particular SC niche but also its proximity to the wound. In SwG cells, for example, where four different unipotent progenitors exist (*53*), luminal and myoepithelial progenitors are mobilized, but these SCs act unipotently in mediating tissue regeneration, at least under conditions where luminal or myoepithelial progenitors are selectively killed (*53*). Although these findings illustrate the ability of different SC compartments to mobilize in response to different types of injuries, each SC niche knows its own job and does not carry out the job of other resident niches.

By monitoring the fate of early IFE progeny during wound repair, signs of transient plasticity begin to surface. Thus, although long-lived IFE SCs are recruited to the wound region and stably contribute to reepithelialization, short-lived *involucrin*+ IFE progenitors also migrate to wound sites. Within a month, most *involucrin*+-derived progeny terminally differentiate (*26*), which suggests that lineage reversion is not sustained long-term (Fig. 5A). The apparent transient nature of the

lineage reversion observed in IFE contrasts with esophagus, where progenitors seem to change their mode of proliferation in repairing incisional wounds (*62*). Whether this difference is attributed to the type and/or severity of wound (incisional versus full thickness) or a fundamental difference in SC behaviors remains to be addressed.

Transient plasticity has also been reported for adult HF SCs in response to injury. In super-

ficial skin wounds, bulge and infundibulum SCs migrate upward, proliferate, and participate in the epidermal repair process (*8*, *19*, *21*, *22*, *63*) (Fig. 5, B and C). Through mechanisms presently unknown, migrating HF SCs lose HF markers and adopt an IFE differentiation program. However, unlike neonatal skin, most of these cells do not seem to persist long-term within IFE (*19*, *63*, *64*). In this regard, they act more like a cellular bandage, which perhaps analogously
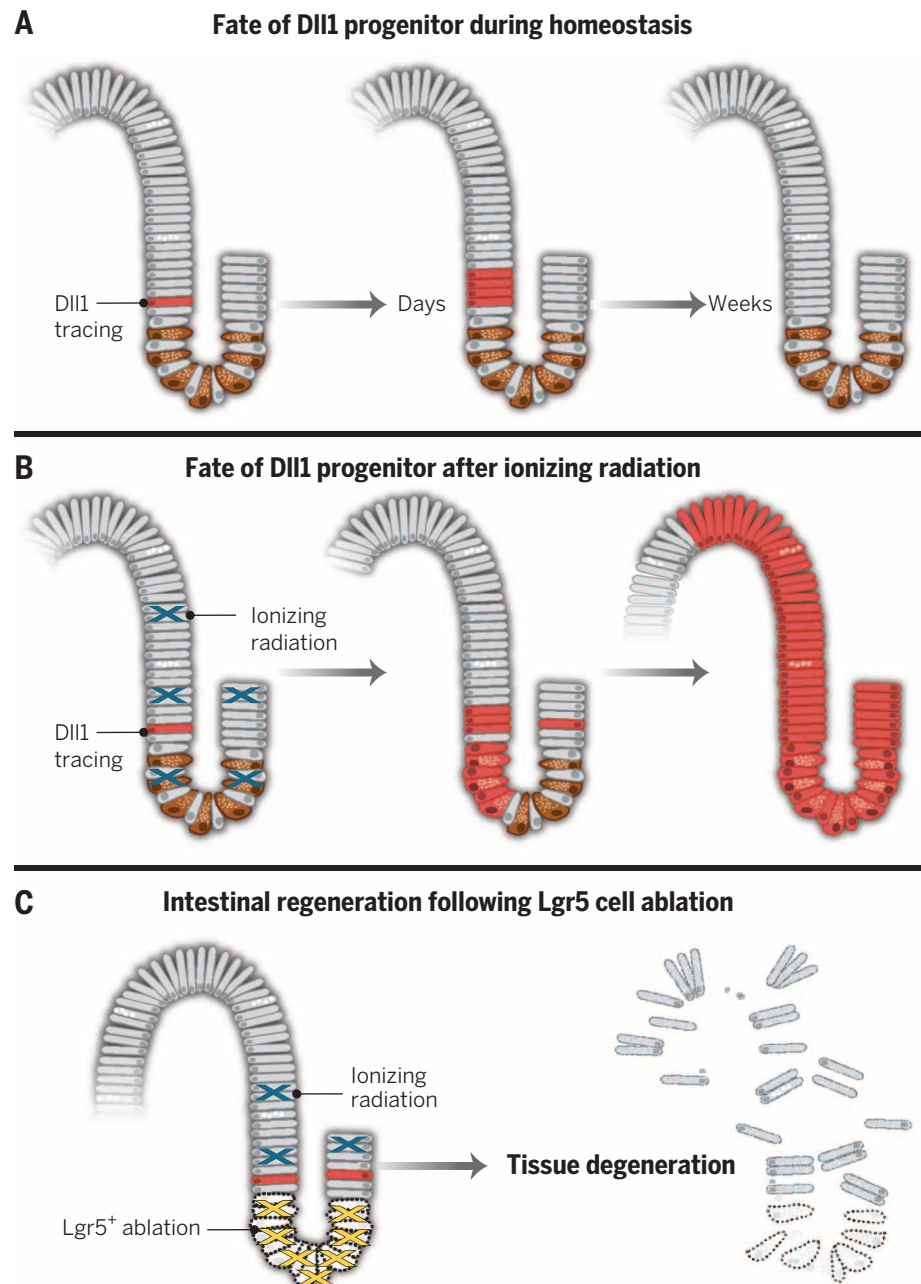

A **Fate of Dll1 progenitor during homeostasis**

Dll1 tracing → Days → Weeks

B **Fate of Dll1 progenitor after ionizing radiation**

Ionizing radiation
Dll1 tracing

C **Intestinal regeneration following Lgr5 cell ablation**

Ionizing radiation
Lgr5+ ablation → **Tissue degeneration**

Fig. 6. Plasticity and interconversion into SCs during intestinal regeneration. (**A** and **B**) Dll1 lineage tracing showing that, although *Dll1*+ cells (red) are transient and typically only differentiate into secretory cells (black; interspersed in villus) during homeostasis (A), upon γ-radiation–induced cell death (blue X's), *Dll1*+ TA cells revert and colonize the crypt (B) (*82*). (**C**) When intestine is depleted of *Lgr5*+ cells (yellow X's) and then exposed to γ-radiation, regeneration is impaired, revealing a critical role for Lgr5+ cells in repairing extensive tissue damage (*83*).

to *involucrin*⁺ IFE progenitors (*26*), are quick to respond but are eventually replaced by IFE SCs and their progeny.

Two relatively recent strategies to kill resident SCs—either laser-ablating them or ablating them through DT expression—have proven to be powerful methods to extricate SCs from their niches and examine the consequences. Initially shown for *Drosophila* germ SCs (*65*, *66*), it is now well established that when mammalian epithelial SCs are ablated, the empty niches can recruit and induce normally committed cells to proliferate and revert back to a stemlike state.

A particularly elegant demonstration of this paradigm was made by coupling live imaging with laser-mediated cell lineage ablation of different HF populations (*14*, *16*). The cellular plasticity within the bulge HF SC niche was documented by illustrating that bulge and HG cells can interconvert when one of these compartments is emptied (*16*, *67*) (Fig. 5D). It is noteworthy that cells located in the upper bulge region, the so-called "junctional zone" SCs, could also replenish the bulge niche after bulge SC ablation (*16*). Although future studies will be necessary to more closely examine the long-term capacity to interconvert into each other's fate and restore tissue function after injury, these findings capture the plasticity displayed by distinct skin epithelial SC compartments after injuries.

The microenvironment controls the fate of epithelial SCs

It has long been observed that when SCs are taken out of context and transplanted, either directly or after cell culture, they exhibit greater multipotency in their new microenvironment. Thus, upon engraftment to immunocompromised mice, freshly isolated bulge cells (*9*, *68*) or clonal progeny of single bulge cells (*69*, *70*) each generate not only HFs, but also IFE and SGs long-term (Fig. 2B). This is also true for isthmus and SG SCs (*71*, *72*). Analogously, when normally unipotent SwG, MG, or prostate basal or myoepithelial SCs are purified and engrafted de novo, they generate entire functional glands (*52*, *53*, *73–76*).

When unipotent MG myoepithelial cells are transplanted into mammary mesenchyme of pregnant mice, they can reform a functional MG (*52*) (Fig. 4B), which demonstrates the plasticity of unipotent myoepithelial cells during regenerative conditions. Note that MG myoepithelial cells can also generate MGs when engrafted to shoulder pads, whereas SwG myoepithelial cells generate SwGs in virgin mammary fat pads (*53*). These findings suggest that for some adult progenitors, once identity is established, they take longer to respond to environmental and systemic programming factors. By contrast, when progenitors form tissue de novo during embryonic development, they have yet to receive the epigenetic marks that restrict their fates.

Similarly, after culture in vitro, marked thymic epithelial cells can be mixed with embryonic thymus and transplanted underneath the kidney capsule, where they integrate into the thymic network and differentiate into functional thymic epithelial cells (*77*). However, when the same cultured thymic epithelial cells are transplanted together with skin mesenchyme onto back skin, they differentiate into all epidermal lineages including HF and IFE (*77*). This plasticity in SC behavior appears to become more permanent with subsequent transplantations, illustrating how the microenvironment can instruct these cells to adopt very different fates.

A hint that adult epithelial cells may be able to undergo permanent fate conversions in vivo comes from monitoring IFE behavior after massive wounding. In this case, the IFE was reported to regenerate HFs, which is something it never does during homeostasis (*78*). It has long been known that transgenic β-catenin stabilization, the output of a Wnt signal, is sufficient to reprogram K14⁺ IFE into HFs replete with their own DP (*79*). Overexpressing the hedgehog pathway also stimulates IFE to HF progenitor reprogramming, but in this case, differentiation becomes suppressed at the expense of hyperproliferation, which leads to basal cell carcinoma (*80*, *81*).

Reversing fates: Converting committed progeny to SCs

Although the ability of adult epithelial SCs to acquire different lineage fates seems remarkable, several studies have recently suggested that committed epithelial lineage cells may have the capacity to acquire stemness. During normal homeostasis in the intestine, Delta-like 1 (*Dll1*)–expressing cells (*82*), or slow-cycling *Lgr5*⁺ cells (*44*), are both short-lived Lgr5-derived progeny committed to the secretory lineage. However, after γ-irradiation–induced tissue damage, these normally committed Dll1⁺ progenitors appear to revert back to ISCs (*82*) and contribute to intestinal regeneration (Fig. 6, A and B). Similarly, when Dll1⁺ progenitors are purified and placed in Wnt3a-supplemented cultures, they form gut organoids containing *Lgr5*⁺ SCs and all intestinal lineages (*82*), which supports the idea that they revert into a stemlike state. How Wnt signaling might influence the reversion process in vivo is a yet-unaddressed intriguing question. Whether these reserve cells are sufficient to be functionally relevant in the context of tissue repair is still unclear, as γ-irradiation–induced intestinal epithelial regeneration does not occur after *Lgr5* ablation (Fig. 6C) (*83*).

Another example of plasticity stems from recent lineage tracing of committed secretory cells in the lung (*84*), which can revert into stable and functional basal SCs in vivo if all airway SCs are ablated (*85*) (Fig. 7). In this case, it was shown that these dedifferentiated cells can respond to epithelial injury and repair injuries equivalently to their endogenous SC counterparts. By contrast, direct contact with a single basal SC was sufficient to prevent secretory cell dedifferentiation, suggestive of negative cross-talk between SCs and committed progeny. Overall, the propensity of committed cells to dedifferentiate is typically inversely correlated to their state of maturity.

The ability of a priori differentiated cells to be reprogrammed and interconvert into SCs has also been
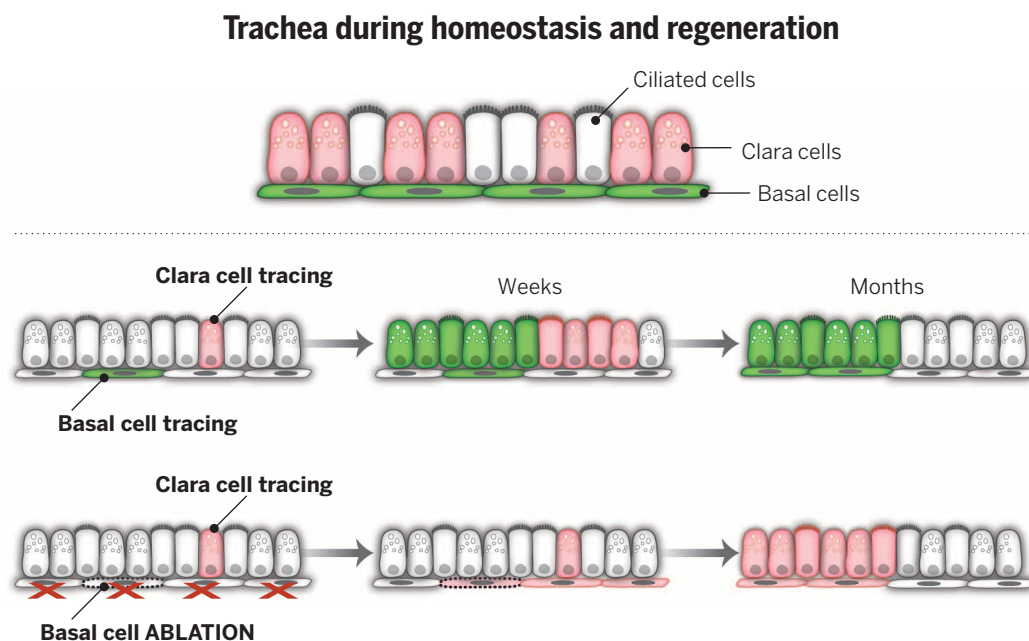
Trachea during homeostasis and regeneration

Fig. 7. Plasticity and interconversion into SCs during tracheal regeneration. During tracheal homeostasis, basal cells (green) give rise to TA Clara cells (pink) and terminally differentiated ciliated cells (white). Lineage ablation of basal cells (red X's) induces the interconversion of Clara and/or ciliated cells into basal SCs (*85*).

illustrated for stomach (*86*). The stomach epithelium is composed of an upper part of rapidly renewing cells, a middle zone, the isthmus that actively proliferates, and a bottom zone that contains two cell types (parietal and chief cells) with very low cellular turnover (*87*). Lineage tracing revealed that *Sox2*-expressing cells in the isthmus region are responsible for the homeostasis of the glandular stomach, giving rise to all stomach lineages (*88*). Through lineage tracing using *Troy*-Cre[ER] to target differentiated parietal and chief cells (*86*), it was reported that progeny of some Troy cells slowly expand and reach the top of the gland after 6 months of chase, which shows that these cells play only a very minor role during homeostasis. However, the Troy cells can be cultured long-term as multipotent organoids in vitro and expand severalfold after tissue damage in vivo, suggestive of their ability to aid in repair of stomach injuries (*86*).

After acute injuries, liver and pancreatic beta-cell regeneration seems to involve self-duplication of differentiated cells (*89*, *90*). In contrast, chronic and severe hepatic injuries stimulate mature hepatocytes and/or biliary cells to dedifferentiate into bipotent progenitor state–expressing SC markers, such as *Lgr5*, that mediate liver regeneration through their proliferation and redifferentiation (*91*).

Altogether, these remarkable studies point to the view that, under certain nonhomeostatic conditions, differentiated cells dedifferentiate, revert back to a SC-like fate, and participate in tissue repair. In particular, this seems to happen after severe injury, a situation where the tissue must respond quickly and creatively to ensure animal survival.

Reversibility of lineage differentiation and SC plasticity during tumorigenesis

The plasticity of epithelial lineage commitment and the ability of committed progeny to revert back to SCs may have important implications for tumorigenesis. In 1990, this notion was initially postulated by Bailleul *et al.*, who observed that mice expressing an oncogenic *Hras* driven by a differentiation-specific promoter develop papillomas after wounding (*92*). In an interesting variation to this theme, normally fate-restricted, unipotent basal and luminal SCs of glandular epithelia reacquire certain features of multipotent SCs during tumor progression. For instance, tumor suppressor inactivation in luminal MG cells can lead to the formation of basal-like breast cancer (*93*), replete with heterogeneous expression of both basal and luminal markers.

Similar observations have been made for prostate cancer, where ablation of a tumor suppressor gene in luminal SCs induces tumor formation (*57*). Basal progenitors seem intrinsically more resistant to tumorigenesis, and even when they undergo a fate transition into luminal cells, the tumorigenic lesions that appear are less aggressive than those originating directly from luminal cells (*56*, *58*).

Irrespective of underlying cause or mechanism, the plasticity within the tissue hierarchical organization is likely to have broader implications for tumor initiation and maintenance. In the intestine, for instance, adenomas arise from activating mutations in the Wnt/β-catenin pathway. After a single oncogenic hit, only *Lgr5/Bmi1/prominin*–expressing ISCs initiate tumor formation (*39*, *94*, *95*), whereas targeting TA progeny have either no effect or induce only microadenomas (*94*). However, concomitant activation of the Wnt pathway and another oncogenic hit cause normally committed TA cells to revert to a SC-like state and induce tumor formation (*96*).

Once initiated, these tumors may display hierarchical organization, replete with tumor-propagating cells (so-called cancer SCs), defined functionally by their ability upon serial transplantation to induce secondary tumors that resemble the parental tumor. Distinct populations of cells with tumor-propagating capacity capable of interconversion have also been identified within cancers (*96*–*99*), which raises the possibility that upon transplantation, more committed cells within a heterogeneous cancer may reacquire SC properties, analogous to the plasticity observed in normal SCs after transplantation. Consistent with this notion, non-SCs of human basal breast cancers can switch to SC state, depending on ZEB1, a regulator of the epithelial-mesenchymal transition (*100*). This result suggests a dynamic model where interconversion between low and high tumorigenic states can occur, which increases the potential for cancer progression. Further studies will be required to define the extent to which extent cell plasticity influences cancer growth and relapse after therapy.

Conclusion

The examples provided in this Review have highlighted the hierarchical and spatial organization of epithelial tissue homeostasis and the important plasticity of progenitors and differentiated cells during regenerative conditions. This cellular plasticity and lineage reversibility may represent adaptive mechanisms for the self-preservation of epithelia to repair body surfaces and linings in whatever ways possible after injuries. Across many different epithelia subjected to a diverse array of injuries, the paradigm emerging is that the minimum number of SCs needed to repair injuries will be activated and recruited during the healing process. As injuries become more severe, and greater numbers of SCs become depleted from their niches, more SCs become mobilized to participate in wound repair. When all SCs are exhausted, early progeny become recruited, until eventually, with massive injuries, the tissue can no longer cope with repair. Although the molecular mechanisms underlying cellular plasticity, fate conversion, and reacquisition of stem cell properties in committed and/or differentiated cells still remain poorly understood, these versatile built-in programs have major implications for regenerative medicine. On the flip side of this coin, however, is that when gone awry,

these repertoires become the curse of epithelial SCs, most of which contribute in major ways to the most life-threatening of human cancers.

REFERENCES AND NOTES

1. C. Blanpain, B. D. Simons, Unravelling stem cell dynamics by lineage tracing. *Nat. Rev. Mol. Cell Biol.* **14**, 489–502 (2013). doi: 10.1038/nrm3625; pmid: 23860235
2. S. J. Morrison, A. C. Spradling, Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* **132**, 598–611 (2008). doi: 10.1016/j.cell.2008.01.038; pmid: 18295578
3. H. Clevers, The intestinal crypt, a prototype stem cell compartment. *Cell* **154**, 274–284 (2013). doi: 10.1016/j.cell.2013.07.004; pmid: 23870119
4. J. G. Rheinwald, H. Green, Serial cultivation of strains of human epidermal keratinocytes: the formation of keratinizing colonies from single cells. *Cell* **6**, 331–343 (1975). doi: 10.1016/S0092-8674(75)80001-8; pmid: 1052771
5. G. G. Gallico 3rd, N. E. O'Connor, C. C. Compton, O. Kehinde, H. Green, Permanent coverage of large burn wounds with autologous cultured human epithelium. *N. Engl. J. Med.* **311**, 448–451 (1984). doi: 10.1056/NEJM198408163110706; pmid: 6379456
6. T. Sato *et al.*, Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **459**, 262–265 (2009). doi: 10.1038/nature07935; pmid: 19329995
7. G. Cotsarelis, T. T. Sun, R. M. Lavker, Label-retaining cells reside in the bulge area of pilosebaceous unit: Implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61**, 1329–1337 (1990). doi: 10.1016/0092-8674(90)90696-C; pmid: 2364430
8. T. Tumbar *et al.*, Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004). doi: 10.1126/science.1092436; pmid: 14671312
9. R. J. Morris *et al.*, Capturing and profiling adult hair follicle stem cells. *Nat. Biotechnol.* **22**, 411–417 (2004). doi: 10.1038/nbt950; pmid: 15024388
10. V. Levy, C. Lindon, B. D. Harfe, B. A. Morgan, Distinct stem cell populations regenerate the follicle and interfollicular epidermis. *Dev. Cell* **9**, 855–861 (2005). doi: 10.1016/j.devcel.2005.11.003; pmid: 16326396
11. Y. V. Zhang, J. Cheong, N. Ciapurin, D. J. McDermitt, T. Tumbar, Distinct self-renewal and differentiation phases in the niche of infrequently dividing hair follicle stem cells. *Cell Stem Cell* **5**, 267–278 (2009). doi: 10.1016/j.stem.2009.06.004; pmid: 19664980
12. Y. C. Hsu, H. A. Pasolli, E. Fuchs, Dynamics between stem cells, niche, and progeny in the hair follicle. *Cell* **144**, 92–105 (2011). doi: 10.1016/j.cell.2010.11.049; pmid: 21215372
13. V. Greco *et al.*, A two-step mechanism for stem cell activation during hair regeneration. *Cell Stem Cell* **4**, 155–169 (2009). doi: 10.1016/j.stem.2008.12.009; pmid: 19200804
14. P. Rompolas *et al.*, Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration. *Nature* **487**, 496–499 (2012). doi: 10.1038/nature11218; pmid: 22763436
15. Y. C. Hsu, L. Li, E. Fuchs, Transit-amplifying cells orchestrate stem cell activity and tissue regeneration. *Cell* **157**, 935–949 (2014). doi: 10.1016/j.cell.2014.02.057; pmid: 24813615
16. P. Rompolas, K. R. Mesa, V. Greco, Spatial organization within a niche as a determinant of stem-cell fate. *Nature* **502**, 513–518 (2013). doi: 10.1038/nature12602; pmid: 24097351
17. V. Horsley *et al.*, Blimp1 defines a progenitor population that governs cellular input to the sebaceous gland. *Cell* **126**, 597–609 (2006). doi: 10.1016/j.cell.2006.06.048; pmid: 16901790
18. H. J. Snippert *et al.*, Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science* **327**, 1385–1389 (2010). doi: 10.1126/science.1184733; pmid: 20233988
19. M. E. Page, P. Lombard, F. Ng, B. Göttgens, K. B. Jensen, The epidermis comprises autonomous compartments maintained by distinct stem cell populations. *Cell Stem Cell* **13**, 471–482 (2013). doi: 10.1016/j.stem.2013.07.010; pmid: 23954751
20. I. Brownell, E. Guevara, C. B. Bai, C. A. Loomis, A. L. Joyner, Nerve-derived sonic hedgehog defines a niche for hair follicle stem cells capable of becoming epidermal stem cells. *Cell Stem Cell* **8**, 552–565 (2011). doi: 10.1016/j.stem.2011.02.021; pmid: 21549329
21. M. Ito *et al.*, Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nat. Med.* **11**, 1351–1354 (2005). doi: 10.1038/nm1328; pmid: 16288281

22. V. Jaks et al., Lgr5 marks cycling, yet long-lived, hair follicle stem cells. Nat. Genet. 40, 1291–1299 (2008). doi: 10.1038/ng.239; pmid: 18849992

23. K. K. Youssef et al., Identification of the cell lineage at the origin of basal cell carcinoma. Nat. Cell Biol. 12, 299–305 (2010).pmid: 20154679

24. A. Marianes, A. C. Spradling, Physiological and stem cell compartmentalization within the Drosophila midgut. eLife 2, e00886 (2013). doi: 10.7554/eLife.00886; pmid: 23991285

25. T. Lechler, E. Fuchs, Asymmetric cell divisions promote stratification and differentiation of mammalian skin. Nature 437, 275–280 (2005). doi: 10.1038/nature03922; pmid: 16094321

26. G. Mascré et al., Distinct contribution of stem and progenitor cells to epidermal maintenance. Nature 489, 257–262 (2012). doi: 10.1038/nature11393; pmid: 22940863

27. E. Clayton et al., A single type of progenitor cell maintains normal epidermis. Nature 446, 185–189 (2007). doi: 10.1038/nature05574; pmid: 17330052

28. D. P. Doupé, A. M. Klein, B. D. Simons, P. H. Jones, The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate. Dev. Cell 18, 317–323 (2010). doi: 10.1016/j.devcel.2009.12.016; pmid: 20159601

29. X. Lim et al., Interfollicular epidermal stem cells self-renew via autocrine Wnt signaling. Science 342, 1226–1230 (2013). doi: 10.1126/science.1239730; pmid: 24311688

30. P. H. Jones, F. M. Watt, Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. Cell 73, 713–724 (1993). doi: 10.1016/0092-8674(93)90251-K; pmid: 8500165

31. A. Rangarajan et al., Notch signaling is a direct determinant of keratinocyte growth arrest and entry into differentiation. EMBO J. 20, 3427–3436 (2001). doi: 10.1093/emboj/20.13.3427; pmid: 11432830

32. C. Blanpain, W. E. Lowry, H. A. Pasolli, E. Fuchs, Canonical notch signaling functions as a commitment switch in the epidermal lineage. Genes Dev. 20, 3022–3035 (2006). doi: 10.1101/gad.1477606; pmid: 17079689

33. M. Moriyama et al., Multiple roles of Notch signaling in the regulation of epidermal development. Dev. Cell 14, 594–604 (2008). doi: 10.1016/j.devcel.2008.01.017; pmid: 18410734

34. S. E. Williams, S. Beronja, H. A. Pasolli, E. Fuchs, Asymmetric cell divisions promote Notch-dependent epidermal differentiation. Nature 470, 353–358 (2011). doi: 10.1038/nature09793; pmid: 21331036

35. H. Cheng, C. P. Leblond, Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. I. Columnar cell. Am. J. Anat. 141, 461–479 (1974). doi: 10.1002/aja.1001410403; pmid: 4440632

36. C. S. Potten, W. J. Hume, P. Reid, J. Cairns, The segregation of DNA in epithelial stem cells. Cell 15, 899–906 (1978). doi: 10.1016/0092-8674(78)90274-X; pmid: 728994

37. A. J. Merritt et al., The role of p53 in spontaneous and radiation-induced apoptosis in the gastrointestinal tract of normal and p53-deficient mice. Cancer Res. 54, 614–617 (1994).pmid: 8306319

38. N. Barker et al., Identification of stem cells in small intestine and colon by marker gene Lgr5. Nature 449, 1003–1007 (2007). doi: 10.1038/nature06196; pmid: 17934449

39. E. Sangiorgi, M. R. Capecchi, Bmi1 is expressed in vivo in intestinal stem cells. Nat. Genet. 40, 915–920 (2008). doi: 10.1038/ng.165; pmid: 18536716

40. R. K. Montgomery et al., Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. Proc. Natl. Acad. Sci. U.S.A. 108, 179–184 (2011). doi: 10.1073/pnas.1013004108; pmid: 21173232

41. N. Takeda et al., Interconversion between intestinal stem cell populations in distinct niches. Science 334, 1420–1424 (2011). doi: 10.1126/science.1213214; pmid: 22075725

42. L. Ritsma et al., Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. Nature 507, 362–365 (2014). doi: 10.1038/nature12972; pmid: 24531760

43. H. Tian et al., A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable. Nature 478, 255–259 (2011). doi: 10.1038/nature10408; pmid: 21927002

44. S. J. Buczacki et al., Intestinal label-retaining cells are secretory precursors expressing Lgr5. Nature 495, 65–69 (2013). doi: 10.1038/nature11965; pmid: 23446353

45. B. A. Ponder et al., Derivation of mouse intestinal crypts from single progenitor cells. Nature 313, 689–691 (1985). doi: 10.1038/313689a0; pmid: 3974703

46. D. F. Griffiths, S. J. Davies, D. Williams, G. T. Williams, E. D. Williams, Demonstration of somatic mutation and

colonic crypt clonality by X-linked enzyme histochemistry. Nature 333, 461–463 (1988). doi: 10.1038/333461a0; pmid: 3374587

47. D. J. Winton, M. A. Blount, B. A. Ponder, A clonal marker induced by mutation in mouse intestinal epithelium. Nature 333, 463–466 (1988). doi: 10.1038/333463a0; pmid: 3163778

48. C. Lopez-Garcia, A. M. Klein, B. D. Simons, D. J. Winton, Intestinal stem cell replacement follows a pattern of neutral drift. Science 330, 822–825 (2010). doi: 10.1126/science.1196236; pmid: 20929733

49. H. J. Snippert et al., Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell 143, 134–144 (2010). doi: 10.1016/j.cell.2010.09.016; pmid: 20887898

50. S. Kozar et al., Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. Cell Stem Cell 13, 626–633 (2013). doi: 10.1016/j.stem.2013.08.001; pmid: 24035355

51. T. H. Kim et al., Broadly permissive intestinal chromatin underlies lateral inhibition and cell plasticity. Nature 506, 511–515 (2014). doi: 10.1038/nature12903; pmid: 24413398

52. A. Van Keymeulen et al., Distinct stem cells contribute to mammary gland development and maintenance. Nature 479, 189–193 (2011). doi: 10.1038/nature10573; pmid: 21983963

53. C. P. Lu et al., Identification of stem cell populations in sweat glands and ducts reveals roles in homeostasis and wound repair. Cell 150, 136–150 (2012). doi: 10.1016/j.cell.2012.04.045; pmid: 22770217

54. R. van Amerongen, A. N. Bowman, R. Nusse, Developmental stage and time dictate the fate of Wnt/β-catenin-responsive stem cells in the mammary gland. Cell Stem Cell 11, 387–400 (2012). doi: 10.1016/j.stem.2012.05.023; pmid: 22863533

55. A. C. Rios, N. Y. Fu, G. J. Lindeman, J. E. Visvader, In situ identification of bipotent stem cells in the mammary gland. Nature 506, 322–327 (2014). doi: 10.1038/nature12948; pmid: 24463516

56. N. Choi, B. Zhang, L. Zhang, M. Ittmann, L. Xin, Adult murine prostate basal and luminal cells are self-sustained lineages that can both serve as targets for prostate cancer initiation. Cancer Cell 21, 253–265 (2012). doi: 10.1016/j.ccr.2012.01.005; pmid: 22340597

57. X. Wang et al., A luminal epithelial stem cell that is a cell of origin for prostate cancer. Nature 461, 495–500 (2009). doi: 10.1038/nature08361; pmid: 19741607

58. Z. A. Wang et al., Lineage analysis of basal epithelial cells reveals their unexpected plasticity and supports a cell-of-origin model for prostate cancer heterogeneity. Nat. Cell Biol. 15, 274–283 (2013). doi: 10.1038/ncb2697; pmid: 23434823

59. S. Šale, D. Lafkas, S. Artavanis-Tsakonas, Notch2 genetic fate mapping reveals two previously unrecognized mammary epithelial lineages. Nat. Cell Biol. 15, 451–460 (2013). doi: 10.1038/ncb2725; pmid: 23604318

60. D. Lafkas et al., Notch3 marks clonogenic mammary luminal progenitor cells in vivo. J. Cell Biol. 203, 47–56 (2013). doi: 10.1083/jcb.201307046; pmid: 24100291

61. M. Ousset et al., Multipotent and unipotent progenitors contribute to prostate postnatal development. Nat. Cell Biol. 14, 1131–1138 (2012). doi: 10.1038/ncb2600; pmid: 23064263

62. D. P. Doupé et al., A single progenitor population switches behavior to maintain and repair esophageal epithelium. Science 337, 1091–1093 (2012). doi: 10.1126/science.1218835; pmid: 22821983

63. V. Levy, C. Lindon, Y. Zheng, B. D. Harfe, B. A. Morgan, Epidermal stem cells arise from the hair follicle after wounding. FASEB J. 21, 1358–1366 (2007). doi: 10.1096/fj.06-6926com; pmid: 17255473

64. J. A. Nowak, L. Polak, H. A. Pasolli, E. Fuchs, Hair follicle stem cells are specified and function in early skin morphogenesis. Cell Stem Cell 3, 33–43 (2008). doi: 10.1016/j.stem.2008.05.009; pmid: 18593557

65. T. Kai, A. Spradling, An empty Drosophila stem cell niche reactivates the proliferation of ectopic cells. Proc. Natl. Acad. Sci. U.S.A. 100, 4633–4638 (2003). doi: 10.1073/pnas.0830856100; pmid: 12676994

66. T. Kai, A. Spradling, Differentiating germ cells can revert into functional stem cells in Drosophila melanogaster ovaries. Nature 428, 564–569 (2004). doi: 10.1038/nature02436; pmid: 15024390

67. M. Ito, K. Kizawa, K. Hamada, G. Cotsarelis, Hair follicle stem cells in the lower bulge form the secondary germ, a biochemically distinct but functionally equivalent progenitor cell population, at the termination of catagen. Differentiation

72, 548–557 (2004). doi: 10.1111/j.1432-0436.2004.07209008.x; pmid: 15617565

68. H. Oshima, A. Rochat, C. Kedzia, K. Kobayashi, Y. Barrandon, Morphogenesis and renewal of hair follicles from adult multipotent stem cells. Cell 104, 233–245 (2001). doi: 10.1016/S0092-8674(01)00208-2; pmid: 11207364

69. C. Blanpain, W. E. Lowry, A. Geoghegan, L. Polak, E. Fuchs, Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche. Cell 118, 635–648 (2004). doi: 10.1016/j.cell.2004.08.012; pmid: 15339667

70. S. Claudinot, M. Nicolas, H. Oshima, A. Rochat, Y. Barrandon, Long-term renewal of hair follicles from clonogenic multipotent stem cells. Proc. Natl. Acad. Sci. U.S.A. 102, 14677–14682 (2005). doi: 10.1073/pnas.0507250102; pmid: 16203973

71. U. B. Jensen et al., A distinct population of clonogenic and multipotent murine follicular keratinocytes residing in the upper isthmus. J. Cell Sci. 121, 609–617 (2008). doi: 10.1242/jcs.025502; pmid: 18252795

72. J. G. Nijhof et al., The cell-surface marker MTS24 identifies a novel population of follicular keratinocytes with characteristics of progenitor cells. Development 133, 3027–3037 (2006). doi: 10.1242/dev.02443; pmid: 16818453

73. M. Shackleton et al., Generation of a functional mammary gland from a single stem cell. Nature 439, 84–88 (2006). doi: 10.1038/nature04372; pmid: 16397499

74. J. Stingl et al., Purification and unique properties of mammary epithelial stem cells. Nature 439, 993–997 (2006).pmid: 16395311

75. A. S. Goldstein et al., Trop2 identifies a subpopulation of murine and human prostate basal cells with stem cell characteristics. Proc. Natl. Acad. Sci. U.S.A. 105, 20882–20887 (2008). doi: 10.1073/pnas.0811411106; pmid: 19088204

76. L. Xin, R. U. Lukacs, D. A. Lawson, D. Cheng, O. N. Witte, Self-renewal and multilineage differentiation in vitro from murine prostate stem cells. Stem Cells 25, 2760–2769 (2007). doi: 10.1634/stemcells.2007-0355; pmid: 17641240

77. P. Bonfanti et al., Microenvironmental reprogramming of thymic epithelial cells to skin multipotent stem cells. Nature 466, 978–982 (2010). doi: 10.1038/nature09269; pmid: 20725041

78. M. Ito et al., Wnt-dependent de novo hair follicle regeneration in adult mouse skin after wounding. Nature 447, 316–320 (2007). doi: 10.1038/nature05766; pmid: 17507982

79. U. Gat, R. DasGupta, L. Degenstein, E. Fuchs, De novo hair follicle morphogenesis and hair tumors in mice expressing a truncated beta-catenin in skin. Cell 95, 605–614 (1998). doi: 10.1016/S0092-8674(00)81631-1; pmid: 9845363

80. S. H. Yang et al., Pathological responses to oncogenic Hedgehog signaling in skin are dependent on canonical Wnt/beta3-catenin signaling. Nat. Genet. 40, 1130–1135 (2008). doi: 10.1038/ng.192; pmid: 19165927

81. K. K. Youssef et al., Adult interfollicular tumour-initiating cells are reprogrammed into an embryonic hair follicle progenitor-like fate during basal cell carcinoma initiation. Nat. Cell Biol. 14, 1282–1294 (2012). doi: 10.1038/ncb2628; pmid: 23178882

82. J. H. van Es et al., Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. Nat. Cell Biol. 14, 1099–1104 (2012). doi: 10.1038/ncb2581; pmid: 23000963

83. C. Metcalfe, N. M. Kljavin, R. Ybarra, F. J. de Sauvage, Lgr5+ stem cells are indispensable for radiation-induced intestinal regeneration. Cell Stem Cell 14, 149–159 (2014). doi: 10.1016/j.stem.2013.11.008; pmid: 24332836

84. E. L. Rawlins et al., The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. Cell Stem Cell 4, 525–534 (2009). doi: 10.1016/j.stem.2009.04.002; pmid: 19497281

85. P. R. Tata et al., Dedifferentiation of committed epithelial cells into stem cells in vivo. Nature 503, 218–223 (2013). pmid: 24196716

86. D. E. Stange et al., Differentiated Troy+ chief cells act as reserve stem cells to generate all lineages of the stomach epithelium. Cell 155, 357–368 (2013). doi: 10.1016/j.cell.2013.09.008; pmid: 24120136

87. N. Barker, S. Bartfeld, H. Clevers, Tissue-resident adult stem cell populations of rapidly self-renewing organs. Cell Stem Cell 7, 656–670 (2010). doi: 10.1016/j.stem.2010.11.016; pmid: 21112561

88. K. Arnold et al., Sox2+ adult stem and progenitor cells are important for tissue regeneration and survival of mice. Cell

Stem Cell **9**, 317–329 (2011). doi: 10.1016/j.stem.2011.09.001; pmid: 21982232

89. A. W. Duncan, C. Dorrell, M. Grompe, Stem cells and liver regeneration. *Gastroenterology* **137**, 466–481 (2009). doi: 10.1053/j.gastro.2009.05.044; pmid: 19470389

90. Y. Dor, J. Brown, O. I. Martinez, D. A. Melton, Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature* **429**, 41–46 (2004). doi: 10.1038/nature02520; pmid: 15129273

91. M. Huch *et al.*, *In vitro* expansion of single Lgr5$^+$ liver stem cells induced by Wnt-driven regeneration. *Nature* **494**, 247–250 (2013). doi: 10.1038/nature11826; pmid: 23354049

92. B. Bailleul *et al.*, Skin hyperkeratosis and papilloma formation in transgenic mice expressing a *ras* oncogene from a suprabasal keratin promoter. *Cell* **62**, 697–708 (1990). doi: 10.1016/0092-8674(90)90115-U; pmid: 1696852

93. G. Molyneux *et al.*, BRCA1 basal-like breast cancers originate from luminal epithelial progenitors and not from basal stem cells. *Cell Stem Cell* **7**, 403–417 (2010). doi: 10.1016/j.stem.2010.07.010; pmid: 20804975

94. N. Barker *et al.*, Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009). doi: 10.1038/nature07602; pmid: 19092804

95. L. Zhu *et al.*, Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation. *Nature* **457**, 603–607 (2009). doi: 10.1038/nature07589; pmid: 19092805

96. S. Schwitalla *et al.*, Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152**, 25–38 (2013). doi: 10.1016/j.cell.2012.12.012; pmid: 23273993

97. E. Quintana *et al.*, Phenotypic heterogeneity among tumorigenic melanoma cells from patients that is reversible and not hierarchically organized. *Cancer Cell* **18**, 510–523 (2010). doi: 10.1016/j.ccr.2010.10.012; pmid: 21075313

98. M. Schober, E. Fuchs, Tumor-initiating stem cells of squamous cell carcinomas and their control by TGF-β and integrin/focal adhesion kinase (FAK) signaling. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10544–10549 (2011). doi: 10.1073/pnas.1107807108; pmid: 21670270

99. G. Lapouge *et al.*, Skin squamous cell carcinoma propagating cells increase with tumour progression and invasiveness.

EMBO J. **31**, 4563–4575 (2012). doi: 10.1038/emboj.2012.312; pmid: 23188079

100. C. L. Chaffer *et al.*, Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61–74 (2013). doi: 10.1016/j.cell.2013.06.005; pmid: 23827675

PALEOCEANOGRAPHY

Onset of Mediterranean outflow into the North Atlantic

F. Javier Hernández-Molina,[1]* Dorrik A. V. Stow,[2] Carlos A. Alvarez-Zarikian,[3]
Gary Acton,[4] André Bahr,[5] Barbara Balestra,[6] Emmanuelle Ducassou,[7] Roger Flood,[8]
José-Abel Flores,[9] Satoshi Furota,[10] Patrick Grunert,[11] David Hodell,[12]
Francisco Jimenez-Espejo,[13] Jin Kyoung Kim,[14] Lawrence Krissek,[15]
Junichiro Kuroda,[16] Baohua Li,[17] Estefania Llave,[18] Johanna Lofi,[19] Lucas Lourens,[20]
Madeline Miller,[21] Futoshi Nanayama,[22] Naohisa Nishida,[22] Carl Richter,[23]
Cristina Roque,[24] Hélder Pereira,[25] Maria Fernanda Sanchez Goñi,[26]
Francisco J. Sierro,[9] Arun Deo Singh,[27] Craig Sloss,[28] Yasuhiro Takashimizu,[29]
Alexandrina Tzanova,[30] Antje Voelker,[24] Trevor Williams,[31] Chuang Xuan[32]

Sediments cored along the southwestern Iberian margin during Integrated Ocean Drilling Program Expedition 339 provide constraints on Mediterranean Outflow Water (MOW) circulation patterns from the Pliocene epoch to the present day. After the Strait of Gibraltar opened (5.33 million years ago), a limited volume of MOW entered the Atlantic. Depositional hiatuses indicate erosion by bottom currents related to higher volumes of MOW circulating into the North Atlantic, beginning in the late Pliocene. The hiatuses coincide with regional tectonic events and changes in global thermohaline circulation (THC). This suggests that MOW influenced Atlantic Meridional Overturning Circulation (AMOC), THC, and climatic shifts by contributing a component of warm, saline water to northern latitudes while in turn being influenced by plate tectonics.

Tectonic activity exerts primary control on sedimentation along continental margins and on seafloor morphology, whereby it also influences bottom current flow pathways. Sea level variation and climate also influence sedimentation along continental margins. Plate tectonic configuration plays an important role in climate (1), although uncertainty persists concerning the specific cause-effect relationships and their relative importance on a temporal scale. From the late Miocene to the Pleistocene, Earth experienced major plate tectonic, climatic, and oceanographic shifts that, along with orbital variation, contributed to present-day global climate dynamics (2). These events include the closure and reopening of gateways between the Atlantic and Mediterranean basins [~5.55 to 5.33 million years ago (Ma) (3)], the closure of the deep Central American Seaway (CAS) [4.5 to 3.0 Ma (4)] and the Indonesian Seaway [~4 Ma (2)], and the onset of Northern Hemisphere glacial (NHG) cycles in the late Pliocene (5). Neogene sedimentary sections also record Pliocene warming trends that culminate in the mid-Pliocene warm period [3.3 to 3.0 Ma (6)], the Panama seaway closure [~3.0 to 2.7 Ma (7)], and NHG intensification around 2.72 Ma [Marine Isotope Stage (MIS) G6 (5)]. During this period, increases in evaporation and precipitation facilitated ice sheet formation, higher albedo, and increased Atlantic Meridional Overturning Circulation (AMOC) (4, 8). The global transition to Quaternary glacial cycles at 2.7 Ma (5, 8), a predicted major glaciation at 2.15 Ma (9), and the recent mid-Pleistocene transition at ~0.9 to 0.7 Ma (10) also relate to the establishment and stabilization of present-day ocean circulation patterns.

Changes in the Mediterranean Outflow Water (MOW) co-occurred with some of these shifts in global ocean circulation and climate, but the exact timing of MOW evolution with respect to major climate events remains unclear. The Strait of Gibraltar strongly influences Mediterranean-Atlantic water mass exchange, contributing warm and highly saline MOW to the Atlantic Ocean. MOW currently enters the Strait of Gibraltar with a density of 29.12 kg/m^3 (8); its density is 27.7 kg/m^3 in the Gulf of Cadiz and off Brittany (11, 12) and 27.4 kg/m^3 further afield along the Rockall Plateau of the northeastern Atlantic extension (12). MOW input thus enhances North Atlantic density and helps drive thermohaline convection. Estimates suggest that without MOW, the AMOC would decline by ~15% and North Atlantic sea surface temperatures would fall by up to 1°C (8). The Gulf of Cadiz and the western margin of Portugal record critical aspects of MOW development as well as the interaction between MOW and AMOC. Upon exiting the Strait of Gibraltar, MOW consists of relatively warm and highly saline water (Fig. 1) that settles into an intermediate bottom current within the mid-slope region, between 400 and 1400 m water depth, at an overflow rate of 0.67 ± 0.28 Sv (11). This region lies beneath Atlantic Water (AW) inflow but rafts above North Atlantic Deep Water (NADW).

This paper interprets the sequence of events that established an important MOW contribution to North Atlantic thermohaline dynamics, and makes inferences concerning how these dynamics relate to Neogene and Quaternary climatic and tectonic events. Our study combines geophysical and drill core data acquired along the southwestern Iberian margin during Integrated Ocean Drilling Program (IODP) Expedition 339 aboard the research vessel JOIDES Resolution (Fig. 1).

The southwestern Iberian margin is located near the Azores-Gibraltar Fracture Zone, which is a section of the convergent plate boundary between Eurasia (Iberia subplate) and Africa (Nubia subplate). The plates currently converge at a rate of ~4 to 5 mm/year in a WNW-ESE direction. Counterclockwise rotation is accommodated by a series of thrusts and dextral strike-slip faults (13, 14) active since at least 1.8 Ma. Starting in the late Miocene, the southwestern Iberian margin evolved by oblique convergence between the Iberia and Nubia subplates (13) and westward rollback subduction of an oceanic lithosphere slab beneath the Gibraltar Arc, both of which facilitate development of the Cadiz Allochthonous Unit (CAU) (15) or accretionary

[1]Department of Earth Sciences, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. [2]Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh EH14 4AS, Scotland, UK. [3]International Ocean Discovery Program, Department of Oceanography, Texas A&M University, College Station, TX 77845, USA. [4]Department of Geography and Geology, Sam Houston State University, Huntsville, TX 77341, USA. [5]Institute of Geosciences, University of Frankfurt, 60438 Frankfurt, Germany. [6]Institute of Marine Sciences, University of California, Santa Cruz, CA 95064, USA. [7]EPOC, Université de Bordeaux, 33615 Pessac Cedex, France. [8]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY 11794, USA. [9]Departamento de Geología, Universidad de Salamanca, 3008 Salamanca, Spain. [10]Department of Natural History Sciences, Hokkaido University, N10W8, Kita-ku, Sapporo 060-0810, Japan. [11]Institute for Earth Sciences, University of Graz, A-8010 Graz, Austria. [12]Godwin Laboratory for Palaeoclimate Research, University of Cambridge, Cambridge CB2 3EQ, UK. [13]Department of Biogeochemistry, JAMSTEC, 237-0061 Yokosuka, Japan. [14]Korea Institute of Ocean Science & Technology, Ansan 426-744, Korea. [15]School of Earth Sciences, Ohio State University, Columbus, OH 43210, USA. [16]Institute for Frontier Research on Earth Evolution (IFREE), JAMSTEC, 2-15 Natsushima-Cho, Yokosuka-city, Kanagawa 237-0061, Japan. [17]Department of Micropalaeontology, Nanjing Institute of Geology and Palaeontology, Chinese Academy of Sciences, Nanjing 210008, P.R. China. [18]Instituto Geológico y Minero de España, Ríos Rosas 23, 28003 Madrid, Spain. [19]Géosciences Montpellier, Université Montpellier II, 34090 Montpellier, France, and Department of Geology, University of Leicester, Leicester LE1 7RH, UK. [20]Institute of Earth Sciences, Utrecht University, 3584 CD Utrecht, Netherlands. [21]Department of Mechanical Engineering, California Institute of Technology, Pasadena, CA 91125, USA. [22]Institute of Geology and Geoinformation, Geological Survey of Japan (AIST), Ibaraki 305-8567, Japan. [23]School of Geosciences, University of Louisiana, Lafayette, LA 70504, USA. [24]Divisão de Geologia e Georecursos Marinhos, IPMA, 1749-077 Lisboa, Portugal. [25]Grupo de Biologia e Geologia, Escola Secundária de Loulé, 8100-740 Loulé, Portugal. [26]École Pratique des Hautes Etudes, EPOC, Université de Bordeaux, 33615 Pessac, France. [27]Department of Geology, Banaras Hindu University, Varanasi 221005, India. [28]School of Earth, Environmental and Biological Sciences, Queensland University of Technology, Brisbane, Queensland 4001, Australia. [29]Department of Geology, Faculty of Education, Niigata University, Niigata 950-2181, Japan. [30]Department of Geological Sciences, Brown University, Providence, RI 02912, USA. [31]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA. [32]Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Waterfront Campus, Southampton SO14 3ZH, UK.
*Corresponding author. E-mail: javier.hernandez-molina@rhul.ac.uk

wedge (16). Distinct periods of crustal deformation, fault reactivation, and halokinesis related to plate motion have contributed to the tectonostratigraphic evolution of the area (13). Sediments within the Gulf of Cadiz record the opening of the Strait of Gibraltar, and event that initiated MOW intrusion into the North Atlantic. Subsequent intensification of MOW established along-slope depositional systems and generated an extensive contourite depositional system (CDS) (17) that provides a sedimentary record of the paleoceanographic events analyzed and interpreted here.

Seismic records and drill cores as tracers of MOW dynamics from the late Miocene to the present

Major regional discontinuities appear as high-amplitude seismic reflections within late Miocene to present-day sediments around the Gulf of Cadiz (Fig. 2 and fig. S1). These discontinuities provide a record of MOW circulation relative to coeval tectonic and climatic events. Pliocene deposits appear in seismic profiles as sheeted drifts

overlying a weakly reflecting Miocene unit that progrades downslope (Fig. 2 and fig. S1). The late Pliocene to lower Quaternary section provides a record of substantial synsedimentary deformation associated with two discontinuities that define a major truncation surfaces (fig. S2). Quaternary deposits are distinguished by high-amplitude seismic reflections and show clear upslope progradation.

Closure of the Strait of Gibraltar and the Betic-Rharb corridors during the Messinian Salinity Crisis (5.56 to 5.33 Ma) prevented Mediterranean-Atlantic water exchange (3). Exchange resembling present-day flows through these channels is thought to have begun with the opening of the Strait of Gibraltar and the related onset of the latest Miocene flooding at ~5.33 Ma (3). Widespread discontinuities evident in seismic records and pervasive hiatuses evident in Expedition 339 drill cores suggest four major phases of MOW evolution after the opening of the Strait of Gibraltar (Fig. 2, Fig. 3, and fig. S1).

The primary phase of downslope activity occurred during the early Pliocene (~5.2 to 4.5 Ma)

as downslope mass transport deposits (slumps and debrites) and turbidites developed in response to widespread tectonic activity. These deposits overlie a regional discontinuity (Fig. 2 and fig. S1) above the late Miocene hemipelagic deposits (Fig. 4 and fig. S3). Drill core results suggest only a limited degree of contourite deposition beginning in the Pliocene. The opening of the Strait of Gibraltar caused slope instability (3), leading to gravitational-collapse breccia deposits (18).

The second phase began with the onset of limited MOW circulation between ~4.5 and 3.5 Ma. Persistent sequences of contourite facies from 4.5 to 4.2 Ma show clear evidence of initial MOW circulation into the Gulf of Cadiz and subsequent interaction between MOW and the North Atlantic.

The third phase is defined by enhanced MOW circulation from 3.2 to 2.1 Ma. Two major hiatuses, at 3.2 to 3.0 Ma and at 2.4 to 2.1 Ma, indicate erosion by bottom currents and enhanced interaction between MOW and the North Atlantic. Major and uninterrupted contourite deposition occurs
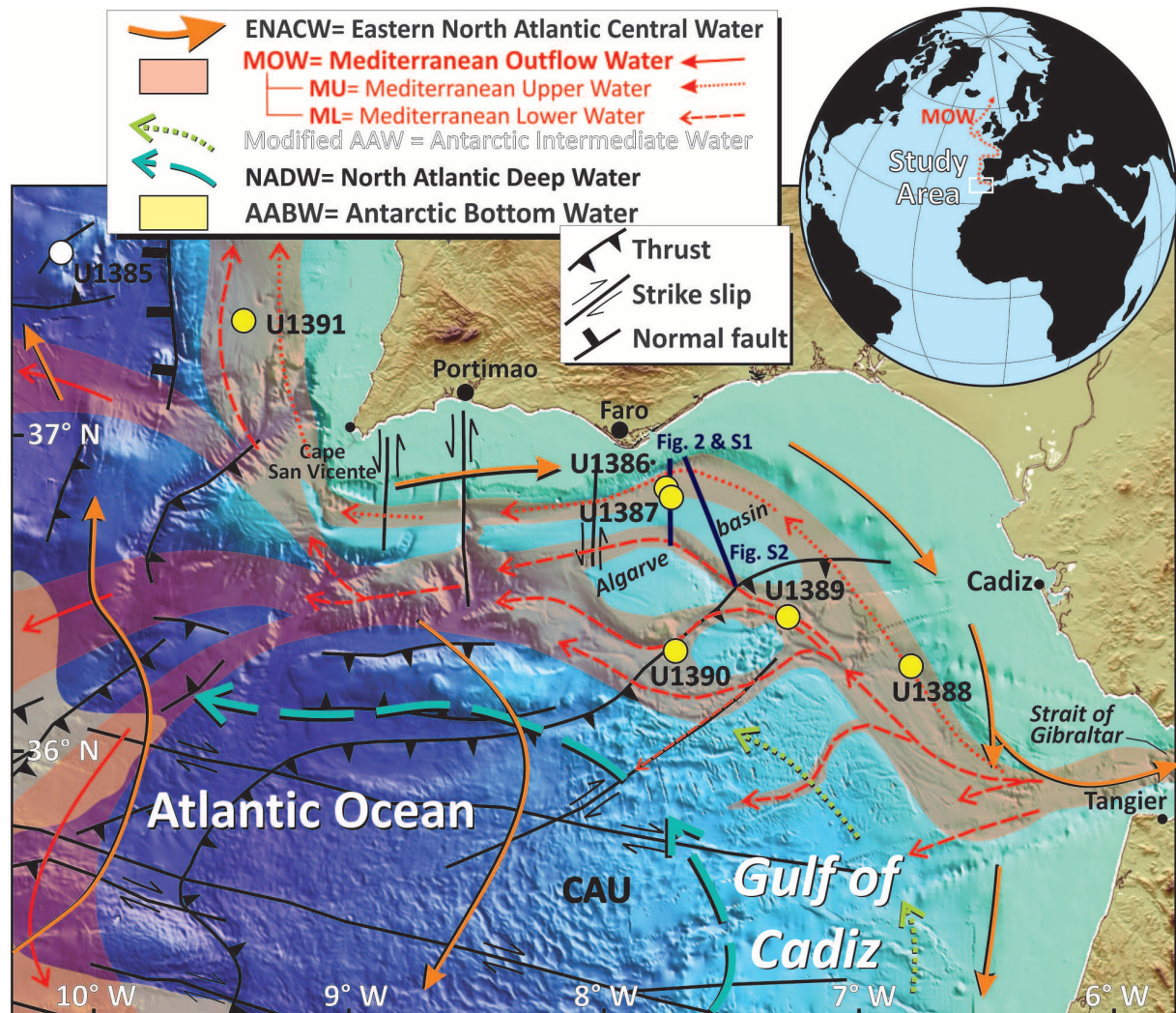

Fig. 1. Regional water masses, major tectonic features, and Gulf of Cadiz CDS site locations sampled during IODP Expedition 339.

only above the upper Pliocene hiatus at 3.2 to 3.0 Ma, where dolostones and debrite deposits occur (Figs. 2 and 4). Interbedded turbidites and contourites are also common in the upper Pliocene section. Dolostone deposits are closely linked to the aforementioned hiatus, wherein they may have precipitated as a result of shallow diagenetic processes (Fig. 2). That hiatus is the first major event evident within the sedimentary stacking pattern throughout the northern Gulf of Cadiz slope (*19*) and within the Strait of Gibraltar (*18*). Quaternary successions show a much more pronounced phase of contourite deposition and drift development. These deposits consist of sandy and muddy contourites with periodic turbidite intercalations. They also show a considerable phase of downslope sedimentation at ~2 Ma. A regional

hiatus, which occurs within the Quaternary section at ~2.4 to 2.1 Ma, is also associated with dolostones and debrite deposits (Figs. 2 and 4). This hiatus appears as an important discontinuity within drift near the base of the Quaternary. It records a period of tectonic instability coeval with erosion by bottom currents and substantially higher volumes of MOW input into the North Atlantic relative to flow onset at 4.5 to 4.2 Ma.

The final phase, beginning with the establishment of present-day circulation at 2.1 Ma and continuing to this day, is evident from increased sedimentation rates and rapid contourite development. A younger major regional discontinuity with local erosion is clearly evident in the seismic records (Fig. 2 and fig. S1) and is attributed to the mid-Pleistocene transition (*17, 20*), with a tentative age

of 0.9 Ma (*17*). A pronounced change in the drift sedimentary stacking pattern appears at this time horizon as enhanced upslope progradation and mounded morphology, indicating an intensification of MOW (*17, 20*). A specific hiatus or condensed section, however, is not found in all of the drill core material at the 0.9 to 0.7 Ma time horizon (Fig. 3). Ongoing analysis seeks to further constrain the exact age and duration of this discontinuity.

Hiatuses, mass transport deposits (debrites, slumps), and tectonic events occurring at the end of the Miocene (*3*), 3.2 to 3.0 Ma (*19, 21*), 2.4 to 2.0 Ma (*13*), or ~0.9 Ma (*17*) around the Gulf of Cadiz appear to coincide with major climatic shifts during this period. Tectonic activity has been a key factor controlling seafloor morphology,
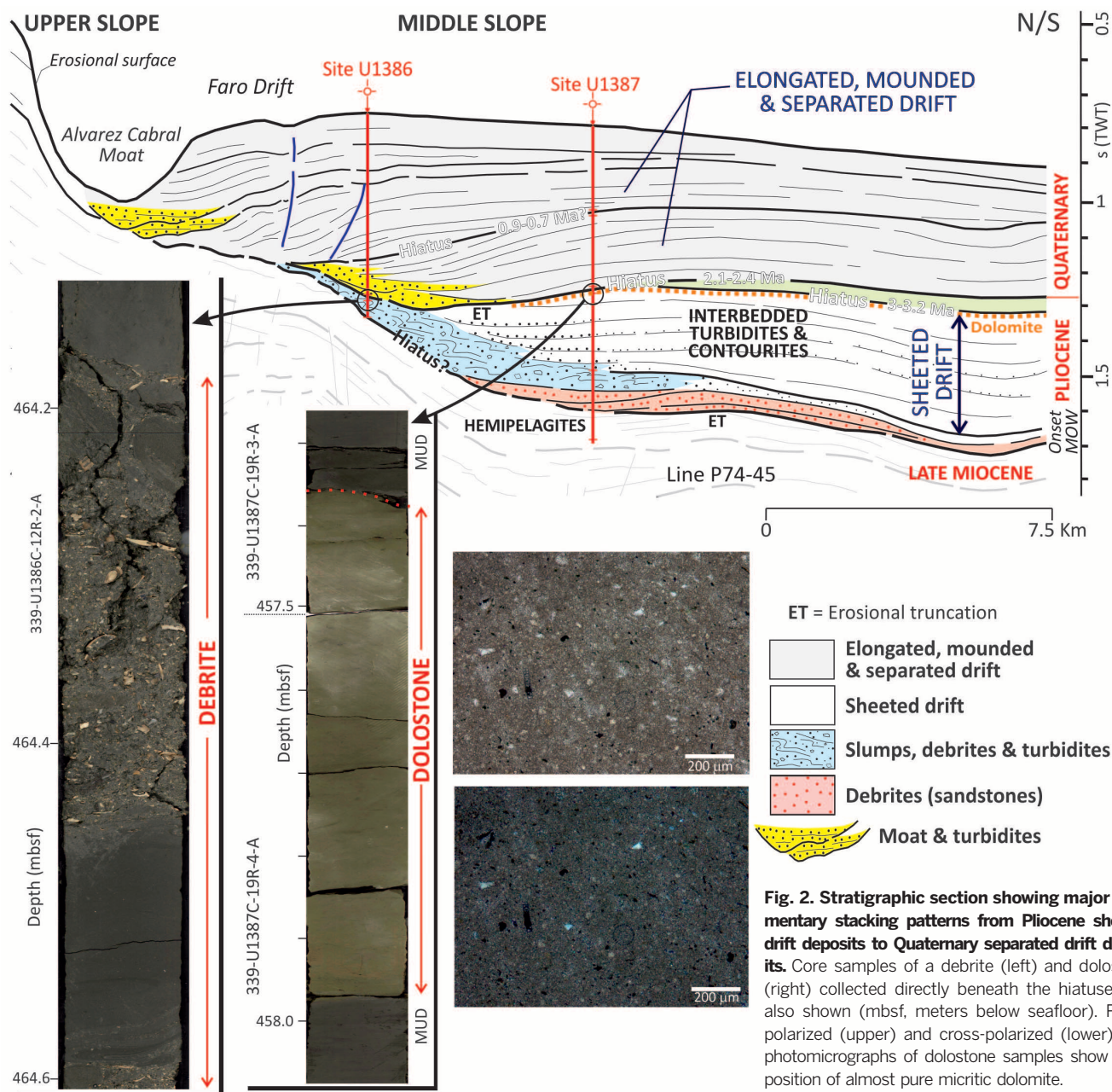

Fig. 2. Stratigraphic section showing major sedimentary stacking patterns from Pliocene sheeted drift deposits to Quaternary separated drift deposits. Core samples of a debrite (left) and dolostone (right) collected directly beneath the hiatuses are also shown (mbsf, meters below seafloor). Plane-polarized (upper) and cross-polarized (lower) light photomicrographs of dolostone samples show composition of almost pure micritic dolomite.

which in turn influences MOW flow pathways and, by extension, deep-water sedimentation. The stratigraphic events thus reflect changes in the Iberia-Nubia subplate kinematics during the Pliocene and Quaternary, which established the present-day transpressional regime. Deformation of the CAU, which initially emplaced in the Algarve basin (fig. S2) in the late Miocene, occurred with reactivation of blind thrusts during the Pliocene and Quaternary. The CAU could therefore record tectonic episodes associated with plate reorganization. Shifts in the Africa-Eurasia plate trajectory after the Messinian Salinity Crisis contributed to the deepening of the gateway (13). The NW-SE to WNW-ESE shift in convergence direction (13, 22) coincides with tectonic events between 3.2 and 2.0 Ma. A hiatus previously described from IODP Expedition 339 core records reflects these events, as do other hiatuses identified near the Strait of Gibraltar (23), in the Alboran Sea (24), and in the eastern Atlantic between Gibraltar and the equator (25). The shift in plate convergence direction for

the southwestern Iberian margin coincides with a decline in westward migration and thrusting activity for the accretionary wedge. Reactivated blind thrusts accommodated shortening and minor sedimentary accretion (16). At ~1.8 Ma, the oblique convergence between the Iberia and Nubia subplates reactivates WNW-ESE dextral strike-slip faults as well as shallower WNW-directed thrusts (14).

Global implications

The spatial distribution of these hiatuses can help to refine our understanding of MOW development and its overall effect on thermohaline circulation (THC) and deep-water sedimentation. The interpreted sequence of events is summarized in Fig. 5.

Lower Pliocene sedimentary deposits show no clear evidence of MOW input into the Atlantic prior to 4.5 Ma. This time constraint indicates a lag between the full opening of the Strait of Gibraltar (~5.33 Ma) and the establishment of

consistent MOW interaction with the North Atlantic (fig. S4). Similarities in the densities of Mediterranean seawater and Eastern North Atlantic Central Water in the beginning of the early Pliocene (8) may have contributed to this lag. By the early Pliocene, Antarctic Bottom Water (AABW) flowed northward and extended into the North Atlantic with greater intensity. This intensification coincided with a weaker AMOC, especially during glacial stages (25).

At ~4.5 Ma, MOW began to circulate into the North Atlantic as an intermediate and deep-water circulation (6). Some authors have reported an additional MOW intensification stage at ~3.8 Ma (23) and 3.5 Ma (12), which coincides with more active circulation in the North Atlantic (12). This second stage demonstrates increasing MOW influence on the North Atlantic and a sedimentary signal shared by MOW- and THC-related phenomena. Notably, the CAS (fig. S4) shoaled by ~4.5 to 3.5 Ma and began to hinder intermediate-layer circulation. These events blocked NADW flow into the Pacific but enhanced NADW flow into the South Atlantic, initiating the modern-day deep-water circulation pattern (6). The modern-day pattern delivers warmer, saltier surface water to more northerly locations in the Arctic Ocean, enhancing evaporation, precipitation, and ice sheet formation. These factors established the initial conditions for the mid-Pliocene warm period (4, 8). Around this same time (~4.5 Ma), Labrador Sea Water (LSW) began to circulate as part of the North East Atlantic Deep Water (NEADW) layer (7). This led to increased AMOC circulation between 4.5 and 4.0 Ma, as well as increased drift sedimentation in the North Atlantic (2).

Two phases of MOW intensification occurred, one at 3.2 to 3.0 Ma and the other at 2.4 to 2.1 Ma. The timing of these events correlates with higher-amplitude phases of temperature and insolation oscillations related to the 400,000-year, eccentricity-modulated cycle (6), suggesting a relationship between orbital changes and MOW production.

The first enhanced MOW period began after 3.2 Ma and coincides with widespread evidence of NEADW circulation. This event contributed to the formation of the present-day NADW (4) through southward deflection of the AABW-NADW contact (south of the Azores). NADW arrival at extreme southerly latitudes around Antarctica in turn cooled the Antarctic water and increased production of AABW, augmenting its flow into the southwest Atlantic (26). A coeval decrease in sedimentary accumulation rates in North Atlantic drift deposits accompanied these events (2). Model experiments indicate that NADW production increased around the time of the final CAS closure at 3 Ma (6). However, the increased strength of the AMOC also suggests additional saline and warm water transport to high latitudes in the North Atlantic. MOW could have supplied this additional saline water component of the AMOC. The present study shows that MOW intensification between 3.2 and 3.0 Ma contributed salt water at intermediate depths to northerly latitudes, thus enhancing AMOC and overall Northern Hemisphere deep-water formation and
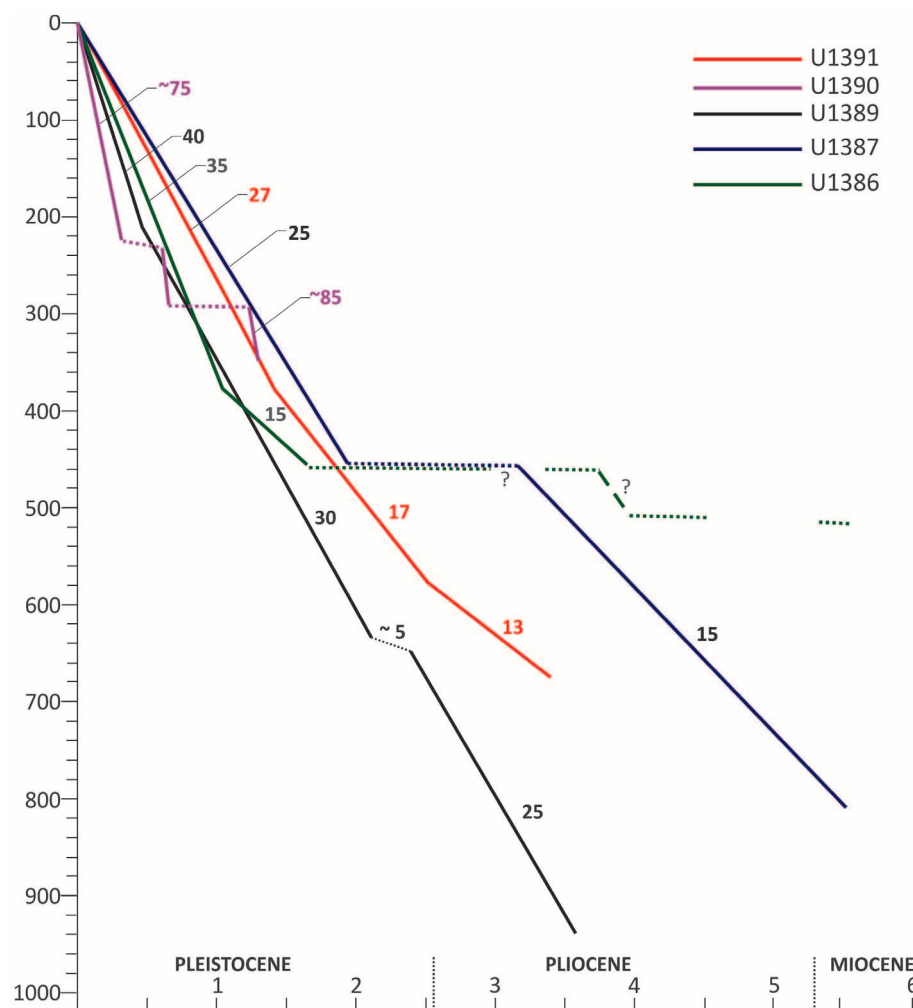

Fig. 3. Age curves and sedimentation rates for sites sampled during IODP Expedition 339. Ages are based on biostratigraphic and magnetostratigraphic dating. These methods identify depositional hiatuses and boundaries between stratigraphic periods and enable the estimation of sedimentary accumulation rates.

reducing pole-to-equator temperature gradients during the mid-Pliocene warm period. Prior to 3.4 Ma, the density of MOW slightly exceeded the present-day value of 27.8 kg/m³ (12). Density increased about 1 kg/m³, maximizing around 3.3 to 3.2 Ma and persisting at this level until ~3.0 Ma (12). The temporal boundary of this density event coincides with the age of the hiatus identified in the Gulf of Cadiz. MOW intensification therefore also contributed salt water at intermediate depths to northerly latitudes, thus enhancing THC, AMOC, and overall Northern Hemisphere deep-water formation.

The second enhanced MOW period (2.4 to 2.1 Ma) coincided with a sizable stratigraphic gap reported for the northeast Atlantic (24), suggesting synchronous reactivation of the THC and AMOC. This new event would suggest an additional increase in MOW density, which is surprisingly coeval with apparent major glaciation at 2.15 Ma (9).

Present-day MOW dynamics (fig. S4) commenced in the earliest Quaternary and coincided with a shift from long-term global cooling trends (27) marked by the final NHG intensification (4). The shift included a decline in atmospheric CO_2 levels, global cooling (27), and a fall in sea level (28). The 41,000-year obliquity cycle replaced the 23,000-year precessional cycle as the primary orbital mechanism influencing climate. This shift is also associated with a progressive increase in the amplitude of Earth's orbital obliquity (4).

While cooler deep waters formed with the subsequent increase in deep ocean stratification, deep-water circulation (and sedimentation), upwelling, and biogenic productivity at low latitudes followed a somewhat asymmetric temporal pattern (29). The present-day NADW configuration deflects the AABW-NADW contact southward during interglacial periods, whereas AABW deflected this contact back in a northward direction during glacial periods (2, 29). These dynamics were reinforced by the mid-Pleistocene revolution or transition (0.9 to 0.7 Ma), which established longer periods for glacial/interglacial cycles. This transition coincided with an increase in the 100,000-year cycle amplitude and a major eustatic drop in sea level (8, 28), which might relate to the MOW intensification and the formation of the mid-Pleistocene discontinuity (17, 20). This increase in MOW coincides with intensification in the AMOC (2). After that transition and subsequent glacial stages, MOW density increased, as it did during the last glaciation when it exceeded (~31.29 kg/m³) the present-day values (30). Rates of MOW flow were also higher (30), as evident from the morphology of the slope (23).

The aforementioned tectonic events in the Gulf of Cadiz are coeval with other events recorded throughout the Northern Hemisphere (31). The temporal overlap of these events suggests that they relate to overall plate reorganization in the North Atlantic (31) or to widespread mantle activity [e.g., (32)]. More general hypotheses (33) affirm this relationship, further corroborating the role of plate tectonics in modulating climate over a wide range
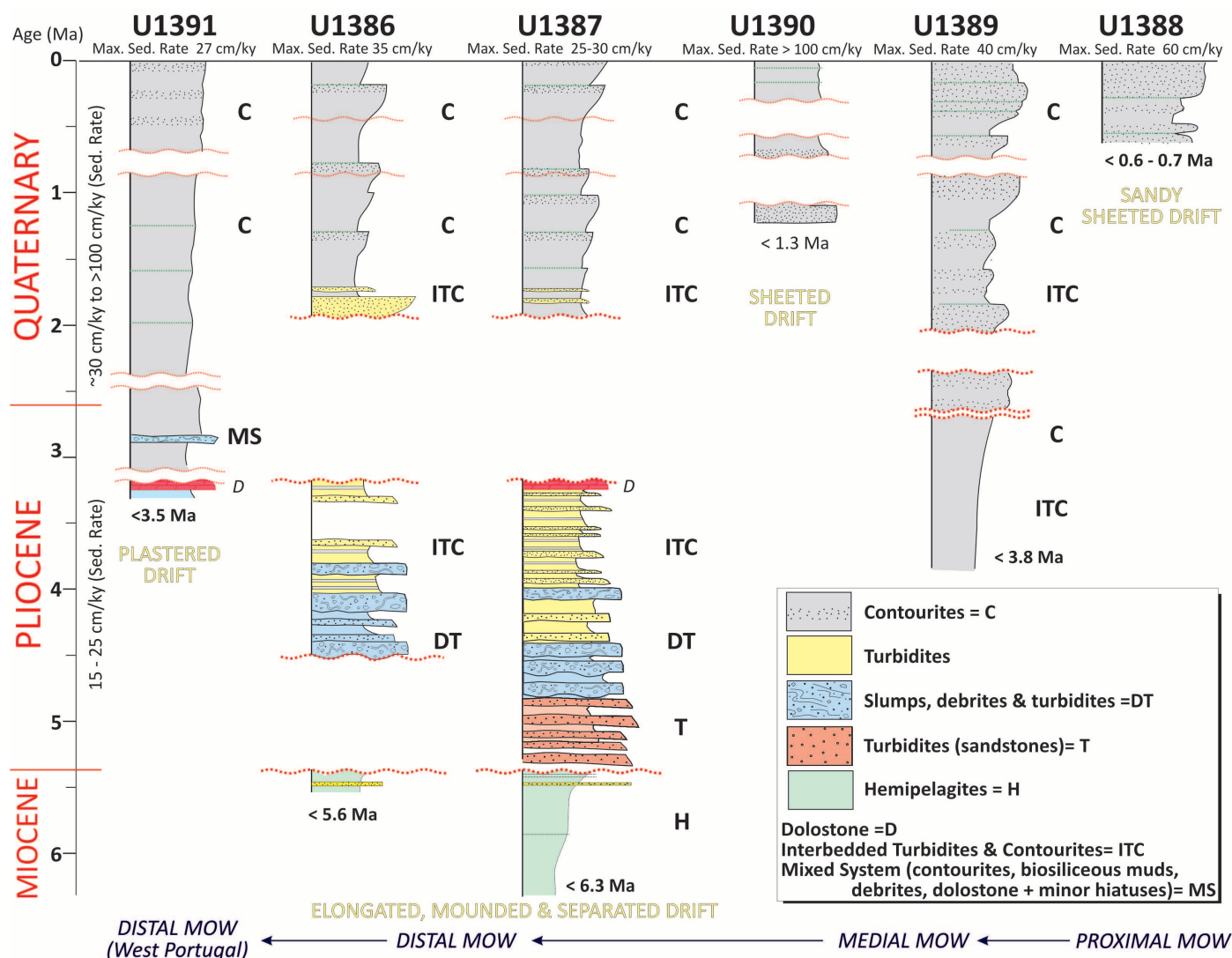

Fig. 4. Lithologic summary of IODP Expedition 339 CDS sample sites, showing major hiatuses. Age model is based on biostratigraphic and magnetostratigraphic data (ky = 1000 years).

Fig. 5. Neogene and Quaternary evolution of the CDS within the Gulf of Cadiz and the western Portuguese margin. Regional evolution is compared with major events affecting the North Atlantic basin and shifts in global climate (*2, 27, 28, 34*).
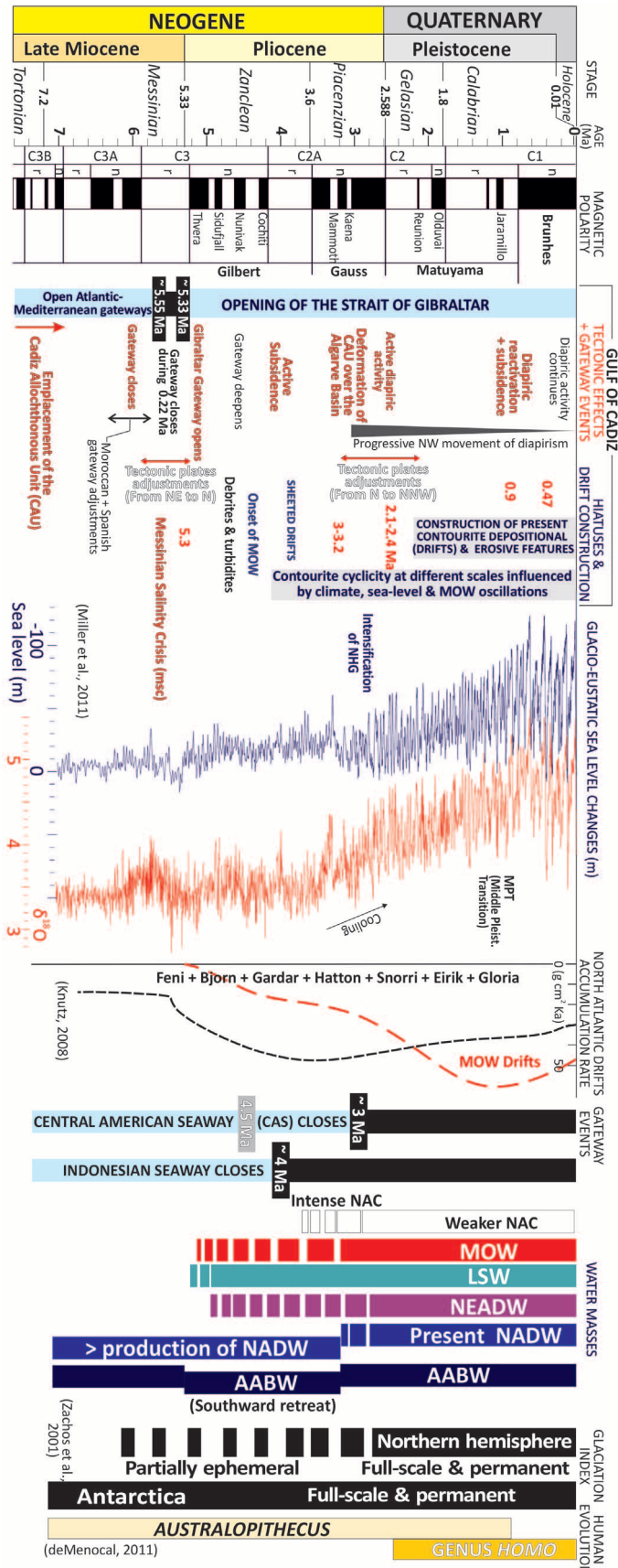
of time scales (*1*). These should be reconsidered in climatic and sea level reconstruction because, for example, they could explain the discrepancy between the benthic $\delta^{18}O$ record and the Gibraltar relative sea level that assumes no tectonic changes in the geometry of the Gibraltar Gateway and adjacent areas (*9*).

Conclusions

Our results show that the initial MOW circulation into the Atlantic after the opening of the Strait of Gibraltar was relatively weak. Meaningful interaction between MOW and the North Atlantic did not begin until the late Pliocene. The establishment of MOW added relatively salty water at intermediate depths and contributed to enhanced THC and AMOC. The addition of the warm, salty MOW component reduced pole-to-equator temperature gradients during the mid-Pliocene warm period (3.2 to 3.0 Ma), during the early Quaternary (2.4 to 2.0 Ma), and at 0.9 to 0.7 Ma. These climatic events coincide with widespread depositional hiatuses, pronounced changes in the sedimentary stacking pattern, and establishment of the present-day seafloor morphology. Hiatuses and shifts in depositional processes are related to regional tectonic events and margin instability. Similar changes in deep-water sedimentation and tectonics have been described in association with other margins and basins around the same time in both the Northern and Southern hemispheres, demonstrating that the relationship between climatic shifts and plate tectonic events operates over a wide range of time scales. The relationship between climate and plate tectonic events is especially relevant for this geographic locality because of its role as the site of major events in hominid migration and evolution (*34*).

REFERENCES AND NOTES

1. W. W. Hay, *Geol. Rundsch.* **85**, 409–437 (1996).
2. P. C. Knutz, Palaeoceanographic significance of contourite drifts. In *Contourites: Developments in Sedimentology, vol. 60*, M. Rebesco, M. A. Camerlenghi, Eds. (Elsevier, Amsterdam, 2008), pp. 511–535.
3. M. Roveri *et al.*, *Mar. Geol.* 10.1016/j.margeo.2014.02.002 (2014).
4. G. Bartoli *et al.*, *Earth Planet. Sci. Lett.* **237**, 33–44 (2005).
5. I. Bailey *et al.*, *Quat. Sci. Rev.* **75**, 181–194 (2013).
6. M. A. Maslin, X. S. Li, M.-F. Loutre, A. Berger, *Quat. Sci. Rev.* **17**, 411–426 (1998).
7. K. H. Nisancioglu, M. E. Raymo, P. H. Stone, *Paleoceanography* **18**, 1006 (2003).
8. M. Rogerson, E. J. Rohling, G. R. Bigg, J. Ramirez, *Rev. Geophys.* **50**, RG2003 (2012).
9. E. J. Rohling *et al.*, *Nature* **508**, 477–482 (2014).
10. H. Elderfield *et al.*, *Science* **337**, 704–709 (2012).
11. N. Serra, I. Ambar, D. Boutov, *Ocean Sci.* **6**, 191–209 (2010).
12. N. Khélifi *et al.*, *Geology* **37**, 811–814 (2009).
13. N. Zitellini *et al.*, *Earth Planet. Sci. Lett.* **280**, 13–50 (2009).
14. J. Duarte *et al.*, *Mar. Geol.* **289**, 135–149 (2011).
15. T. Medialdea *et al.*, *Mar. Geol.* **209**, 173–198 (2004).
16. M.-A. Gutscher *et al.*, *Geology* **30**, 1071–1074 (2002).
17. E. Llave *et al.*, in *Economic and Paleoceanographic Importance of Contourites*, A. Viana, M. Rebesco, Eds. (Geological Society of London Special Publication, 2007), vol. 276, pp. 49–79.
18. M. Esteras *et al.*, *Rev. Soc. Geol. Esp.* **13**, 539–550 (2000).

19. R. A. Brackenridge, F. J. Hernández-Molina, D. A. V. Stow, E. Llave, *Mar. Pet. Geol.* **46**, 36–50 (2013).
20. C. Roque et al., *Mar. Geol.* **303–306**, 42–62 (2012).
21. F. J. Flinch, P. R. Vail, in *Mesozoic and Cenozoic Sequence Stratigraphy of European Basins, SEPM Special Publication Series no. 60*, P. C. de Graciansky, J. Hardenbol, T. Jacquin, P. R. Vail, Eds. (Society for Sedimentary Geology, Tulsa, OK, 1998), pp. 199–208.
22. P. Martínez-García, M. Comas, J. I. Soto, L. Lonergan, A. B. Watts, *Basin Res.* **25**, 361–387 (2013).
23. F. J. Hernández-Molina et al., *Geology* **42**, 227–230 (2014).
24. K. Billups, A. C. Ravelo, J. C. Zachos, *Paleoceanography* **13**, 84–95 (1998).
25. R. Stein et al., in *North Atlantic Paleoceanography*, C. P. Summerhayes, N. J. Shackleton, Eds. (Geological Society, London, 1986), pp. 103–118.
26. M. T. Ledbetter, D. F. Williams, B. B. Ellwood, *Nature* **272**, 237–239 (1978).
27. J. Zachos, M. Pagani, L. Sloan, E. Thomas, K. Billups, *Science* **292**, 686–693 (2001).
28. K. G. Miller, G. Mountain, J. Wright, J. Browning, *Oceanography* **24**, 40–53 (2011).
29. G. S. Dwyer et al., *Science* **270**, 1347–1351 (1995).
30. M. Rogerson, G. R. Bigg, E. J. Rohling, J. Ramirez, *Clim. Dyn.* **39**, 589–598 (2012).
31. S. Cloetingh, F. M. Gradstein, H. Kooi, A. C. Grant, M. Kaminski, *J. Geol. Soc. London* **147**, 495–506 (1990).
32. N. White, B. Lovell, *Nature* **387**, 888–891 (1997).
33. P. E. Potter, P. Szatmari, *Earth Sci. Rev.* **96**, 279–295 (2009).
34. P. B. deMenocal, *Science* **331**, 540–542 (2011).

NEURONAL REPAIR

Asynchronous therapy restores motor control by rewiring of the rat corticospinal tract after stroke

A. S. Wahl,[1,2]* W. Omlor,[2] J. C. Rubio,[3] J. L. Chen,[2] H. Zheng,[3] A. Schröter,[4] M. Gullo,[1,2] O. Weinmann,[1,2] K. Kobayashi,[5] F. Helmchen,[2] B. Ommer,[3] M. E. Schwab[1,2]*

The brain exhibits limited capacity for spontaneous restoration of lost motor functions after stroke. Rehabilitation is the prevailing clinical approach to augment functional recovery, but the scientific basis is poorly understood. Here, we show nearly full recovery of skilled forelimb functions in rats with large strokes when a growth-promoting immunotherapy against a neurite growth–inhibitory protein was applied to boost the sprouting of new fibers, before stabilizing the newly formed circuits by intensive training. In contrast, early high-intensity training during the growth phase destroyed the effect and led to aberrant fiber patterns. Pharmacogenetic experiments identified a subset of corticospinal fibers originating in the intact half of the forebrain, side-switching in the spinal cord to newly innervate the impaired limb and restore skilled motor function.

Stroke is a major cause of severe disability in the elderly population, and recovery after large strokes is limited (*1*, *2*). Current strategies to improve long-term outcome in humans include mostly rehabilitative training and, in experimental models, electrical stimulation and pharmacological interventions (*3*). However, all of these treatment options have had only limited success thus far (*4*, *5*). Here, we show that rehabilitative training, if preceded by a nerve growth–promoting antibody therapy, almost completely restored skilled forelimb functions after cortical strokes in adult rats. Sequential application of the treatments was essential: When growth promotion by blockade of the neurite growth–inhibitory protein Nogo-A was simultaneously applied with intensive forced-use training of the forelimb during the first 2 weeks after the stroke, functional outcome was poorer compared with training, immunotherapy alone, or no treatment at all (*6*). Anatomically, Nogo-A neutralization promoted growth of corticospinal fibers from the intact forebrain motor cortex across the midline of the cervical spinal cord. In rats with simultaneous antibody treatment and training, fiber branching was abundant with anatomically aberrant terminations. In contrast, in animals trained for forelimb function subsequent to antibody treatment, axonal fibers originally terminating in the intact spinal hemicord crossed the midline and innervated the ventral motor regions of the spinal hemicord that had lost its input from the motor cortex. To prove the functional relevance of these newly grown, "side-switched" descending corticospinal tract (CST) fibers, we selectively and temporarily blocked these fibers by two different pharmacogenetic techniques, both of which suppressed the restored forelimb function. Our results demonstrate that a sequential strategy of first promoting fiber growth to enhance the low endogenous plastic potential of the brain and spinal cord, followed by rehabilitative training–induced selection and stabilization of functionally meaningful connections can lead to much higher levels of functional restoration after the formation of large brain lesions than currently obtained in conventional rehabilitation medicine.

Success of rehabilitation depends on timing

We compared four different therapy and rehabilitation schedules for promoting functional recovery of fine motor skills of forelimbs in a thrombotic stroke model in rats. Using the well-established technique of photothrombosis (*6*), we induced blood vessel blockade by multiple microthrombi, which destroyed >90% of the sensory-motor cortex of adult rats. Rats were then treated with intrathecal anti-Nogo-A or control antibody for 2 weeks (*6*, *7*). In addition, we trained rats intensely in skilled forelimb reaching (100 reaches per day), either simultaneously with antibody application (parallel groups) or during the 2 weeks after antibody treatment (sequential groups) (Fig. 1A). To avoid a training effect of testing itself, we did not reassess the sequential groups (anti-Nogo-A/sequential and control sequential) during the first 2 weeks after lesion formation. When the growth-enhancing anti-Nogo-A treatment was followed by the rehabilitative training (anti-Nogo-A/sequential group), animals improved their performance from day 16 poststroke onwards, and their skilled reaching abilities almost completely recovered [reaching 86.3 ± 2.0% of prestroke level; significantly better than all other groups; $P < 0.001$, two-way repeated measures analysis of variance (ANOVA) with post hoc Bonferroni] (Fig. 1, A and B). This group also performed best in two skilled forelimb use tasks tested at the end of the experiment

[1]Department of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland. [2]Brain Research Institute, University of Zurich, Zurich, Switzerland. [3]Computer Vision Group, Heidelberg Collaboratory for Image Processing and Interdisciplinary Center for Scientific Computing (IWR), University of Heidelberg, Heidelberg, Germany. [4]Institute for Biomedical Engineering, ETH Zurich, Zurich, Switzerland. [5]National Institute for Physiological Sciences, National Institute of Natural Sciences Myodaiji, Okazaki, Japan.
*Corresponding author. E-mail: schwab@hifo.uzh.ch (M.E.S.); wahl@hifo.uzh.ch (A.S.W.)

[Montoya staircase grasping: success rate 34.1 ± 5.1% (Fig. 1C); horizontal ladder crossing: success rate 65.8 ± 3.7% (Fig. 1D)]. As animals had not been exposed to these tasks before, these results indicate a generalization of the recovery of forelimb function in the sequential training group that is transferable to nontrained motor skills.

In contrast to these results, rats receiving intensive forelimb training concurrently to the anti-Nogo-A antibody treatment (anti-Nogo-A/parallel group) performed worse than all other groups in the single-pellet grasping task (success rate 10.0 ± 5.2%) (Fig. 1, A and B). In the tasks evaluated at the end of the experiment, these animals showed either no significant improvement (Fig. 1C) or a tendency to decline over the course of trials ($P = 0.1$, two-way repeated measures ANOVA with post hoc Bonferroni) (Fig. 1D). The control antibody-treated groups reached low levels of recovery (35 to 40% success rate in pellet

grasping) (Fig. 1, A to D), with an early training effect visible in the group trained during the first 2 poststroke weeks (control/parallel) (Fig. 1A). Final success scores did not correlate with stroke volume, as determined ex vivo from either whole-brain magnetic resonance images ($r = 0.04$, Spearman correlation) (Fig. 1F) or histological Nissl stains (fig. S1), suggesting no specific neuroprotective effect by any of the four therapeutic schedules. We conclude that applying a nerve fiber growth-promoting cellular therapy and rehabilitative physical training in sequence delivers greater functional recovery than when the same protocols are applied concurrently.

Rehabilitative schedules induce distinct neuronal fiber patterns

We investigated the neuroanatomical correlates of functional recovery by labeling the intact,

contralesional CST, which normally innervates the spinal cord half opposite to the one that has lost its cortical input, with only few fibers crossing the spinal cord midline. Each of the four experimental groups presented a distinct rehabilitation-induced pattern of fiber sprouting in the cervical spinal cord (Fig. 2C). We counted labeled fibers originating in the cortex of the intact side opposite to the stroke and crossing the midline of the spinal cord (Fig. 2A). We also quantified their elongation and branching within the gray matter of cervical spinal cord, that is, the cord region containing the motor control circuits of the forelimb and paw (Fig. 2, A and B). The greatest number of midline-crossing fibers was seen in the anti-Nogo-A/sequential treatment group, which also had the best functional outcome. In contrast, the anti-Nogo-A/parallel group, showed extensive branching of the midline crossing corticospinal fibers
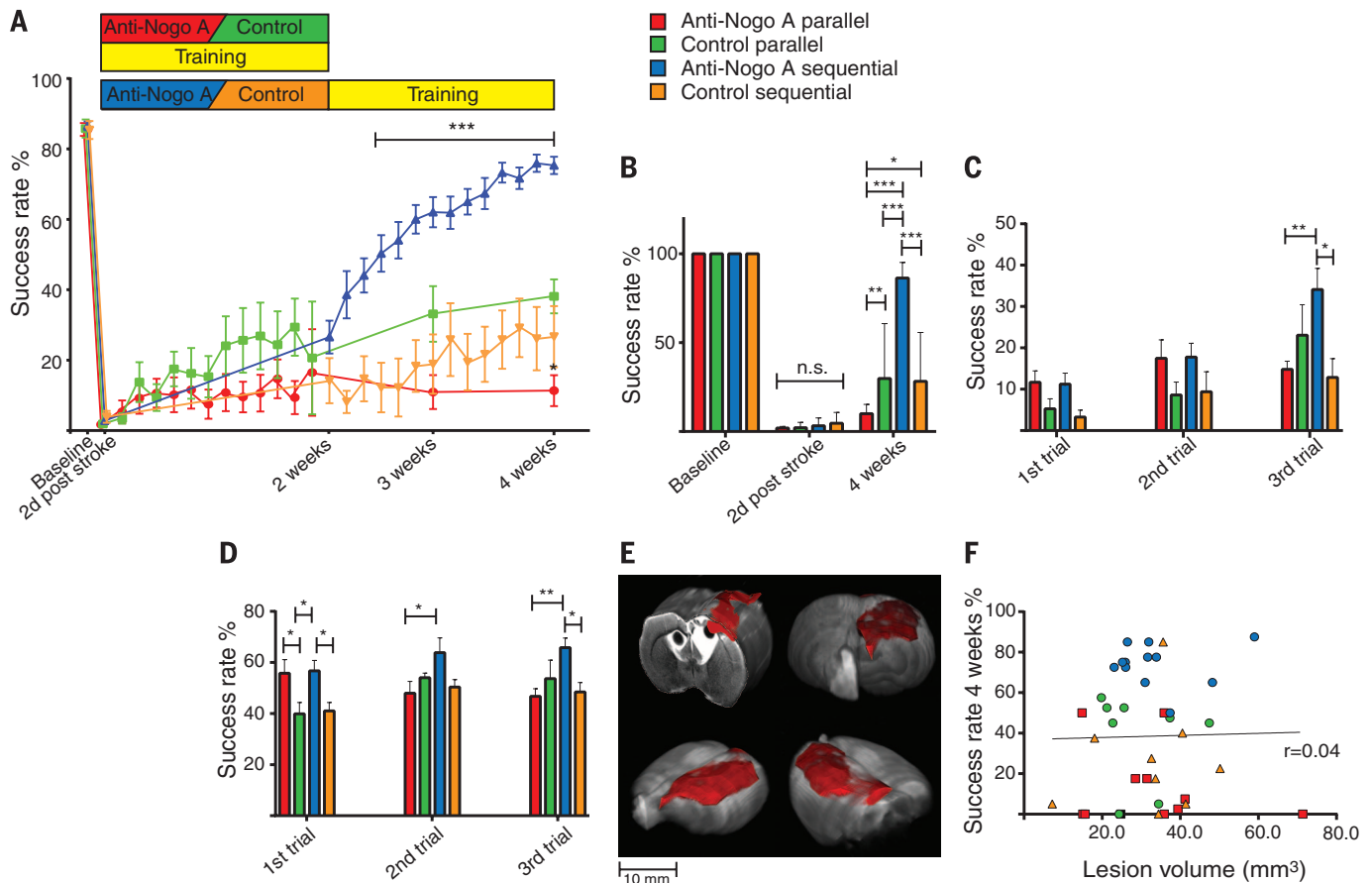

Fig. 1. Timing matters when a growth-promoting anti-Nogo-A immunotherapy is combined with training. (A) Success rates in the single-pellet grasping task at baseline (intact, trained), 2 days after a large, unilateral photothrombotic stroke to the sensorimotor cortex of the preferred paw, and during training and retesting sessions until 4 weeks post-insult. The anti-Nogo-A/sequential group showed significant improvement compared with all other groups, whereas the performance of the anti-Nogo-A/parallel group was significantly worse (anti-Nogo-A/parallel, n = 16; control/parallel, n = 8; anti-Nogo-A/sequential, n = 16; and control/sequential, n = 9). (B) Recovery rates expressed as success rates of the last testing session normalized to baseline performance (100%). n.s., not significant. (C and D) Animals in the anti-

Nogo-A/sequential group also performed significantly better in grasping tasks, such as the Montoya staircase test (C) or the horizontal ladder crossing task (D), introduced after the completion of the rehabilitation schedules for three consecutive trials. In (A) to (D), data are presented as means ± SEM. Statistical evaluation was carried out with two-way ANOVA repeated measure, followed by Bonferroni post hoc test. Asterisks indicate significances: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$. (E) Representative picture of an ex vivo magnetic resonance image stack for three-dimensional stroke reconstruction. (F) There was no correlation between stroke volume and endpoint success rate in the single-pellet grasping task among all rehabilitation groups ($r = 0.04$, Spearman correlation).

(Fig. 2B) ($P < 0.05$, two-way repeated measures ANOVA with post hoc Bonferroni).

A quantitative analysis of the distribution and density of ipsilaterally projecting corticospinal fibers using pattern-recognition algorithms to analyze both single corticospinal fibers and related fiber-growth parameters (see supplementary materials and methods) confirmed overshooting fiber growth and aberrant termination patterns in the anti-Nogo-A/parallel group. In the anti-Nogo-A/sequential group, midline-crossing, sprouting corticospinal fibers displayed a radial organization with few branches and a preference for the premotor and motor spinal cord (laminae 6 to 9) (Fig. 2, D and G). In contrast, fibers in the anti-Nogo-A/parallel group appeared less organized with more than double the number of branches and a different laminar distribution including the dorsal, predominantly sensory laminae 1 to 5. We also assessed the connectivity of the ipsilaterally projecting corticospinal fibers by quantifying the density of axonal boutons recognized morphologically in the premotor interneuron lamina 7: We detected a significantly higher bouton density in the anti-Nogo-A/parallel group compared with the anti-Nogo-A/sequential group (Fig. 2H) ($P < 0.05$, Student's t test, two-tailed, unpaired). The anti-Nogo-A/parallel group showed a greater tendency of axons to grow beyond the gray matter–white matter boundary, as well as a highly aberrant growth pattern (Fig. 2, H and I). In the medio-ventral funiculus, such fibers are probably intermixed with sprouts of the small, uncrossed ipsilateral CST.

Nerve cells from the intact forebrain cortex are responsible for recovery

Our results suggest that the recovery of rat forelimb function after stroke in the anti-Nogo-A/sequential group originates from extensive and precise reinnervation of the stroke-denervated spinal hemicord by midline-crossing fibers from the intact motor cortex and CST. We tested this hypothesis in the animals of the anti-Nogo-A/sequential group, all of which showed excellent functional recovery, by using two different experimental approaches for inducible, selective, and reversible inactivation of the ipsilaterally projecting corticospinal fibers on the long and short term, respectively. For long-term blockade, we used a virus to deliver a doxycyclin-inducible tetanus toxin to temporarily inactivate the synaptic release mechanism (8). We injected the highly efficient retrograde gene transfer lentivector HiRet carrying enhanced tetanus neurotoxin light chain (eTeNT) with an enhanced green fluorescent protein (EGFP) downstream of a tetracycline-responsive element (TRE) into the stroke-denervated side of the cervical spinal cord at level C5-C6, and we injected the adeno-associated serotype 2.2 (AAV2) vector carrying the reverse tetracycline transactivator (rtTAV16, Tet-on) into the contralesional, intact premotor and motor cortex ($n = 6$ animals) (Fig. 3, A and B). Only cortical neurons with axons projecting to the stroke-denervated spinal
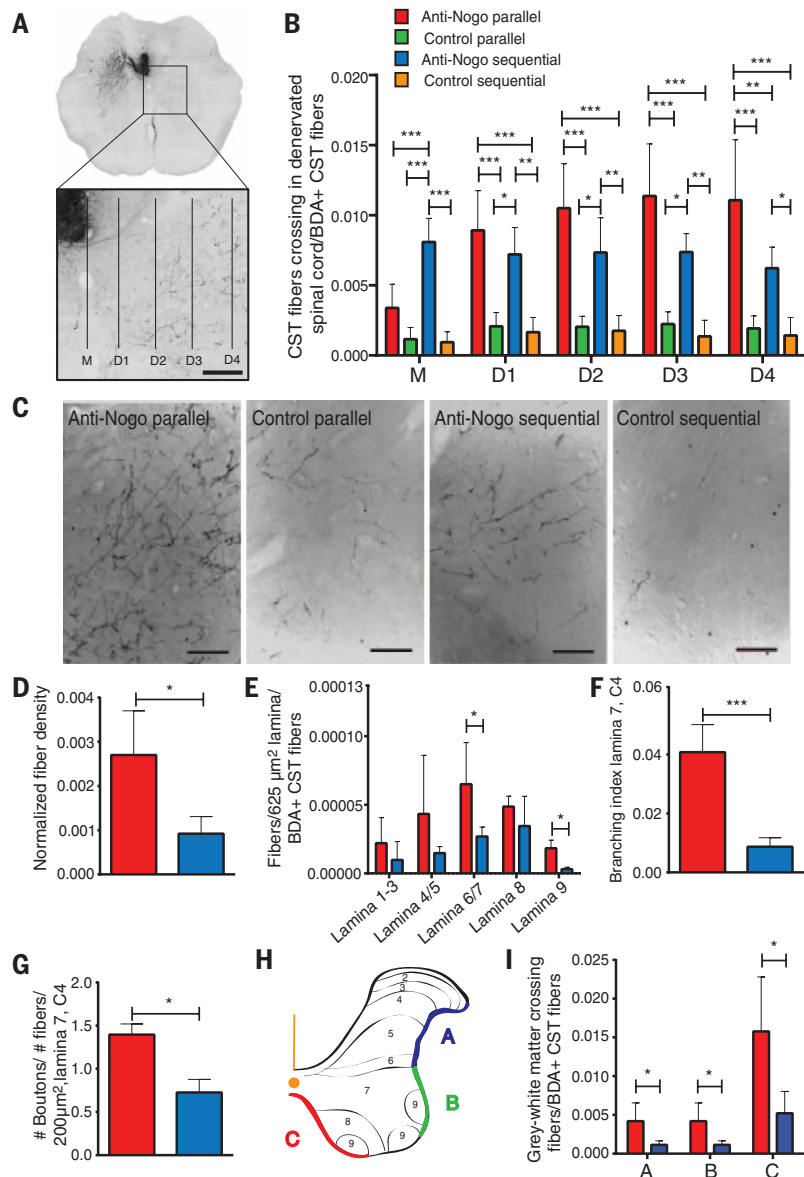

Fig. 2. Corticospinal tract sprouting depends on timing of rehabilitative training correlating with functional recovery. The four rehabilitation schedules (anti-Nogo-A/parallel, $n = 10$; control/parallel, $n = 8$; anti-Nogo-A/sequential, $n = 10$; and control/sequential, $n = 8$) differently influenced the sprouting of corticospinal fibers from the intact side of the spinal cord across the spinal cord midline (M). (**A**) Low- and high-magnification micrographs of biotinylated dextran amine (BDA)–labeled corticospinal fibers in intact spinal hemicord (left) growing into the stroke-denervated hemicord (right; inset) at spinal cord level C4. D1 to D4: Lines for intersection counts with corticospinal fibers. Scale bar, 200 μm. (**B**) Fibers crossing the midline (M) and branching in the gray matter at distances D1 to D4 were counted and normalized to the number of BDA-positive labeled fibers in the main tract. (**C**) Micrographs showing different sprouting patterns of corticospinal fibers from the ipsilateral cortex in the denervated cervical spinal cord (C4) in lamina 7 in the different treatment groups. Scale bars, 200 μm. (**D**) Combining anti-Nogo-A immunotherapy with simultaneous training (anti-Nogo-A/parallel) results in a significantly higher density of ipsilateral CST fibers in the stroke-denervated cervical spinal cord than does anti-Nogo-A/sequential treatment. (**E**) The most significant difference in fiber density between anti-Nogo-A/parallel and anti-Nogo-A/sequential animals was detected in lamina 6/7 and lamina 9 of the denervated cervical hemicord. Lamina 7 was also significant for increased fiber branching (**F**) (branching index = branches per fiber per BDA-positive fibers in the intact CST) and bouton numbers (**G**) in the anti-Nogo-A/parallel group. (**H** and **I**) Significantly more fibers cross the gray matter–white matter boundaries in the dorsolateral (labeled with "A"), the ventrolateral (label "B"), and the ventro-medial funiculus [label "C," scheme shown in (H)] in the anti-Nogo-A/parallel group. Data are presented as means ± SEM. Statistical evaluation was carried out with two-way ANOVA repeated measure, followed by Bonferroni post hoc (B) and Student's t test (two-tailed, unpaired) (D to G and I). Asterisks indicate significances: $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

cord would contain both transgenes and activate tetanus toxin in response to doxycycline. We applied the same procedure in four control animals, except these animals received injections of the HiRet lentivector coding for EGFP only.

After recovery from the surgery and the reassessment of regained grasping skills, doxycycline
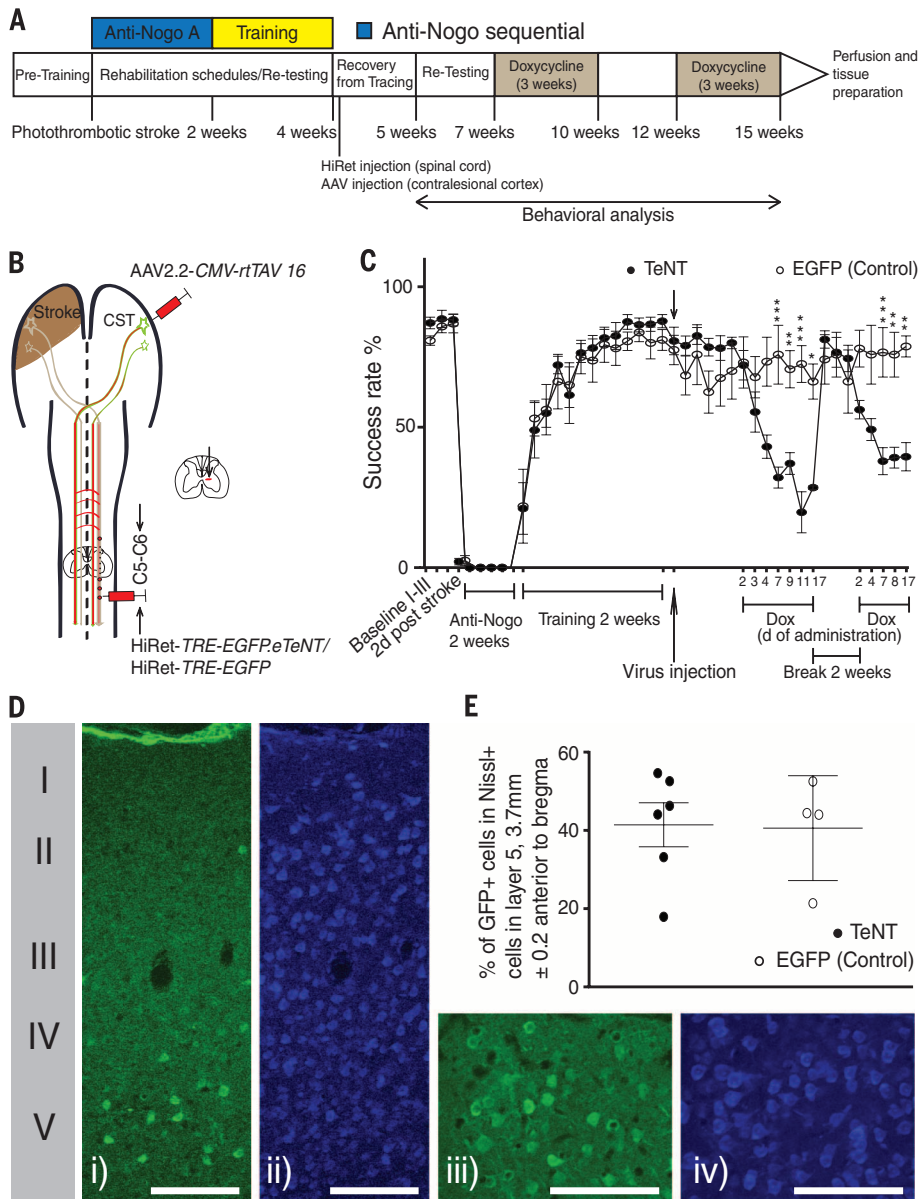

Fig. 3. Long-term reversible blockade of midline-crossing corticospinal fibers abolishes the functional recovery after stroke. (**A**) Experimental schedule. All animals had anti-Nogo-A immunotherapy followed by intensive reaching training, leading to almost full functional restoration of skilled forelimb functions. (**B**) Schematic diagram of vector injections into the contralesional motor cortex (AAV2.2-CMV-rtTAV-16) and stroke-denervated cervical spinal cord (HiRet-TRE-EGFP-TeNT or HiRet-TRE-EGFP as control). (**C**) Induction of tetanus toxin by doxycycline leads to strong impairment of reach and grasping movements within 7 days. The effect was fully reversible within 7 to 10 days of dox removal and could be reinduced by reapplication of dox. TeNT animals, *n* = 6; control animals, *n* = 4. (**D**) Examples of GFP-positive cells (i) in the rostral sensorimotor cortex, 3.7 mm anterior to bregma in comparison to Nissl-positive cells (ii) [scale bars for (i) and (ii), 100 μm], and zooming in to GFP-positive cells in layer five (iii) in relation to Nissl-positive cells (iv) [scale bars for (iii) and (iv), 50 μm]. Roman numerals I to V represent five out of six histological layers in the motor cortex. (**E**) Quantification of GFP-positive cells in percentage of Nissl-positive cells in the layer five motor cortex. Data are presented as means ± SEM. Statistical evaluation was carried out with two-way ANOVA repeated measure, followed by Bonferroni post hoc test. Asterisks indicate significances: *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

was orally administered for 2 weeks (Fig. 3, A and C). In the experimental group grasping performance declined within a few days, reaching a very low level from day 7 after doxycycline initiation ($P < 0.05$, statistical comparison TeNT versus control group, two-way repeated measures ANOVA with post hoc Bonferroni) (Fig. 3C). When the drug administration was ceased, the lost fuction was regained within 2 weeks. All animals again showed a loss of skilled food-pellet grasping movements when oral doxycycline intake was restarted for a second time over the course of another 3 weeks (Fig. 3C). No deterioration of the poststroke recovered skilled grasping was observed in control animals (EGFP instead of eTeNT) under the same dosage of doxycycline and within the same time frame (98.1 ± 0.9% of pre-doxycycline grasping performance). As the retrogradely transported virus was EGFP-tagged, corticospinal neurons projecting to the ipsilateral cervical spinal cord segments C5 and C6 could be quantified. These neurons were concentrated in layer five of a specific region of the rostral, premotor, and primary (M1) forelimb motor cortex (3.7 ± 0.2 mm anterior to bregma). In the center of the labeled region, 41.4 ± 5.6% of Nissl-positive cells of layer five contained the transgene (Fig. 3, D and E). These results demonstrate that midline-crossing corticospinal fibers from the intact hemisphere opposite to the stroke lesion anatomically and functionally switch sides, which is crucial for the recovery of skilled forelimb movements. These fibers maintain their new function, even if they are functionally blocked for weeks. Evidently, their role cannot be compensated by other cortical or subcortical motoneuronal pathways in the injured system.

Temporally blocking rewired corticospinal fibers results in decline of regained function

For short-term reversible inactivation of the midline-crossing corticospinal fibers, we used a pharmacogenetic approach involving virus-mediated gene transfer to express engineered $G_{i/o}$-coupled DREADD receptors ("designer receptor exclusively activated by designer drug"). These receptors are only activated by the otherwise pharmacologically inert synthetic ligand clozapine-*N*-oxide (CNO), resulting in increased intracellular $G_{i/o}$-mediated signaling, which leads to membrane hyperpolarization and silencing of the infected neurons (*9–11*). We used only rats that had undergone the treatment schedule of growth promotion by anti-Nogo-A antibodies, followed by another 2 weeks of rehabilitative training. The virus injections started at the end of the training period (fig. S2A). To selectively manipulate the corticospinal fibers projecting from the intact, contralesional motor cortex to the denervated cervical spinal cord, we first injected an AAV2.9 vector carrying the Cre sequence into segments C5 and C6 of the stroke-denervated cervical spinal cord, followed by injections of the Cre-dependent AAV2.1 vector carrying the $G_{i/o}$-coupled DREADD (hM$_4$Di)

receptor into the contralesional pre- and sensorimotor cortex ($n = 6$) (fig. S2B). The DREADD receptor hM$_4$Di was tagged with mCherry, which allowed neuroanatomical confirmation that only double-infected cells expressed the hM$_4$Di receptor (fig. S2C): We found mCherry-positive cells concentrated in layer five of the same region of the contralesional premotor and motor cortex (3.7 ± 0.2 mm anterior to bregma), as in the tetanus toxin experiment. A portion of Nissl-positive layer five cells in this region ($31.6 \pm 2.4\%$) were positive for mCherry (Fig. 4, A and B), with very limited numbers of mCherry-positive cells outside of this area. In control animals that received injection of the Cre-dependent AAV2.1-hM$_4$Di vector in the contralesional cortex without

AAV-Cre virus injection in the spinal cord, only $4.3 \pm 0.5\%$ of Nissl-positive cells in layer five were positive for mCherry, indicating low background noise ($n = 4$) (Fig. 4A).

Three weeks after virus injection, both hM$_4$Di/+Cre and control hM4Di/–Cre animals performed at >80% success rate in single-pellet grasping (Fig. 4C). Animals were then injected intraperitoneally with the channel-activating drug CNO. Control animals showed no change in reaching and grasping abilities over the 50 min of observation time. However, animals of the hM$_4$Di/+Cre group lost their grasping abilities over 10 to 30 min, with performance declining to $38.9 \pm 6.0\%$ success rate, which is significantly lower compared with the control group (Fig. 4C) ($P <$

0.001, two-way repeated measures ANOVA with post hoc Bonferroni). The defective movements were characterized by a specific failure to target the paw to the pellet and to close the paw around the pellet (Fig. 4D), but with little modifications in overall grasping trajectories (fig. S3). The abatement of distal motor functions was confirmed by performing discriminative classification based on a nonparametric representation (12) of paw posture and its change over time ($P = 0.02$, grasps at baseline versus after 30 min CNO, Kolmogorov-Smirnov test). This mainly distal impairment may also be due to the fact that only the cervical segments C5 and C6 were injected with the Cre virus. These functional defects were fully reversible, with performance

Fig. 4. Short-term reversible blockade of ipsilateral corticospinal fibers abolishes recovered grasping function and forelimb EMGs. (**A** and **B**) Quantification of mCherry-positive cells in percentage of Nissl-positive cells in layer five of contralesional premotor (M2) and motor cortex (M1), 3.7 mm anterior to bregma in the hM$_4$Di/+Cre group ($n = 6$) and the hM$_4$Di/–Cre control group ($n = 4$) (A), and the illustration of its location using Neurolucida software reconstruction (B). (**C**) Activation of the DREADD receptor in the hM$_4$Di/+Cre group induced a rapid and massive but fully reversible impairment of grasping performance 10 to 40 min after CNO application. Reaching success rates were unchanged in the control hM$_4$Di/–Cre CNO-treated group over the same time frame. (**D**) CNO application disturbed fine motor function

of closing the paw around the pellet in the hM$_4$Di/+Cre group ($P = 0.02$, Kolmogorov-Smirnov test): The figure shows the spatial probability densities for the location of hand closure relative to the sugar pellet during grasping at baseline and after CNO application. "x" represents the position of the sugar pellet relative to the forelimb position during grasping. Scale bar, 10 mm. (**E**) CNO application leads to a decrease of EMG responses in the hM$_4$Di/+Cre group ($n = 5$) after ICMS of the contralesional motor cortex compared with ICMS stimulation at baseline and 30 min after CNO application in control animals ($n = 3$). Heat maps of

the cortical stimulation grid are shown (60 stimulation points, 80 µA, +3 to −3 mm anterior-posterior and 1 to 3.5 mm medio-lateral relative to bregma). Each stimulation point is color coded with the mean value of EMG response for a muscle group (wrist, elbow, shoulder) at this stimulation point (in millivolts) normalized to the mean of all stimulation points at the baseline. Data are presented as means ± SEM. Statistical evaluation was carried out with two-way ANOVA repeated measure, followed by Bonferroni post hoc test. Asterisks indicate significances: ***$P < 0.001$.

returning to preinjection levels ~40 to 50 min after the CNO injection (Fig. 4, C and D).

Pharmacogenetic inhibition of regained EMG activity

We confirmed the CNO-specific blockade of neuronal firing of ipsilaterally projecting, partially midline-crossing corticospinal fibers of the intact, contralesional motor cortex by electrophysiology using intracortical microstimulation (ICMS) at the end of the behavioral testing: We used a 5-by-12–point stimulation grid (positioned at +3 to –3 mm anterio-posterior and 1 to 3.5 mm medio-lateral relative to bregma) (fig. S4A) and electromyogram (EMG) recordings of wrist, elbow, and shoulder muscles of the impaired paw as readouts (fig. S4B). In all animals, each cortical position within the exploration grid was stimulated twice, first as baseline stimulation, then again 30 min after CNO injection. In all hM_4Di/–Cre control animals, the EMG responses at baseline and 30 min after CNO injection were not significantly different for the 60 stimulation points [for wrist, 95% confidence interval between baseline and 30 min CNO was for animal one: 0 (0 to 0.001), animal two: 0 (0 to 0), animal three: 0 (0 to 0.001); one-sample Wilcoxon signed rank test] (Fig. 4E and fig. S4C). In contrast, in four out of five hM_4Di/+Cre animals, the ICMS-evoked EMG responses significantly decreased 30 min after CNO injection compared with baseline [for wrist, 95% confidence interval was for animal one: –0.002 (–0.002 to –0.001), animal two: –0.004 (–0.005 to –0.003), animal three: –0.002 (–0.002 to –0.002), animal four: 0 (0 to 0.001), animal five: –0.012 (–0.016 to –0.009); one-sample Wilcoxon signed rank test] (Fig. 4E and fig. S4D). Calculating the mean EMG response for every stimulation point in hM_4Di/+Cre animals versus hM_4Di/–Cre controls revealed a significant decline of EMG responses in wrist and elbow muscles in the hM_4Di/+Cre group 30 min after CNO application compared with controls ($P < 0.05$, Mann-Whitney test) (Fig. 4E). No significant abatement of EMG responses occurred in shoulder muscles of hM_4Di/+Cre animals ($P = 0.4$, Mann-Whitney test) (Fig. 4E). CNO application resulted in the largest difference in EMG responses of wrist recordings when premotor and rostral forelimb areas of the contralesional motor cortex were stimulated in the hM_4Di/+Cre group compared with controls (fig. S2D). These areas also expressed the highest concentration of mCherry-expressing cells (Fig. 4A).

Discussion

Our study shows that in a rat model of large forebrain cortex strokes, timing of rehabilitative training relative to timing of a nerve fiber growth-promoting therapy affects the recovery of lost motor function and the pattern of fiber sprouting. When rats received their first anti-Nogo-A immunotherapy followed by 2 weeks of specific, intense rehabilitative training, forelimb function was almost fully restored (Fig. 1, A and B), indicating a far more extensive recovery rate relative to stroke size than previously obtained

by training (13–15) or growth-promoting therapy alone (6, 16). Not only did these animals outperform the other rehabilitation groups in the single-pellet grasping task, but they were also able to better transfer their regained skills to novel forelimb tasks (15). The behavioral recovery was associated with crossing of CST fibers from the lesion-spared, intact motor cortex to the stroke-denervated side of the spinal cord. This observation is supported by various stroke and spinal cord injury models (15–19) that relate midline-crossing corticospinal fibers to functional outcome. Our analysis shows that not only the quantity of newly out-sprouting corticospinal fibers is relevant but also their termination pattern: Intensive training, when applied too early, induced hyperinnervation and aberrant growth even beyond the gray matter–white matter boundary and into dorsal sensory laminae. Such widespread sprouting may result in wrong circuit connectivity involving cervical interneurons, V2a propriospinal neurons, and motoneurons, thus impairing grasping function, for example, by coactivation of agonistic and antagonistic muscles (20, 21).

Our data suggest the presence of critical time windows during which the brain is most responsive to the application of growth-promoting agents and to training-dependent plasticity. A correct, timed sequence of interventions is required to maximize the effectiveness of rehabilitative therapies after stroke. In a first step, suppression of the action of the endogenous growth-inhibitory factor Nogo-A by immunotherapy may diminish constraints on lesion-induced structural plasticity through mechanisms such as neurotrophic factors, modified electrical properties of motoneurons, alteration in neuronal energy balance (22), and recruitment of new circuits leading to hyperexcitability and prolonged responses to external stimuli (1). In analogy to development, many of these newly formed connections may be weak and imprecise. Training in a second step may then help to shape the spared and new circuits by selection and stabilization of functional connections and pruning of the nonfunctional ones. This second step might involve Hebbian learning rules, in the sense that Hebbian plasticity redistributes synaptic strength to favor functionally relevant pathways that are coincidently active (1). The high degree of recovery of important, cortically controlled motor functions in rats with large ischemic strokes, as demonstrated here, points to a possible avenue to explore growth-inhibitor blockade in combination with rehabilitative training as a treatment strategy for humans with motor cortex stroke. Antibodies against Nogo-A are currently used in clinical trials in humans for amyotrophic lateral sclerosis, multiple sclerosis, and spinal cord injury (www.clinicaltrials.gov). Careful consideration of rehabilitation onset times—particularly with regard to windows of sprouting and circuit plasticity, but also vulnerability of the injured brain—and tailored training adapted to the type and extent of stroke and the patient's history will be essential for future approaches in the clinic (2, 23).

REFERENCES AND NOTES

1. T. H. Murphy, D. Corbett, *Nat. Rev. Neurosci.* **10**, 861–872 (2009).
2. M. A. Dimyan, L. G. Cohen, *Nat. Rev. Neurol.* **7**, 76–85 (2011).
3. P. Langhorne, J. Bernhardt, G. Kwakkel, *Lancet* **377**, 1693–1702 (2011).
4. S. T. Carmichael, *Ann. Neurol.* **59**, 735–742 (2006).
5. S. R. Zeiler, J. W. Krakauer, *Curr. Opin. Neurol.* **26**, 609–616 (2013).
6. N. T. Lindau et al., *Brain* **137**, 739–756 (2014).
7. T. Oertle et al., *J. Neurosci.* **23**, 5393–5406 (2003).
8. M. Kinoshita et al., *Nature* **487**, 235–238 (2012).
9. B. R. Conklin et al., *Nat. Methods* **5**, 673–678 (2008).
10. G. M. Alexander et al., *Neuron* **63**, 27–39 (2009).
11. S. M. Ferguson et al., *Nat. Neurosci.* **14**, 22–24 (2011).
12. B. Ommer, T. Mader, J. M. Buhmann, *Int. J. Comput. Vis.* **83**, 57–71 (2009).
13. D. L. Adkins, T. A. Jones, *Neurosci. Lett.* **380**, 214–218 (2005).
14. M. Alaverdashvili, A. Foroud, D. H. Lim, I. Q. Whishaw, *Behav. Brain Res.* **188**, 281–290 (2008).
15. M. L. Starkey, C. Bleul, I. C. Maier, M. E. Schwab, *Exp. Neurol.* **232**, 81–89 (2011).
16. S. Y. Tsai et al., *Exp. Brain Res.* **182**, 261–266 (2007).
17. G. García-Alías, S. Barkhuysen, M. Buckle, J. W. Fawcett, *Nat. Neurosci.* **12**, 1145–1151 (2009).
18. C. Wiessner et al., *J. Cereb. Blood Flow Metab.* **23**, 154–165 (2003).
19. I. C. Maier et al., *J. Neurosci.* **28**, 9386–9403 (2008).
20. C. O. Asante, J. H. Martin, *PLOS ONE* **8**, e74454 (2013).
21. E. Azim, J. Jiang, B. Alstermark, T. M. Jessell, *Nature* **508**, 357–363 (2014).
22. J. C. Petruska et al., *J. Neurosci.* **27**, 4460–4471 (2007).
23. J. W. Krakauer, S. T. Carmichael, D. Corbett, G. F. Wittenberg, *Neurorehabil. Neural Repair* **26**, 923–931 (2012).

REPORTS

QUANTUM NONLOCALITY

Detecting nonlocality in many-body quantum states

J. Tura,[1] R. Augusiak,[1]* A. B. Sainz,[1] T. Vértesi,[2] M. Lewenstein,[1,3] A. Acín[1,3]

Intensive studies of entanglement properties have proven essential for our understanding of quantum many-body systems. In contrast, much less is known about the role of quantum nonlocality in these systems because the available multipartite Bell inequalities involve correlations among many particles, which are difficult to access experimentally. We constructed multipartite Bell inequalities that involve only two-body correlations and show how they reveal the nonlocality in many-body systems relevant for nuclear and atomic physics. Our inequalities are violated by any number of parties and can be tested by measuring total spin components, opening the way to the experimental detection of many-body nonlocality, for instance with atomic ensembles.

ocal measurements by different observers on entangled composite quantum systems may lead to correlations that are not of classical nature (*1*). This phenomenon is known as nonlocality (*2*) and is detected by means of the so-called Bell inequalities (*1*). These are linear inequalities formulated in terms of the probabilities that result from the local measurements performed by the observers, and their violation signals nonlocality. Apart from its fundamental interest, nonlocality has also become a resource for certain information-theoretic tasks, such as secure key distribution (*3–5*) or certified quantum randomness generation (*6–8*). Hence, revealing the nonlocality of a given composite quantum state is one of the central problems of quantum information theory.

An intensive theoretical effort has been devoted in recent years to the study of entanglement in many-body systems (*9–15*), but the role of nonlocality in such systems has hardly been explored. Entanglement and nonlocality are known to be inequivalent quantum resources. In principle, a generic many-body state—say, a ground state of a local Hamiltonian—is pure, entangled, and because all pure entangled states violate a Bell inequality (*16*), also nonlocal. However, this result cannot be easily verified because the known Bell inequalities (*17–21*) usually involve products of observables of all observers (also referred to as parties). Unfortunately, measurements of such observables are out of reach in many-body systems, in which one has access only to few-body correlations, often two-body. Thus, the physically relevant question concerning the nonlocality of many-body quantum

states is whether its detection is possible by using only two-body correlations.

When tackling this problem, one has to face the following technical challenges. First of all, finding Bell inequalities that are valid for an arbitrary number of parties is usually a difficult task because the complexity of characterizing the set of classical correlations scales exponentially with the number of parties. Second, one can expect that high-order correlations—that is, those among several measurements made by the observers—carry most of the information about the correlations in the system. Consequently, Bell inequalities based on them reveal the nonlocality most efficiently. Whereas Bell inequalities involving correlations among all but one parties have been constructed in (*22–24*), thus proving that all-partite correlations are not necessary to reveal nonlocality, here we have to consider the more demanding question of whether nonlocality detection is possible for systems of an arbitrary number of parties from the minimal information achievable in a Bell test—two-body correlations. We provide a positive answer to this question by proposing classes of Bell inequalities constructed from one- and two-body expectation values that are violated by many-body quantum states.

We considered a Bell experiment in which N spatially separated observers perform measurements on their shares of some N-partite composite quantum state ρ. We focused on the simplest case, in which each party freely chooses one of two dichotomic measurements, whose outcomes are ±1; we denote these measurements by $\mathcal{M}_j^{(i)}$ with $j = 0, 1$ and $i = 1, ..., N$. In this scenario, a convenient way of describing the established correlations is to use the collection of expectation values

$$\langle \mathcal{M}_{j_1}^{(i_1)} ... \mathcal{M}_{j_k}^{(i_k)} \rangle \tag{1}$$

with $i_l = 1, ..., N$ and $j_l = 0, 1$ for $l = 1, ..., k$ ($k = 1, ..., N$). Each of the expectation values in Eq. 1 involving

more than one party is called a correlator. One then constructs a real vector of dimension $3^N - 1$ with components given by all the expectation values in Eq. 1; in what follows, by "correlations" we mean the corresponding vector.

Within this framework, we say that the correlations represented by Eq. 1 are classical (or local) whenever they can be simulated by the observers with shared classical information as the only resource. Such correlations form a polytope $\mathbb{P}$ that is a bounded convex set with a finite number of extreme points (Fig. 1). The polytope $\mathbb{P}$ of classical correlations can be fully determined by a finite number of tight Bell inequalities, those corresponding to the facets of $\mathbb{P}$. Correlations that fall outside of $\mathbb{P}$ are nonlocal.

The general form of a Bell inequality involving only one- and two-body mean values is

$$\sum_{i=1}^{N} (\alpha_i \langle \mathcal{M}_0^{(i)} \rangle + \beta_i \langle \mathcal{M}_1^{(i)} \rangle) + \sum_{i<j}^{N} \gamma_{ij} \langle \mathcal{M}_0^{(i)} \mathcal{M}_0^{(j)} \rangle +$$
$$\sum_{i \neq j}^{N} \delta_{ij} \langle \mathcal{M}_0^{(i)} \mathcal{M}_1^{(j)} \rangle + \sum_{i<j}^{N} \varepsilon_{ij} \langle \mathcal{M}_1^{(i)} \mathcal{M}_1^{(j)} \rangle +$$
$$\beta_C \geq 0 \tag{2}$$

where α_i, β_i, γ_{ij}, δ_{ij}, and ε_{ij} are real parameters and β_C is the so-called classical bound. The corresponding polytope $\mathbb{P}_2$ of classical correlations is the one constructed from the elements of $\mathbb{P}$ by neglecting correlators of order higher than two. Analogously, the vertices of $\mathbb{P}_2$ are those collections of correlators for which $\langle \mathcal{M}_k^{(i)} \mathcal{M}_l^{(j)} \rangle = \langle \mathcal{M}_k^{(i)} \rangle \cdot \langle \mathcal{M}_l^{(j)} \rangle$, while the individual mean values are ±1.

The complete characterization of $\mathbb{P}_2$ reduces to finding all of its facets—tight two-body Bell inequalities. Although the dimension of $\mathbb{P}_2$ is $2N^2$, which is much smaller than the one of $\mathbb{P}$, $3^N - 1$, it still grows with N, thus making difficult the task of determining the facets of $\mathbb{P}_2$ for large N. One way to reduce the complexity of this problem is to consider Bell inequalities that obey some symmetries. In the spirit of (*20*), we focus on Bell inequalities with one- and two-body correlators that are symmetric under permutations of the parties.

By imposing the permutational symmetry, one requires that the expectation values $\langle \mathcal{M}_k^{(i)} \rangle$ and $\langle \mathcal{M}_k^{(i)} \mathcal{M}_l^{(j)} \rangle$, with fixed k, l and different i, j, appear in the Bell inequality Eq. 2 with the same "weights," that is, $\alpha_i = \alpha$ or $\beta_i = \beta$, and so on. This means that the general form of a symmetric Bell inequality with one- and two-body correlators is

$$I := \alpha \mathcal{S}_0 + \beta \mathcal{S}_1 + \frac{\gamma}{2} \mathcal{S}_{00} + \delta \mathcal{S}_{01} +$$
$$\frac{\varepsilon}{2} \mathcal{S}_{11} \geq -\beta_C \tag{3}$$

where α, β, γ, δ, and ε are real parameters. By $\mathcal{S}_k$ and $\mathcal{S}_{kl}$ with k, $l = 0, 1$, we denote the one- and two-body correlators symmetrized over all observers

$$\mathcal{S}_k = \sum_{i=1}^{N} \langle \mathcal{M}_k^{(i)} \rangle, \quad \mathcal{S}_{kl} = \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \langle \mathcal{M}_k^{(i)} \mathcal{M}_l^{(j)} \rangle \tag{4}$$

[1]ICFO–Institut de Ciències Fotòniques, Avenida Carl Friedrich Gauss, 3, 08860 Castelldefels, Barcelona, Spain. [2]Institute for Nuclear Research, Hungarian Academy of Sciences, H-4001 Debrecen, Post Office Box 51, Hungary. [3]ICREA–Institució Catalana de Recerca i Estudis Avançats, Lluis Companys 23, 08010 Barcelona, Spain.
*Corresponding author. E-mail: remigiusz.augusiak@icfo.es

Geometrically, under this symmetry the polytope $\mathbb{P}_2$ is mapped to a simpler one $\mathbb{P}_2^S$, which, independently of N, is always five-dimensional, and its elements are vectors $[\mathcal{S}_0, \mathcal{S}_1, \mathcal{S}_{00}, \mathcal{S}_{01}, \mathcal{S}_{11}]$. $\mathbb{P}_2^S$ is fully characterized if one knows all its facets, which we call tight symmetric two-body Bell inequalities. Moreover, the number of vertices is reduced from 2^{2N} of $\mathbb{P}_2$ to $2(N^2 + 1)$ of $\mathbb{P}_2^S$, and the vertices of the latter can be conveniently parameterized by three natural numbers [(25), section 1].

We could now search for particular Bell inequalities violated by multipartite quantum states. For a sufficiently low number of parties, all Bell inequalities corresponding to the facets of $\mathbb{P}_2^S$ can be listed by using a computer algorithm [(25), section 4]. In what follows, we focus on a general class of few-parameter symmetric Bell inequalities and show that they reveal nonlocality in quantum states for any N.

To this end, in Eq. 3 we substituted $\gamma = x^2$ and $\varepsilon = y^2$ with x, y being positive natural numbers, and $\delta = \sigma xy$, where $\sigma = \pm 1$ stands for the sign of δ. Moreover, we let $\alpha_\pm = x[\sigma\mu \pm (x + y)]$ with $\mu \equiv \beta/y$, and assumed that μ is an integer with opposite parity to ε (γ) for odd N (even N). Exploiting the above parameterization, we prove [(25), section 2.1] that the classical bound of the resulting Bell inequality is

$$\beta_C = \frac{1}{2}[N(x + y)^2 + (\sigma\mu \pm x)^2 - 1] \quad (5)$$

A particular example of a Bell inequality of this form arises from $x = y = -\sigma = 1$ and $\alpha_- = -2$. According to Eq. 5, $\beta_C^N = 2N$, and the resulting Bell inequality is

$$-2\mathcal{S}_0 + \frac{1}{2}\mathcal{S}_{00} - \mathcal{S}_{01} + \frac{1}{2}\mathcal{S}_{11} + 2N \geq 0 \quad (6)$$

To search for quantum violations of this inequality, we considered the case in which all parties measure the same pair of observables (our numerical results suggest that this does not imply any loss of optimality)

$$\mathcal{M}_0^{(i)} = \mathcal{M}_0 = \sigma_z,$$
$$\mathcal{M}_1^{(i)} = \mathcal{M}_1 = \cos\theta\sigma_z + \sin\theta\sigma_x \quad (7)$$

for $i = 1, ..., N$ and $\theta \in [0, \pi]$, where σ_z and σ_x are the standard Pauli matrices. For these measurements, we associate a Bell operator $\mathcal{B}_N(\theta)$ with the left-hand side of Eq. 6 [(25), section 2.2]. The latter displays quantum violation if $\mathcal{B}_N(\theta)$ has a negative eigenvalue for some θ. Exploiting the permutational invariance of $\mathcal{B}_N(\theta)$, we have numerically determined its minimal negative eigenvalue over $\theta \in [0, \pi]$, which corresponds to the maximal quantum violations Q_V^N of Eq. 6, for any $2 \leq N \leq 10^4$. The obtained values of Q_V^N relative to the classical bound $\beta_C^N = 2N$— that is, Q_V^N/β_C^N, along with the corresponding values of the

optimal angle θ_N for which these violations are achieved—are presented in Fig. 2A. In Fig. 2B, the minimal eigenvalue of $\mathcal{B}_N(\theta)$ is also plotted as a function of θ for few values of N. The relative maximal violation grows with N. Additionally, the ranges of θ for which our inequality is violated increases with N, which reflects its robustness against measurement misalignment. Noticeably, it can also be robust against white noise and particle losses [(25), section 2.3]. We then found that the quantum state giving the maximal violation is always symmetric—that is, it belongs to the symmetric subspace of an N-qubit Hilbert space (there are also states violating this inequality that are orthogonal to the symmetric subspace). Because any N-qubit symmetric state is entangled if, and only if, it is genuinely multipartite entangled (26), our Bell inequalities detect states that have genuine multipartite entanglement. Moreover, because the state is symmetric under permutation of the parties, it follows from (27) that all the two-body reduced states are local in the considered scenario.

The next question we analyzed is whether the derived Bell inequalities are powerful enough to reveal nonlocality in "physically relevant" states, such as ground states of spin models that naturally appear in many-body physics. We show that this is the case by constructing a class of two-body symmetric Bell inequalities that are violated by the Dicke states (28). These are symmetric N-qubit states that read

$$|D_N^k\rangle = \mathcal{S}(|\{0, N - k\}, \{1, k\}\rangle) \quad (8)$$

with $k = 0, ..., N$, where $|\{0, N-k\}, \{1, k\}\rangle$ is any pure product vector with $N - k$ qubits in the state $|0\rangle$ and k in the state $|1\rangle$, and $\mathcal{S}$ denotes symmetrization over all parties. It is worth mentioning that $|D_N^k\rangle$ are genuinely multipartite entangled for any $k \neq 0, N > 1$, and the state $|D_6^3\rangle$ was recently generated experimentally (29, 30).

In many-body physics, the Dicke states arise naturally as the lowest-energy eigenstates of the isotropic Lipkin-Meshkov-Glick Hamiltonian (31)

$$H = -\frac{\lambda}{N}\sum_{\substack{i,j=1 \\ i<j}}^{N}(\sigma_x^{(i)}\sigma_x^{(j)} + \sigma_y^{(i)}\sigma_y^{(j)}) - h\sum_{i=1}^{N}\sigma_z^{(i)} \quad (9)$$

which describes N spins interacting through the two-body ferromagnetic coupling ($\lambda > 0$), embedded into the magnetic field of strength $h \geq 0$ acting along the z direction. Again, $\sigma_a^{(i)}$ ($a = x, y, z$) are the Pauli matrices acting at site i. We then considered the case of a weak applied magnetic field, where $h \leq \lambda/N$. Then, the ground state of H is $|D_N^{N/2}\rangle$ for even N and $|D_N^{\lceil N/2 \rceil}\rangle$ for odd N, except for the case of $h = 0$ and odd N, for which the lowest energy is twofold degenerate and the corresponding subspace is spanned by $|D_N^k\rangle$, with $k = \lfloor N/2 \rfloor$ and $k = \lceil N/2 \rceil$.

Fig. 1. Local correlations. The set of local, or classical, correlations defines the polytope $\mathbb{P}$ (blue region). Its vertices (red points) correspond to those vectors constructed from Eq. 1 in which every correlator factorizes $\langle \mathcal{M}_{j_1}^{(i_1)}...\mathcal{M}_{j_k}^{(i_k)} \rangle = \langle \mathcal{M}_{j_1}^{(i_1)} \rangle \cdots \langle \mathcal{M}_{j_k}^{(i_k)} \rangle$, with every local mean value being ± 1; these are correlations in which every observable has a well-defined value. The yellow point indicates an example of nonlocal correlations outside the local polytope; the thick line represents a tight Bell inequality (facet of the polytope).
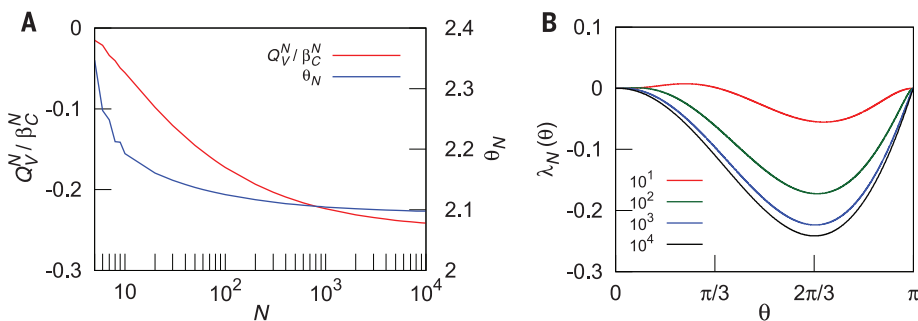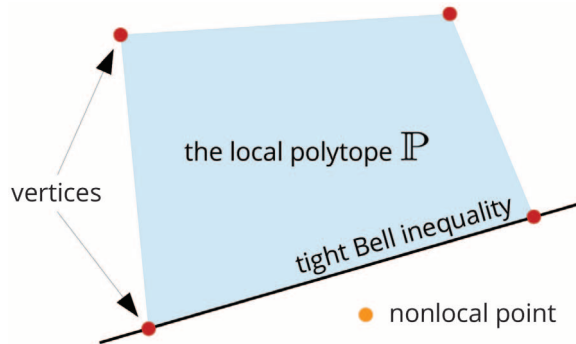

the local polytope $\mathbb{P}$

vertices

tight Bell inequality

● nonlocal point

Fig. 2. Bell quantum violations. (**A**) Maximal relative quantum violation Q_V^N/β_C^N of the inequality in Eq. 6 (red line), and the corresponding optimal angle θ_N in $\mathcal{M}_1$ (blue line) as functions of N. (**B**) Minimal negative eigenvalue $\lambda_N(\theta)$ of $\mathcal{B}_N(\theta)$ as a function of θ for $N = 10^k$ with $k = 1, 2, 3, 4$. The violation is more robust against misalignments of $\mathcal{M}_1$ for large N.
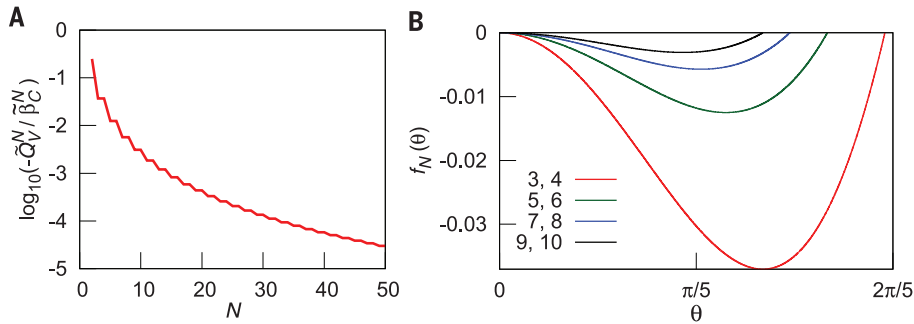
Fig. 3. Bell violations for Dicke states. (**A**) Relative quantum violation $\tilde{Q}_V^N/\tilde{\beta}_C^N$ of our Bell inequality by the state $|D_N^{\lceil N/2 \rceil}\rangle$ as a function of N. It decays with N as $1/N^3$. (**B**) The relative quantum violation $f_N(\theta) := \langle \tilde{\mathcal{B}}_N(\theta)\rangle/\tilde{\beta}_C^N$ as a function of θ for few values of N.

The class of tight two-body symmetric Bell inequalities that we use to detect nonlocality of the above Dicke states is obtained by taking $\alpha_N = N(N-1)(\lceil N/2 \rceil - N/2)$, $\beta_N = \alpha_N/N$, $\gamma_N = N(N-1)/2$, $\delta_N = N/2$, and $\varepsilon_N = -1$ in Eq. 3. This choice of parameters allows us to compute analytically the classical bounds of the resulting Bell inequalities for any N [(25), section 3.1]

$$\tilde{\beta}_C^N = \frac{1}{2}N(N-1)\left\lceil \frac{N+2}{2}\right\rceil \qquad (10)$$

These Bell inequalities are independent of the class presented previously, and for $N = 2$, they reproduce the Clauser-Horne-Shimony-Holt Bell inequality (32).

To prove that these Bell inequalities are indeed violated by the Dicke states with $k = \lfloor N/2 \rfloor, \lceil N/2 \rceil$, we again let all parties have the same pairs of observables given in Eq. 7. Denoting then by $\tilde{\mathcal{B}}_N(\theta)$ the associated Bell operator, we analytically minimize the mean value $\langle \tilde{\mathcal{B}}_N(\theta)\rangle = \langle D_N^{\lceil N/2 \rceil}|\tilde{\mathcal{B}}_N(\theta)|D_N^{\lceil N/2 \rceil}\rangle$ over all $\theta \in [0, \pi]$. This gives the following maximal quantum violation for the above measurements [details provided in (25), section 3.2]

$$\tilde{Q}_V^N = -\frac{\lfloor N/2 \rfloor}{\lceil N/2 \rceil + 1} \qquad (11)$$

The same violation can be obtained for $k = \lfloor N/2 \rfloor$ by renaming (flipping) the outcomes of $\mathcal{M}_j$ ($j = 0, 1$) because $|D_N^{\lfloor N/2 \rfloor}\rangle$ is obtained from $|D_N^{\lceil N/2 \rceil}\rangle$ by swapping the elements of the computational basis $\{|0\rangle, |1\rangle\}$. Presented in Fig. 3, A and B, respectively, are the relative quantum violation $\tilde{Q}_V^N/\tilde{\beta}_C^N$ as a function of N and $\tilde{\mathcal{B}}_N(\theta)/\tilde{\beta}_C^N$ as a function of θ for a few values of N. A violation is obtained for all number of parties, although $\tilde{Q}_V^N/\tilde{\beta}_C^N$ decays with N as $1/N^3$ (Fig. 3A).

Although the above choice of measurements allows one to analytically find the quantum violation for any N, this is not the optimal one. In fact, by taking $\mathcal{M}_0$ to be the most general real measurement, we can dramatically increase the violation of our Bell inequalities; in particular, we can improve the scaling with N to $1/N$ [(25), section 3.2].

We conclude by discussing the possible implementation of the previous results. Our main

aim here is not to perform a proper loophole-free Bell experiment but to provide feasible methods with which to witness whether many-body systems feature nonlocality. Our inequalities, which involve only two-body correlators, are violated by symmetric states. Such states appear naturally as ground states of models that can be realized with ultracold atoms or ions, such as Lipkin-Meshkov-Glick–like models with long-range interactions [ionic spin-1/2 and spin-1 realizations are available in (33) and references therein; cold atoms in nanophotonic waveguides are available in (34)] or degenerate ground states of the ferromagnetic Heisenberg model (35).

Last, in the case in which all observers measure the same pair of observables, as discussed above, the one- and two-body symmetric expectation values, and therefore the Bell violation, can be estimated via collective measurements of total spin operators $S_\alpha = (1/2)\sum_{i=1}^N \sigma_\alpha^{(i)}$ with $\alpha = x, y, z$ and their combinations $\boldsymbol{n} \cdot \boldsymbol{S}$ in any direction $\boldsymbol{n}$, where $\boldsymbol{S} = [S_x, S_y, S_z]$. To be more explicit, consider a pair of qubit observables $\mathcal{M}_i = \boldsymbol{m}_i \cdot \boldsymbol{\sigma}$ ($i = 0, 1$) (such as those we used to test our inequalities) with $\boldsymbol{m}_i \in \mathbb{R}^3$ such that $|\boldsymbol{m}_0| = |\boldsymbol{m}_1| = 1$ and $\boldsymbol{\sigma} = [\sigma_x, \sigma_y, \sigma_z]$. One directly finds that $\mathcal{S}_i = 2\langle \boldsymbol{m}_i \cdot \boldsymbol{S}\rangle$, $\mathcal{S}_{ii} = 4\langle(\boldsymbol{m}_i \cdot \boldsymbol{S})^2\rangle - N$ with $i = 0, 1$, and last, $\mathcal{S}_{01} = \langle((\boldsymbol{m}_0 + \boldsymbol{m}_1) \cdot \boldsymbol{S})^2\rangle - \langle((\boldsymbol{m}_0 - \boldsymbol{m}_1) \cdot \boldsymbol{S})^2\rangle - N(\boldsymbol{m}_0 \cdot \boldsymbol{m}_1)$. Such quantities are routinely measured in atomic systems such as atomic ensembles and optical lattices with current experimental technologies, such as spin polarization spectroscopy (36, 37). Such collective measurements were already proposed (38) and implemented (36) for entanglement detection in many-body systems. Therefore, our work opens the way to the theoretical and experimental study of the nonlocal correlations of many-body states.

REFERENCES AND NOTES

1. J. S. Bell, *Physics* **1**, 195–200 (1964).
2. N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, S. Wehner, *Rev. Mod. Phys.* **86**, 419–478 (2014).
3. A. K. Ekert, *Phys. Rev. Lett.* **67**, 661–663 (1991).
4. J. Barrett, L. Hardy, A. Kent, *Phys. Rev. Lett.* **95**, 010503 (2005).
5. A. Acín et al., *Phys. Rev. Lett.* **98**, 230501 (2007).
6. S. Pironio et al., *Nature* **464**, 1021–1024 (2010).
7. R. Colbeck, thesis, University of Cambridge (2007).
8. R. Colbeck, A. Kent, *J. Phys. A* **44**, 095305 (2011).
9. A. Osterloh, L. Amico, G. Falci, R. Fazio, *Nature* **416**, 608–610 (2002).
10. T. J. Osborne, M. A. Nielsen, *Phys. Rev. A* **66**, 032110 (2002).
11. G. Vidal, J. I. Latorre, E. Rico, A. Kitaev, *Phys. Rev. Lett.* **90**, 227902 (2003).
12. F. Verstraete, M. M. Wolf, D. Pérez-Garcia, J. I. Cirac, *Phys. Rev. Lett.* **96**, 220601 (2006).
13. M. Levin, X.-G. Wen, *Phys. Rev. Lett.* **96**, 110405 (2006).
14. A. Kitaev, J. Preskill, *Phys. Rev. Lett.* **96**, 110404 (2006).
15. H. Li, F. D. M. Haldane, *Phys. Rev. Lett.* **101**, 010504 (2008).
16. S. Popescu, D. Rohrlich, *Phys. Lett. A* **166**, 293–297 (1992).
17. R. F. Werner, M. M. Wolf, *Phys. Rev. A* **64**, 032112 (2001).
18. M. Żukowski, Č. Brukner, *Phys. Rev. Lett.* **88**, 210401 (2002).
19. C. Śliwa, *Phys. Lett. A* **317**, 165–168 (2003).
20. J.-D. Bancal, N. Gisin, S. Pironio, *J. Phys. A Math. Theor.* **43**, 385303 (2010).
21. L. Aolita, R. Gallego, A. Cabello, A. Acín, *Phys. Rev. Lett.* **108**, 100401 (2012).
22. N. Brunner, J. Sharam, T. Vértesi, *Phys. Rev. Lett.* **108**, 110501 (2012).
23. M. Wieśniak, M. Nawareg, M. Żukowski, *Phys. Rev. A* **86**, 042339 (2012).
24. L. E. Würflinger, J.-D. Bancal, A. Acín, N. Gisin, T. Vértesi, *Phys. Rev. A* **86**, 032117 (2012).
25. Materials and methods are available as supplementary materials on Science Online.
26. R. Augusiak, J. Tura, J. Samsonowicz, M. Lewenstein, *Phys. Rev. A* **86**, 042316 (2012).
27. B. Toner, F. Verstraete, http://arxiv.org/abs/quant-ph/0611001.
28. R. H. Dicke, *Phys. Rev.* **93**, 99–110 (1954).
29. R. Prevedel et al., *Phys. Rev. Lett.* **103**, 020503 (2009).
30. W. Wieczorek et al., *Phys. Rev. Lett.* **103**, 020504 (2009).
31. H. J. Lipkin, N. Meshkov, A. J. Glick, *Nucl. Phys.* **62**, 188–198 (1965).
32. J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, *Phys. Rev. Lett.* **23**, 880–884 (1969).
33. T. Graß, B. Juliá-Díaz, M. Kuś, M. Lewenstein, *Phys. Rev. Lett.* **111**, 090404 (2013).
34. D. E. Chang, J. I. Cirac, H. J. Kimble, *Phys. Rev. Lett.* **110**, 113606 (2013).
35. S. Sachdev, *Quantum Phase Transitions* (Cambridge Univ. Press, Cambridge, 1999).
36. K. Hammerer, A. S. Sørensen, E. S. Polzik, *Rev. Mod. Phys.* **82**, 1041–1093 (2010).
37. K. Eckert et al., *Nat. Phys.* **4**, 50–54 (2008).
38. A. Sørensen, L.-M. Duan, J. I. Cirac, P. Zoller, *Nature* **409**, 63–66 (2001).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/344/6189/1256/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S2
Tables S1 to S6
References (39, 40)

QUANTUM GASES

Observation of many-body dynamics in long-range tunneling after a quantum quench

Florian Meinert,[1] Manfred J. Mark,[1] Emil Kirilov,[1] Katharina Lauber,[1] Philipp Weinmann,[1] Michael Gröbner,[1] Andrew J. Daley,[2,3] Hanns-Christoph Nägerl[1]*

Quantum tunneling is at the heart of many low-temperature phenomena. In strongly correlated lattice systems, tunneling is responsible for inducing effective interactions, and long-range tunneling substantially alters many-body properties in and out of equilibrium. We observe resonantly enhanced long-range quantum tunneling in one-dimensional Mott-insulating Hubbard chains that are suddenly quenched into a tilted configuration. Higher-order tunneling processes over up to five lattice sites are observed as resonances in the number of doubly occupied sites when the tilt per site is tuned to integer fractions of the Mott gap. This forms a basis for a controlled study of many-body dynamics driven by higher-order tunneling and demonstrates that when some degrees of freedom are frozen out, phenomena that are driven by small-amplitude tunneling terms can still be observed.

Quantum tunneling is ubiquitous in physics and forms the basis for a multitude of fundamental effects (1) related to electronic transport, nuclear motion, light propagation, and superfluidity in lattice systems (2). Whereas for weakly interacting particles tunneling at a rate J will occur as an individual process for each particle, in strongly interacting systems the behavior of each particle is correlated with the behavior of other particles. Such correlated processes are believed to play an important role, for example, in superconductivity of the cuprate systems (3–5). Second-order tunneling has been observed in cold atom experiments as driven resonances (6) or directly as a dynamical process for pairs of strongly interacting particles in arrays of double-well potentials (7). That process results in an effective nearest-neighbor super-exchange interaction (8, 9), which forms the basis of important forms of quantum magnetism (10), and provides a starting point for the formation of quantum many-body phases. Such tunneling processes have also recently been observed for electrons in systems of quantum dots (11).

Higher-order processes involving correlated tunneling across multiple lattice sites can give rise to longer-range effective interaction terms and more complex many-body critical phenomena (12), as well as marked changes in out-of equilibrium dynamics. Parallels can be drawn between long-range tunneling processes in tilted lattices and multiphoton electron-positron creation in strong electric fields, with connections to relativistic phenomena such as the Sauter-Schwinger effect

in tilted Mott insulators (13), and also to long-distance electron transport in molecular systems, for example (14, 15). However, although single-particle tunneling loss via higher-band resonances (16) has been demonstrated, it has been difficult to observe coherent quantum dynamics due to higher-order tunneling processes

because the small amplitude driving these terms places challenging upper limits on the energy scales for required temperatures and allowed disorder.

Our experiment is based on an array of one-dimensional (1D) Mott-insulating "Ising" chains of bosons in an optical lattice near zero temperature (17–21). We model the system by a single-band Bose-Hubbard (BH) Hamiltonian (22, 23). For large on-site interaction energy $U >> J$, the many-body ground state is a Mott insulator with unit occupation at commensurate filling (Fig. 1A). This phase is characterized by exponentially localized atoms and highly suppressed tunneling. In addition, we superimpose a linear gradient potential, which introduces a site-to-site constant energy shift E. We perform a quantum quench to a highly nonequilibrium situation by rapidly tilting the initial Mott state to an integer fraction of the Mott gap $E \approx U/n$. The quench initiates resonantly enhanced long-range tunneling to the nth neighbor for all sites simultaneously (lower part of Fig. 1 A). For $n = 1$, one couples to nearest-neighbor dipole states and observes strong coherent oscillations in the number of doubly occupied sites (doublons) with a characteristic frequency $4J$ (20). For $n > 1$, resonant tunnel coupling occurs across $n – 1$ intermediate lattice sites. The process involves up to n other particles, giving rise to occupation-dependent nth-order tunneling, with a characteristic rate of atom-pair formation set by $\alpha_n \times J^n/(U/n)^{n-1}$
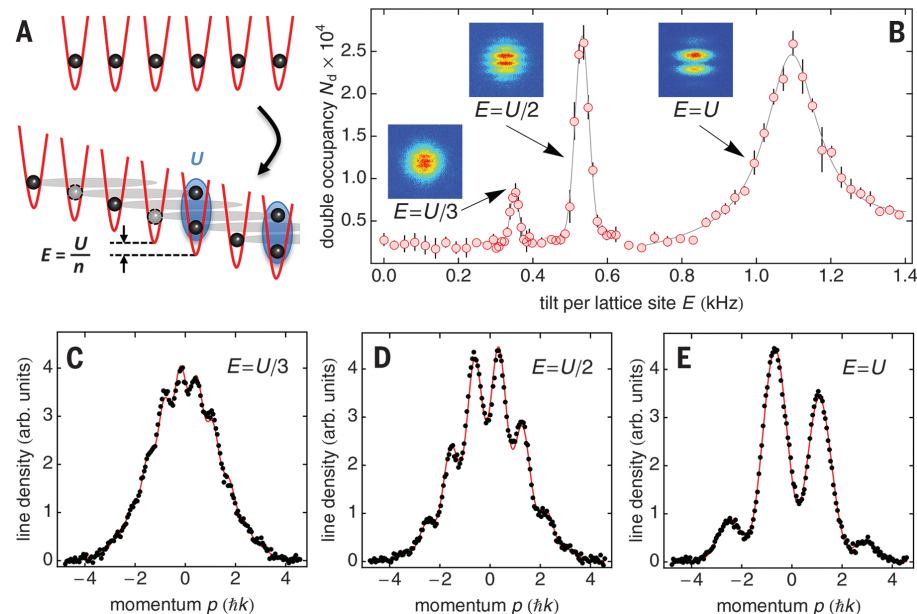
Fig. 1. Tunneling resonances in a tilted 1D Mott insulator. (A) Schematic view of the long-range correlations across n sites for a tilt of $E = U/n$ after the quench from the initial 1D one-atom Mott insulator (top) to the tilted configuration (bottom). Here, $n = 3$. **(B)** Number of doublons N_d as a function of E at $t_h = 200$ ms after the quench. Here, $V_z = 10E_R$ and $a_s = 252(5)$ a_0, giving $U = 1.077(20)$ kHz for $V_{x,y} = 20E_R$. The solid lines are Lorentzian (for $E = U$) and Gaussian (for $E = U/2$ and $E = U/3$) fits to the data to determine the center positions and widths. The insets show matter-wave interference patterns obtained in TOF at $E_1 = U$, $E_2 = U/2$, and $E_3 = U/3$ taken after $t_h = 1$ ms, 9 ms, and 28 ms, respectively. **(C to E)** Integrated line densities of the TOF images shown in the insets in (B). The solid lines are fits according to double-slit interference patterns with Gaussian envelopes (32). Error bars in all figures reflect ±1 SD.

[1]Institut für Experimentalphysik und Zentrum für Quantenphysik, Universität Innsbruck, 6020 Innsbruck, Austria. [2]Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA. [3]Department of Physics and the Scottish Universities Physics Alliance, University of Strathclyde, Glasgow G4 0NG, UK.
*Corresponding author. E-mail: christoph.naegerl@uibk.ac.at

in nth order perturbation theory. Here, α_n is a proportionality factor that includes the effect of Bose enhancement (23). Because all particles participate in a tunnel process across n sites, one expects the build-up of massive correlations in the interacting many-body system. As we discuss in (23), the quench onto the critical point in the many-body system results in many-body dephasing of oscillations in the atom-pair number, corresponding to a characteristic growth in many-body entanglement (24–26) in our numerical simulations.

We prepare an ensemble of 1D Mott insulators (20) starting from a 3D Bose-Einstein condensate (BEC) of typically 8.5×10^4 Cs atoms without detectable uncondensed fraction. The BEC is levitated against gravity by a magnetic field gradient of $|\nabla B| \approx 31.1$ G/cm and initially held in a crossed optical dipole trap (20, 27). We load the sample adiabatically into a cubic 3D optical lattice generated by laser beams at a wavelength of $\lambda_l = 1064.5$ nm, thereby creating a singly occupied 3D Mott insulator for a lattice depth of $V_q = 20E_R$ (28) in each direction ($q = x, y, z$) with less than 4% residual double occupancy. Here, $E_R = \hbar^2 k_l^2/(2m)$ is the photon-recoil energy, with $k_l = 2\pi/\lambda_l$ and m the mass of the Cs atom. The optical lattice results in a residual harmonic confinement of $\nu_z = 11.9(2)$ Hz in the z direction of gravity. A broad Feshbach resonance allows us to set the atomic scattering length a_s, and thus U independently of J, by means of an offset magnetic field B (20).

Tunneling resonances are observed by quickly tilting the lattice in the z direction through a reduction of $|\nabla B|$ and then lowering V_z to 10 E_R within 1 ms, giving $J \approx 25$ Hz (22, 28). All dynamics are now restricted along 1D Mott chains with an average length of 40 sites (20). The chains, in total ≈ 2000, are decoupled from each other on the relevant experimental time scales. We let the systems evolve for a hold time t_h of up to 200 ms in the tilted configuration and then quickly ramp back V_z to its original value and remove the tilt. The ensemble is characterized by measuring the number of doubly occupied sites N_d through Feshbach molecule formation with an overall efficiency of 80(3)% (20). Alternatively, we detect the emergence of momentum-space coherence in time-of-flight (TOF) by quickly turning off all trapping potentials and allowing for 20 ms of free levitated expansion at $a_s = 0$ (27) before taking an absorption image.

The experimental result for a specific choice of $U = 1077(0)$ Hz is shown in Fig. 1B. For a hold time of $t_h = 200$ ms, the transient response as discussed below has settled to a steady-state value. Besides a broad resonance at $E = 1095(2)$ Hz with full width at half maximum (FWHM) = 172(9) Hz, two narrower resonances at $E = 532(1)$ and 351(1) Hz with FWHM = 44(2) and 27(2) Hz can be seen. Whereas the broad resonance is the result of resonant tunnel coupling to nearest-neighbor dipole states at $E_1 = U$ (20), the positions of the narrower resonances are consistent with $E_2 = U/2$ and $E_3 = U/3$, and we hence interpret them to emerge from tunnel processes
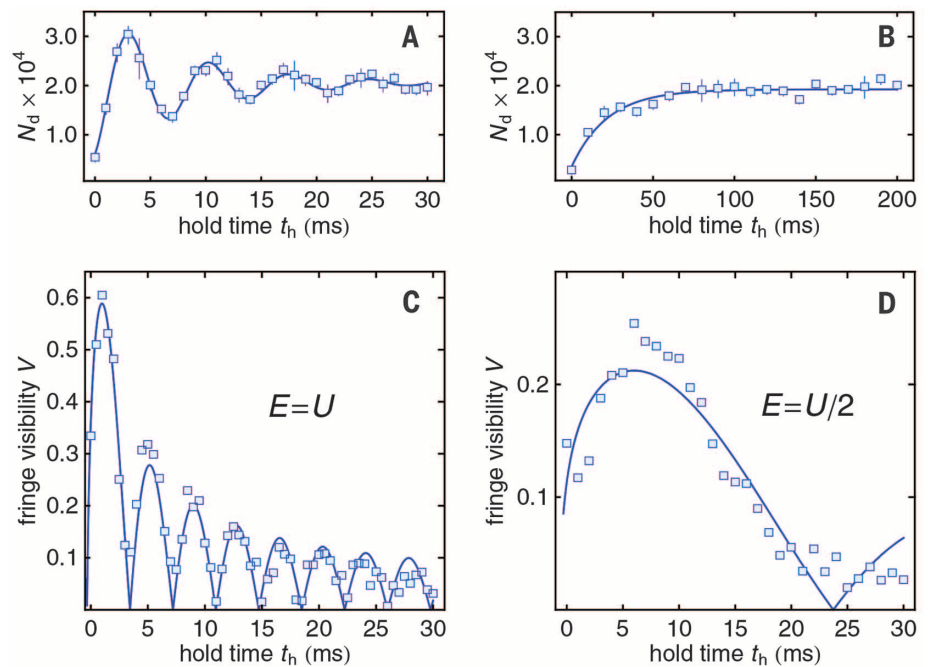
Fig. 2. Comparison of the tunneling dynamics to nearest and second-nearest neighbors. (**A** and **B**) Double occupancy N_d and (**C** and **D**) fringe visibility V in the TOF images as a function of hold time t_h after the quench (symbols). Coherent oscillations in N_d at $E_1 = U$ in (A) are contrasted to overdamped dynamics at $E_2 = U/2$ (B) for $V_z = 10E_R$ and $a_s = 253(5)$ a_0. The evolution of N_d is fitted (lines) by an exponentially damped sinusoid (A) and a saturated growth (B). The solid lines in (C) and (D) are fits to guide the eye based on the modulus of an algebraically decaying sinusoid.

extending over a distance of two and three lattice sites, respectively. The reduced widths reflect the smaller amplitude of the higher-order tunnel processes. We believe that the resonances are slightly broadened inhomogeneously by the external harmonic confinement. The assignment of the resonance features to tunneling processes over multiple lattice sites is supported by TOF images (insets to Fig. 1B) taken for each resonance E_n in the course of the transient response. The images clearly exhibit matter-wave interference patterns, indicating delocalization of the atoms during the tunnel processes. The integrated line densities are presented in Fig. 1, C to E. The periodicity of the sinusoidal density modulation, found to be $2\hbar k_l/n$, is in agreement with spatial coherence of the atomic wave function over a distance of n sites.

We now investigate the transient dynamics after the quantum quench. Figure 2, A and C (B and D), shows the on-resonance response of N_d and the fringe visibility V in the TOF images for E_1 (E_2). The quench to E_1 results in large-amplitude oscillations for N_d; calculations show that the decay is due to many-body dephasing, which plays an increased role for larger chain lengths (20). The oscillatory response at E_1 is clearly reflected in the dynamics for V, as each local minimum coincides with an extremum for N_d. The dynamics for E_2 are, in contrast, highly overdamped and fit to a saturated growth function of the form $\propto (1 - e^{-t_h/\tau})$, with a characteristic rate $1/\tau$; a simple three-site BH model predicts

oscillations at frequency $\nu_2 = 4(2\sqrt{2} + \sqrt{2})J^2/U$ (23). In the experiment, we find a single maximum for V before it decays. The dephasing here results from more complicated dynamics in BH chains longer than three sites (23).

We now focus on the scaling of the resonant doublon growth rate $1/\tau$ with J and U for the resonance E_2. Example data sets (Fig. 3A) clearly demonstrate that $1/\tau$ depends not only on V_z but also on U when V_z and thereby J are kept constant. In Fig. 3B, we plot the same data with the time axis rescaled by the energy scale $J^2/(U/2)$ for a second-order tunneling process. The data collapse onto a single curve, demonstrating that indeed second-order tunneling dominates the transient dynamics after the quench. The numerical data for 10 to 30 site BH chains (23) show similar rise characteristics and reveal the same scaling collapse (Fig. 3 C). The values for $1/\tau$ from measurements taken at different combinations of V_z and U have a linear dependence on $J^2/(U/2)$ (Fig. 3D) with a surprisingly large prefactor $\alpha_2 = 38(2)$, which we analyze in two ways. First, we compare to the frequency of coherent doublon oscillations in the simple three-site model. The role of many-body dephasing faster than a full second-order tunneling cycle is estimated by assuming τ as a quarter of the full tunneling period. The value $1/\tau \approx 4 \times \nu_2$ is indicated by the solid line in Fig. 3D. Second, we extract a characteristic growth rate from the numerical data, indicated by the dashed regions in Fig. 3, C and D, revealing quantitative agreement with the experiment.
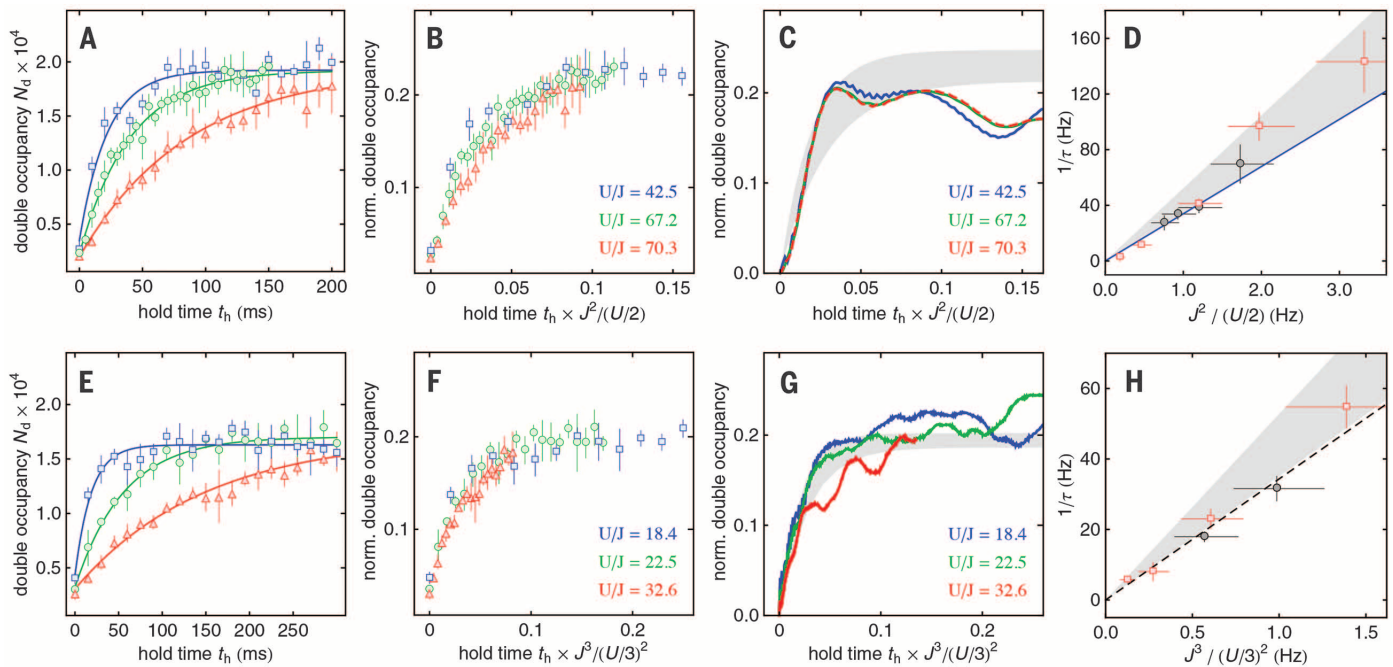
Fig. 3. Second- and third-order tunneling dynamics. (**A**) Double occupancy at E_2 for V_z/E_R, $a_s/a_0 = 10, 253(5)$ (squares), 12, 253(5) (triangles), and 10, 400(5) (circles). (**E**) Double occupancy at E_3 for V_z/E_R, $a_s/a_0 = 7, 253(5)$ (squares), 9, 253(5) (triangles), and 9, 175(5) (circles). The solid lines are fits to the data with saturated growth functions. (**B** and **F**) Collapse of the data shown in (A) and (E) for rescaled time axes. (**C** and **G**) Result of a numerical simulation of the resonant response at E_2 and E_3, respectively. (**D** and **H**) Growth rates $1/\tau$ for E_2 and E_3, respectively. In (D), the data for $V_z = (8,9,10,12,14)E_R$ with fixed $a_s = 253\ a_0$

(squares) and $a_s = (175,253,325,400)\ a_0$ with fixed $V_z = 10E_R$ (circles) are plotted as a function of $J^2/(U/2)$. The solid line gives the prediction from a three-site BH model. In (H), the data for $V_z = (7,8,9,10)E_R$ with $a_s = 253\ a_0$ (squares) and for $a_s = 175\ a_0$ at $V_z = 9E_R$ and $a_s = 300\ a_0$ at $V_z = 7E_R$ (circles) are plotted as a function of $J^3/(U/3)^2$. The dashed line is a linear fit to the experimental data. The shaded areas in (C), (D), (G), and (H) indicate the spread in the growth rate extracted from the numerical data with fixed steady-state values from the experiment (23).
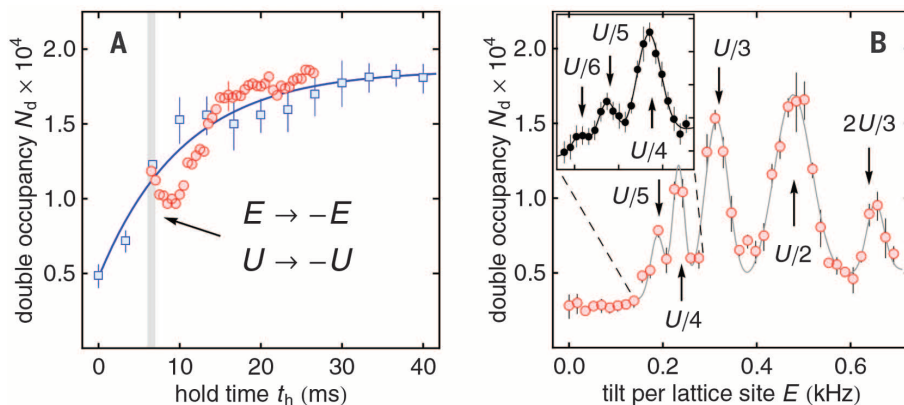
Fig. 4. Many-body echo and higher-order tunneling resonances. (**A**) Double-occupancy N_d as a function of hold time t_h at $E = U/2$ for $V_z = 8E_R$ and $a_s = -250(5)\ a_0$, giving $U = -994(20)$ Hz (squares). Partial time reversal of the many-body dynamics (circles) occurs after switching a_s to $+250(5)\ a_0$ and simultaneously reversing E to $-E$ at $t_h = 6$ ms within 1 ms (gray bar). For the echo data (circles), a typical error bar is given for the data point at $t_h = 16$ ms. (**B**) N_d as a function of E after $t_h = 200$ ms at $V_z = 7E_R$ with $U = 959(20)$ Hz, for $a_s = 252(5)\ a_0$. The arrows indicate the expected positions of the tunneling resonances at $E_n = U/n$. An additional resonance at $2U/3$ appears. The inset gives a fine scan of the $U/4$ and $U/5$ resonances. The solid line is a fit based on the sum of multiple Gaussians to guide the eye.

A similar behavior in the dynamical scaling of the resonant response at the resonance $E_3 = U/3$ is seen in Fig. 3, E to H. Scaling collapse is observed when rescaling time by $J^3/(U/3)^2$, indicating a third-order tunneling process. Residual oscillations after the initial growth period

in the numerical data (Fig. 3G) relate to the finite system size and the lack of averaging over positions in the trap (23). From the linear fit to the growth rate $1/\tau$ in Fig. 3H, we obtain a slope of $\alpha_3 = 34(2)$, in good agreement with a characteristic growth rate determined from the nu-

merical data, which we indicate by the dashed region as before. We note that the signature of the third-order process is not masked by the presence of second-order energy shifts (23).

To what extent can one reverse this many-body dephasing dynamics? In Fig. 4A, we show the result of a many-body echo experiment for which we switch the sign of U and E at the $E_2 = U/2$ resonance in the course of the transient response. A clear, although only partial, reversal in the time evolution for N_d can be seen before N_d reaches the same steady-state value as before. It would be interesting to test whether the revival could be improved by switching the sign of J as well. Naively, the second-order process scaling with J^2 should not depend on the sign of J. Switching J by means of modulation techniques (29) may allow a detailed benchmarking of many-body damping versus the presence of mere inhomogeneous broadening in our system.

Finally, in Fig. 4B, we show resonances corresponding to many-body tunneling across four and five lattice sites. For these data, the lattice depth was reduced to $V_z = 7E_R$ to speed up the processes; the system was initially in the Mott-insulating regime. With decreasing V_z, the resonances at $U/2$ and $U/3$ slightly broaden, which we attribute to the increase of the second- and third-order tunneling rates. The new resonances at $U/4$ and $U/5$ are clearly detectable. We note that these fourth- and fifth-order tunneling processes greatly benefit from substantial Bose enhancement

(23) and speculate that even higher-order processes should become accessible when one eliminates residual parabolic energy shifts due to the trapping laser beams.

Our results underline the utility of cold atoms in optical lattices for the investigation of fundamental physical processes driven by small-amplitude terms and specifically higher-order tunneling. By partly freezing the motion in the deep lattice, these sensitive processes can be observed here despite finite initial temperatures (which here are converted into defects and missing atoms in an ensemble of initial states). This will motivate further investigation of quantum phases and critical properties near these higher-order resonances, which are presently unknown, including systems with tilts along multiple axes (19, 30). Our initial studies of parameter reversals also open the door to the study of many-body dephasing and echo-type experiments on a quantum many-body system, as well as investigations into the nature of the many-body dephasing and (apparent) thermalization (31). Parallels can be drawn with arrays of quantum dots, opening further possibilities to model electron tunneling over multiple sites (11) by using fermionic atoms.

REFERENCES AND NOTES

1. J. Ankerhold, *Quantum Tunneling in Complex Systems: The Semiclassical Approach* (Springer, Berlin, 2010).
2. O. Morsch, M. Oberthaler, *Rev. Mod. Phys.* **78**, 179–215 (2006).
3. A. H. MacDonald, S. M. Girvin, D. Yoshioka, *Phys. Rev. B* **37**, 9753–9756 (1988).
4. P. W. Anderson *et al.*, *J. Phys. Condens. Matter* **16**, R755–R769 (2004).
5. M. R. Norman, *Science* **332**, 196–200 (2011).
6. R. Ma *et al.*, *Phys. Rev. Lett.* **107**, 095301 (2011).
7. S. Fölling *et al.*, *Nature* **448**, 1029–1032 (2007).
8. S. Trotzky *et al.*, *Science* **319**, 295–299 (2008).
9. D. Greif, T. Uehlinger, G. Jotzu, L. Tarruell, T. Esslinger, *Science* **340**, 1307–1310 (2013).
10. E. Dagotto, *Rev. Mod. Phys.* **66**, 763–840 (1994).
11. F. R. Braakman, P. Barthelemy, C. Reichl, W. Wegscheider, L. M. K. Vandersypen, *Nat. Nanotechnol.* **8**, 432–437 (2013).
12. S. Sachdev, *Nat. Phys.* **4**, 173–185 (2008).
13. F. Queisser, P. Navez, R. Schützhold, *Phys. Rev. A* **85**, 033625 (2012).
14. G. Sedghi *et al.*, *Nat. Nanotechnol.* **6**, 517–523 (2011).
15. J. R. Winkler, H. B. Gray, *J. Am. Chem. Soc.* **136**, 2930–2939 (2014).
16. C. Sias *et al.*, *Phys. Rev. Lett.* **98**, 120403 (2007).
17. J. Simon *et al.*, *Nature* **472**, 307–312 (2011).
18. W. S. Bakr *et al.*, *Nature* **480**, 500–503 (2011).
19. S. Sachdev, K. Sengupta, S. M. Girvin, *Phys. Rev. B* **66**, 075128 (2002).
20. F. Meinert *et al.*, *Phys. Rev. Lett.* **111**, 053003 (2013).
21. M. Kolodrubetz, D. Pekker, B. K. Clark, K. Sengupta, *Phys. Rev. B* **85**, 100505 (2012).
22. D. Jaksch, C. Bruder, J. I. Cirac, C. W. Gardiner, P. Zoller, *Phys. Rev. Lett.* **81**, 3108–3111 (1998).
23. Materials, methods, and additional theoretical background are available as supporting material on *Science* Online.
24. L. Amico, R. Fazio, A. Osterloh, V. Vedral, *Rev. Mod. Phys.* **80**, 517–576 (2008).
25. S. Trotzky *et al.*, *Nat. Phys.* **8**, 325–330 (2012).
26. C. Kollath, A. M. Läuchli, E. Altman, *Phys. Rev. Lett.* **98**, 180601 (2007).
27. T. Weber, J. Herbig, M. Mark, H.-C. Nägerl, R. Grimm, *Science* **299**, 232–235 (2003).
28. The lattice depth V_q is calibrated by Kapitza-Dirac diffraction. The statistical error for V_q is 1%, although the systematic error can reach up to 5%. We give all energies in frequency units.
29. H. Lignier *et al.*, . *Phys. Rev. Lett.* **99**, 220403 (2007).
30. S. Pielawa, T. Kitagawa, E. Berg, S. Sachdev, *Phys. Rev. B* **83**, 205135 (2011).
31. M. Rigol, V. Dunjko, M. Olshanii, *Nature* **452**, 854–858 (2008).
32. The matter-wave interference pattern is fit by $n(p) = N_0 e^{-(p-p_0)^2/w^2}(1 + V\cos[k(p-p_0) + \phi])$ with the fringe visibility V, the wave vector k, and a phase ϕ.

QUANTUM MECHANICS

Mechanically detecting and avoiding the quantum fluctuations of a microwave field

J. Suh,[1,2] A. J. Weinstein,[1,2] C. U. Lei,[1,2] E. E. Wollman,[1,2] S. K. Steinke,[1,3] P. Meystre,[3] A. A. Clerk,[4] K. C. Schwab[1,2]*

Quantum fluctuations of the light field used for continuous position detection produce stochastic back-action forces and ultimately limit the sensitivity. To overcome this limit, the back-action forces can be avoided by giving up complete knowledge of the motion, and these types of measurements are called "back-action evading" or "quantum nondemolition" detection. We present continuous two-tone back-action evading measurements with a superconducting electromechanical device, realizing three long-standing goals: detection of back-action forces due to the quantum noise of a microwave field, reduction of this quantum back-action noise by 8.5 ± 0.4 decibels (dB), and measurement imprecision of a single quadrature of motion 2.4 ± 0.7 dB below the mechanical zero-point fluctuations. Measurements of this type will find utility in ultrasensitive measurements of weak forces and nonclassical states of motion.

During the theoretical investigation of the ultimate sensitivity of gravitational wave detectors in the 1970s, it was understood how quantum physics places limits on measurements of motion (1, 2). For continuous position measurements, the quantum back-action forces from vacuum fluctuations of the detection light field limit the sensitivity at the standard quantum limit (SQL) (2). Recently, the detection of motion has advanced to the point that quantum back-action engineering has become necessary to improve measurement sensitivity. Position imprecision below the SQL has been achieved in microwave (3) and optical (4) systems, force sensitivity near the quantum limit has been demonstrated with cold atoms (5), and quantum back-action forces from electronic currents (6) or optical fields (7, 8) have been observed. The quantum back-action can be manipulated by modulating the coupling between the detection field and mechanical element, where the coupling modulation can be implemented by either sudden stroboscopic measurement (9, 10) or continuous two-tone back-action evading (BAE) measurement (11–14).

Our system is a microwave-frequency parametric transducer of motion (Fig. 1A): A mechanical resonator ($\omega_m = 2\pi \times 4.0074$ MHz) modulates the capacitance of a superconducting electrical resonator ($\omega_c = 2\pi \times 5.4481$ GHz), and modifies ω_c by 13.8 ± 0.3 Hz ($g_0/2\pi$) per x_{zp}, where $x_{zp} = \sqrt{\hbar/(2m\omega_m)} \approx 1.8$ fm is the amplitude of zero-point fluctuations of the mechanical resonator with mass m (15–17). The intrinsic damping rate of the mechanical resonator is $\Gamma_{m0} = 2\pi \times 10$ Hz at 20 mK. The damping rate of the electrical resonator is $\kappa = 2\pi(869 \pm 9)$ kHz, which places this system into the sideband resolved limit ($\omega_m > \kappa$) required to realize BAE dynamics (14, 18). The excitation and read-out of the electrical resonator is achieved via input and output ports, with coupling rates of $\kappa_L = 2\pi(0.15 \pm 0.04)$ MHz and $\kappa_R = 2\pi(0.5 \pm 0.1)$ MHz, respectively.

We investigate the device by monitoring the response when driven by a carefully engineered microwave field (Fig. 1B). When driving the device with a coherent microwave tone ("pump") at $\omega_p = \omega_c - \omega_m$, the electromechanical coupling together with mechanical motion results in frequency up-conversion of pump photons to ω_c in a Raman-like process at a rate $n_m\Gamma_{opt}$, where n_m is the occupation factor of the mechanical mode, and $\Gamma_{opt} = 4g_0^2 n_p/\kappa$, with n_p as the number of

[1]Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA. [2]Kavli Nanoscience Institute, California Institute of Technology, Pasadena, CA 91125, USA. [3]Department of Physics, University of Arizona, Tucson, AZ 85721, USA. [4]Department of Physics, McGill University, Montreal, Quebec, H3A 2T8 Canada.
*Corresponding author. E-mail: schwab@caltech.edu

pump photons in the electrical resonator. Similarly, when driving at $\omega_p = \omega_c + \omega_m$, photons are down-converted at a rate $(n_m+1)\Gamma_{opt}$ (19). These up- or down-converted sidebands are the signals analyzed; examples of the measured power spectra of these sidebands are shown in Fig. 3.

The sideband powers are calibrated by measuring the thermomechanical fluctuations at calibrated sample temperatures (20, 21) (Fig. 1C). During the thermal calibration, we observe that the thermalization of the weakly pumped device at ultralow temperature is excellent, evidenced by the motional sideband corresponding to 7.2 ± 0.2 mK at the base temperature of our refrigera-

tor (Fig. 1C, red cross and inset). Also, with a back-action damping measurement (19, 21, 22), we calibrate n_p (Fig. 1D). In all the following measurements, we feedback-control the sample temperature at 20.0 ± 0.2 mK with our calibrated RuO resistance thermometer to ensure an accurate and stable thermal environment.

In this system, two-tone BAE is accomplished by modulating the microwave pump field in the resonator: $a_{pump}(t) = \sqrt{2n_p}\cos \omega_m t \cdot \cos \omega_c t = \sqrt{n_p/2}[\cos(\omega_c - \omega_m)t + \cos(\omega_c + \omega_m)t]$. The modulation has the effect of coupling to a single mechanical quadrature, X_1, where $x(t) = X_1(t)\cos \omega_m t + X_2(t)\sin \omega_m t$. The BAE nature

can be understood by noting that the back-action force produced at the mechanical frequency by the beating between voltage noise at the microwave resonance and the large pump field produces displacements exclusively in the X_2 quadrature (18). In this way, one gains information about X_1 and places both the classical and quantum back-action into X_2. Here, the classical back-action is due to thermal and technical noise in the microwave field, quantified by the effective occupation factor n_c. In contrast, the quantum back-action is due to the vacuum fluctuations of the electromagnetic field when the electrical resonator is in the quantum ground state ($n_c = 0$). This back-action
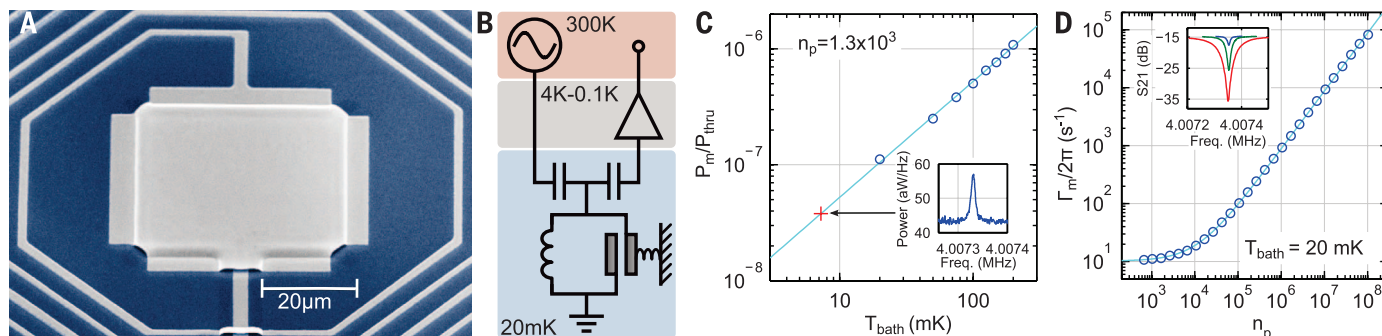
Fig. 1. Device, measurement scheme, and sample characterization. (**A**) Device electron micrograph. A parallel plate capacitor is connected to a spiral inductor, forming a lumped-element microwave resonator. The top plate of the capacitor is a compliant membrane, and we study its fundamental mechanical resonance (blue, silicon; gray, aluminum). (**B**) Measurement scheme. Shot-noise-limited microwave tones are applied to the device, and the output microwave field from the device is amplified by a low-noise amplifier at 4.2 K

for analysis (21). (**C**) Calibration of motional sidebands (blue circles and line). (Inset) Motional sideband at base temperature. (**D**) Back-action damping. In addition to a red-detuned pump, a weak probe sweeping near the electrical resonance is applied, and its absorption shows the resonant mechanical response. Blue circle, mechanical damping rate; blue line, back-action damping theory fit (19). (Inset) Examples of absorption spectra at $n_p \approx 5 \times 10^3$, 3×10^4, and 1×10^5 from top to bottom.
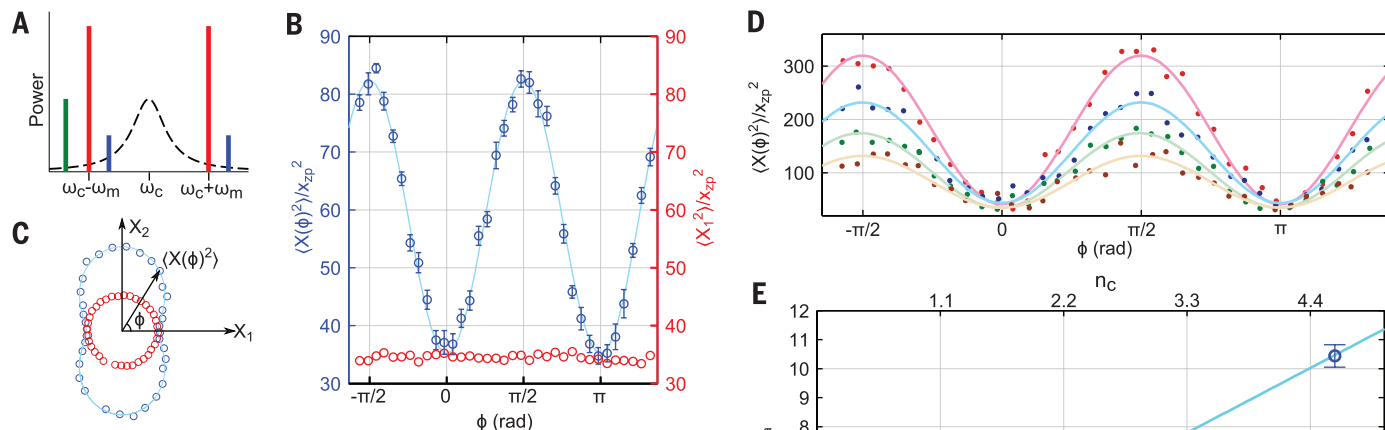
Fig. 2. Measurement of the quadrature-dependent back-action. (**A**) Microwave set-up. Strong BAE pumps (red) are applied symmetrically about the microwave resonance (dashed line). Weak BAE probes (blue) with a small detuning (+30 kHz) from the pumps are used to measure the back-action from the BAE pumps. A weak cooling tone with $\Gamma_{opt} = 2\pi(88 \pm 3)$ Hz is applied at the same time (green). (**B**) An example of measured mechanical fluctuations along the BAE pump axis (red circles) and the probe axis (blue circles). ϕ is the angle between these two axes. The blue line is a fit to $A \sin^2\phi + B$ (21). (**C**) Polar plot of (B), defining X_1 and X_2 along the direction of minimum and maximum fluctuation, respectively. (**D**) Mechanical fluctuations along the probe axis ($\langle X(\phi)^2 \rangle$) at different microwave noise powers: $\Delta\eta = 5.71$, 9.17, 13.41, and 20.99 (±0.04) aW/Hz (brown, green, blue, and red dots, respectively).

(**E**) Back-action in the X_2 quadrature normalized by quantum back-action $\langle X_2^2 \rangle_{qba} = 2(\Gamma_{opt}/\Gamma_m)x_{zp}^2$. Inset, the microwave noise spectrum at microwave occupation factor $n_c \approx 0.9$. $\Delta\eta$ is the noise density from the noise floor at ω_c, proportional to n_c. The sharp peak at the center is the mechanical sideband at $n_p = 1.3 \times 10^6$.

evading mechanism can be also understood by observing that X_1 and X_2 are constants of motion and thus quantum nondemolition observables (12).

The mechanical quadrature fluctuations are computed as

$$\langle X_1^2 \rangle / x_{zp}^2 = 1 + 2n_m + 2n_{bad}$$

$$\langle X_2^2 \rangle / x_{zp}^2 = 1 + 2n_m + 2n_{ba}^{BAE} + 2n_{bad} \qquad (1)$$

where $n_{ba}^{BAE} = (\Gamma_{opt}/\Gamma_m)(2n_c + 1)$, $n_{bad} = [\kappa^2/(32\omega_m^2)]n_{ba}^{BAE}$, and Γ_m is the mechanical damping rate (14). One sees that the uncoupled X_2 quadrature experiences the measurement back-action n_{ba}^{BAE}, proportional to n_p, but the back-action onto X_1 is reduced by $[\kappa^2/(32\omega_m^2)]$, which is expected to be small (1.5×10^{-3}) in our experiment due to the sideband resolution realized in our device ($\kappa/\omega_m = 0.217 \pm 0.002$).

We first study the quadrature-dependent back-action using two sets of BAE tones, denoted as "pump" and "probe" (Fig. 2). The pump tones ($n_p = 1.1 \times 10^6$) are set 20 dB stronger than the probe tones, and we control the relative phase (ϕ) between the beat tones of these two sets. This set-up allows us to control the phase of the mechanical quadrature measured by the probe tones, relative to the strong pump tones. The resulting output spectrum of the probe tones measures $\langle X(\phi)^2 \rangle = \langle X_1^2 \rangle \cos^2 \phi + \langle X_2^2 \rangle \sin^2 \phi$ (21).

Figure 2, B and C, compare the signals from the two sets of BAE tones. The fluctuations at $\phi = \pi/2$ along the X_2 quadrature show maximal increase in fluctuations due to back-action from

the BAE pumps, and the minimum fluctuations at $\phi = 0$ agree with $\langle X_1^2 \rangle$ measured by the pump tones, demonstrating the avoidance of the back-action from the noise in the detection microwave field. The back-action noise into X_2, $\langle X_2^2 \rangle_{ba} = 2n_{ba}^{BAE} x_{zp}^2$, is expected to be proportional to $2n_c + 1$ (Eq. 1), where the first term (n_c) represents classical microwave noise and the second term ($+1$) is from the quantum fluctuations of microwave field. In Fig. 2, D and E, we explore $\langle X_2^2 \rangle_{ba}$ as a function of n_c by injecting microwave noise from room temperature into the device. In the measurement of quadrature-dependent back-action shown in Fig. 2D, we clearly observe an increase of back-action in X_2, $\langle X_2^2 \rangle_{ba} = \langle X(\phi = \pi/2)^2 \rangle - \langle X(\phi = 0)^2 \rangle$, as n_c increases.

To determine n_c, we concurrently measure the noise power density above the amplifier noise floor at ω_c ($\Delta\eta$), which is linearly related to n_c, using a cryogenic microwave switch (21) (Fig. 2E, inset). Figure 2E shows the observed $\langle X_2^2 \rangle_{ba}$ versus $\Delta\eta$. We see that the relationship is linear but does not extrapolate to zero back-action for zero classical microwave noise at $\Delta\eta = 0$: The mechanical device experiences a fluctuating force in excess of the classical noise contribution. The offset we observe is 1.1 ± 0.1, which is calibrated with high accuracy through the mechanical thermal calibration (21). This agrees well with the "+1" in n_{ba}^{BAE} expected from the quantum fluctuations in the microwave field, with a minor contribution of thermal $n_c \approx 0.1$ (21). In addition, the slope in the fit to Eq. 1 yields the calibration of classical microwave noise: $n_c/\Delta\eta = 0.22 \pm 0.02$

$(aW/Hz)^{-1}$. Most important, this measurement constitutes the first demonstration of mechanical detection of the quantum fluctuations of a microwave field (23, 24).

The back-action from both the classical and quantum noise of the detection electromagnetic field is avoided in a single-quadrature measurement of motion with two-tone BAE. However, in general, the mechanical motion can also experience measurement back-action from sources other than the noise in the detection field. For instance, the increase in mechanical fluctuations due to thermal dissipation in the device is not avoided in BAE. To quantify these other back-action effects, and measure the total back-action into X_1, we measure the initial thermal motion and observe the increase in $\langle X_1^2 \rangle$. We also measure $\langle x^2 \rangle$ in a non-BAE configuration to compare the back-action (21). As expected, we observe that the BAE spectrum exhibits reduced back-action compared to the non-BAE case (Fig. 3, A and B). In Fig. 3C, we plot the initial thermal motion $\langle x^2 \rangle_0$, $\langle x^2 \rangle$ from non-BAE, and $\langle X_1^2 \rangle$ from BAE at various pump strengths. In Fig. 3D, the total measurement back-action, $\langle x^2 \rangle_{ba}$ and $\langle X_1^2 \rangle_{ba}$, and the measurement imprecision, $\langle x^2 \rangle_{imp}$ and $\langle X_1^2 \rangle_{imp}$, are plotted. Here, we define n_p as $n_p = |a_{pump}^2|/2$ for BAE and $n_p = |a_{pump}^2|/4$ for non-BAE in order to compare the imprecision and back-action in these two pump configurations properly (21).

In non-BAE, as the imprecision decreases, the fluctuations due to back-action increase (blue dots in Fig. 3D). At $n_p = 2.3 \times 10^6$, the mechanical occupation increases from 13.0 ± 0.5 to $68.5 \pm$

Fig. 3. Back-action evading measurement. (**A**) An example of measured motional sidebands in two-tone non-BAE (blue) and BAE (red) configurations. (Inset) The blue area under the motional sideband represents the mechanical fluctuations $\langle x^2 \rangle$. Imprecision ($\langle x^2 \rangle_{imp}$) is the additive noise inferred from the measurement noise floor (S_x^0), equal to the gray area under a Lorentzian with its peak height S_x^0 and linewidth Γ_m. (**B**) The increase in the motional sideband area either in non-BAE (blue) or in BAE (red), compared to the no-pump case (green), identifies the total measurement back-action. The spectra are centered at the mechanical resonance. (**C**) The mechanical fluctuations measured in non-BAE (blue) and BAE (red) at various pump strengths. (**D**) Measurement imprecision (circles) and back-action (dots) of non-BAE (blue) and BAE (red). The solid blue line represents a fit to the measured back-action, including classical noise in the electrical resonator. The solid green line is the expected quantum back-action from microwave shot noise. The dashed blue line shows a fit to the measured imprecision, and the dashed green line is the imprecision at the quantum limit. (Inset) The spectrum indicating BAE imprecision below the zero point.

0.1, consistent with a small finite microwave occupation factor ($n_c \approx 0.6 \pm 0.1$) in addition to the quantum fluctuations. In contrast, we do not observe a large increase in the mechanical fluctuations in BAE as the imprecision decreases (red dots in Fig. 3D). The expected back-action into the measured quadrature due to the finite sideband resolution (14) is $0.12 x_{zp}^2$ at $n_p = 4.7 \times 10^6$. The measured back-action of $\langle X_1^2 \rangle_{ba} \approx 10 x_{zp}^2$ is likely due to thermal dissipation in our device. Nonetheless, we demonstrate avoidance of the back-action noise by 10.7 ± 0.3 dB at $n_p = 2.3 \times 10^6$ compared with non-BAE. Most important, we show that the back-action $\langle X_1^2 \rangle_{ba}$ is 8.5 ± 0.4 dB below the level set by quantum fluctuations of the microwave field, $2(\Gamma_{opt}/\Gamma_m)x_{zp}^2$, at $n_p = 4.7 \times 10^6$.

In addition, the quadrature imprecision is below x_{zp}^2 at this point: $\langle X_1^2 \rangle_{imp} = (0.57 \pm 0.09)x_{zp}^2$ (Fig. 3D, inset). This is approximately a factor of 200 above that of quantum-limited imprecision, which is consistent with the detection efficiency determined by $\kappa_R/\kappa \approx 0.5$, the microwave loss between the device and amplifier ($\approx$2 dB), and the noise temperature ($\approx$4 K) of the amplifier at 4 K stage.

Quantum fluctuations of a single mode in a microwave field can be sensed with a mechanical system, and, with the proper pump field, these fluctuations can be steered into an unobserved aspect of the motion. These results lead the way toward measuring and manipulating the quantum noise of a mechanical resonator. As described in (14), feedback on the motion may be applied to produce a squeezed state from a thermal state. Using a nearly quantum-limited amplifier (25), we expect the generation of a squeezed state ($\langle X_1^2 \rangle / x_{zp}^2 < 1$) with $n_p \approx 10^5$, which would be useful for detection of weak forces and fundamental studies of quantum decoherence (26). Finally, we note that the mechanical mode reaching 7.2 mK demonstrates a new application of a micromechanical resonator as a primary ultralow-temperature thermometer.

REFERENCES AND NOTES

1. R. W. P. Drever, J. Hough, W. A. Edelstein, J. R. Pugh, W. Martin, *Proc. Int. Mtg on Experimental Gravitation, Pavia, Italy*, B. Bertotti, Ed. (Accademia Nazionale Dei Lincei, Rome, 1976).
2. C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, M. Zimmermann, *Rev. Mod. Phys.* **52**, 341–392 (1980).
3. J. D. Teufel, T. Donner, M. A. Castellanos-Beltran, J. W. Harlow, K. W. Lehnert, *Nat. Nanotechnol.* **4**, 820–823 (2009).
4. G. Anetsberger et al., *Phys. Rev. A* **82**, 061804 (2010).
5. S. Schreppler et al., Optically measuring force near the standard quantum limit. arXiv:1312.4896 (2013).
6. A. Naik et al., *Nature* **443**, 193–196 (2006).
7. K. W. Murch, K. L. Moore, S. Gupta, D. M. Stamper-Kurn, *Nat. Phys.* **4**, 561–564 (2008).
8. T. P. Purdy, R. W. Peterson, C. A. Regal, *Science* **339**, 801–804 (2013).
9. V. B. Braginsky, F. Y. Khalili, *Quantum Measurement* (Cambridge Univ. Press, Cambridge, 1992).
10. R. Ruskov, K. Schwab, A. N. Korotkov, *Phys. Rev. B* **71**, 235407 (2005).
11. V. B. Braginsky, Y. I. Vorontsov, *Sov. Phys. Usp.* **17**, 644–650 (1975).
12. K. S. Thorne, R. W. P. Drever, C. M. Caves, M. Zimmermann, V. D. Sandberg, *Phys. Rev. Lett.* **40**, 667–671 (1978).
13. V. B. Braginsky, Y. I. Vorontsov, K. S. Thorne, *Science* **209**, 547–557 (1980).
14. A. A. Clerk, F. Marquardt, K. Jacobs, *New J. Phys.* **10**, 095010 (2008).
15. J. D. Teufel et al., *Nature* **475**, 359–363 (2011).
16. J. Suh, M. D. Shaw, H. G. Leduc, A. J. Weinstein, K. C. Schwab, *Nano Lett.* **12**, 6260–6265 (2012).
17. J. Suh, A. J. Weinstein, K. C. Schwab, *Appl. Phys. Lett.* **103**, 052604 (2013).
18. M. F. Bocko, R. Onofrio, *Rev. Mod. Phys.* **68**, 755–799 (1996).
19. F. Marquardt, J. P. Chen, A. A. Clerk, S. M. Girvin, *Phys. Rev. Lett.* **99**, 093902 (2007).
20. J. B. Hertzberg et al., *Nat. Phys.* **6**, 213–217 (2010).
21. Materials and methods are available as supplementary materials on *Science* Online.
22. T. Rocheleau et al., *Nature* **463**, 72–75 (2010).
23. R. H. Koch, D. J. Van Harlingen, J. Clarke, *Phys. Rev. B* **26**, 74–87 (1982).
24. A. Fragner et al., *Science* **322**, 1357–1360 (2008).
25. J. Y. Mutus et al., *Appl. Phys. Lett.* **103**, 122602 (2013).
26. B. L. Hu, Y. Zhang, *Mod. Phys. Lett. A* **8**, 3575–3584 (1993).

EARTH'S INTERIOR

Dehydration melting at the top of the lower mantle

Brandon Schmandt,[1]* Steven D. Jacobsen,[2]* Thorsten W. Becker,[3] Zhenxian Liu,[4] Kenneth G. Dueker[5]

The high water storage capacity of minerals in Earth's mantle transition zone (410- to 660-kilometer depth) implies the possibility of a deep H_2O reservoir, which could cause dehydration melting of vertically flowing mantle. We examined the effects of downwelling from the transition zone into the lower mantle with high-pressure laboratory experiments, numerical modeling, and seismic P-to-S conversions recorded by a dense seismic array in North America. In experiments, the transition of hydrous ringwoodite to perovskite and (Mg,Fe)O produces intergranular melt. Detections of abrupt decreases in seismic velocity where downwelling mantle is inferred are consistent with partial melt below 660 kilometers. These results suggest hydration of a large region of the transition zone and that dehydration melting may act to trap H_2O in the transition zone.

The water content of the upper mantle as sampled by mid-ocean ridge basalts is 0.005 to 0.02 weight % (wt %) (1), but a potentially much larger, deep reservoir of water may exist in the mantle transition zone between 410- and 660-km depth owing to the 1 to 3 wt % H_2O storage capacity of the major mineral phases wadsleyite and ringwoodite (2, 3). Convective mass transfer across the boundaries of the transition zone could cause dehydration melting, and consequently filtering of incompatible elements, if water contents in the transition zone exceed that of the shallower or deeper mantle (4). An open question is whether transition zone water contents are sufficient to cause dehydration melting where there is downward flow into the lower mantle.

Dehydration melting due to downward flow across the 660-km discontinuity (660) would require both hydration of ringwoodite in the transition zone and low water storage capacity at the top of the lower mantle. The recent discovery of a ~1.5 wt % H_2O hydrous ringwoodite inclusion in a diamond (5) demonstrates that, at least locally, the mantle transition zone may be close to water saturation. Regional detections of high seismic attenuation (6) and electrical conductivity (7) in the transition zone suggest hydration at larger scales. However, high-pressure experiments on the incorporation of H_2O into silicate perovskite vary widely from 0.0001 wt % (8) to 0.4 wt % H_2O (9), with other estimates in between (10). Recent experiments on coexisting phase assemblages indicate a high H_2O partition coefficient between ringwoodite and silicate perovskite of 15:1 (11), suggesting a large contrast in water storage capacity at the boundary between the base of the transition zone and the top of the lower mantle.

We integrated laboratory experiments, seismic imaging, and numerical models of mantle flow to investigate mass transfer and melting at the interface between the transition zone and lower mantle beneath North America (12). We conducted in situ

[1]Department of Earth and Planetary Science, University of New Mexico, Albuquerque, NM, USA. [2]Department of Earth and Planetary Sciences, Northwestern University, Evanston, IL, USA. [3]Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA. [4]Geophysical Laboratory, Carnegie Institution of Washington, Washington DC, USA. [5]Department of Geology and Geophysics, University of Wyoming, Laramie, WY, USA.
*Corresponding author. E-mail: bschmandt@unm.edu (B.S.); steven@earth.northwestern.edu (S.D.J.)

laser heating experiments to directly transform hydrous ringwoodite to form silicate perovskite and (Mg,Fe)O in a diamond-anvil cell (DAC) and analyzed the recovered sample with synchrotron–Fourier transform infrared (FTIR) spectroscopy and transmission electron microscopy (TEM) (Fig. 1). FTIR spectra taken away from the laser-heated spots are typical for hydrous Fe–bearing ringwoodite (13). Within the laser-heated spots where perovskite and (Mg,Fe)O formed, there is still strong absorption in the OH-stretching region, although notably different from absorption that occurred before heating (Fig. 1B). Within laser-heated spots, the maximum in OH-stretching absorbance occurs at ~3400 cm^{-1}, and there is a sharp peak at

~3680 cm^{-1} associated with brucite (8). Both features are common to the spectra of previous studies (9, 10) that reported 0.1 to 0.4 wt % H_2O in perovskite. Recovery of the sample from the DAC allowed detailed study by TEM (Fig. 1C), which shows that nanoscale, intergranular silicate melt was formed around single crystals of perovskite. The broad, asymmetric absorption band observed in Fig. 1B thus likely represents OH in the melt phase, because the partition coefficient of H_2O between ringwoodite and silicate perovskite is about 15:1 (11).

Partial melt in the mantle strongly affects seismic velocities and can create sharp velocity gradients where it is adjacent to subsolidus mantle. If

dehydration melting occurs where ringwoodite is entrained downward across the 660, then in situ detection would be feasible in areas with dense seismic sampling. A major component of the Earth-Scope project (14) is the deployment of broadband seismometers with ~70-km spacing across the United States (Fig. 2A). These data enable imaging of geographic variations in seismic structure near the 660. We isolated conversions of earthquake seismic waves from P-to-S (Ps) as a result of sharp vertical gradients in seismic velocity with receiver function analysis (15). Ps receiver functions were then mapped to depth using P and S tomography models, creating a high-fidelity, common conversion point (CCP) image of vertical velocity gradients near
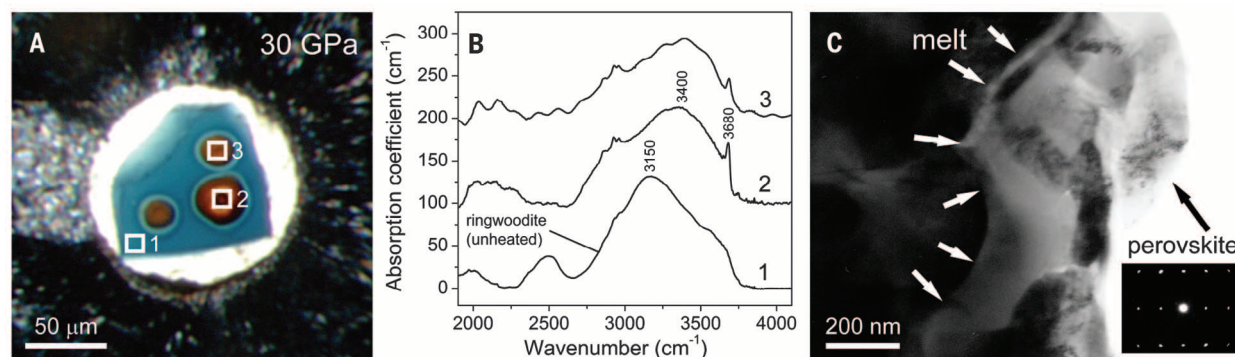
Fig. 1. Laboratory experiments on hydrous ringwoodite. (**A**) Single-crystal of hydrous ringwoodite (blue crystal) containing 1 wt % H_2O inside a DAC at 30 GPa. The sample was laser heated to 1600°C in several spots (orange circles) to perform direct transformation to perovskite and (Mg,Fe)O. (**B**) Synchrotron-FTIR spectra of the recovered sample in three locations: an unheated part of the crystal (spectrum 1) and two locations within laser-heated spots (spectra 2 and 3). FTIR spectra were collected with a 10 μm by 10 μm aperture, illustrated and numbered by white boxes in (A). (**C**) TEM within a laser-heated spot (position 2) shows crystals of perovskite and intergranular amorphous quench (melt).
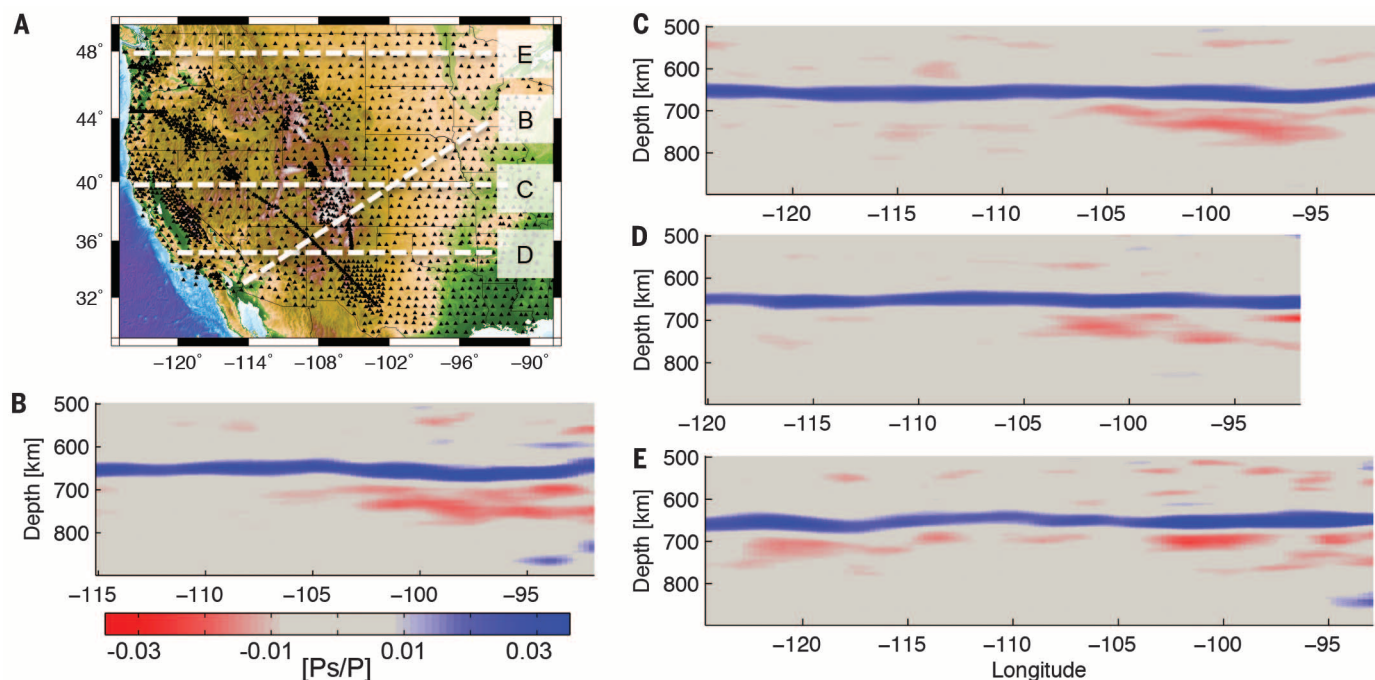
Fig. 2. Vertical cross-sections through the CCP image. (**A**) Map of the study region with broadband seismometers denoted by black triangles and the locations of vertical cross-sections shown in (B) to (E) denoted by white dashed lines. (**B**) Vertical cross-section through the CCP image. The location of the eastern end of the cross-section is shown as B′ in (A). The 660, which separates the transition zone and lower mantle, is clearly imaged by 2.5 to 4% amplitude positive Ps/P arrivals in all cross-sections. Negative arrivals beneath the 660 have amplitudes ≤2.2%. (**C** to **E**) Vertical cross-sections through the CCP image beneath the three other lines labeled in (A).

the 660 beneath much of the continental United States (12, 16).

Positive-amplitude Ps conversions clearly define the abrupt velocity increase with depth at the 660 beneath the entire array (Fig. 2), and other, weaker discontinuities are sporadically detected in the mantle beneath the EarthScope array (16, 17). Common secondary features in the CCP image near the 660 are negative-amplitude Ps conversions in the uppermost 120 km of the lower mantle. A map of the locations of sub-660 negative Ps conversions with magnitude >1.25% of the direct P-wave shows that these features are prevalent beneath a large area including the Great Plains and near the northern margin of the array (Figs. 2 and 3). In contrast, negative Ps conversions near the top of the lower mantle are absent beneath the southwestern United States. The amplitude of the negative Ps conversions from beneath the 660 is up to 2.2% of the direct P-wave

amplitude. For a negative Ps conversion near the top of the lower mantle, 730 km, an amplitude of 2% is consistent with a decrease in shear velocity of 2.6% over a depth interval of ≤20 km based on a synthetic calculation of receiver functions (12). The areas where negative Ps conversions are detected are not correlated with volumes of anomalously low-velocity mantle in tomography images (16, 18); hence, a thermal origin for the reduced velocities is unlikely.

To assess the correlation between the locations of abrupt velocity decreases near the top of the lower mantle and convective flow patterns, we used numerical models of mantle circulation. Computations were solved for flow using radially varying viscosity and prescribed plate motions at the surface (12, 19). Regional mantle flow rates are controlled by the poorly constrained ratio between density anomalies and viscosity, but flow directions are mainly determined by density

anomaly patterns. To explore this sensitivity, we used multiple tomography models (15, 17, 19) to infer density fields. For length scales less than ~500 km, the resulting vertical flow patterns at 660-km depth vary widely between models, but circulation models yield a common long-wavelength pattern of vertical flow across the 660 (Fig. 3). Downward flow through the 660 is dominant beneath the Great Plains and along the U.S.-Canada border, and upward flow is dominant beneath the southwestern United States (Fig. 3). This pattern is driven by large volumes of anomalously high-velocity lower mantle, which are inferred to be subducted slabs sinking beneath central and eastern North America (20–22).

Comparing the results of two flow models shows that nearly all the locations of negative Ps conversions beneath the 660 coincide with downward flow (Fig. 3). One flow model is based on an inversion of travel-time data from the EarthScope stations using the TX-2008 global tomography model (20) as a starting model (Fig. 3A, SH11-TX). The other flow model is based on density structure from the global tomography model S40RTS (18), which does not include EarthScope seismic data and hence only recovers long-wavelength structure (Fig. 3B). Consequently, the first flow model includes stronger short-wavelength variations (19), but the long-wavelength pattern is consistent. Comparing the distributions of vertical velocities across the 660 in the entire CCP image and the subset of the CCP image where sub-660 negative Ps conversions are detected demonstrates that the latter area is biased toward regions of downward flow (Fig. 4). For both flow models, less than 3% of sub-660 negative Ps conversions are found in areas with upward flow across the 660. Comparisons with smooth three-dimensional velocity structure at 660-km depth, rather than vertical flow velocity, show some bias toward higher tomographically imaged shear velocities in areas with sub-660 negative Ps conversions, but less similarity between the two models than is observed for the comparison with vertical flow.

The correlation between abrupt seismic velocity decreases near the top of the lower mantle and
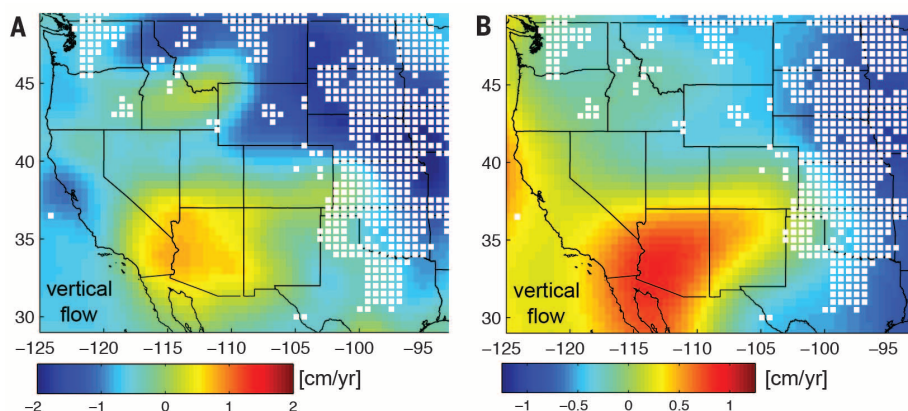
Fig. 3. Vertical flow between the transition zone and lower mantle. (A) The background color shows the vertical flow velocity at the boundary between the transition zone and lower mantle predicted by a mantle circulation model using density structure inferred from the SH11-TX tomography model. White squares denote locations where the seismic CCP image detects velocity decreases with depth in the depth range of 670 to 800 km. (B) The background color shows the vertical flow velocity predicted using the S40RTS tomography model. In both models, velocity decreases near the top of the lower mantle are absent beneath the southwestern United States where upward flow is predicted and prevalent in areas where downward flow into the lower mantle is predicted.
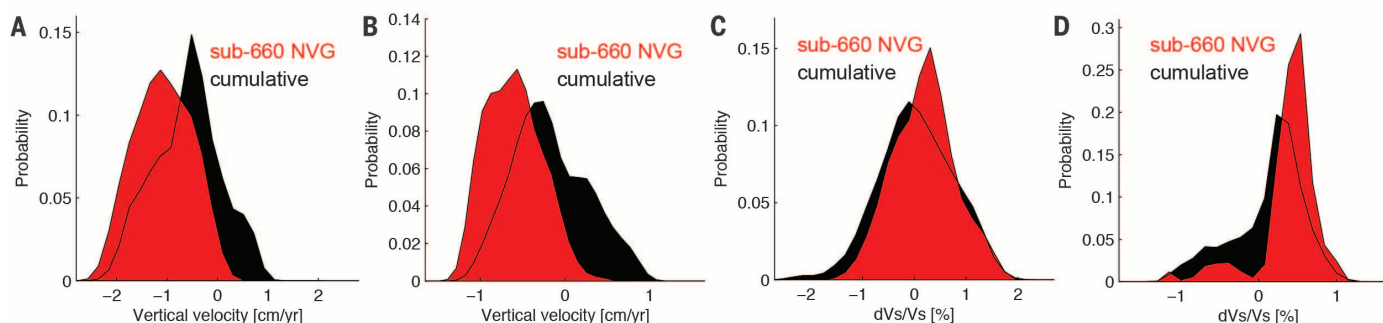
Fig. 4. Distributions of vertical flow and seismic velocity variations at 660 km. (A) Probability density functions (PDFs) of vertical flow velocity at 660-km depth from the convection model using SH11-TX tomography. The PDF in red represents the subset of the CCP image where negative velocity gradients are detected beneath the 660 (sub-660 NVG), and the PDF in black represents the cumulative area from the CCP image. The

cumulative area of the CCP image includes the area within 200 km of the seismometers shown in Fig. 2A. (B) PDFs of vertical flow velocity from the convection model using S40RTS tomography. (C) PDFs of mantle shear velocity variations near the 660 from the tomography model SH11-TX. (D) PDFs of mantle shear velocity variations from the tomography model S40RTS.

areas of downwelling across the 660 is consistent with the occurrence of dehydration melting as observed in our laboratory experiments. An alternative bulk-compositional origin of low velocities near the top of the lower mantle is segregated basalt that may be neutrally buoyant (*23*) and would reduce seismic velocities (*24*).

However, long-term accumulation of basalt near the top of the lower mantle is not expected to be preferentially present where there is downwelling across the 660 and absent where there is not. The areas of downward flow across 660 do not all coincide with local presence of subducted slabs, so a direct link to composition of the sinking Farallon slab cannot explain the negative velocity gradients below 660. Assuming that the velocity reductions result from partial melt, and that the shear-velocity decrease per percent of melt is between 2.6 and 3.8%, as predicted for partial melt near 400-km depth (*25*), then 0.68 to 1% melt could explain a 2.6% shear velocity reduction indicated by negative *Ps* conversions with amplitude of 2% in the CCP image.

Prediction of partial melt percentages at 660-km depth for various H_2O contents requires knowledge of water partition coefficients between minerals and melts at relevant pressure-temperature (*P-T*) conditions in the peridotite-saturated compositional system. At present, experiments in the hydrous peridotite system at conditions near the 660 have not been performed. However, using experimental results for partial melting near the 410-km discontinuity (410) in a bulk peridotite system with 1 wt % H_2O indicates that ~5% partial melt at 410 km is expected (*26, 27*) where the partition coefficient of H_2O between wadsleyite and olivine is at least 5:1 (*11*). We can expect at least 5% partial melt in a bulk 1 wt % H_2O peridotite system where the partition coefficient between ringwoodite and silicate perovskite is 15:1 (*11*). Thus, production of up to 1% melt by dehydration melting of hydrous ringwoodite viscously entrained into the lower mantle is feasible.

The density of hydrous melt near the top of the lower mantle is uncertain, but it is likely buoyant with respect to the top of the lower mantle (*28*). Hence, we expect that the velocity decreases imaged beneath the 660 are transient features resulting from ongoing downward flow through the 660 that is driven by sinking slabs in the lower mantle. Eventually, the slightly buoyant hydrous melt would percolate upward, returning H_2O to the transition zone (*4*). Dehydration melting has also been suggested to occur where hydrous wadsleyite upwells across the 410 and into the olivine stability field (*3, 27*). Experiments indicate that hydrous melt is gravitationally stable atop the 410 (*28*), so once melt is generated, it may remain or spread laterally rather than maintaining a clear correlation with ongoing vertical flow patterns. Seismic detections of a low-velocity layer atop the 410 are common but laterally sporadic beneath North America and globally (*29, 30*). The combination of dehydration melting driven by downwelling across the 660 and upwelling across the 410 could create a long-term H_2O trap in the transition zone (*4*).

REFERENCES AND NOTES

1. A. E. Saal, E. H. Hauri, C. H. Langmuir, M. R. Perfit, *Nature* **419**, 451–455 (2002).
2. J. R. Smyth, *Am. Mineral.* **72**, 1051 (1987).
3. D. L. Kohlstedt, H. Keppler, D. C. Rubie, *Contrib. Mineral. Petrol.* **123**, 345–357 (1996).
4. D. Bercovici, S. Karato, *Nature* **425**, 39–44 (2003).
5. D. G. Pearson et al., *Nature* **507**, 221–224 (2014).
6. H. Zhu, E. Bozdag, T. S. Duffy, J. Tromp, *Earth Planet. Sci. Lett.* **381**, 1–11 (2013).
7. A. Kelbert, A. Schultz, G. Egbert, *Nature* **460**, 1003–1006 (2009).
8. N. Bolfan-Casanova, H. Keppler, D. C. Rubie, *Geophys. Res. Lett.* **30**, 1905 (2003).
9. M. Murakami, K. Hirose, H. Yurimoto, S. Nakashima, N. Takafuji, *Science* **295**, 1885–1887 (2002).
10. K. Litasov et al., *Earth Planet. Sci. Lett.* **211**, 189–203 (2003).
11. T. Inoue, T. Wada, R. Sasaki, H. Yurimoto, *Phys. Earth Planet. Inter.* **183**, 245–251 (2010).
12. Materials and methods are available on *Science* Online.
13. S. D. Jacobsen, J. R. Smyth, H. A. Spetzler, C. M. Holl, D. J. Frost, *Phys. Earth Planet. Inter.* **143-144**, 47–56 (2004).
14. R. A. Kerr, *Science* **340**, 1283–1285 (2013).
15. J. P. Mercier, M. G. Bostock, A. M. Baig, *Geophysics* **71**, SI95–SI102 (2006).
16. B. Schmandt, K. G. Dueker, E. D. Humphreys, S. M. Hansen, *Earth Planet. Sci. Lett.* **331-332**, 224–236 (2012).
17. B. Tauzin, R. D. van der Hilst, G. Wittlinger, Y. Ricard, *J. Geophys. Res. Solid Earth* **118**, 2307–2322 (2013).
18. J. Ritsema, A. Deuss, H. J. van Heijst, J. H. Woodhouse, *Geophys. J. Int.* **184**, 1223–1236 (2011).
19. T. W. Becker, C. Faccenna, E. D. Humphreys, A. R. Lowry, M. S. Miller, *Earth Planet. Sci. Lett.* (2014).
20. N. A. Simmons, A. M. Forte, S. P. Grand, *Geophys. J. Int.* **177**, 1284–1304 (2009).
21. B. Steinberger, *Phys. Earth Planet. Inter.* **118**, 241–257 (2000).
22. A. M. Forte, R. Moucha, N. A. Simmons, S. P. Grand, J. X. Mitrovica, *Tectonophysics* **481**, 3–15 (2010).
23. C. T. A. Lee, W. P. Chen, *Earth Planet. Sci. Lett.* **255**, 357–366 (2007).
24. W. B. Xu, C. Lithgow-Bertelloni, L. Stixrude, J. Ritsema, *Earth Planet. Sci. Lett.* **275**, 70–79 (2008).
25. S. Hier-Majumder, A. Courtier, *Earth Planet. Sci. Lett.* **308**, 334–342 (2011).
26. M. M. Hirschmann, T. Tenner, C. Aubaud, A. C. Withers, *Phys. Earth Planet. Inter.* **176**, 54–68 (2009).
27. T. J. Tenner, M. M. Hirschmann, A. C. Withers, P. Ardia, *Contrib. Mineral. Petrol.* **163**, 297–316 (2012).
28. T. Sakamaki, A. Suzuki, E. Ohtani, *Nature* **439**, 192–194 (2006).
29. B. Schmandt, K. G. Dueker, S. M. Hansen, J. J. Jasbinsek, Z. Zhang, *Geochem. Geophys. Geosyst.* **12**, Q08014 (2011).
30. B. Tauzin, E. Debayle, G. Wittlinger, *Nat. Geosci.* **3**, 718–721 (2010).

DINOSAUR PHYSIOLOGY

Evidence for mesothermy in dinosaurs

John M. Grady,[1]* Brian J. Enquist,[2,3] Eva Dettweiler-Robinson,[1] Natalie A. Wright,[1] Felisa A. Smith[1]

Were dinosaurs ectotherms or fast-metabolizing endotherms whose activities were unconstrained by temperature? To date, some of the strongest evidence for endothermy comes from the rapid growth rates derived from the analysis of fossil bones. However, these studies are constrained by a lack of comparative data and an appropriate energetic framework. Here we compile data on ontogenetic growth for extant and fossil vertebrates, including all major dinosaur clades. Using a metabolic scaling approach, we find that growth and metabolic rates follow theoretical predictions across clades, although some groups deviate. Moreover, when the effects of size and temperature are considered, dinosaur metabolic rates were intermediate to those of endotherms and ectotherms and closest to those of extant mesotherms. Our results suggest that the modern dichotomy of endothermic versus ectothermic is overly simplistic.

Over the past few decades, the original characterization of dinosaurs by early paleontologists as lumbering, slow-metabolizing ectotherms has been challenged. Recent studies propose that dinosaurs were capable of an active lifestyle and were metaboli-cally similar to endothermic mammals and birds (*1–3*). This debate is of more than heuristic interest; energy consumption is closely linked to life history, demographic, and ecological traits (*4*). Extant endothermic mammals and birds possess metabolic rates ~5 to 10 times higher than

those of reptiles and fish (5, 6), but characterizing the metabolic rates of dinosaurs has been difficult.

A promising method for inferring paleoenergetics comes from studies of ontogenetic growth, in which age is determined from annual rings in bone cross sections and mass is determined from bone dimensions. Ultimately, growth is powered by metabolism, and rates of growth and energy use should correspond. Pioneering work by Erickson and others has led to a growing body of literature on dinosaur growth and generated important insights (7, 8). However, many analyses were hampered by small samples, an outdated comparative data set, and the lack of an appropriate energetic framework. Increasing data availability permits a reassessment of dinosaur growth against a broader spectrum of animals, standardized for environmental temperature. Further, recent advances in metabolic theory provide a theoretical framework for evaluating metabolic rate on the basis of growth.

We used a comparative approach to characterize the energetics of dinosaurs and other extinct taxa. We examined the empirical and theoretical relationship between growth and resting metabolic rate, using a broad database of major vertebrate clades (9), and used our results to examine the energetics of Mesozoic dinosaurs. From empirical studies, we constructed ontogenetic growth curves and determined a maximum rate of growth for each species. Environmental temperature was standardized by only considering growth rates in ectotherms from tropical and subtropical climates or from laboratory settings between 24° and 30°C, comparable to temperatures experienced by dinosaurs during the Mesozoic (10). Data for dinosaur growth were taken from published reports that provided a minimum of five measurements of size and age. All metabolic rates were converted to watts (W). Where multiple metabolic or maximum growth rates for a species were recorded, the geometric mean was determined. Overall, our data set includes ~30,000 values and was used to characterize growth for 381 species, including 21 species of Mesozoic dinosaurs, 6 extinct crocodilians, and a Cretaceous shark (table S1). Dinosaurs are well represented both temporally (late Triassic to end-Cretaceous) and taxonomically (Theropoda, Sauropodomorpha, Ornithopoda, and Ceratopsia). Values for resting metabolic rates were compiled from the literature and standardized to a common temperature of 27°C (table S1). We performed phylogenetic independent contrasts (PICs) in addition to conventional ordinary least-squares regression (OLS) and standardized major axis regression (table S2).

Data show, within and across species, that resting metabolic rate B scales with body mass m as a power function, $B = B_0 m^\alpha$, where B_0 is

[1]Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA. [2]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA. [3]The Santa Fe Institute, USA, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.
*Corresponding author. E-mail: jgrady@unm.edu

Fig. 1. The scaling of maximum growth rate in vertebrates. (**A**) Growth rates of thermoregulatory guilds. Red indicates endothermy; blue, ectothermy, gray, dinosaurs; and black, mesothermy. (**B**) Vertebrate taxa scaling with 95% confidence bands. The red dashed line indicates marsupials, and the black dashed line is tuna; all other taxa are labeled. See table S2 for regression parameters and statistics.
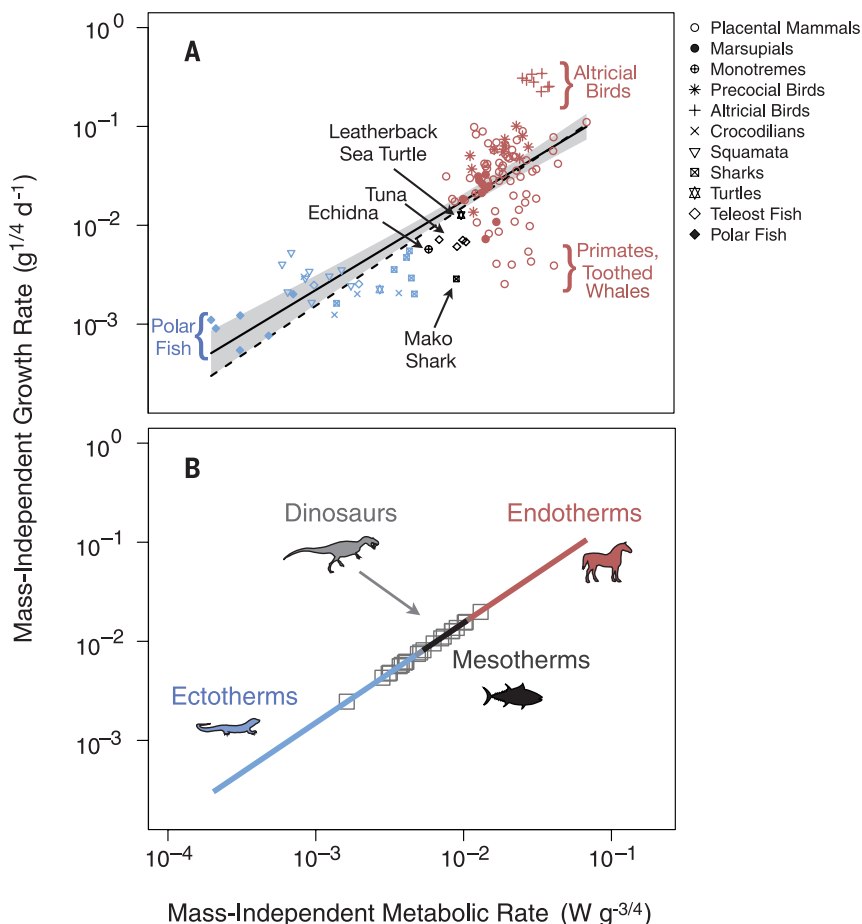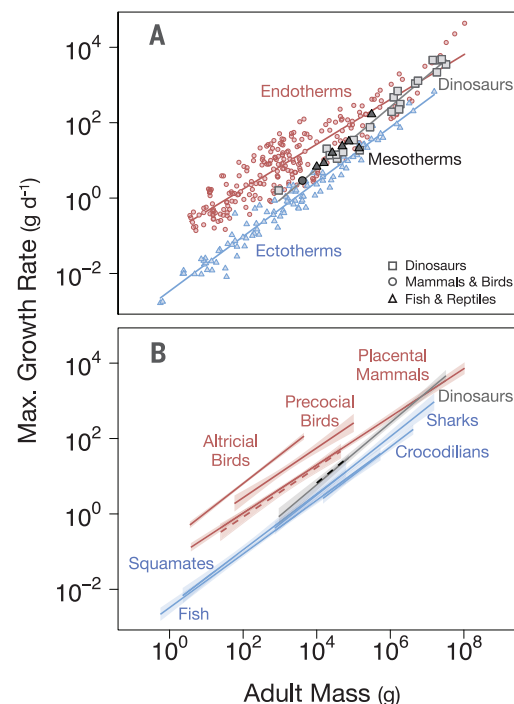
Fig. 2. Vertebrate growth energetics. (**A**) Relationship between growth and resting metabolic rate for vertebrates. The dashed line is the theoretical prediction; the solid line represents an OLS fitted regression with 95% confidence bands. (**B**) Predicted energetics of dinosaurs. Dinosaur rates (open squares) from Eq. 2 are plotted on the theoretical line. The ranges in metabolic rates occupied by extant endotherms, mesotherms, and ectotherms are indicated by color.

a normalization constant representing mass-independent metabolic rate, and α is ~3/4 and ranges from 0.65 to 0.85 (*11, 12*). Growth rate varies over ontogeny, but use of the maximum growth rate (G_{max}) standardizes growth and permits interspecific comparisons. Empirical evidence (*13*) indicates that G_{max} scales similarly to B, where $G_{max} = G_0 M^\alpha$. This suggests that $B \propto G_{max}^1$ and thus that metabolic rate may be inferred from growth. However, the relationship between G_{max} and B across major vertebrate taxa has received little attention, and many uncertainties exist. For instance, Case (*13*) reported that fish G_{max} was an order of magnitude lower than that of reptiles, despite similarities in metabolic and thermoregulatory lifestyle (*6*).

Theoretical assessments of growth complement a strictly empirical approach and can strengthen paleontological inferences. An ontogenetic growth model based on metabolic scaling theory (MST) quantifies the linkages between G_{max} and metabolic rate from first principles of allometry and conservation of energy (*14, 15*). According to MST (*9*), the relationship between B (W) and G_{max} (g day^{-1}) at final adult mass M is

$$B_M = cG_{max}^1 \qquad (1)$$

where $c \approx 0.66$ (W g^{-1} day). To observe the mass-independent relationship and compare energetic groups, we divide both sides by M^α, yielding

$$B_0 = cG_0 \qquad (2)$$

To calculate metabolic rate at any ontogenetic mass m from the observed maximum growth rate, we combine Eqs. 1 and 2

$$B_m = cG_0 m^{3/4} \qquad (3)$$

Fig. 3. Resting metabolic rates in vertebrates. (**A**) Predicted metabolic rates compared to observed rates. The solid line is the fitted regression, with shaded 95% confidence bands; the dashed line is the theoretical fit. (**B**) Metabolic scaling of vertebrates. Dinosaur resting metabolic rates are predicted from growth (dashed line); all other fits are predicted from empirical data. Endotherms: $y = 0.019x^{0.75}$, $r^2 = 0.98$, $n = 89$; Ectotherms (27°C): $y = 0.00099x^{0.84}$, $r^2 = 0.95$, $n = 22$; Dinosaurs: $y = 0.0020x^{0.82}$, $r^2 = 0.96$, $n = 21$. $P < 0.001$ for all regressions.
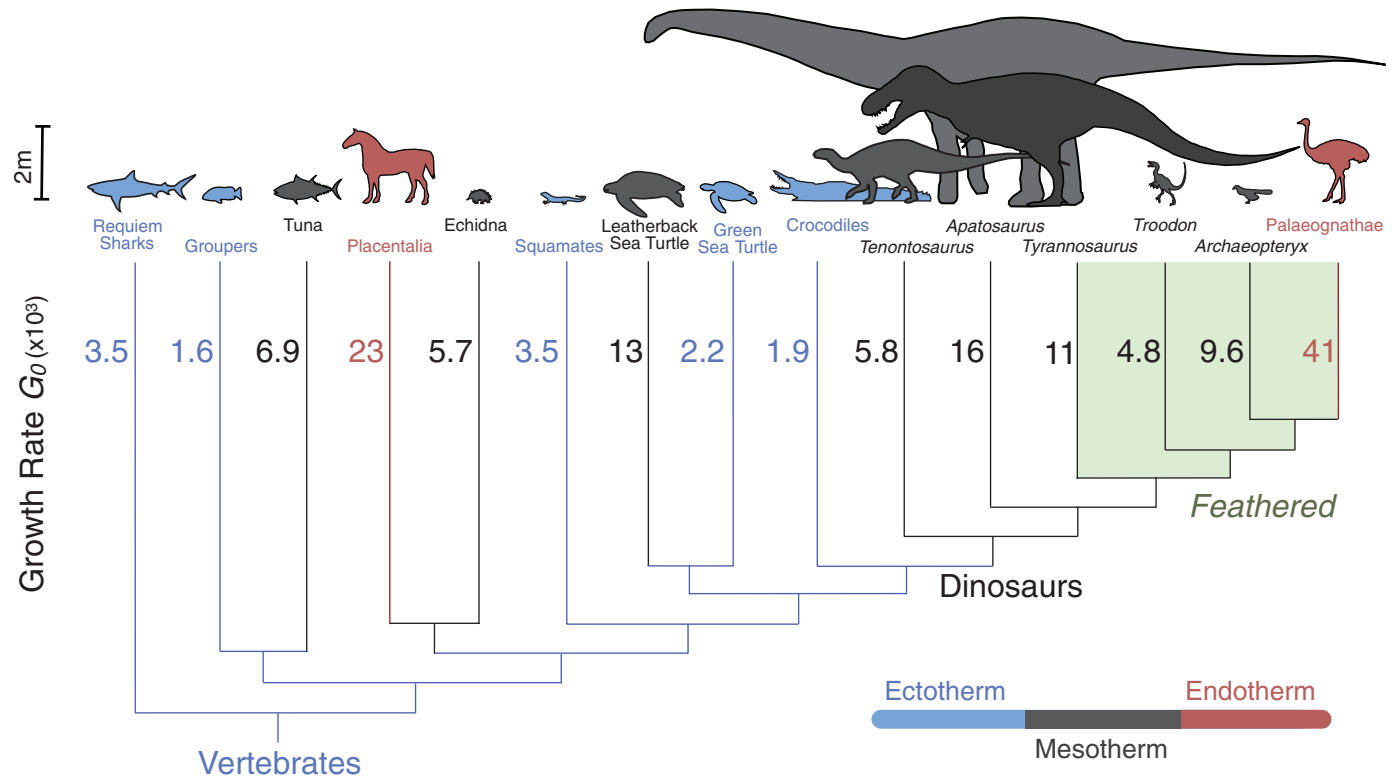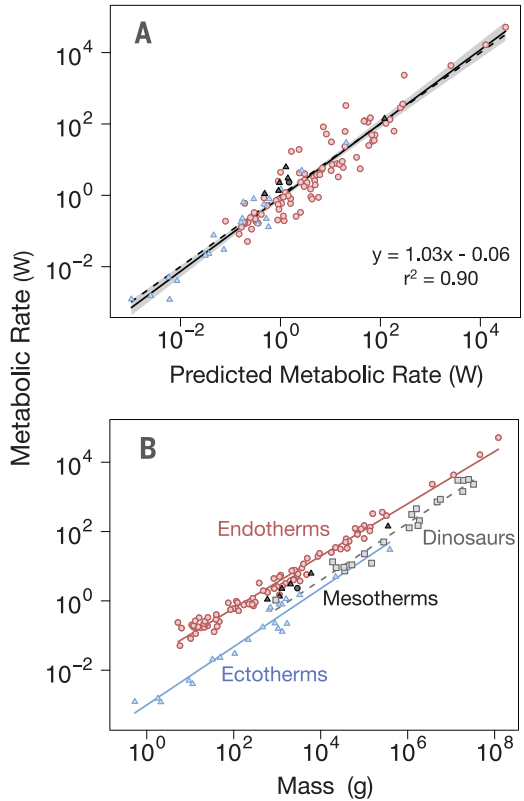
Fig. 4. Phylogeny of mass-independent growth rates (g$^{1/4}$ day^{-1}). Color signifies thermoregulatory state; branch lengths are not standardized for divergence times. Green shading indicates feathered coelurosaurian dinosaurs.

MST makes the following theoretical predictions regarding growth and metabolic rate:

(1) G_{max} scales as M^{α}, where $\alpha \sim 3/4$.

(2) B scales isometrically with G_{max} if masses are standardized (9). Regression of B against G_{max} yields a slope of 1 and an intercept of ≈ 0.66.

(3) Plotting G_0 against B_0 will reveal distinct energetic clusters corresponding to endotherms and ectotherms. High-power endotherms will exhibit an elevated G_0 and B_0, and ectothermic organisms the converse. Thermally intermediate taxa, termed mesotherms, such as tuna and lamnid sharks (16), should fall between the upper and lower quadrats. The predicted slope and intercept are 1 and 1.52, respectively. Similar clustering is observed if G_{max} and B residuals are plotted.

(4) $B_{predicted} = B_{observed}$ in extant animals, where $B_{predicted}$ is calculated from Eq. 3.

Our analyses find broad support for all four predictions. First, growth scales with mass as $\sim 3/4$, although taxonomic variation is observed (Fig. 1 and fig. S1, mean $\alpha_{OLS} = 0.73$; mean $\alpha_{PIC} = 0.69$, table S2). This indicates that larger species acquire their bulk by accelerating their maximum growth rate proportionate to $\sim M^{3/4}$. Second, G_{max} is a strong predictor of B, where $B_M = 0.56 G_{max}^{1.03}$, which is close to theoretical predictions [figs. S3 and S4; slope confidence interval (CI) = 0.97 to 1.10; intercept CI = 0.47 to 0.97; coefficient of determination (r^2) = 0.90, $n = 118$]. Third, we find that the observed relationship between mass-independent growth and metabolic rates corresponds closely to predicted values (slope = 0.90, CI = 0.77 to 1.03; intercept = 1.10, CI = 0.59 to 2.06, $r^2 = 0.61$, $n = 124$). Ectothermic species fall in the lower left quadrat; endotherms in the upper right; and thermally intermediate taxa, including tuna, a lamnid shark, the leatherback turtle, and a prototherian mammal, fall between values for endo- and ectotherms (Fig. 2 and figs. S1, S2, and S5). These results are robust; the inclusion of cold-water fish, with reduced growth and metabolic rates, simply extends the lower portion of the regression line. Furthermore, the ratio G_0/B_0 (g J^{-1}), a measure of efficiency in converting energy to biomass, does not differ significantly between endo- and ectotherms, indicating that energy allocation to growth does not vary with thermoregulatory strategy (t statistic = 0.46, $P = 0.64$, fig. S6). Finally, regression of observed against calculated metabolic rates does not differ significantly from unity (Fig. 3A; slope CI = 0.97 to 1.10; intercept CI = −0.14 to 0.02). We can therefore predict dinosaur resting metabolic rates from growth rate, using either a theoretical model (Eq. 3) or an empirically determined equation (9)

$$B_M = 0.6\, G_{max} \qquad (4)$$

Our analyses are robust to variation in the scaling exponent, phylogenetic correction, inclusion of captive versus wild animals, critiques of dinosaur growth studies, and uncertainty in estimating M and metabolic temperature (9).

Our results find that mass-independent growth rates in dinosaurs were intermediate to, and significantly different from, those of endothermic and ectothermic taxa (table S2). Although some dinosaur growth rates overlap with high-power ectotherms or low-power endotherms, they cluster closest to energetically and thermally intermediate taxa, such as tuna (Fig. 2). Further, our analyses uphold the somewhat surprising finding that feathered dinosaurs, including protoavian *Archaeopteryx* (17), did not grow markedly differently from other dinosaurs (Fig. 4). It appears that modern avian energetics did not coincide with feathers or flight, which is consistent with fossil evidence that modern bone histology in birds did not appear until the late Cretaceous (18).

At the largest body masses, the growth rates of the largest dinosaurs and mammals overlap (Fig. 1B). This pattern is driven by two factors. First, dinosaurs have a relatively high slope ($\alpha_{OLS} = 0.82$, but $\alpha_{PIC} = 0.76$). This value is consistent with suggestions of thermal inertia for larger taxa; the removal of sauropods yields a reduced OLS slope of 0.77. Second, significantly reduced growth rates are observed in several large mammalian taxa, particularly primates, elephants, and toothed whales, whereas small shrews and rodents have relatively high rates, leading to a low overall slope for placental mammals ($\alpha_{OLS} = 0.64$, $\alpha_{PIC} = 0.63$; table S2 and fig. S11). The slow growth of many large endothermic mammals is associated with large brain size and low juvenile mortality (19, 20); this is unlikely to be relevant to most dinosaurs.

Our results highlight important similarities and differences from previous studies. For example, our work agrees with assessments by Erickson (7, 17) that dinosaurs grew at rates intermediate to most endo- and ectotherms. However, we find considerably more similarity in ectothermic growth rates than reported by Case (13) and significantly higher growth rates for fish (~seven times higher), marsupials (~four times higher) and precocial birds (~two times higher; fig. S8). We attribute these differences to enhanced sampling and standardization of the thermal environment for ectotherms (e.g., Case included temperate fish). Moreover, our expansion of the comparative growth framework indicates that dinosaurs grew and metabolized at rates most similar to those of active sharks and tuna (Fig. 2 and fig. S1), rather than those of endothermic marsupials, as has been suggested (17).

Past work has often struggled to fit dinosaurs into a simple energetic dichotomy; our work suggests that an intermediate view (17, 21) is more likely. Although dinosaur growth rates vary, they cluster most closely to those of thermally intermediate taxa (Figs. 1 and 2), which we term mesotherms. Mesothermic tuna, lamnid sharks, and the leatherback turtle rely on metabolic heat to raise their body temperature (T_b) above the ambient temperature (T_a) but do not metabolically defend a thermal set point as endotherms do (16, 22). This reliance on metabolic heat distinguishes them from other large homeothermic reptiles, such as crocodiles (23), which bask to elevate T_b.

The echidna, while maintaining a set point of ~31°C, shows remarkable lability, because T_b values can range over 10°C while it is active (24). Unlike hibernating mammals or torpid hummingbirds, this variability is externally imposed. Collectively, these animals are distinguished from endotherms and ectotherms by a weak or absent metabolic defense of a thermal set point but sufficient internal heat production to maintain $T_b > T_a$ when T_a is low [see (9) for further discussion]. Although some feathered dinosaurs may have been endotherms, they would have been uniquely low-powered compared to extant birds and mammals. We suggest that mesothermy may have been common among dinosaurs, ranging from modest metabolic control of T_b, as seen in furred echidnas, to the absent metabolic defense observed in tuna and leatherback turtles. Analysis of fossil isotopes, which can shed light on body temperatures, will be useful in testing this hypothesis. In particular, attention to neonate and juvenile dinosaurs in seasonally cool environments, such as polar regions, may help distinguish among thermoregulatory states.

Dinosaurs dominated the flux of matter and energy in terrestrial ecosystems for more than 135 million years. Consequently, our results have important implications for understanding ancient Mesozoic ecosystems. We emphasize the primary importance of comparative energetics for integrating form, function, and diversity. Knowing only two facts from the fossil record—adult mass and maximum growth rate—we show that the metabolic rates of extinct clades can be predicted with accuracy. Such an approach will be useful in resolving the energetics of metabolically ambiguous taxa, such as pterosaurs, therapsids, and Mesozoic birds.

REFERENCES AND NOTES

1. K. Padian, A. J. de Ricqlès, J. R. Horner, *Nature* **412**, 405–408 (2001).
2. P. M. Sander *et al.*, *Biol. Rev. Camb. Philos. Soc.* **86**, 117–155 (2011).
3. R. A. Eagle *et al.*, *Science* **333**, 443–445 (2011).
4. J. H. Brown, J. F. Gillooly, A. P. Allen, V. M. Savage, G. B. West, *Ecology* **85**, 1771–1789 (2004).
5. P. Else, A. Hulbert, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **240**, R3–R9 (1981).
6. J. F. Gillooly, J. H. Brown, G. B. West, V. M. Savage, E. L. Charnov, *Science* **293**, 2248–2251 (2001).
7. G. M. Erickson, K. C. Rogers, S. A. Yerby, *Nature* **412**, 429–433 (2001).
8. A. H. Lee, S. Werning, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 582–587 (2008).
9. See the supplementary materials.
10. F. Seebacher, *Paleobiology* **29**, 105–122 (2003).
11. R. H. Peters, *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, MA, 1983).
12. C. R. White, N. F. Phillips, R. S. Seymour, *Biol. Lett.* **2**, 125–127 (2006).
13. T. J. Case, *Q. Rev. Biol.* **53**, 243–282 (1978).
14. G. B. West, J. H. Brown, B. J. Enquist, *Nature* **413**, 628–631 (2001).
15. C. Hou *et al.*, *Science* **322**, 736–739 (2008).
16. D. Bernal, K. A. Dickson, R. E. Shadwick, J. B. Graham, *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **129**, 695–726 (2001).
17. G. M. Erickson *et al.*, *PLOS ONE* **4**, e7390 (2009).
18. A. Chinsamy, *Mesozoic Birds: Above the Heads of Dinosaurs* (Univ. of California Press, Berkeley, CA, 2002), pp. 421–431.

19. E. L. Charnov, D. Berrigan, *Evol. Anthropol.* **1**, 191–194 (1993).
20. K. Isler, C. P. van Schaik, *J. Hum. Evol.* **57**, 392–400 (2009).
21. R. Reid, in *The Complete Dinosaur*, M. K. Brett-Surman, T. R. Holtz, J. O. Farlow, Eds. (Indiana Univ. Press, Bloomington, IN, 2012), pp. 873–24.
22. F. V. Paladino, M. P. O'Connor, J. R. Spotila, *Nature* **344**, 858–860 (1990).
23. F. Seebacher, G. C. Grigg, L. A. Beard, *J. Exp. Biol.* **202**, 77–86 (1999).
24. P. Brice, G. C. Grigg, L. A. Beard, J. A. Donovan, *Aust. J. Zool.* **50**, 461–475 (2002).

NONHUMAN GENETICS

Strong male bias drives germline mutation in chimpanzees

Oliver Venn,[1] Isaac Turner,[1] Iain Mathieson,[1]* Natasja de Groot,[2] Ronald Bontrop,[2] Gil McVean[1]†

Germline mutation determines rates of molecular evolution, genetic diversity, and fitness load. In humans, the average point mutation rate is 1.2×10^{-8} per base pair per generation, with every additional year of father's age contributing two mutations across the genome and males contributing three to four times as many mutations as females. To assess whether such patterns are shared with our closest living relatives, we sequenced the genomes of a nine-member pedigree of Western chimpanzees, *Pan troglodytes verus*. Our results indicate a mutation rate of 1.2×10^{-8} per base pair per generation, but a male contribution seven to eight times that of females and a paternal age effect of three mutations per year of father's age. Thus, mutation rates and patterns differ between closely related species.

Accurate determination of the rate of de novo mutation in the germ line of a species is central to the dating of evolutionary events. However, because mutations are rare events, efforts to measure the rate in humans have typically been indirect, calculated from the incidence of genetic disease or sequence divergence (*1–4*). However, high-throughput sequencing technologies have enabled direct estimates of the mutation rate from comparison of the genome sequence of family members (*5–8*). Unexpectedly, these studies have indicated a mutation rate of, on average, ~1.2×10^{-8} per base pair per generation, or ~0.5×10^{-9} per base pair per year, approximately half that inferred from phylogenetic approaches (*1, 9*). Moreover, they have demonstrated a substantial male bias to mutation, such that three to four times as many autosomal mutations occur in the male compared to the female germ line (*6, 7*). Male bias is largely caused by an increase in the rate of paternal but not maternal mutation with the age of the parent; approximately two additional mutations per year of father's age at conception (*7*). This difference is consistent with ongoing cell division in the male germ line but not in females (*10*).

An alternative approach for estimating the extent of male bias is to compare rates of sequence divergence on the autosomes (which spend equal time in the male and female germ lines) and the X chromosome (which spends two-thirds of the time in females) (*2, 11*). Such indirect approaches broadly agree with direct estimates in humans, but suggest that male bias may be stronger in chimpanzees (*12*). To test this hypothesis, we sequenced the genomes of nine members of a three-generation pedigree of Western chimpanzees, *Pan troglodytes verus* (Fig. 1A and fig. S1). One trio was sequenced at high depth (average 51×), while other family members were sequenced to an average of 27× (table S1). We inferred the structure of recombination and transmission across the pedigree (Fig. 1B), which enabled us to detect de novo point mutations in regions of high sequence complexity and to remove artifacts caused by mismapping, sequence that is absent from the reference genome, and reference misassembly (*13*).

We used a probabilistic approach that, at a given site, compared the likelihood of different models for genetic variation inconsistent with the inferred transmission: genotyping error at a segregating variant, de novo mutation, single-gene conversion event, segregating deletion and erroneous call (Fig. 1C). The design was expected to enable haplotype phasing through transmission for 99.2% of sites that were heterozygous in the founders and 87.5% of de novo mutation events inherited by chimpanzee F (Fig. 1A). Read-based phasing was used to phase de novo events in other offspring, and we performed independent validation to assess the accuracy of de novo variant calls. The false-negative rate was estimated from allele-dropping simulations (*13*).

Across the genomes of the nine pedigree members, we called 4.1 million variants [single-nucleotide polymorphisms (SNPs) and short insertions and deletions (indels)] using a mapping-based approach and 3.0 million variants using an assembly-based approach (*14*). Genotype data confirmed expected pedigree relationships (fig. S2). The intersection of call sets (1.6 million sites with a transition-transversion ratio of 2.2) established the underlying structure of recombination and transmission across the pedigree with a robust version of the Lander-Green algorithm (fig. S3). Briefly, this is a two-stage strategy of identifying dominant inheritance vectors over 1-Mb intervals, followed by fine-mapping of cross-over breakpoints, which guards against problems caused by false-positive variants and genotyping errors (*13*). Across the pedigree, we identified 375 cross-over events, with a distribution similar to that of human homologs, with the exception of human chromosome 2, which is a fusion of chimpanzee chromosomes 2A and 2B (*15*) (Fig. 2A, fig. S4, and tables S2 and S3).

Overall, we estimate the sex-averaged autosomal genetic map length to be 3150 cM [95% equal-tailed probability interval (ETPI) 2850 to 3490], compared to 3505 cM in humans (*16, 17*). On the X chromosome, we detected nine cross-over events in the non-pseudoautosomal (non-PAR) region, indicating a female-specific genetic map length of 160 cM (95% ETPI 83 to 300), compared to 180 cM in humans. On the pseudoautosomal region (PAR), we detected four male cross-overs, giving a male-specific estimate of 34 cM (95% ETPI 28 to 180; tables S4 and S5), in agreement with estimates in humans (*13*). Males have 58% of the autosomal cross-over events of females and, unlike females, show an increase in cross-over frequency toward the telomere (Fig. 2B), similar to humans (fig. S5). We also observed a decrease in cross-over frequency with maternal (2.65 cM per year, linear model *P* = 0.025), but not paternal age (Fig. 2C). However, this observation could be explained by between-female variation (linear model *P* = 0.13, allowing for a maternal effect). The median interval size to which cross-over events can be localized is 7.0 kb, with 95% of all intervals localized to within 80 kb (excluding complex cross-over events), with cross-over events enriched in regions inferred to have

[1]Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford, OX3 7BN, UK. [2]Biomedical Primate Research Centre, Lange Kleiweg 161, 2288 GJ Rijswijk, Netherlands.
*Present address: Department of Genetics, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115 USA.
†Corresponding author. E-mail: mcvean@well.ox.ac.uk

high rates of recombination from patterns of linkage disequilibrium (*18*) (fig. S6).

Conditional on the inferred transmission, we used a probabilistic approach to identify candidate de novo mutations among all variants called by the mapping approach, incorporating uncertainty in the inferred genotype through the use of genotype likelihoods (*13*). Across the pedigree, we identified 204 autosomal de novo mutations (2 of which are multinucleotide variants) that pass thresholds for evidence (fig. S7), purity, and consistency (fig. S8 and table S6).

Several lines of evidence indicate a low false-positive rate. First, none of these sites were called as variants in the genomes of 10 unrelated chimpanzees from the same subspecies (*18*). Second, the transition-transversion ratio of the candidate de novo events (2.16) is comparable to that for segregating variants. Third, the transmission of candidate de novo events in chimpanzee F to her offspring is consistent with expectations. Finally, we used a genotyping platform to validate all de novo events identified in chimpanzees F and I. Of the 61 sites with valid assays (18 failed design), 1 is a false positive, indicating a false discovery rate of ~2% (table S7). To estimate false-negative rates, we used allele-dropping simulations, with empirical distributions of coverage and allele balance. Within the F_1 generation, the false-

negative rate is estimated to be 3.4%. However, the F_2 generation has a higher rate (23%), arising from lower coverage (25.6×) in founder chimpanzee C. False-negative rates were used to correct subsequent regression analyses (table S8). On the X chromosome, we identified three de novo point mutations.

As expected (*1*), a high fraction of de novo mutations are C>T transitions at CpG sites [24% of all point mutations, compared to 17% in humans (*7*); likelihood ratio test (LRT), $P = 0.03$], and even after accounting for such mutations, we see a trend toward AT bases (73 G/C>A/T, 55 A/T>G/C; ratio 1.3:1; LRT, $P = 0.11$; Fig. 3A). We also found that point mutations tend to cluster within individuals at nearby locations, similar to observations in humans (*8*). For example, 17% of all point mutations are within 1 Mb of at least one other mutation event in the pedigree, and in 41% of such cases, these all occur in the same individual (compared to an expectation of 13%; permutation $P = 0.001$). Notably, we validated all variants in clusters of 1 Mb or less in E and F, indicating that these are not false positives. The excess of within-sample clustering extends up to ~200 kb (fig. S9A) and does not correspond to a single mutation type (e.g., CpG mutation). Moreover, the effect remains after increasingly stringent filters for specificity are applied (fig.

S9B) (*13*). The finding of clustered point mutation events, which may potentially arise from correlated exposure to mutagens or variation in the efficacy of DNA repair, implies non-independence in the way novel variation enters a species and has consequences for interpretation of patterns of genetic variation. We do not observe enrichment around genes, repeat elements, or gaps in the assembly (fig. S10).

To assess whether the rate of de novo mutation is affected by parental age, we used Poisson regression, allowing for family effects and separate linear relationships between age and mutation rate for males and females. Despite the small sample size, we find no evidence for either familial or maternal age effects (linear model $P > 0.05$), but we do find evidence for a paternal age effect (linear model $P = 0.006$) and consistency in effect on repeat and nonrepeat DNA backgrounds (fig. S11). Although we cannot formally exclude the possibility of familial effects (*6*), our results are consistent with observations (*7*) that paternal age explains nearly all variation in mutation rate in humans. Bayesian linear regression, allowing for a paternal effect only and accommodating variation in false-negative rate, indicates a posterior mean paternal age of effect of 3.02 mutations per year (Fig. 3B; 95% ETPI 1.35 to 4.68). In contrast, the paternal age effect in humans is estimated to be 1.95 mutations per year [reanalysis of data from (*6*); 95% ETPI 1.65 to 2.26].

We ascertained the parent of origin of de novo mutations for chimpanzee F through transmission to the F_2 generation, finding that 30 of the 35 autosomal mutations occurred in the paternal lineage. We also found that 25% of de novo events could be phased directly from read-pairs spanning the mutation and a nearby heterozygous site that could be phased through transmission. Across the pedigree, we assigned 31 paternal and 6 maternal autosomal mutations (Fig. 3C). Overall, we estimate the aggregate male-to-female mutation ratio, $\alpha = 5.5$ (95% ETPI = 3.0 to 10). The point estimate is 40% higher than that reported for humans (*7*), though it is close to estimates from chimpanzee-specific divergence rates on the X and autosomes (*12*). In contrast to indirect approaches (*3*), we find no evidence that different types of mutation have different values of α. For example, α at CpG sites is 5.3, compared to 5.6 at non-CpG sites. Combining data across all mutation types and using available parent-of-origin information in the Bayesian regression model indicates that, on average, mothers contribute 6.7 de novo mutations (95% ETPI = 3.5 to 10.3) and each additional year of paternal age generates 3.0 mutations (95% EPTI = 1.2 to 4.4; fig. S12), with the onset of mutation occurring at 8.1 years of age (95% ETPI = 0 to 12 years), consistent with the onset of spermarche in chimpanzees of 7.5 years (*19*).

Within the pedigree studied, the average number of autosomal de novo mutations occurring in each generation is 35, lower than current estimates for humans of 74.4 (*7*, *9*). However, the parental ages in the Western chimpanzee pedigree
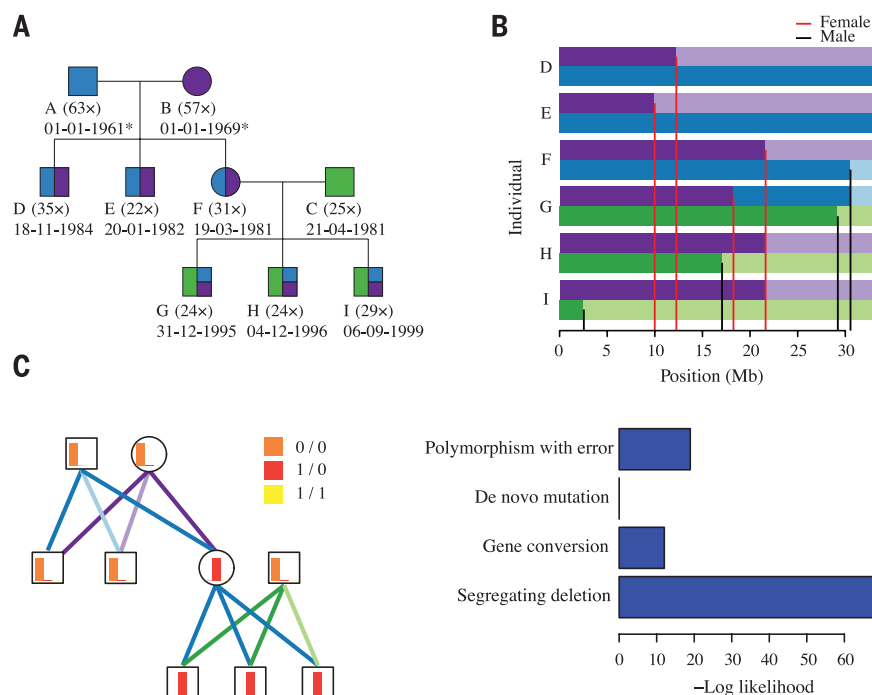
Fig. 1. De novo mutations in a three-generation chimpanzee pedigree. (**A**) Structure of the chimpanzee pedigree, indicating sequencing depth and date of birth (estimated for wild founders). Blue, purple, and green colors indicate expected inheritance proportions from founders. (**B**) Visual representation of transmission across chromosome 21 following color assignment in (A); grandparental origin distinguished by intensity. Red and black lines indicate position of female- and male-specific crossover events, respectively. (**C**) Site classification at a candidate de novo mutation on chromosome 2A, showing (left) the relative genotype likelihoods for each individual (as red, orange, and yellow barplots; with height representing relative likelihood and color representing genotype), the resolved transmission structure [as purple, blue, and green lines; following color and intensity assignment in (B)], and (right) the relative likelihoods for different models for the observed genetic variation.

(averages of 18.9 for males and 15.0 for females) are lower than estimates of parental ages in the wild (24.3 for males and 26.3 for females), which are lower than estimates for humans (31.5 for males and 25.6 for females) (20). Using the fitted model for mutation rates, we predict that the average number of autosomal de novo mutations per offspring in the wild should be ~69. We estimate the length of the autosomal genome accessible in our study to be 2360 Mb across the autosomes (table S9), indicating a mutation rate of 1.2×10^{-8} per base pair per generation and α to be 7.8 (table S10).

Under a model in which the mutation rate increases linearly with parental age, the rate of neutral substitution is the ratio of the average number of mutations inherited per generation to the average parental age. We predict the neutral substitution rate to be ~0.46×10^{-9} per base pair (bp) per year in chimpanzees, compared to estimates in humans of ~0.51×10^{-9} bp^{-1} year^{-1} (9). These results are consistent with near-identical levels of lineage-specific sequence divergence (12) but surprising given the differences in paternal age effect. In the intersection of the autosomal genome accessible in this study and regions where human and chimpanzee genomes can be aligned with high confidence, the rate is slightly lower (0.45×10^{-9} bp^{-1} year^{-1}) and the level of divergence is 1.2% (13), implying an average time to the most common ancestor of 13 million years, assuming uniformity of the mutation rate over this time (95% ETPI 11 to 17 million years; table S11).

Increased male bias can explain low levels of diversity on the chimpanzee X chromosome (21, 22). Taking into account differences in gen-

eration time and effective population size, we predict that X-chromosome diversity should be 56% that of autosomes (assuming equal and constant effective population sizes for males and females; table S10), comparable to empirical estimates (21, 22). Similarly, our results predict that the X-chromosome rate of divergence is lower in chimpanzees than humans (74% of the autosomal rate in chimpanzees, 85% of the autosomal rate in humans). Previous explanations for unusual patterns of X-chromosome diversity and divergence include a complex speciation event (23), extensive natural selection on the X chromosome (22), or, as supported by this study, a greater male mutational bias in chimpanzees (12). This is likely related to differences in mating system between the species, with chimpanzees showing higher levels of sperm
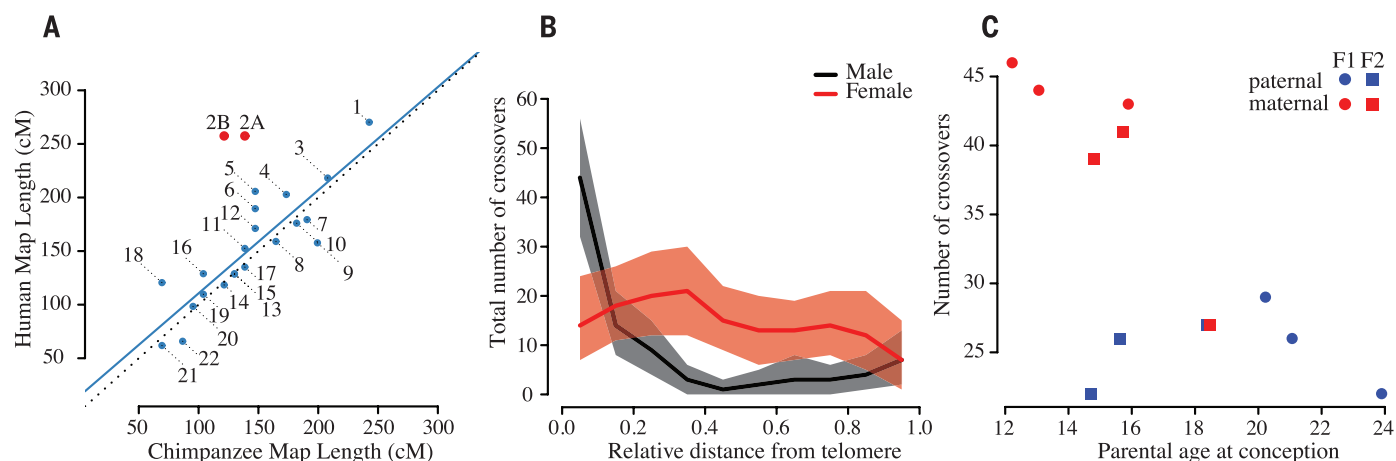
Fig. 2. Parameters of cross-over. (**A**) Scatterplot showing the relationship between total genetic map length inferred for chromosomes and their homologs in humans. Black dotted line indicates the identity relationship; blue line indicates best linear fit (excluding chromosomes 2A and 2B). (**B**) Relationship of recombination rate to relative telomere proximity demonstrates a telomeric effect in males, which is absent in females. (**C**) Relationship between parental age and the number of cross-over events in mothers and fathers for the two family groups.
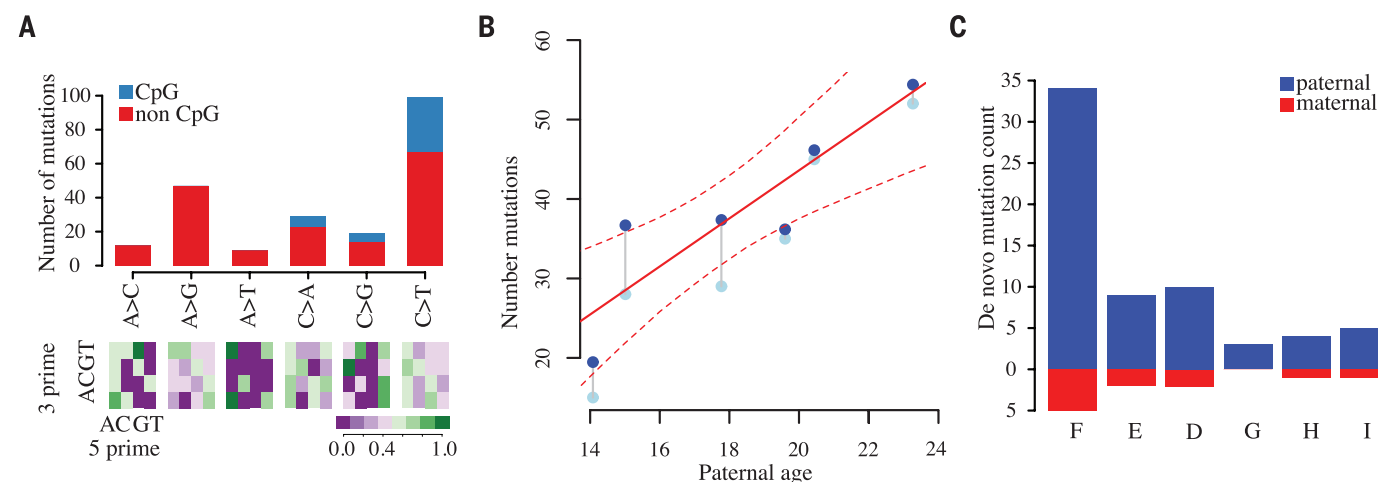
Fig. 3. Characteristics of de novo mutations. (**A**) (Top) Numbers of different types of de novo single-base substitutions (multiple nucleotide events not shown). (Bottom) Relative abundance of each mutation type as a function of the adjacent nucleotides. (**B**) Relationship between paternal age and total number of de novo mutations; gray points: observed counts; blue points: correction for false-negative rate. The posterior mean and marginal 95% equal-tailed probability intervals for the Bayesian Poisson regression are shown as solid and dotted lines, respectively. (**C**) The number of de novo mutation events confirmed as paternal or maternal through either transmission to the F_2 generation or from direct phasing within read-pairs. Offspring ordered left to right by increasing paternal age.

competition through multiple mating and a higher relative testes mass than humans (0.27% of average adult male weight versus 0.079%) and higher levels of sperm production (*24, 25*). If differences in male mutational bias are to explain observed patterns of divergence, then gorillas would have a male mutational bias lower than that of humans arising from decreased sperm competition (*12*). Our results suggest that variation in mating patterns between species can affect the sex bias of mutation and motivate the wider study of mutation rates and relationship to parental age across species.

REFERENCES AND NOTES

1. M. W. Nachman, S. L. Crowell, *Genetics* **156**, 297–304 (2000).
2. K. D. Makova, W. H. Li, *Nature* **416**, 624–626 (2002).
3. J. Taylor, S. Tyekucheva, M. Zody, F. Chiaromonte, K. D. Makova, *Mol. Biol. Evol.* **23**, 565–573 (2006).
4. F. A. Kondrashov, A. S. Kondrashov, *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **365**, 1169–1176 (2010).
5. J. C. Roach *et al.*, *Science* **328**, 636–639 (2010).
6. D. F. Conrad *et al.*, *Nat. Genet.* **43**, 712–714 (2011).
7. A. Kong *et al.*, *Nature* **488**, 471–475 (2012).
8. J. J. Michaelson *et al.*, *Cell* **151**, 1431–1442 (2012).
9. A. Scally, R. Durbin, *Nat. Rev. Genet.* **13**, 745–753 (2012).
10. J. F. Crow, *Nat. Rev. Genet.* **1**, 40–47 (2000).
11. T. Miyata, H. Hayashida, K. Kuma, K. Mitsuyasu, T. Yasunaga, *Cold Spring Harb. Symp. Quant. Biol.* **52**, 863–867 (1987).
12. D. C. Presgraves, S. V. Yi, *Trends Ecol. Evol.* **24**, 533–540 (2009).
13. Supplementary materials are available on *Science* Online.
14. Z. Iqbal, M. Caccamo, I. Turner, P. Flicek, G. McVean, *Nat. Genet.* **44**, 226–232 (2012).
15. J. W. IJdo, A. Baldini, D. C. Ward, S. T. Reeders, R. A. Wells, *Proc. Natl. Acad. Sci. U.S.A.* **88**, 9051–9055 (1991).
16. A. Kong *et al.*, *Nature* **467**, 1099–1103 (2010).
17. A. Kong *et al.*, *Nat. Genet.* **31**, 241–247 (2002).
18. A. Auton *et al.*, *Science* **336**, 193–198 (2012).
19. J. Marson, S. Meuris, R. W. Cooper, P. Jouannet, *Biol. Reprod.* **44**, 448–455 (1991).
20. K. E. Langergraber *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15716–15721 (2012).
21. G. H. Perry, J. C. Marioni, P. Melsted, Y. Gilad, *Mol. Ecol.* **19**, 5332–5344 (2010).
22. C. Hvilsom *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 2054–2059 (2012).
23. N. Patterson, D. J. Richter, S. Gnerre, E. S. Lander, D. Reich, *Nature* **441**, 1103–1108 (2006).
24. R. V. Short, *Adv. Stud. Behav.* **9**, 131–158 (1979).
25. A. P. Møller, *J. Hum. Evol.* **17**, 479–488 (1988).

NEURAL MIGRATION

Structures of netrin-1 bound to two receptors provide insight into its axon guidance mechanism

Kai Xu,[1]* Zhuhao Wu,[2]* Nicolas Renier,[2]* Alexander Antipenko,[1] Dorothea Tzvetkova-Robev,[1] Yan Xu,[1] Maria Minchenko,[1] Vincenzo Nardi-Dei,[1]† Kanagalaghatta R. Rajashankar,[3] Juha Himanen,[1] Marc Tessier-Lavigne,[2]‡ Dimitar B. Nikolov[1]‡

Netrins are secreted proteins that regulate axon guidance and neuronal migration. Deleted in colorectal cancer (DCC) is a well-established netrin-1 receptor mediating attractive responses. We provide evidence that its close relative neogenin is also a functional netrin-1 receptor that acts with DCC to mediate guidance in vivo. We determined the structures of a functional netrin-1 region, alone and in complexes with neogenin or DCC. Netrin-1 has a rigid elongated structure containing two receptor-binding sites at opposite ends through which it brings together receptor molecules. The ligand/receptor complexes reveal two distinct architectures: a 2:2 heterotetramer and a continuous ligand/receptor assembly. The differences result from different lengths of the linker connecting receptor domains fibronectin type III domain 4 (FN4) and FN5, which differs among DCC and neogenin splice variants, providing a basis for diverse signaling outcomes.

N etrins, acting as both attractants and repellents, regulate neuronal migration, axon guidance, and synaptogenesis (*1–4*). In nonneural tissues, netrins have a variety of functions, including promoting cell adhesion and invasion, leukocyte migration, angiogenesis, and cell survival (*5*). Netrins contain an N-terminal laminin domain (LN, also known as domain VI), followed by three cysteine-rich LN-type epidermal growth factor (EGF)–like modules (LE1, LE2, and LE3; also known as domain V), and a small positively charged C-terminal domain (LC). In mammals, the secreted netrins-1, -3 and -4 are only distantly related to the glycosylphosphatidylinositol–anchored G netrins (*6, 7*).

Netrin actions are mediated by distinct receptors (*4*). Deleted in colorectal cancer (DCC) mediates attractive responses to netrin-1, whereas Unc5 proteins, alone or with DCC, are required for its repulsive effects (*4, 8*). The ectodomain of DCC is composed of four immunoglobulin-like domains and six fibronectin type III (FNIII) domains. The DCC FNIII repeats mediate interactions with netrin-1 through its LN-LE region (*9–11*). That region, when added as an Fc-fusion protein, is sufficient to mimic the axon outgrowth activity of full-length netrin-1 (*12*). Neogenin is structurally similar to DCC and binds netrin-1 and -3 (*12, 13*), but also binds the structurally distinct repulsive guidance molecule (RGM) (*14, 15*). Knockdown analysis in zebrafish supports a role for neogenin

in mediating axonal attraction to netrin (*16*), but this role has not been established in mammals, where it has mostly been studied as an adhesive factor (*17*) and a putative guidance receptor for RGM (*15*).

We revisited the role of neogenin in netrin attraction while studying commissural axon attraction to a netrin-1 source at the spinal cord midline. Defects in this guidance in human leads to neurological syndromes, some of which result from mutations in *DCC* (*18–20*). Prior analysis suggested that DCC mediates the entire attractive effect of netrin-1 because the phenotype observed with a commissural axon marker, antibody 4D7 to TAG-1, appeared stronger in *Dcc* mutant than in *netrin-1* mutant embryos (*21*). As new markers became available (*22*), we reevaluated embryos mutant for *Dcc* or *netrin-1*. Commissural projections develop between embryonic days 10.5 (E10.5) and E12.5, when spinal cord shape changes rapidly. To minimize artifacts from variation in embryo size and stage across litters, we compared size-matched embryos that were littermates from intercrosses of compound heterozygous animals. By using an antibody to Robo3 (*22*) in E11.5 *Dcc* mutant embryos, we observed only a 55% reduction in width of the ventral commissure compared with wild-type littermates, less than in *netrin-1^{−/−}* embryos, which had a 78% reduction (Fig. 1, A and E). The same was seen with a new antibody to TAG-1 (Fig. 1A) and with antibodies to the axonal marker neurofilament-M (Fig. 1A). The difference with the prior study appears to result from 4D7 giving weaker labeling of commissural axons that is also influenced by *Dcc* expression (fig. S1). Thus, the guidance phenotype is actually less severe than in *Dcc* than in *netrin-1* mutants, suggesting an additional netrin-1 receptor(s) contributes to residual attraction in *Dcc^{−/−}* embryos.

[1]Structural Biology Program, Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA. [2]Laboratory of Brain Development and Repair, Rockefeller University, New York, NY 10065, USA. [3]Department of Chemistry and Chemical Biology, Cornell University and Northeastern Collaborative Access Team, Advanced Photon Source, Argonne, IL 60439, USA.
*These authors contributed equally to this work. †Present address: Novartis Vaccines, Siena, Italy. ‡Corresponding author. E-mail: nikolovd@mskcc.org (D.B.N.); marctl@mail.rockefeller.edu (M.T.-L.)

To test this, we examined whether *Dcc* mutant commissural axons retain a response to netrin-1 in vitro. We cultured dorsal spinal cord explants from E11 wild-type and *Dcc*−/− embryos (Fig. 1, B and C). In control explants, netrin-1 application induced robust axonal outgrowth that peaked at 250 ng/ml. The peak response was reduced significantly (by ~97%) when explants from *Dcc*−/− embryos were used, confirming DCC's central role as a netrin-1 receptor, but a dose-dependent response of *Dcc* mutant axons was still consistently observed (Fig. 1, B and C). To determine which receptor mediates the residual netrin-1 response, we screened known and putative netrin-1 receptors by in situ hybridization and immunohistochemistry in E11.5 spinal cord. We observed neogenin immunoreactivity on commissural axons (*23*), which was lost in neogenin (*Neo1*) mutant spinal cords (fig. S2A), suggesting that neogenin might collaborate with DCC in guiding these axons. Consistent with this, whereas commissural axon trajectories

in transverse sections from *Neo1*−/− embryos were apparently normal (Fig. 1D), removing neogenin as well as DCC in *Dcc*−/−;*Neo1*−/− double mutants resulted in an 84% reduction in ventral commissure size, that is, greater than *Dcc*−/− but comparable to *netrin-1*−/− embryos (Fig. 1, D and E). Moreover, we observed abnormal Robo3+ commissural axons in the motor column of *netrin-1*−/− embryos; fewer are seen in *Dcc*−/− single mutants, but a comparable number was seen in *Dcc*−/−; *Neo1*−/− double mutants (fig. S2, B to D). Although the *Neo1*−/− and *netrin-1*−/− alleles are severely hypomorphic rather than complete null alleles (fig. S2A for *Neo1*−/−), the *Dcc* allele is a null allele, so our finding that commissural axon guidance defects in *Dcc*−/−;*Neo1*−/− embryos are greater than in *Dcc*−/− mutants but comparable to those in *netrin-1*−/− mutants is consistent with the model that neogenin is a functional netrin-1 receptor that acts in concert with DCC to direct commissural axons to the midline netrin source.

To study how neogenin and DCC function as netrin-1 receptors, we investigated the structural basis of the netrin-1/neogenin and netrin-1/DCC interactions. There are conflicting reports regarding which DCC FNIII domains mediate interactions with the netrin-1 LN-LE region (*9–11*), so we conducted biolayer interferometry binding studies (materials and methods) to clarify this. Our results (Fig. 2A) show that domains FN4 and FN5 both interact with this ligand and that they account for the full in vitro binding affinity. Accordingly, in our structural studies we used a netrin construct that contains the LN and LE(1-3) domains and neogenin/DCC constructs containing FN4 and FN5. We did not include LC [also known as C345C, suggested to bind heparan sulfate (*24*)], because it is attached via a flexible linker and not strictly required for receptor binding (*9–11*) and because a netrin-1–Fc fusion construct lacking this domain induces similar axon outgrowth in vitro as full-length netrin-1 (*12*). Splice variants
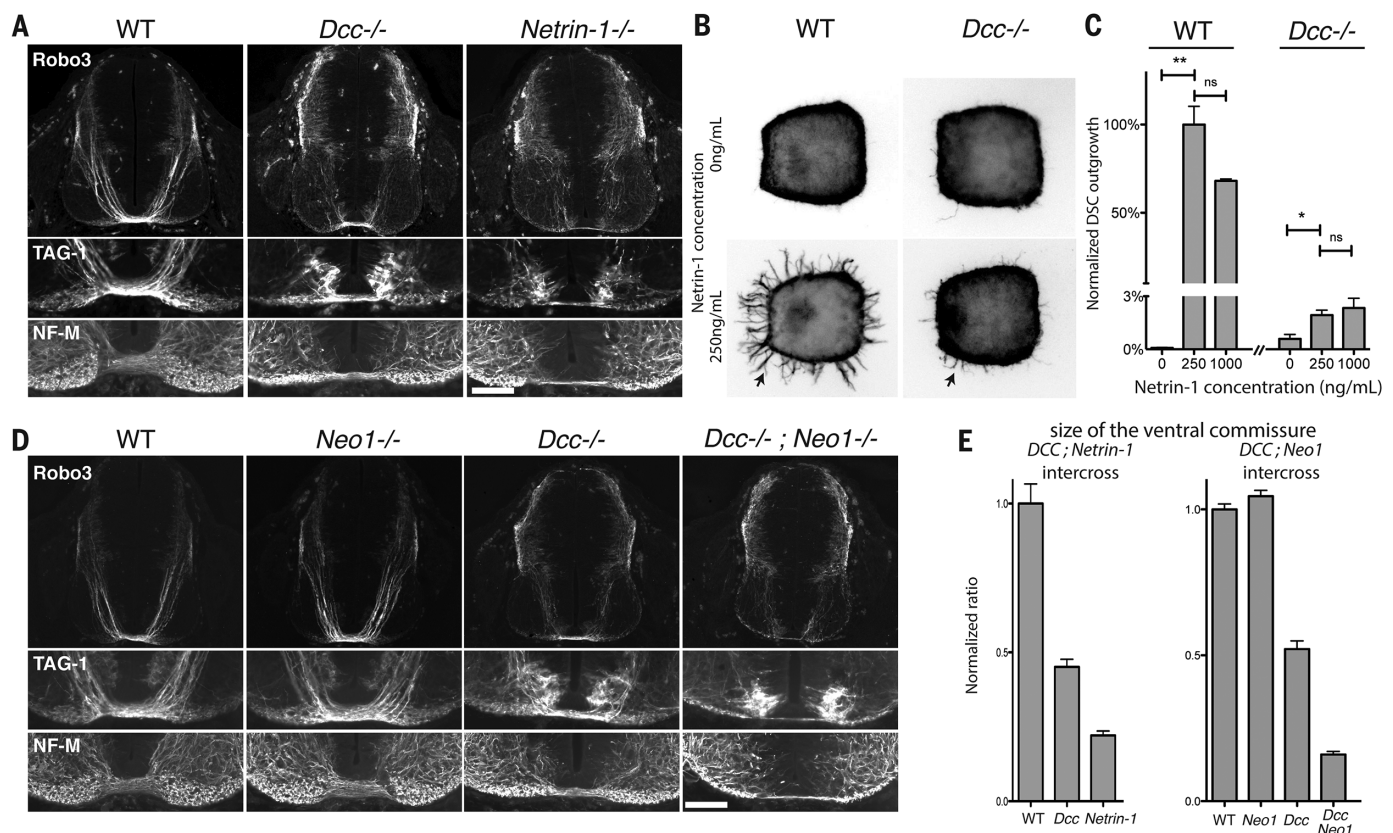
Fig. 1. Neogenin and Dcc collaborate to mediate the attraction of commissural neurons by netrin-1. (**A**) Cross sections of E11 wild-type (WT), *Dcc*−/− and *Netrin-1*−/− littermate mouse embryos at the level of brachial spinal ganglia, stained for TAG-1, Robo3, and neurofilament, medium chain (NF-M). Bottom two rows show details of the ventral commissural axon bundle. *Dcc* mutants have a reduced ventral commissure, but a large number of axons still cross. *Netrin-1*−/− embryos have a much-reduced ventral commissure. (**B**) Axon outgrowth (arrows) in E11 wild-type or *Dcc*−/− littermate mouse dorsal spinal cord explants cultured in three-dimensional collagen gels with increasing concentrations of netrin-1 and stained for Tuj1. (**C**) Quantification of the response of axon outgrowth from wild-type and *Dcc*−/− dorsal spinal cord explants to increasing netrin-1 concentrations, normalized to wild type at 250 ng/m of netrin-1. *Dcc*−/− mutants show a residual response to netrin-1 application (arrow). ns,

not significant. (**D**) Cross sections of E11 wild-type, *Dcc*−/−, *Neo1*−/− and *Dcc*−/−; *Neo1*−/− littermate mouse embryos at the level of brachial spinal ganglia, stained for TAG-1, Robo3, and NF-M. Bottom two rows show details of the motor column and the ventral commissural axon bundle. The *Dcc*−/−; *Neo1*−/− double mutant has a much reduced ventral commissure and numerous axons in the motor column. (**E**) Ratio of the commissural axon bundle size to the dorsoventral spinal cord length of wild-type, *Dcc*−/−, and *Netrin1*−/− embryos, normalized to wild types (left). Ratio of commissural axon bundle size to the dorsoventral spinal cord length of *Dcc*−/−, *Neo1*−/−, and *Dcc*−/−; *Neo1*−/− embryos normalized to wild type (right). The quantification shows the mean and SEM of five sections taken in brachial spinal cord in littermates and is representative of three litters. Scale bars are 200 μm (Robo3) and 100 μm (TAG-1 and NF-M).

(isoforms) of both neogenin and DCC with different length of the FN4-FN5 linker have been reported in most species. Both shorter and longer isoforms bind netrin-1 with high affinity (Fig. 2B). For our structural studies, we used the shorter isoforms (25).

The structure of the netrin-1 LN-LE region was determined at 2.8 Å resolution (table S1 and figs. S3 to S7), revealing an elongated molecule with the same flowerlike shape as laminin and netrin-G (Fig. 2C). The LN domain forms the head, and LE the stalk. The disulfide bond network throughout the molecule and the short linkers between the individual netrin domains result in a rigid molecular architecture with little interdomain flexibility. The globular LN domain has the canonical laminin LN fold, including a conserved Ca^{2+} binding site. The LE region contains three EGF repeats, and its structure is similar to those of laminin-$\alpha5$, laminin-$\gamma1$ (26, 27) and netrin-G (6, 7), although the latter lacks the third EGF repeat (LE3).

The structure of the netrin/neogenin complex (Fig. 3A) was determined at 3.2 Å resolution and reveals a 2:2 heterotetramer, consistent with its gel-filtration elution profile. At the heart of the

complex are two netrin molecules forming a head-to-head X-shaped dimer and interacting via an extensive LE2/LE2 interface. This dimer brings together two neogenin molecules, with receptor molecules arranged parallel to each other and their C-termini facing the same direction, presumably toward the neuronal membrane. The two receptor binding regions are located about 90 Å apart on the two ends of the rigid netrin structure, but the distance between netrin-binding surfaces of neogenin FN4 and FN5 domains cannot exceed 55 Å, so the two receptor binding sites on netrin must interact with two different receptor molecules. Netrin does not undergo any significant conformational changes upon receptor binding, and the bound and unbound netrin structures could be superposed with root mean square deviation (RMSD) of 0.9 Å over 353 Cα atoms.

The neogenin FN4 and FN5 domains share the canonical FNIII folding topology (fig. S7) and are arranged linearly, with the linker between them in a fully extended conformation. This linker region would be flexible in the absence of bound ligand.

The netrin/neogenin 2:2 complex (Fig. 3) contains five protein-protein interfaces that fall in

three categories: Interface-1, between neogenin-FN4 and netrin-LN, buries ~680 Å² in each interacting domain and is dominated by van der Waals interactions between two largely hydrophobic surfaces. In addition, there are several peripheral hydrogen bonds and a salt bridge (Fig. 3B and figs. S3, S4, and S8). The LN Ca^{2+} binding site is immediately adjacent to interface-1, and bound Ca^{2+} would be required to maintain its proper conformation. Indeed, EDTA reduces the netrin/receptor binding affinities (Fig. 2B).

Interface-2, between neogenin-FN5 and netrin-LE3, buries ~610 Å² in each interacting domain and contains a mix of hydrophobic and polar contacts, including hydrogen bonds involving main-chain carbonyls and nitrogens (Fig. 3C and figs. S3, S4, and S8). This interface is centered around Met[985] of neogenin, the hydrophobic side chain of which is buried in a netrin surface hydrophobic pocket. Interface-1 is slightly larger and more hydrophobic than interface-2, consistent with our binding affinity measurements (Fig. 2A).

Interface-3 is the netrin dimerization interface in the netrin/neogenin 2:2 complex, burying ~1020 Å² in each interacting molecule. Notably, LE2 is longer than the other netrin LE domains, containing an

Fig. 2. Structure and receptor-binding affinity of the netrin-1 LN-LE region. (**A**) Binding of netrin-1 (LN-LE1-LE2-LE3) to different DCC constructs documenting that the receptor FN4-FN5 region is necessary and sufficient for netrin binding. K_d, dissociation constant (in μM). (**B**) Binding of netrin-1 (LN-LE1-LE2-LE3) to the FN4-FN5 region of the different neogenin and DCC isoforms (fig. S4). To evaluate the role of the netrin-bound Ca^{++}, 10 mM EDTA was added in one of the measurements. (**C**) Structure of unbound netrin-1. The individual netrin domains are labeled and colored in blue (LN), green (LE1), pink (LE2), and red (LE3). The glycosylation moieties at the three glycosylated Asn residues are shown as gray spheres. The N- and C-termini are labeled. (Inset) A close-up view of the calcium-binding sire in the LN domain. The calcium ion is drawn in magenta and two bound water molecules in red. D, Asp; F, Phe; S, Ser; T, Thr; and W, Trp.

A

Receptor	DCC FN1-5	DCC FN4	DCC FN4-5	DCC FN5	DCC FN4-5-6	DCC FN5-6
K_d (μM)	0.042	0.22	0.038	5.6	0.046	6.1

B

Receptor FN4-5	Neo-Short	Neo-Long	DCC-Short	DCC-Long	Neo-Short + EDTA
K_d (μM)	0.027	0.024	0.037	0.023	0.15

C
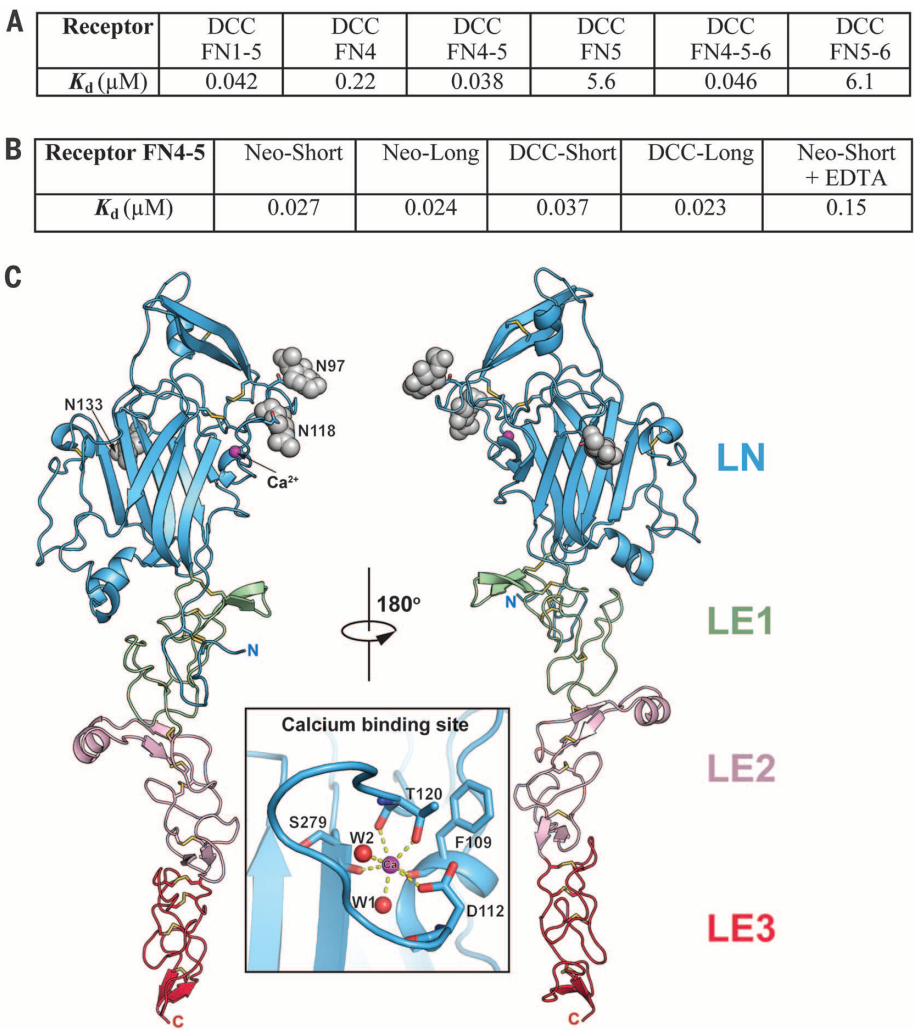
extra strand-helix-strand motif, which provides most of the dimerization contacts. This netrin region is conserved between the canonical netrins (netrin 1 to 5) but is very different in the G netrins, suggesting that the latter might not support this netrin-dimer architecture. Interface-3 is twofold symmetric, although it is not on a crystallographic symmetry axis because the crystal asymmetric unit contains the full 2:2 heterotetramer. Unlike interface-1 and -2, the vast majority of the interface-3 residues are polar, forming several hydrogen bonds and four salt bridges (Fig. 3D and figs. S3 and S8).

The netrin LN-LE region is positively charged [isoelectric point (pI) ~8.5], as is the neogenin/DCC FN4-FN5 region (pI ~9.2). The main positively charged surfaces on netrin (on its LE2 domain) and receptors (on FN5) are exposed to solvent in the complex, making them potentially available for interactions with negatively charged entities like proteoglycans (28, 29).

The structure of the netrin-1/DCC complex (Fig. 4A) was determined at 2.9 Å resolution and shows a different overall architecture than the netrin-1/neogenin structure, namely a continuous -DCC-netrin-DCC-netrin-DCC- assembly. Each netrin molecule still interacts, via its two receptor binding sites on the LN and LE3 domains, with two different DCC molecules. At the same time, each DCC receptor interacts with two netrin molecules via its two distinct netrin-binding sites on

FN4 and FN5, but these two netrins are shared with two other DCC neighbors (Fig. 4D). The reason the netrin/neogenin complex architecture cannot be replicated in the netrin/DCC complex is that the FN4-FN5 linker is slightly shorter in DCC than in neogenin (fig. S4) and the DCC linker also forms a short α helix. Formation of the 2:2 netrin-1/neogenin signaling complex around the X-shaped netrin dimer requires full extension of neogenin FN4-FN5 linkers, but the DCC FN4-FN5 linker is not long enough to accommodate this architecture. As in the netrin/neogenin complex, the DCC receptor molecules are parallel to each other, and their C-termini face the same direction.

The individual DCC and neogenin FNIII domains share about 70% sequence identity and their structures are very similar, with RMSDs between equivalent Cα positions of 0.47 Å for FN4 and 0.38 Å for FN5. The netrin-1 structure is also very similar in the complexes with its two receptors, with RMSDs between Cα positions of 0.73 Å. The netrin/DCC interfaces are nearly identical to the netrin/neogenin interfaces (Fig. 4 and fig. S4). Interface-1 is again formed between the netrin LN domain and the DCC FN4 domain, whereas interface-2 is formed between the netrin LE3 domain and the DCC FN5 domain. Figure 4, B and C, illustrates the similarly of the interacting LN-FN4 and LE3-FN5 domains in the two complexes.

On the basis of the structures reported here, we propose that netrin induces signaling by binding to and bringing together receptor molecules via its two binding sites, thus creating ligand/receptor signaling assemblies at the neuronal surface. One netrin molecule cannot form a bivalent complex with a single receptor molecule, even with the long isoforms, because the distance between the two receptor binding sites on netrin is larger than the distance between the two ligand-binding sites on receptors. Our structures illustrate the potential for netrins to cross-link different receptor types via distinct receptor binding sites: for example, DCC or neogenin on one end and Unc5 the other, or DCC on one end and neogenin the other. The netrin LC domain might further concentrate or cluster assemblies, for example, via interactions with heparan sulfate, because the degree of axon outgrowth seen with the netrin LN-LE region linked to a dimeric Fc tag is similar to that with full-length netrin-1 and much greater than with the LN-LE region alone (12).

The differences in the two signaling architectures result from different lengths of the linker connecting the receptor FN4 and FN5 domains, which differs between the DCC and neogenin isoforms studied here. In most species with these molecules, two different isoforms for each receptor, short and long, arise from alternative splicing of the FN4-FN5 linker sequence (25) (fig. S4). The

Fig. 3. Structure of the netrin-1/neogenin complex. (**A**) Structure of the 2:2 netrin-1/neogenin complex. The netrin molecules are in green and blue and the neogenin in orange and magenta. The N- and C-termini of the molecules are labeled. (**B**) Close-up view of the netrin-LN/neogenin-FN4 interface (interface-1). Interacting residues and the netrin-bound calcium are labeled. (**C**) Close-up view of the netrin-LE3/neogenin-FN5 interface (interface-2). Interacting residues are labeled. (**D**) Close-up view of the netrin-LE2/netrin-LE2 interface (interface-3). Interacting residues are labeled. C, Cys; G, Gly; H, His; M, Met; P, Pro; Q, Gln; V, Val.
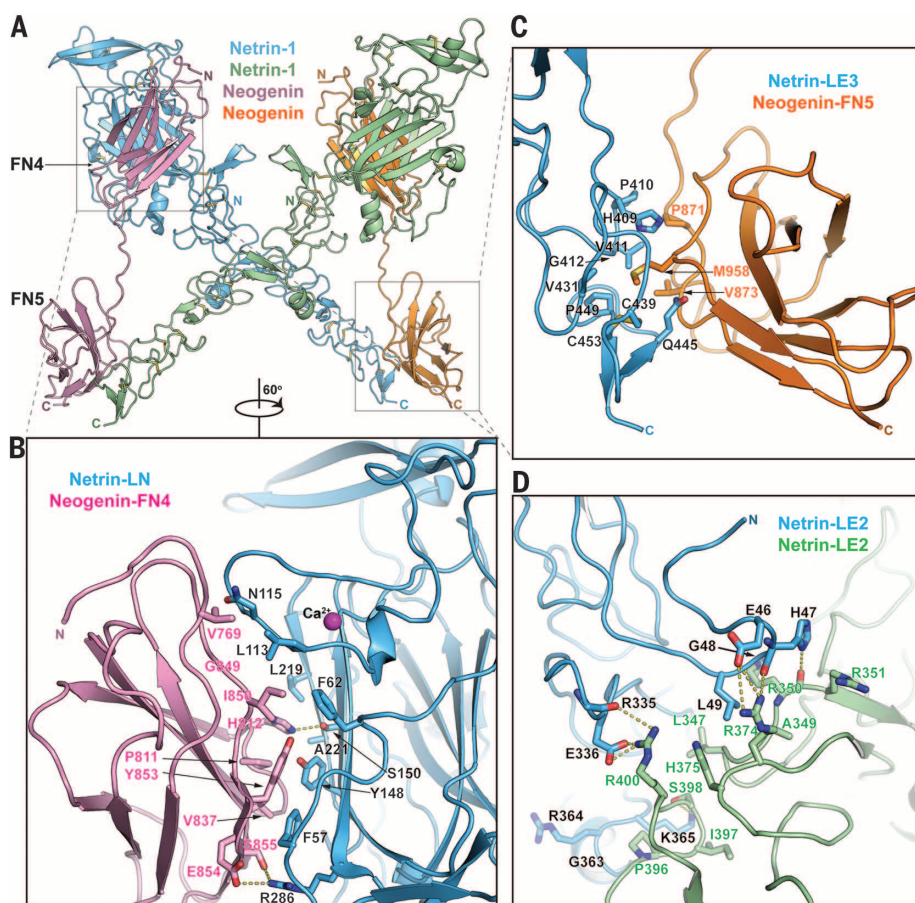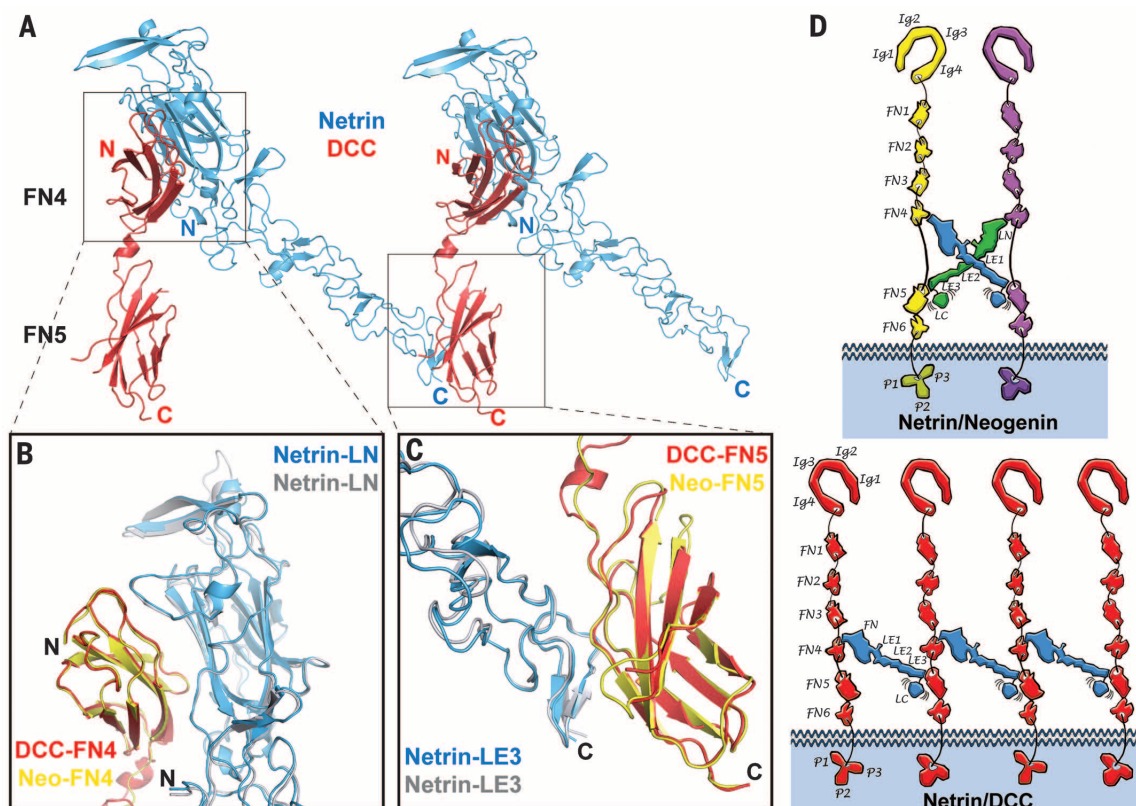
Fig. 4. Structure of the netrin-1/DCC complex. (**A**) Structure of the netrin-1/DCC complex. Netrin-1 is colored in blue, and DCC in red. (**B**) Superposition of the LN-FN4 interaction site (interface-1) in the netrin-1/DCC (blue/red) and netrin-1/neogenin (gray/yellow) complexes. (**C**) Superimposition of the LE3-FN5 interaction site (interface-2) in the netrin-1/DCC (blue/red) and netrin-1/neogenin (gray/yellow) complexes. (**D**) Schematic drawing comparing the distinct netrin-1/neogenin and netrin-1/DCC signaling assemblies. The individual netrin and receptor domains are labeled. Ig, immunoglobulin. (Top)

Schematic representation of the 2:2 netrin-1/neogenin complex. The netrin molecules are colored in blue and green, and the neogenin in yellow and magenta. (Bottom) Schematic representation of the continuous netrin-1/DCC assembly. The netrin molecules are colored in blue, and DCC in red. In both netrin-1/DCC and netrin-1/neogenin assemblies, the positively charged netrin LC domain would be positioned toward the negatively charged plasma membrane, thus potentially further stabilizing the signaling complexes at the neuronal surface.

two neogenin isoforms and the long DCC isoform all contain linkers long enough to support the 2:2 signaling-complex architecture, whereas the shorter DCC isoform does not allow this formation. Our binding experiments (Fig. 2B) suggest that the 2:2 assembly is energetically favored over the continuous assembly, presumably owing to the additional large interface between the two netrin molecules that it permits. The short neogenin and DCC isoforms are reported to predominate in many tissues (*30*) and could have distinct signaling properties mediated by distinct signaling-complex architectures (Fig. 4D). However, our results also highlight the need to evaluate which isoform of DCC is used in any given cell type to determine the potential of distinct assemblies to elicit particular cellular responses to netrin.

REFERENCES AND NOTES

1. T. Serafini *et al.*, *Cell* **78**, 409–424 (1994).
2. T. Serafini *et al.*, *Cell* **87**, 1001–1014 (1996).
3. C. Forcet *et al.*, *Nature* **417**, 443–447 (2002).
4. S. W. Moore, M. Tessier-Lavigne, T. E. Kennedy, *Adv. Exp. Med. Biol.* **621**, 17–31 (2007).
5. S. Rajasekharan, T. E. Kennedy, *Genome Biol.* **10**, 239 (2009).
6. J. Brasch, O. J. Harrison, G. Ahlsen, Q. Liu, L. Shapiro, *J. Mol. Biol.* **414**, 723–734 (2011).
7. E. Seiradake *et al.*, *EMBO J.* **30**, 4479–4488 (2011).
8. M. Tessier-Lavigne, C. S. Goodman, *Science* **274**, 1123–1133 (1996).
9. R. P. Kruger, J. Lee, W. Li, K. L. Guan, *J. Neurosci.* **24**, 10826–10834 (2004).
10. B. V. Geisbrecht, K. A. Dowd, R. W. Barfield, P. A. Longo, D. J. Leahy, *J. Biol. Chem.* **278**, 32561–32568 (2003).
11. F. Mille *et al.*, *Cell Death Differ.* **16**, 1344–1351 (2009).
12. K. Keino-Masu *et al.*, *Cell* **87**, 175–185 (1996).
13. H. Wang, N. G. Copeland, D. J. Gilbert, N. A. Jenkins, M. Tessier-Lavigne, *J. Neurosci.* **19**, 4938–4947 (1999).
14. C. H. Bell *et al.*, *Science* **341**, 77–80 (2013).
15. S. Rajagopalan *et al.*, *Nat. Cell Biol.* **6**, 756–762 (2004).
16. N. H. Wilson, B. Key, *Dev. Biol.* **296**, 485–498 (2006).
17. K. Srinivasan, P. Strickland, A. Valdes, G. C. Shin, L. Hinck, *Dev. Cell* **4**, 371–382 (2003).
18. E. C. Engle, *Cold Spring Harb. Perspect. Biol.* **2**, a001784 (2010).
19. S. Vulliemoz, O. Raineteau, D. Jabaudon, *Lancet Neurol.* **4**, 87–99 (2005).
20. M. Srour *et al.*, *Science* **328**, 592 (2010).
21. A. Fazeli *et al.*, *Nature* **386**, 796–804 (1997).
22. C. Sabatier *et al.*, *Cell* **117**, 157–169 (2004).
23. E. Palmesino, P. C. Haddick, M. Tessier-Lavigne, A. Kania, *J. Neurosci.* **32**, 411–416 (2012).
24. J. Kappler *et al.*, *Biochem. Biophys. Res. Commun.* **271**, 287–291 (2000).
25. H. Shen, H. Illges, A. Reuter, C. A. Stuermer, *Mech. Dev.* **118**, 219–223 (2002).
26. S. A. Hussain, F. Carafoli, E. Hohenester, *EMBO Rep.* **12**, 276–282 (2011).
27. F. Carafoli, S. A. Hussain, E. Hohenester, *PLOS ONE* **7**, e42473 (2012).
28. E. Kastenhuber *et al.*, *J. Neurosci.* **29**, 8914–8926 (2009).
29. Y. Matsumoto, F. Irie, M. Inatani, M. Tessier-Lavigne, Y. Yamaguchi, *J. Neurosci.* **27**, 4342–4350 (2007).
30. C. Manitt, K. M. Thompson, T. E. Kennedy, *J. Neurosci. Res.* **77**, 690–700 (2004).

HUMAN GENETICS

The genetics of Mexico recapitulates Native American substructure and affects biomedical traits

Andrés Moreno-Estrada,[1]*† Christopher R. Gignoux,[2]†‡ Juan Carlos Fernández-López,[3]†
Fouad Zakharia,[1] Martin Sikora,[1] Alejandra V. Contreras,[3] Victor Acuña-Alonzo,[4,5]
Karla Sandoval,[1] Celeste Eng,[6] Sandra Romero-Hidalgo,[3] Patricia Ortiz-Tello,[1]
Victoria Robles,[1] Eimear E. Kenny,[1]§ Ismael Nuño-Arana,[7] Rodrigo Barquera-Lozano,[4]
Gastón Macín-Pérez,[4] Julio Granados-Arriola,[8] Scott Huntsman,[6] Joshua M. Galanter,[6,9]
Marc Via,[6]‖ Jean G. Ford,[10] Rocío Chapela,[11] William Rodriguez-Cintron,[12]
Jose R. Rodríguez-Santana,[1,3] Isabelle Romieu,[14] Juan José Sienra-Monge,[15]
Blanca del Rio Navarro,[15] Stephanie J. London,[16] Andrés Ruiz-Linares,[5]
Rodrigo Garcia-Herrera,[3] Karol Estrada,[3]¶ Alfredo Hidalgo-Miranda,[3]
Gerardo Jimenez-Sanchez,[3]# Alessandra Carnevale,[3] Xavier Soberón,[3]
Samuel Canizales-Quinteros,[3,17] Héctor Rangel-Villalobos,[7] Irma Silva-Zolezzi,[3]**
Esteban Gonzalez Burchard,[6,9]* Carlos D. Bustamante[1]*

Mexico harbors great cultural and ethnic diversity, yet fine-scale patterns of human genome-wide variation from this region remain largely uncharacterized. We studied genomic variation within Mexico from over 1000 individuals representing 20 indigenous and 11 mestizo populations. We found striking genetic stratification among indigenous populations within Mexico at varying degrees of geographic isolation. Some groups were as differentiated as Europeans are from East Asians. Pre-Columbian genetic substructure is recapitulated in the indigenous ancestry of admixed mestizo individuals across the country. Furthermore, two independently phenotyped cohorts of Mexicans and Mexican Americans showed a significant association between subcontinental ancestry and lung function. Thus, accounting for fine-scale ancestry patterns is critical for medical and population genetic studies within Mexico, in Mexican-descent populations, and likely in many other populations worldwide.

Understanding patterns of human population structure, where regional surveys are key for delineating geographically restricted variation, is important for the design and interpretation of medical genetic studies. In particular, we expect rare genetic variants, including functionally relevant sites, to exhibit little sharing among diverged populations (1). Native Americans display the lowest genetic diversity of any continental group, but there is high divergence among subpopulations (2). As a result, present-day American indigenous populations (and individuals with indigenous ancestry) may harbor local private alleles rare or absent elsewhere, including functional and medically relevant variants (3, 4). Mexico serves as an important focal point for such analyses, because it harbors one of the largest sources of pre-Columbian diversity and has a long history of complex civilizations with varying contributions to the present-day population.

Previous estimates of Native Mexican genetic diversity examined single loci or were limited to a reduced number of populations or small sample sizes (5–8). We examined local patterns of variation from nearly 1 million genome-wide autosomal single-nucleotide polymorphisms (SNPs) for 511 Native Mexican individuals from 20 indigenous groups, covering most geographic regions across Mexico (table S1). Standard principal components analysis (PCA) summarizes the major axes of genetic variation in the data set [see (9)]. Whereas PC1 and PC2 separate Africans and Europeans from Native Mexicans, PC3 differentiates indigenous populations within Mexico, following a clear northwest-southeast cline (Fig. 1A). A total of 0.89% of the variation is explained by PC3, nearly three times as much as the variation accounted for by the north-south axis of differentiation within Europe [0.30%, in (10)]. The northernmost (Seri) and southernmost (Lacandon) populations define the extremes of the distribution, with very clear clustering of individuals by population, indicating high levels of divergence among groups (fig. S1). Seri and Lacandon show the highest level of population differentiation as measured by Wright's fixation index F_{ST} (0.136, Fig. 1B and table S4), higher than the F_{ST} between Europeans and Chinese populations in HapMap3 (0.11) (11). Other populations within Mexico also show extreme F_{ST} values; for example, the Huichol and Tojolabal have a pairwise F_{ST} of 0.068, similar to that observed between the Gujarati Indians and the Chinese in HapMap3 (0.076).

The high degree of differentiation between populations measured by F_{ST} argues that these populations have experienced high degrees of isolation. Indeed, when autozygosity is inferred using runs of homozygosity (ROH), all populations on average have long homozygous tracts, with the Huichol, Lacandon, and Seri all having on average over 10% of the genome in ROH [figs. S2 and S3 (9)]. These populations are relatively small, increasing the effects of genetic drift and driving some of the high F_{ST} values. In contrast, the Mayan and Nahuan populations have much smaller proportions of the genome in ROH, consistent with ROH levels found in Near Eastern populations in HGDP (12). These populations are the descendants of large Mesoamerican civilizations, and concordant with large historical populations, have relatively low proportions of ROH. The high degree of variance in ROH among populations is an additional indicator of population substructure and suggests a large variance in historical population sizes. Comparing the observed ROH patterns to those derived from coalescent simulations, we find that Native American groups within Mexico are characterized by small effective population sizes under a model with a strong bottleneck, in agreement with other studies of Native American populations (13). The degree of population size recovery to the current day is consistent with the degree of isolation of the extant populations, ranging from 1196 chromosomes [95% confidence interval (CI) 317 to 1548] for the Seri in the Sonora desert, to 3669 (95% CI 2588 to 5522) for the Mayans from Quintana Roo (figs. S4 to S6; (9)).

Isolation also correlates with the degree of relatedness within and between ethnic groups, ultimately shaping the pattern of genetic relationships among populations. We built a relatedness graph (Fig. 1C) of individuals sharing >13 cM of the genome identically by descent (IBD) (corresponding to third/fourth cousins or closer relatives). Almost all the connections are within- versus among-population, consistent with the populations being discrete rather than exhibiting large-scale gene flow [figs. S7 and S8 (9)].

As seen with the ROH calculations, the Mayan and Nahuan groups have fewer internal connections. The few between-population connections appear in populations close to each coastline, such as the connections between the Campeche Mayans and populations to the west along the Gulf of Mexico.

The long-tract ROH and IBD analyses are especially relevant to the recent history of isolation of Native American populations. We ran TreeMix (14) to generate a probabilistic model of diver-

gence and migration among the Native American populations (Fig. 1D). The inferred tree with no migration paths recapitulates the north/south and east/west gradients of differentiation from the PCA and IBD analyses, with populations with high ROH values also exhibiting longer tip branches. The primary branches divide populations by geography. All northern populations (dark blue) branch from the same initial split at the root. We also find two additional major clades: a grouping of populations from the southern states of Guerrero and Oaxaca (green labels) and a "Mayan clade" composed of Mayan-speaking populations from Chiapas and the Yucatan peninsula in the southeast (orange labels). Introducing migratory edges to the model connects the Maya in Yucatan to a branch leading to the Totonac, whose ancestors occupied the large pre-Columbian city of El Tajin in Veracruz (15). This result points to an Atlantic coastal corridor of gene flow between the Yucatan peninsula and central/northern Mexico (fig. S9), consistent with our IBD analysis. Indeed, the only Mayan language outside the Mayan territory is spoken by the Huastec, nearby in northern Veracruz, supporting a shared history (16).

These signals remain today as a legacy of the pre-Columbian diversity of Mexican populations. Over the past 500 years, population dynamics have changed drastically. Today, the majority of Mexicans are admixed and can trace their ancestry back not only to indigenous groups but also to Europe and Africa. To investigate patterns of admixture, we combined data from continental source populations (including the 20 native Mexican groups, 16 European populations, and 50 West African Yorubas) with 500 admixed mestizo individuals from 10 Mexican states recruited by the National Institute of Genomic Medicine (INMEGEN) for this study, Mexicans from Guadalajara in the POPRES collection (17), and individuals of Mexican descent from Los Angeles in the HapMap Phase 3 project (table S1). We ran the unsupervised mixture model algorithm ADMIXTURE (18) to estimate ancestry proportions for individuals in our combined data set (Fig. 2, fig. S10, and table S5). Allowing for three ancestral clusters ($K = 3$), we find that most individuals have a large amount of Native and European ancestry, with a small (typically <5%) amount of African ancestry. At the best-fit model for $K = 9$, the Native American cluster breaks down into six separate components (Fig. 2B). Three of these are mostly restricted to isolated populations (Seri, navy blue; Lacandon, yellow; and Tojolabal, brown). The other three show a wider but geographically well-defined distribution: A northern component (light blue) represented by Tarahumara, Tepehuano, and Huichol, gradually decreases southward. Correspondingly, a southern component (blue), which includes Triqui, Zapotec, and Mazatec, gradually decreases northward. In the Yucatan peninsula and the neighboring state of Chiapas, we found what we termed the "Mayan component" (orange in Fig. 2B, bottom panel), found primarily in Mayan-speaking groups. This Mayan component is also
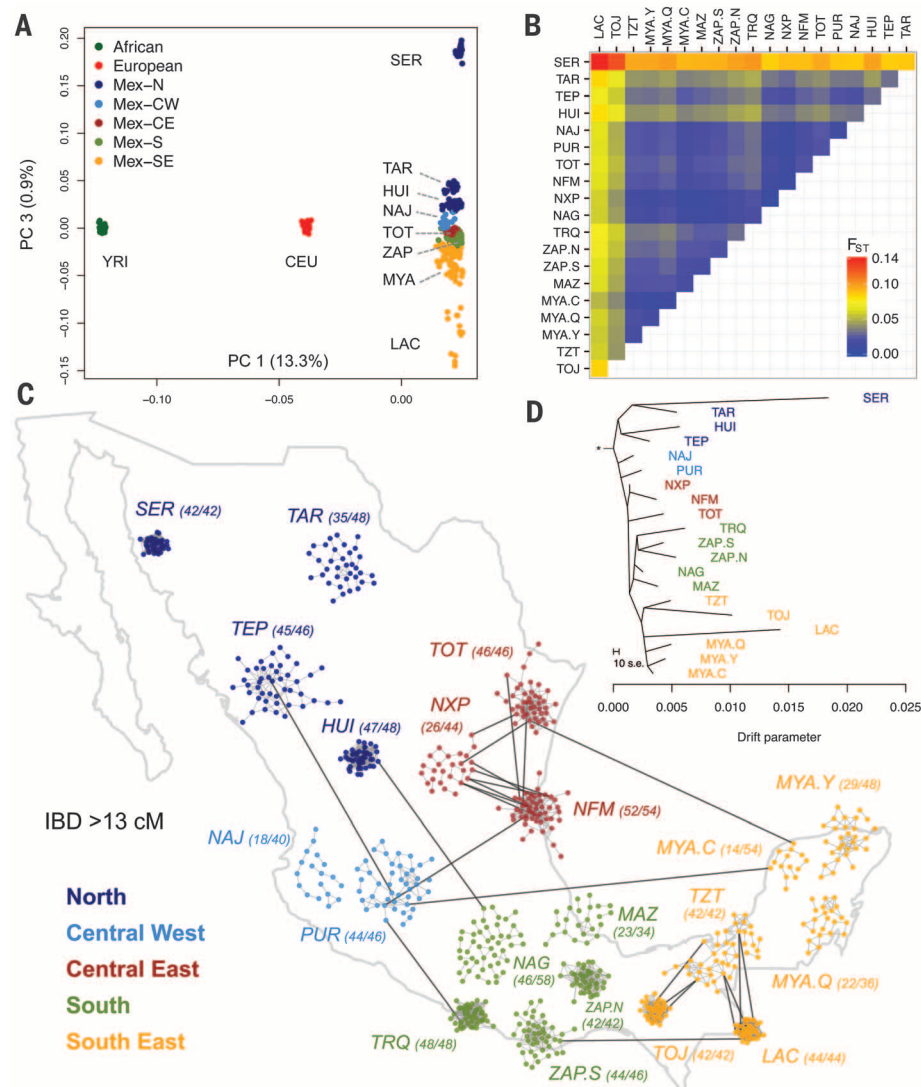
Fig. 1. Genetic differentiation of Native Mexican populations. (**A**) PCA of Native Mexicans with HapMap YRI and CEU samples. Population labels as in table S1. (**B**) Pairwise F_{ST} values among Native Mexican populations ordered geographically (see also table S4). (**C**) Relatedness graph of individuals sharing more than 13 cM of the genome as measured by the total of segments IBD. Each node represents a haploid genome, and edges within clusters attract nodes proportionally to shared IBD. The spread of each cluster is thus indicative of the level of relatedness in each population, as determined by a force-directed algorithm. Only the layout of nodes within each cluster is the result of the algorithm, as populations are localized to their approximate sampling locations to ease interpretation. Parentheses indicate the number of individuals represented out of the total sample size (2N). The full range of IBD thresholds are shown in fig. S8. (**D**) TreeMix graph representing population splitting patterns of the 20 Native Mexican groups studied. The length of the branch is proportional to the drift of each population. African, European, and Asian samples were used as outgroups to root the tree (fig. S9).

Fig. 2. Mexican population structure. (**A**) Map of sampled populations (detailed in table S1) and admixture average proportions (table S5). Dots correspond to Native Mexican populations color-coded according to $K = 9$ clusters identified in (B) (bottom), and shaded areas denote states in which cosmopolitan populations were sampled. Pie charts summarize per-state average proportions of cosmopolitan samples at $K = 3$ (European in red, West African in green, and Native American in gray). Bars show the total Native American ancestry decomposed into average proportions of the native subcomponents identified at $K = 9$. (**B**) Global ancestry proportions at $K = 3$ (top) and $K = 9$ (bottom) estimated with ADMIXTURE, including African, European, Native Mexican, and cosmopolitan Mexican samples (tables S1 and S2). From left to right, Mexican populations are displayed north-to-south. (**C**) Interpolation maps showing the spatial distribution of the six native components identified at $K = 9$. Contour intensities are proportional to ADMIXTURE values observed in Native Mexican samples, with crosses indicating sampling locations. Scatter plots with linear fits show ADMIXTURE values observed in cosmopolitan samples versus a distance metric summarizing latitude and longitude (long axis) for the sampled states. From left to right: Yucatan, Campeche, Oaxaca, Veracruz, Guerrero, Tamaulipas, Guanajuato, Zacatecas, Jalisco, Durango, and Sonora. Values are adjusted relative to the total Native American ancestry of each individual (9).
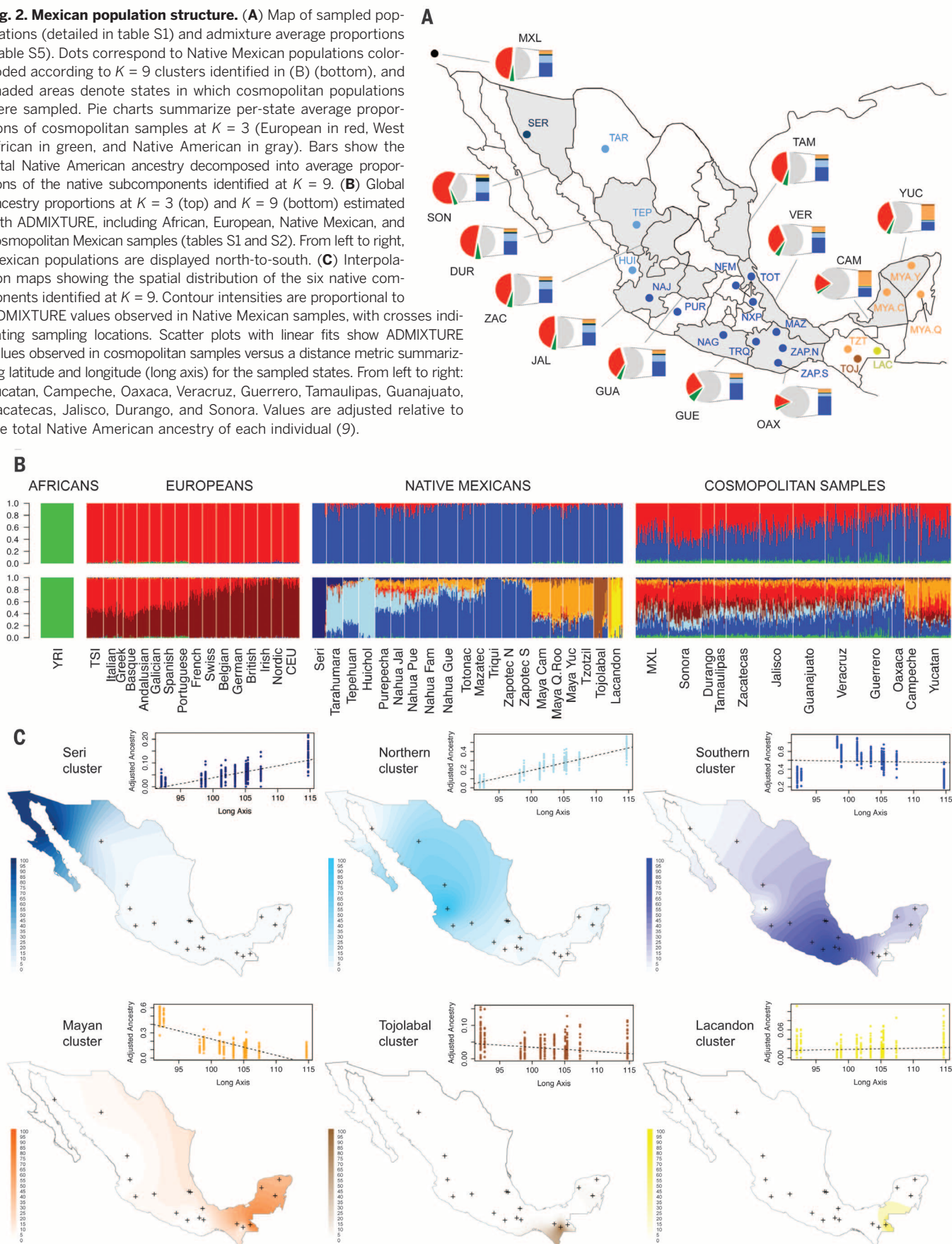
present at ~10 to 20% in central Mexican natives, consistent with the IBD and migration edges connecting the regions. This relationship between the Yucatan peninsula and central Mexico, seen in both recent shared IBD and genetic drift–based models of allele frequencies (TreeMix and ADMIXTURE), suggests that gene flow between the two regions has been ongoing for a long time. In contrast, Mayan admixture is not found at appreciable levels in highlanders of the southern state of Oaxaca (Triqui and Zapotec), where mountain ranges may have acted as geographic barriers to gene flow.

Patterns of Native American population substructure are recapitulated in the genomes of Mexican mestizos from cosmopolitan populations throughout Mexico. Sonora and neighboring northern states show the highest average proportions of the northern native component (15%, light blue in Fig. 2B, bottom), whereas only traces are detected in Oaxaca and the Yucatan peninsula. Conversely, the southern native component is the most prevalent across states, reaching maximum values in Oaxaca and decreasing northward. Cosmopolitan samples from the Yucatan peninsula have Native American fractions of the genome dominated by the Mayan component, which diminishes in northward populations. Likewise, Mayan-related local components, Tojolabal and Lacandon, are detected above 1% exclusively among individuals from the states neighboring the Yucatan peninsula. In contrast, Mexican-Americans sampled in Los Angeles (MXL) do not show a homogeneous pattern, consistent with their diverse origins within Mexico. Overall, the continuous geographic distribution of each Native American component across Mexico (fig. S12) demonstrates a high correlation of individual admixture proportions with geography, even in individuals of mixed ancestry (Fig. 2C, NW-SE axis F-test for all native clusters, $P < 10^{-16}$).

To further test whether ancestral population structure is recapitulated in the genomes of mestizos, we used an ancestry-specific PCA (ASPCA) approach [fig. S13 (9, 19)]. We estimated local ancestry using PCAdmix (20) to identify segments of the genome belonging to Native American, European, or African ancestries. We focused on the European and Native American components of ancestry, given the low proportions of African ancestry overall. We would expect the history of Spanish occupation and colonization in Mexico to be reflected in the European segments of Mexican mestizos, as has been seen previously (21). ASPCA of the European haplotypes in present-day Mexicans confirms this, as individuals cluster tightly with present-day Iberians even with a dense set of European populations (17, 22) (fig. S14).

In contrast, given the complex demographic history of Native Americans, high isolation, and limited characterization of regional ancestry patterns (23, 24), it remains unknown whether the correlation between genes and geography observed in Europe (10) can be similarly recapitulated within Mexico. We used ASPCA to uncover

hidden population structure within Native American ancestry beyond that found solely in extant indigenous groups (Fig. 3A). Consistent with the previous PCA analyses, we observed the most diverged indigenous populations defining the extremes of the top PCs due to high levels of genetic drift and isolation. However, including all the indigenous groups in the plot masks the signal contained in the indigenous segments of the Mexican mestizos. When plotting the ASPCA values for the admixed individuals only, we discovered a strong correlation between Native American ancestry and geography within Mexico (Fig. 3B), with ASPC1 representing a west-to-east dimension and ASPC2 one from north to south. Both of these correlations are highly significant and linearly predictive of geographic location (Pearson's r^2 of 72% and 38% for ASPC1 and 2, respectively, both P values $< 10^{-5}$). The correlation is strong enough that the overall distribution of mestizo-derived indigenous haplotypes in ASPCA space resembles a geographic map of Mexico (Fig. 3B and fig. S15). This finding suggests that the genetic composition of present-day Mexicans recapitulates ancient Native American substructure, despite the potential homogenizing effect of postcolonial admixture. Fine-scale population structure going back centuries is not merely a property of isolated or rural indigenous communities. Cosmopolitan populations still reflect the underlying genetic ancestry of local native populations, arguing for a strong relationship between the indigenous and the Mexican mestizo population, albeit without the extreme drift exhibited in some current indigenous groups.

Having found these hidden patterns of ancestry in the native component of Mexican mestizos, we investigated whether this structure could have potential biomedical applications. Over the past decade, genetic ancestry has been associated with numerous clinical endpoints and disease risks in admixed populations, including neutrophil counts (25), creatinine levels (26), and breast cancer susceptibility (27). Similarly, ancestral background is especially important in pulmonary medicine, where different reference equations are used for different ethnicities, defining normative predicted volumes and identifying thresholds for disease diagnosis in standard clinical practice (28). That is, depending on one's ethnic background, the same value of forced expiratory volume in 1 s (FEV$_1$, a standard measure of lung function) could be either within the normal range or indicative of pulmonary disease. Previous work has shown that the proportion of African and European ancestry was associated with FEV$_1$ in African Americans (29) and Mexicans (30), respectively, establishing the importance of genomic ancestry in lung function prediction equations.

To investigate possible associations between ancestral structure in Mexicans and FEV$_1$, we applied our ASPCA approach to two studies measuring lung function in Mexican or Mexican-American children: the Mexico City Childhood Asthma Study (MCCAS) (31) and the Genetics of Asthma in Latino Americans (GALA I) Study

(32). Due to differences in protocols and genotyping platforms, we calculated ASPCA values for the two studies independently (fig. S17) using the same reference populations described above, then used fixed-effects meta-analysis to combine the results (9).

First, in GALA I we looked for significant ancestry-specific differences between Mexico City and the San Francisco Bay Area, the two recruitment sites. ASPCA values were associated with recruitment location, with the receiver-operator characteristic curve from the Native American ancestry dimensions resulting in an area under the curve (AUC) of 80% (fig. S17). After we adjusted for overall ancestry proportions (here both African and Native American), both ASPCs were significant in a logistic regression: ASPC1 OR per SD: 0.44 (95% CI 0.22 to 0.68), $P = 3.8 \times 10^{-4}$, ASPC2 OR per SD: 1.68 (95% CI 1.03 to 2.76), $P = 0.039$. The ASPCs defined similar east-west and north-south axes as in the previous analysis (fig. S17) and show that Mexican-Americans in the San Francisco Bay Area tend to have increased Native American ancestry from northwest Mexico as compared to individuals from Mexico City (joint logistic regression likelihood ratio test $P = 6.4 \times 10^{-5}$).

We then used the ASPCA values for both studies to test for an association with FEV$_1$ as transformed to percentile of predicted "normal" function via the standard set of reference equations (28) for individuals of Mexican descent. These equations use population-specific demographic characteristics to account for age, sex, and height in estimates of lung function. Adjusting for overall ancestry proportions in linear regressions, we observed a significant association between FEV$_1$ and the east-west component (ASPC1) in both studies, with a meta-analysis P value of 0.0045 (2.2% decrease in FEV$_1$ per 1 SD, 95% CI 0.69 to 3.74). The effect sizes were homogeneous (Fig. 3C and table S6) despite differences in recruitment strategy, geography, and genotyping platform (9). In contrast, ASPC2 showed no association with FEV$_1$. Whereas FEV$_1$ has previously been associated with overall ancestry in several populations, the effect seen here is not correlated with overall admixture proportions, because we adjusted for those in the regression model. The combined results here indicate that subcontinental ancestry as measured by ASPCA is important for characterizing clinical measurements.

To estimate how variation in genetic ancestry within Mexico may affect FEV$_1$, we used the results from GALA I and MCCAS to predict trait values by state (Fig. 3D) for the INMEGEN mestizo samples. We found that difference in subcontinental Native American ancestry as measured by ASPC1 results in an expected 7.3% change in FEV$_1$, moving from the state of Sonora in the west to the state of Yucatan in the east. These results suggest that fine-scale patterns of native ancestry alone could have significant impacts on clinical measurements of lung function in admixed individuals within Mexico.

These changes due to ancestry are comparable to other factors affecting lung function. Comparing the expected effect of ancestry across Mexico with the known effects of age in the standard Mexican-American reference equations (28), the inferred 7.3% change in FEV_1 associated with subcontinental ancestry is similar to the decline in FEV_1 that a 30-year-old Mexican-American individual of average height would experience by aging 10.3 years if male and 11.8 years if female. Similarly, comparing our results from the Mexican data with the model incorporating ancestry in African Americans, a difference of 7.3% in FEV_1 would correspond to a 33% difference in African ancestry (29). The association between FEV_1 and ASPC1 is not an indicator of impaired lung function on its own–rather, it contributes to the distribution of FEV_1 values and would modify clinical thresholds. This finding indicates that diagnoses of diseases such as asthma and chronic obstructive pulmonary disease (COPD) relying on specific lung function thresholds may benefit from taking finer-scale ancestry into consideration.

An important implication of our work is that multi- and transethnic mapping efforts will benefit from including individuals of Mexican ancestry, because the Mexican population harbors rich amounts of genetic variation that may underlie important biomedical phenotypes. A key question in this regard is whether existing catalogs of human genome variation capture the genetic variation present in the samples analyzed here. We performed targeted SNP tagging and genome-wide haplotype sharing analysis within 100-kb sliding windows to assess the degree to which haplotype diversity in the Mexican mestizo samples could be captured by existing reference panels [figs. S18 to 20 (9)]. Although Mexican-American samples (MXL) were included in both the HapMap and 1000 Genomes catalogs, average haplotype sharing for the INMEGEN mestizo samples is limited to 81.2 and 90.5% when combined with all continental HapMap populations. It is only after including the Native American samples genotyped here that nearly 100% of haplotypes are shared, maximizing the chances of capturing most of the variation in Mexico.

Much effort has been invested in detecting common genetic variants associated with complex disease and replicating associations across populations. However, functional and medically relevant variation may be rare or population-specific, requiring studies of diverse human populations to identify new risk factors (4). Without detailed knowledge of the geographic
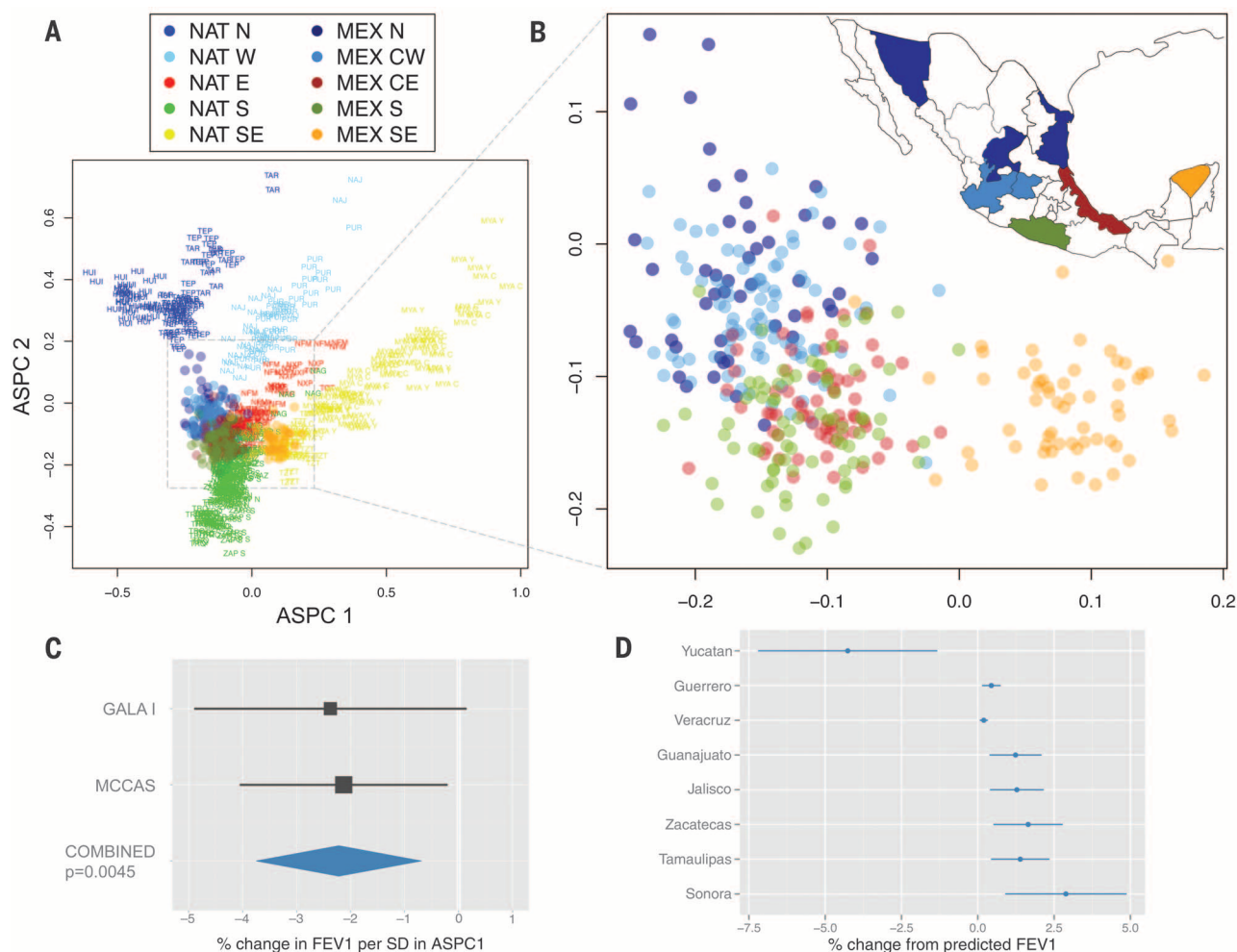
Fig. 3. Subcontinental ancestry of admixed Mexican genomes and biomedical implications. (**A**) ASPCA of Native American segments from Mexican cosmopolitan samples (colored circles) together with 20 indigenous Mexican populations (population labels). Samples with >10% of non-native admixture were excluded from the reference panel, as well as population outliers such as Seri, Lacandon, and Tojolabal. (**B**) Zoomed detail of the distribution of the Native American fraction of cosmopolitan samples throughout Mexico. Native ancestral populations were used to define PCA space (prefixed by NAT) but removed from the background to highlight the subcontinental origin of admixed genomes (prefixed by MEX). Each circle represents the combined set of haplotypes called Native American along the haploid genome of each sample with >25% of Native American ancestry. The inset map shows the geographic origin of cosmopolitan samples per state, color-coded by region (9). (**C**) Coefficients and 95% CIs for associations between ASPC1 and lung function (FEV_1) from Mexican participants in the GALA I study, and the MCCAS, as well as both studies combined (table S6 and fig. S17) (9). (**D**) Means and CIs of predicted change in FEV_1 by state, extrapolated from the model in (C).

stratification of genetic variation, negative results and lack of replication are likely to dominate the outcome of genetic studies in uncharacterized populations. Here we have demonstrated a high degree of fine-scale genomic structure across Mexico, shaped by pre-Columbian population dynamics and affecting the present-day genomes of Mexican mestizos, which is of both anthropological and biomedical relevance. Studies such as this one are crucial for enabling precision medicine, providing novel data resources, empowering the next generation of genetic studies, and demonstrating the importance of understanding and measuring fine-scale population structure and its associations with biomedical traits.

REFERENCES AND NOTES

1. S. Gravel et al., Proc. Natl. Acad. Sci. U.S.A. 108, 11983–11988 (2011).
2. S. Wang et al., PLOS Genet. 3, e185 (2007).
3. V. Acuña-Alonzo et al., Hum. Mol. Genet. 19, 2877–2885 (2010).
4. A. L. Williams et al., Nature 506, 97–101 (2014).
5. R. Lisker, E. Ramírez, V. Babinsky, Hum. Biol. 68, 395–404 (1996).
6. K. Sandoval et al., Am. J. Phys. Anthropol. 148, 395–405 (2012).
7. A. Gorostiza et al., PLOS ONE 7, e44666 (2012).
8. D. Reich et al., Nature 488, 370–374 (2012).
9. See supplementary materials on Science Online.
10. J. Novembre et al., Nature 456, 98–101 (2008).
11. D. M. Altshuler et al., Nature 467, 52–58 (2010).
12. B. M. Henn et al., PLOS ONE 7, e34267 (2012).
13. J. Hey, PLOS Biol. 3, e193 (2005).
14. J. K. Pickrell, J. K. Pritchard, PLOS Genet. 8, e1002967 (2012).
15. A. Pascual Soto, El Tajín. En Busca de los Orígenes de una Civilización (UNAM-INAH, Mexico, 2006).
16. L. Campbell, T. Kaufman, Annu. Rev. Anthropol. 14, 187–198 (1985).
17. M. R. Nelson et al., Am. J. Hum. Genet. 83, 347–358 (2008).
18. D. H. Alexander, J. Novembre, K. Lange, Genome Res. 19, 1655–1664 (2009).
19. A. Moreno-Estrada et al., PLOS Genet. 9, e1003925 (2013).
20. A. Brisbin et al., Hum. Biol. 84, 343–364 (2012).
21. N. A. Johnson et al., PLOS Genet. 7, e1002410 (2011).
22. L. R. Botigué et al., Proc. Natl. Acad. Sci. U.S.A. 110, 11791–11796 (2013).
23. S. Wang et al., PLOS Genet. 4, e1000037 (2008).
24. I. Silva-Zolezzi et al., Proc. Natl. Acad. Sci. U.S.A. 21, 8611–8616 (2009).
25. M. A. Nalls et al., Am. J. Hum. Genet. 82, 81–87 (2008).
26. C. A. Peralta et al., Am. J. Nephrol. 31, 202–208 (2010).
27. L. Fejerman et al., Cancer Res. 68, 9723–9728 (2008).
28. J. L. Hankinson, J. R. Odencrantz, K. B. Fedan, Am. J. Respir. Crit. Care Med. 159, 179–187 (1999).
29. R. Kumar et al., N. Engl. J. Med. 363, 321–330 (2010).
30. K. Salari et al., Genet. Epidemiol. 29, 76–86 (2005).
31. D. B. Hancock et al., PLOS Genet. 5, e1000623 (2009).
32. D. G. Torgerson et al., J. Allergy Clin. Immunol. 130, 76, e12 (2012).

TRANSCRIPTION

Interactions between RNA polymerase and the "core recognition element" counteract pausing

Irina O. Vvedenskaya,[1]* Hanif Vahedian-Movahed,[2]* Jeremy G. Bird,[1,2]* Jared G. Knoblauch,[1] Seth R. Goldman,[1] Yu Zhang,[2] Richard H. Ebright,[2]† Bryce E. Nickels[1]†

Transcription elongation is interrupted by sequences that inhibit nucleotide addition and cause RNA polymerase (RNAP) to pause. Here, by use of native elongating transcript sequencing (NET-seq) and a variant of NET-seq that enables analysis of mutant RNAP derivatives in merodiploid cells (mNET-seq), we analyze transcriptional pausing genome-wide in vivo in Escherichia coli. We identify a consensus pause-inducing sequence element, $G_{-10}Y_{-1}G_{+1}$ (where –1 corresponds to the position of the RNA 3′ end). We demonstrate that sequence-specific interactions between RNAP core enzyme and a core recognition element (CRE) that stabilize transcription initiation complexes also occur in transcription elongation complexes and facilitate pause read-through by stabilizing RNAP in a posttranslocated register. Our findings identify key sequence determinants of transcriptional pausing and establish that RNAP-CRE interactions modulate pausing.

Regulation of gene expression during transcription elongation often involves sequences in DNA that cause the transcription elongation complex (TEC) to pause. Pausing can affect gene expression by facilitating engagement of regulatory factors, influencing formation of RNA secondary structures, and enabling synchronization of transcription and translation.

Several lines of evidence suggest that pausing involves specific sequence signals that inhibit nucleotide addition (1–11). To define key sequence determinants for pausing, we used native elongating transcript sequencing (NET-seq), which permits occupancies of TECs to be mapped genome-wide with base-pair resolution (12, 13) (fig. S1). The occupancy of the TEC at a given position is correlated with the tendency of the TEC to pause at the position. Accordingly, NET-seq analysis enables identification of pause sites. To perform NET-seq in Escherichia coli, cells carrying a chromosomal rpoC-3xFLAG gene, encoding RNAP β′ subunit with a C-terminal 3xFLAG tag were grown to midexponential phase; cells were flash-frozen and lysed; 3xFLAG-tagged TECs were immunoprecipitated with an antibody against FLAG; RNAs were extracted from TECs; and RNA 3′ ends were converted to cDNAs and analyzed using high-throughput sequencing. We defined pause sites as positions where TEC occupancy exceeded TEC occupancy at each position 25 base pairs (bp) upstream and downstream. We identified 15,553 pause sites, which corresponds to ~19,800 total pause sites, given the estimated ~78% saturation of the analysis (tables S1 to S7). Alignment of pause-site sequences revealed a clear consensus pause element (PE): $G_{-10}Y_{-1}G_{+1}$, where position –1 corresponds to the position of the RNA 3′ end (Fig. 1A and fig. S2). Of the

[1]Department of Genetics and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA. [2]Department of Chemistry and Waksman Institute, Rutgers University, Piscataway, NJ 08854, USA.
*These authors contributed equally. †Corresponding author. E-mail: bnickels@waksman.rutgers.edu (B.E.N.); ebright@waksman.rutgers.edu (R.H.E.)

identified pause sites, ~35% exhibited a 3-of-3 match to the consensus PE, and ~42% exhibited a 2-of-3 match to the consensus PE (tables S4 and S5). Comparing the total number of pause sites with a 3-of-3 match (~6900) to the total number of sequences in the transcribed portion of the genome with a 3-of-3 match (~43,500 to 58,000) (*14*) indicates that, under these conditions, functional pausing occurs at ~12 to 16% of sequences with a 3-of-3 match.

To validate the NET-seq results, we selected two consensus PEs for analysis in vitro. One, located in *yrbL*, has −1C, and the other, located in *gltP*, has −1T (Fig. 1B and fig. S3). Transcription assays show that the *yrbL* PE and *gltP* PE induce pausing in vitro (Fig. 1B and fig. S4) and that consensus base pairs at positions +1, −1, and −10 are required for efficient pause capture in vitro (Fig. 1B).

We next determined the contribution of individual bases of the consensus PE to pausing at the *yrbL* PE (Fig. 2A). The results indicate that introduction of each of the three nonconsensus base pairs at position +1 (C, A, and T) reduces pausing; introduction of each of the two nonconsensus base pairs at position −1 (A and G) reduces pausing; and introduction of two of three nonconsensus base pairs at position −10 (A and T) reduces pausing (Fig. 2A). In each case, the effect on pausing is manifest at the level of pause capture efficiency.

The consensus PE comprises sequence determinants that previously have been implicated as important for pausing: G at +1 (*10, 15*), T or C

at −1 (*3, 5, 6, 10*), and G at −10 (*4, 10*). The consensus PE is especially similar to a consensus pause–inducing sequence identified by Herbert *et al.* in single-molecule studies (*10*).

The sequence determinants in the consensus PE ($G_{-10}Y_{-1}G_{+1}$) can be understood in terms of the structure and mechanism of the TEC (Fig. 2B). In each nucleotide-addition cycle in transcription elongation, RNAP translocates between a "pretranslocated state," in which the RNAP active center "*i*" and "*i* + 1" sites interact with RNA positions −2 and −1, and a "posttranslocated state,"

Fig. 1. Identification of consensus PE.
(**A**) Sequence logo for consensus PE from NET-seq. Red, bases with ≥0.2 bit sequence-information content.
(**B**) In vitro transcription assays with consensus PEs and mutant PEs. (Top) Templates; (bottom) results. +29, RNA before addition of uridine 5'-triphosphate (UTP); +46, RNA in TEC at PE; RT, read-through RNA; red, consensus PE bases.
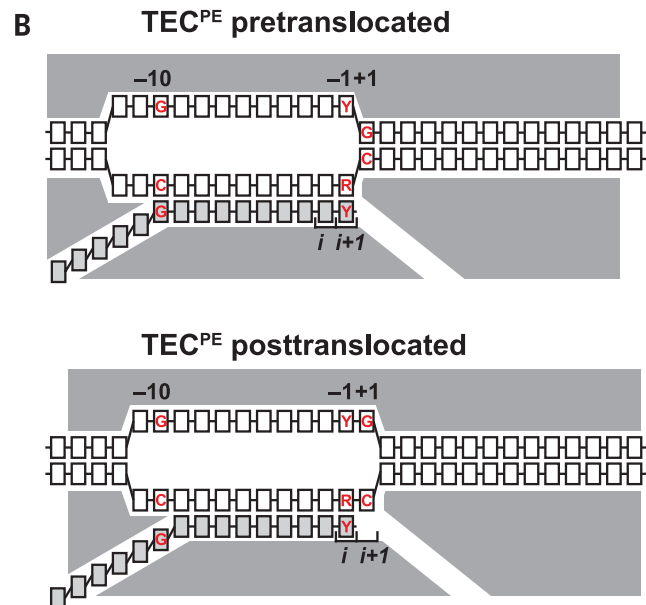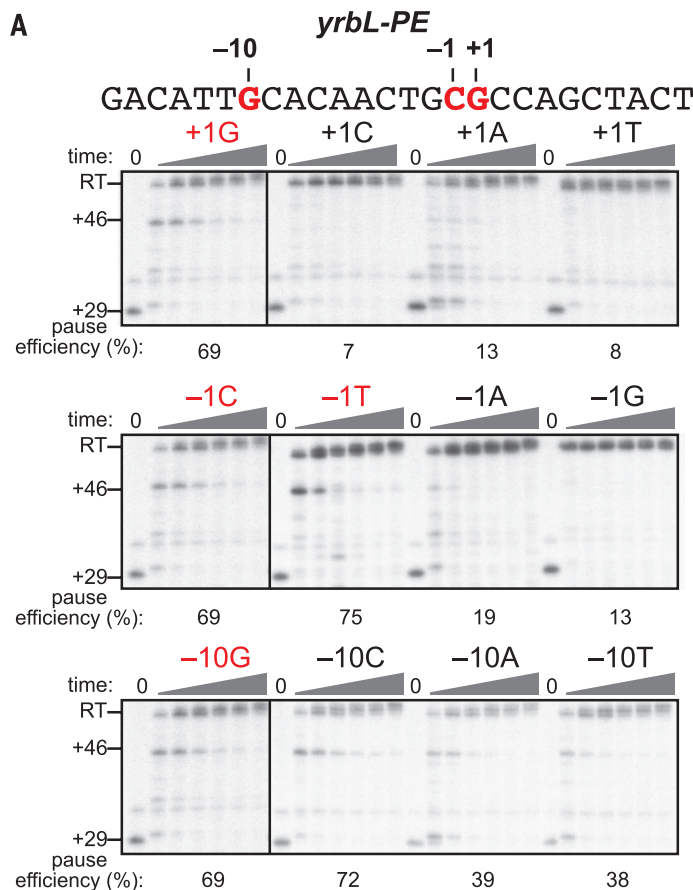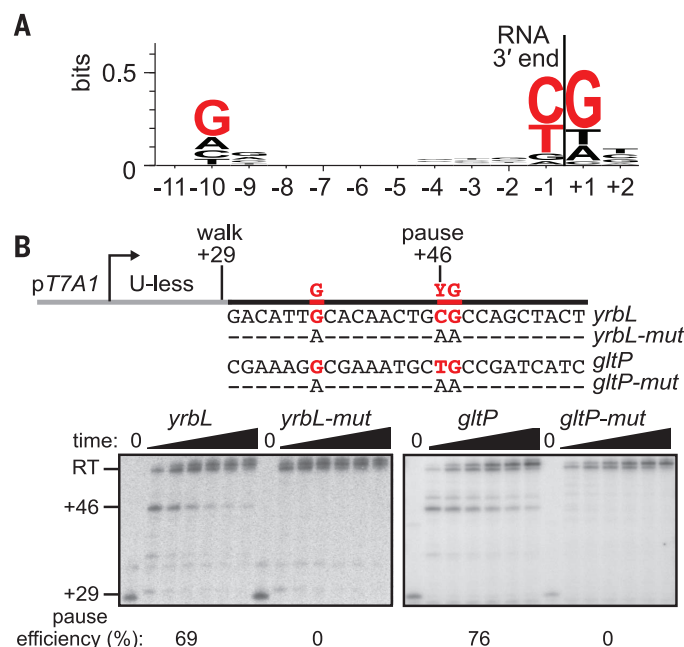
Fig. 2. Contributions of individual base pairs of consensus PE. (**A**) In vitro transcription assays with *yrbL*-PE derivatives. Red, consensus PE bases. (**B**) Schematic representation of TEC in pretranslocated state at consensus PE (top) and TEC in posttranslocated state at consensus PE (bottom). White boxes, DNA; gray boxes, RNA; gray shading, RNAP; red, consensus PE bases; *i* and *i* + 1, RNAP active center *i* and *i* + 1 sites.

in which the RNAP i site interacts with RNA position –1 and the RNAP $i + 1$ site is unoccupied and available for binding of a nucleoside triphosphate (NTP) (1, 2, 16). Translocation requires breaking the DNA base pair at position +1 and breaking the RNA-DNA base pair at position –10 (17, 18). Because the DNA base pair at position +1 must be broken for forward translocation, G/C at +1 will disfavor forward translocation relative to the less stable A/T. Similarly, because the RNA-DNA base pair at position –10 must be broken for forward translocation, G/C at –10 will disfavor forward translocation relative to the less stable A/T. Furthermore, because 5′-rG-rN-3′/3′-dC-dN-5′ is more stable than 5′-rC-rN-3′/3′-dG-dN-5′ (19), rG:dC at position –10 also will disfavor translocation over rC:dG at –10. Translocation requires movement of the template-strand DNA nucleotide and base-paired RNA nucleotide at position –1 from the RNAP active center $i + 1$ site to the i site and movement of the template-strand DNA nucleotide at position +1 to the $i + 1$ site. Because available evidence indicates that the RNAP active center i and $i + 1$ sites preferentially interact with 5′-rR-rY-3′/3′-dY-dR-5′ (3), a nontemplate-strand Y at position –1 and nontemplate-strand G at position +1 will disfavor forward translocation relative to all other sequences (the former by stabilizing the pre-

translocated state; the latter by destabilizing the posttranslocated state). Thus, each position of the consensus sequence is predicted to favor the pretranslocated state over the posttranslocated state (–10 through effects on duplex stability, –1 through effects on active-center binding, and +1 through both). Accordingly, each position of the consensus PE would be predicted to increase the opportunity for the TEC to enter an "elemental pause" state (a state that, according to one view, is accessed from the pretranslocated state and serves as an obligatory intermediate for pausing) (7, 8, 10, 11, 20) or a "backtracked" state (a state that, according to another view, is accessed from the pretranslocated state and serves as the primary state for pausing) (2, 16).

A TEC in a posttranslocated state at a PE will contain an unpaired G at the downstream end of the nontemplate strand of the unwound "transcription bubble" (Fig. 3A). In transcription initiation complexes, RNAP core enzyme makes sequence-specific interactions with an unpaired G at the downstream end of the transcription bubble ("core recognition element" CRE) (21) (Fig. 3A). In transcription initiation, interaction between RNAP and the unpaired G of the CRE (G_{CRE}) facilitates promoter unwinding to form a stable, catalytically competent, RNAP-promoter open complex (RPo) (21). It has been proposed that

RNAP-G_{CRE} interaction may be functionally important in transcription elongation as well as in transcription initiation (21), but this has not been previously documented.

The observation that the consensus PE contains the sequence feature required for establishment of RNAP-G_{CRE} interaction raises the possibility that RNAP-G_{CRE} interaction may mediate or modulate pausing. To test this possibility, we constructed and analyzed a mutant RNAP defective in sequence-specific recognition of G_{CRE}. The crystal structure of RPo shows that RNAP βD446 makes H bonds with Watson-Crick atoms of G_{CRE} and suggests that D446 reads the identity of G_{CRE} (21). Replacement of βD446 by alanine results in the loss of the ability to distinguish G, A, T, or an abasic site at the position corresponding to G_{CRE} (Fig. 3B and fig. S5), confirming that βD446 reads the identify of G_{CRE} and providing a reagent to assess functional significance of sequence-specific RNAP-G_{CRE} interactions.

To establish whether RNAP-CRE interactions affect pausing, we used a variant of NET-seq, "merodiploid NET-seq" (mNET-seq), that enables analysis of mutant RNAP derivatives, including mutant RNAP derivatives that do not support viability in haploid (fig. S1). mNET-seq uses merodiploid cells containing a plasmid-encoded, epitope-tagged RNAP and a chromosome-encoded,

Fig. 3. Sequence-specific RNAP-G_{CRE} interactions modulate pausing. (**A**) Structural organization of TEC in posttranslocated state at consensus PE (left) and RPo in transcription initiation at promoter containing consensus CRE (right) (21). The presence in each case of an unpaired G at downstream end of nontemplate strand of transcription bubble suggests the possibility of equivalent sequence-specific interactions between RNAP core and the G. Red at left, bases of the consensus PE; red at right, G_{CRE}. (**B**) Results of fluorescence-detection equilibrium assays (top) and kinetic (bottom) assays of interactions of RNAP derivatives with nucleic acid scaffolds containing G, A, T, or an abasic site (X) at position corresponding to G_{CRE}. ntDNA, nontemplate-strand DNA; tDNA template-strand DNA. (**C**) Sequence logos for consensus PE with RNAP-β[WT] [(top) 0.4 bit sequence-information content at G_{CRE}] and RNAP-β[D446A] [(bottom) 1 bit of sequence-information content at G_{CRE}], as defined by mNET-seq. Red, bases with ≥0.2 bit sequence information content. (**D**) Pause-capture efficiencies of RNAP β[WT] (left panels) and RNAP β[D446A] (right panels) at yrbL PE and gltP PE in vitro. +29, RNA before addition of UTP; +46, RNA in TEC at PE; asterisks, RNA in TECs at additional sites where RNAP β[D446A] exhibits higher pause-capture efficiency than RNAP β[WT]; RT, read-through RNA; red, consensus PE bases.

B

ntDNA TATAATGGGAGCT**N**
              TCACGGATGCAGG
tDNA   GGGAGCTGAGTGCCTACGTCC

N	RNAP-β[WT]		RNAP-β[D446A]	
	K_d (μM)	$K_{d,X}/K_{d,G}$	K_d (μM)	$K_{d,X}/K_{d,G}$
G	0.13	[1]	0.93	[1]
X	0.57	4	0.56	0.6
A	0.85	6	0.79	0.8
T	0.81	6	0.90	1

N	RNAP-β[WT]		RNAP-β[D446A]	
	k_{off} (h⁻¹)	$k_{off,X}/k_{off,G}$	k_{off} (h⁻¹)	$k_{off,X}/k_{off,G}$
G	1.4	[1]	11	[1]
X	4.8	3	5.9	0.5
A	6.5	5	16	1
T	7.4	5	14	1
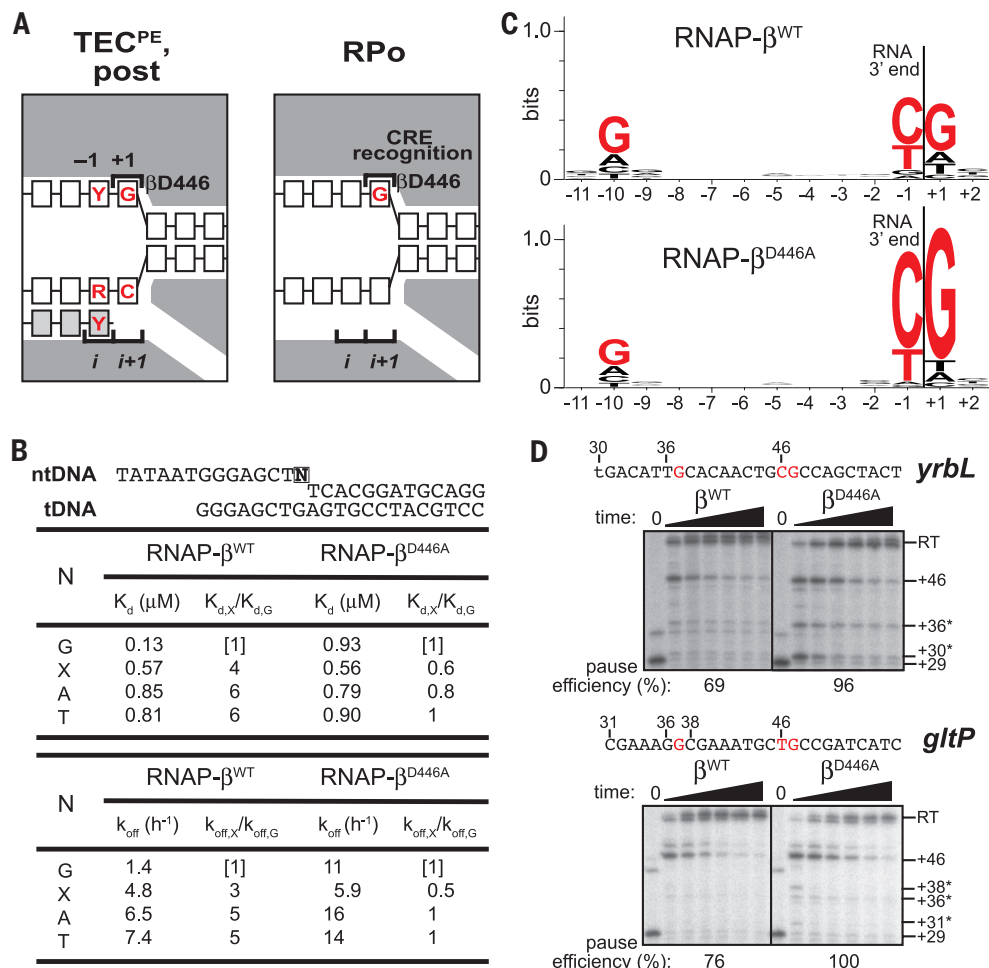
Fig. 4. Sequence-specific RNAP-G$_{CRE}$ interactions modulate translocation bias. (**A**) Structural organization of TEC in pretranslocated state at consensus PE (left; unfavorable RNAP-CRE interaction) and TEC in posttrans-located state at consensus PE (right; favorable RNAP-CRE interaction). PPi, pyrophosphate. (**B**) Nucleic-acid scaffolds used for translocation-bias assays. Asterisk, radiolabel on RNA 5′ end; box, position corresponding to G$_{CRE}$; red, consensus PE base. (**C**) Translocational bias for RNAP β^{WT} and RNAP β^{D446A} on nucleic-acid scaffolds containing G or T at position corresponding to G$_{CRE}$. Gel images show pyrophosphorolysis reaction progress from 0 to 30 min. Labels: 9-nucleotide (nt), scaffold; 8-nt, product of pyrophosphorolysis; 10-nt, product of "chase" reaction with GTP (+1G template) or UTP (+1T template). Graph shows fraction of unaltered scaffold (means ± SEM; three measurements) as function of time.

untagged RNAP, and it involves selective analysis of transcripts associated with epitope-tagged RNAP in the presence of a mixed population of epitope-tagged RNAP and untagged RNAP. We introduced into cells a plasmid encoding 3xFLAG-tagged wild-type RNAP β subunit (β^{WT}) or 3xFLAG-tagged RNAP β subunit containing D446A (β^{D446A}), we isolated RNAs associated with RNAP-β^{WT} or RNAP-β^{D446A} by immunoprecipitation, and we identified pause sites. For RNAP-β^{WT}, alignment of pause sites revealed a consensus sequence matching the consensus PE, which validated mNET-seq as an effective system for analysis of pausing (Figs. 1A and 3C; figs. S2 and S6; and tables S4 to S9). For RNAP-β^{D446A} we identified ~60 to 90% more pause sites than with RNAP-β^{WT} (tables S4 to S10). Alignment of the pause sites revealed that ~30% more of the pause sites carried a G at position +1 (Fig. 3C, fig. S6, and table S6). We conclude that RNAP-β^{D446A} is more susceptible than RNAP-β^{WT} to pausing at sites with G at position +1.

We next compared pausing properties of RNAP-β^{D446A} and RNAP-β^{WT} in vitro, using templates carrying the *yrbL* PE and *gltP* PE (Fig. 3D and fig. S7). RNAP-β^{D446A} enhances pausing at *yrbL* PE and *gltP* PE. RNAP-β^{D446A} also enhances pausing at other positions where the next nucleotide to be added to the transcript is G (positions with asterisks in Fig. 3D). The results indicate that a substitution that disrupts sequence-specific RNAP-CRE interaction increases pausing, both in vivo and in vitro, at positions where the post-translocated state contains G$_{CRE}$. We conclude that sequence-specific RNAP-G$_{CRE}$ interaction occurs during elongation and counteracts pausing.

To explore why disruption of sequence-specific RNAP-G$_{CRE}$ interaction enhances pausing, we assessed whether RNAP-G$_{CRE}$ interactions affect the translocational register of the TEC by

assessing sensitivity of TECs to pyrophosphorolysis (Fig. 4). Sensitivity to pyrophosphorolysis provides a measure of TEC translocation because a TEC in a pretranslocated state is sensitive to pyrophosphorolysis but a TEC in a posttrans-located state is resistant (Fig. 4A) (*3*). We performed assays with RNAP-β^{WT} or RNAP-β^{D446A} on templates containing G or T at position +1 (Fig. 4B). We found that TECs with RNAP-β^{WT} were ~5 times as sensitive to pyrophosphorolysis when the template contained +1G as when the template contained +1T, which indicated that a greater proportion of TECs on templates containing +1G were in a pretranslocated state than of TECs on templates containing +1T (Fig. 4C). The results directly demonstrate the effect of G at position +1 on translocation bias. TECs with RNAP-β^{D446A} were ~4 times as sensitive to pyrophosphorolysis as TECs with RNAP-β^{WT} when the template contained +1G (Fig. 4C), whereas, in contrast, TECs with RNAP-β^{D446A} and RNAP-β^{WT} exhibited identical sensitivities to pyrophosphorolysis when the template contained +1T (Fig. 4C). The results indicate that, on templates containing +1G, a greater proportion of TECs with RNAP-β^{D446A} are in a pretranslocated state as compared to TECs with RNAP-β^{WT}. We conclude that sequence-specific RNAP-G$_{CRE}$ interactions stabilize the TEC posttranslocated state, which provides a mechanistic explanation for the finding that RNAP-G$_{CRE}$ interactions counteract pausing.

Our findings define the key sequence determinants of transcriptional pausing as G$_{-10}$Y$_{-1}$G$_{+1}$ (Figs. 1 and 2). The consensus PE promotes pausing by disfavoring translocation of the TEC to the posttranslocated state (Fig. 2B), which increases the opportunity for the TEC to enter an "elemental pause" state and/or a "backtracked" state. We further show that sequence-specific RNAP-G$_{CRE}$ interactions counteract pausing by stabi-

lizing the TEC in a posttranslocated state (Figs. 3 and 4). Because residues of RNAP core that mediate sequence-specific RNAP-CRE interaction are conserved in RNAP from all living organisms, we suggest that RNAP-CRE interaction counteracts pausing in all multisubunit RNAPs. The consensus PE will, on average, be encountered by RNAP every ~32 bp during transcription elongation for organisms, such as *E. coli*, with ~50% G/C content and will be encountered even more frequently for organisms with higher G/C content. RNAP-G$_{CRE}$ interactions may help overcome a barrier to forward translocation that occurs each time RNAP encounters a G$_{-10}$Y$_{-1}$G$_{+1}$ sequence during transcription elongation. A major function of RNAP-G$_{CRE}$ interactions may be to suppress noise during transcription elongation by smoothing the sequence-dependent energy landscape for transcription elongation.

REFERENCES AND NOTES

1. M. H. Larson, R. Landick, S. M. Block, *Mol. Cell* **41**, 249–262 (2011).
2. M. Dangkulwanich, T. Ishibashi, L. Bintu, C. Bustamante, *Chem. Rev.* **114**, 3203–3223 (2014).
3. P. P. Hein, M. Palangat, R. Landick, *Biochemistry* **50**, 7002–7014 (2011).
4. S. Kyzer, K. S. Ha, R. Landick, M. Palangat, *J. Biol. Chem.* **282**, 19020–19028 (2007).
5. C. L. Chan, R. Landick, *J. Mol. Biol.* **233**, 25–42 (1993).
6. V. A. Aivazashvili, R. Sh. Bibilashvili, R. M. Vartikian, T. A. Kutateladze, *Mol. Biol. (Moscow)* **15**, 915–929 (1981).
7. A. Weixlbaumer, K. Leon, R. Landick, S. A. Darst, *Cell* **152**, 431–441 (2013).
8. R. Landick, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8797–8798 (2009).
9. M. L. Kireeva, M. Kashlev, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 8900–8905 (2009).
10. K. M. Herbert *et al.*, *Cell* **125**, 1083–1094 (2006).
11. K. C. Neuman, E. A. Abbondanzieri, R. Landick, J. Gelles, S. M. Block, *Cell* **115**, 437–447 (2003).
12. L. S. Churchman, J. S. Weissman, *Nature* **469**, 368–373 (2011).
13. Materials and methods are available as supplementary materials on *Science* Online.
14. B. J. Haas, M. Chin, C. Nusbaum, B. W. Birren, J. Livny, *BMC Genomics* **13**, 734 (2012).

15. D. N. Lee, L. Phung, J. Stewart, R. Landick, *J. Biol. Chem.* **265**, 15145–15153 (1990).
16. M. Dangkulwanich *et al.*, *eLife* **2**, e00971 (2013).
17. L. Bai, A. Shundrovsky, M. D. Wang, *J. Mol. Biol.* **344**, 335–349 (2004).
18. V. R. Tadigotla *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 4439–4444 (2006).
19. N. Sugimoto *et al.*, *Biochemistry* **34**, 11211–11216 (1995).
20. J. Zhou, K. S. Ha, A. La Porta, R. Landick, S. M. Block, *Mol. Cell* **44**, 635–646 (2011).
21. Y. Zhang *et al.*, *Science* **338**, 1076–1080 (2012).

DISEASE ECOLOGY

Ecological and evolutionary effects of fragmentation on infectious disease dynamics

Jussi Jousimo,[1]* Ayco J. M. Tack,[1]* Otso Ovaskainen,[1] Tommi Mononen,[1,2] Hanna Susi,[1] Charlotte Tollenaere,[1] Anna-Liisa Laine[1]†

Ecological theory predicts that disease incidence increases with increasing density of host networks, yet evolutionary theory suggests that host resistance increases accordingly. To test the combined effects of ecological and evolutionary forces on host-pathogen systems, we analyzed the spatiotemporal dynamics of a plant (*Plantago lanceolata*)–fungal pathogen (*Podosphaera plantaginis*) relationship for 12 years in over 4000 host populations. Disease prevalence at the metapopulation level was low, with high annual pathogen extinction rates balanced by frequent (re-)colonizations. Highly connected host populations experienced less pathogen colonization and higher pathogen extinction rates than expected; a laboratory assay confirmed that this phenomenon was caused by higher levels of disease resistance in highly connected host populations.

Although infectious diseases pose serious threats to human health and food security, pathogens are also a fundamental component of natural biodiversity (*1*, *2*). Infection prevalence fluctuates through space and time in natural plant-pathogen associations (*3–5*), but there is little evidence of devastating epidemics in native interactions. This is in stark contrast to the boom-and-bust dynamics of many fungal pathogens attacking commercially important crops under agricultural settings (*6*, *7*) and epidemics caused by introduced species (*8*). With most of the focus in epidemiology targeted toward understanding conditions generating infectious disease outbreaks (*9–11*), the mechanisms that enable long-term disease persistence in wild communities remain elusive.

Spatially explicit metapopulation theory has identified habitat size and connectivity to other habitat patches as the key parameters affecting species distributions (*12*). The limited data available have confirmed metapopulation theory to be relevant also for disease dynamics, with colonizations and extinctions of local host populations

generating the observed patterns of infection (*13*). Far less is known about how evolutionary dynamics vary through space and what the resulting effects on disease dynamics are (*14*). There is substantial theory demonstrating that the spatial scale of host-pathogen interactions and resulting rates of gene flow are critical for how coevolutionary dynamics play out between hosts and their pathogens (*15*, *16*). Hence, although host availability for pathogens increases with increasing density of host networks, the evolutionary potential of host populations may also change as a function of connectivity. Increased host gene flow into the local populations generates higher variation in resistance, allowing resistance to be selected for (*17*, *18*). A comparison across species found the relative rate of gene flow in hosts versus parasites to be the strongest predictor of pathogen local adaptation to host resistance. When host dispersal rates exceeded those of their parasites, the potential for parasite maladaptation emerged (*19*). However, such species-level approaches provide little insight into how coevolutionary dynamics and resulting infectious disease dynamics are played out across the spatially heterogeneous landscapes that most species inhabit. Indeed, the spatial and the genetic complexities that natural host-pathogen interactions support have rendered epidemiological predictions challenging in wild communities, yet identifying the effect of such complexity on epidemiological dynamics

may provide key insights into the battle against diseases.

To understand how spatial and evolutionary processes interact to shape disease dynamics, we analyzed the spatial and environmental factors that drive metapopulation dynamics of an obligate fungal pathogen occupying a highly fragmented host population network. In our analyses, we targeted the three key steps of metapopulation dynamics: presence-absence of disease within local host populations (i.e., occupancy); colonization by the pathogen of previously unoccupied host populations; and extinction, whereby a pathogen population previously found infecting a local host population is deemed lost. We also analyzed factors affecting within-population abundance of disease. Our spatial variables included host population size, presence of road, host population connectivity (S^H), and pathogen population connectivity at the beginning (S^P) and the end (S^M) of epidemics. The environmental variables were percentage of dry host plants within populations, habitat boundary type, light availability, and precipitation in July and August. We used a hierarchical spatiotemporal Bayesian logistic regression model that allows estimation of the residual spatial and temporal autocorrelation.

Our data were collected from *Podosphaera plantaginis*, a specialist powdery mildew naturally infecting *Plantago lanceolata* in the Åland archipelago, southwestern Finland. *P. lanceolata* is an obligate outcrossing perennial herb that reproduces both sexually and clonally via side rosettes. The pollen of *P. lanceolata* is wind-dispersed, and seeds typically drop close to the mother plant. In Åland, the host populations are highly fragmented, forming a network of several thousand populations spanning an area of about 50 km by 70 km. Locations of the local *P. lanceolata* populations in the Åland Islands have been annually mapped since the 1990s. Because of clonal reproduction and a seed bank, host populations rarely go extinct, and hence the spatial configuration of the host populations remains relatively constant (*12*). Visible signs of infection by *P. plantaginis* begin to appear in late June in those host populations where the pathogen has successfully survived the winter as resting spores. The resting spores remain attached to the senescent host leaves throughout the winter, and in the spring the spores may reinfect the same host plant or new host plants in close vicinity (*20*, *21*). The epidemic starts from these initial disease foci when the pathogen is transmitted among hosts by clonally produced dispersal spores that are passively carried by wind (*5*, *22*). Some six to eight clonally produced generations follow one another in quick

[1]Metapopulation Research Group, Department of Biosciences, University of Helsinki, Post Office Box 65 (Viikinkaari 1), FI-00014 University of Helsinki, Helsinki, Finland. [2]Department of Biomedical Engineering and Computational Science, Aalto University School of Science, FI-00076 Aalto, Finland.
*These authors contributed equally to this work. †Corresponding author. E-mail: anna-liisa.laine@helsinki.fi
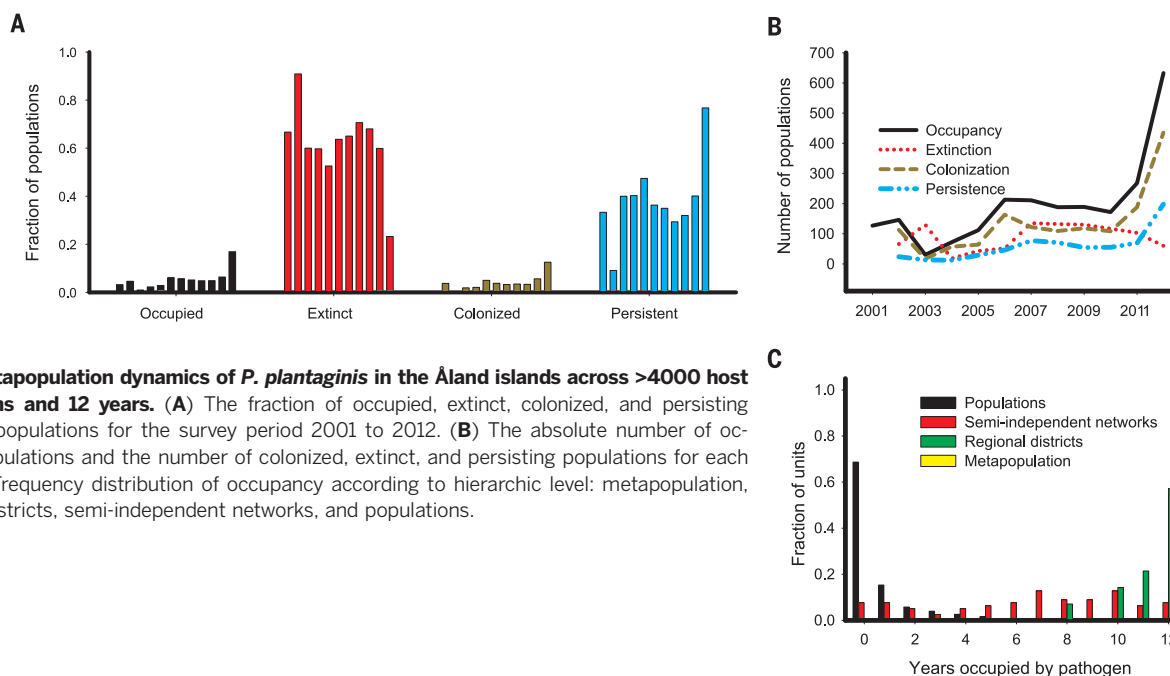
Fig. 1. Metapopulation dynamics of *P. plantaginis* in the Åland islands across >4000 host populations and 12 years. (**A**) The fraction of occupied, extinct, colonized, and persisting pathogen populations for the survey period 2001 to 2012. (**B**) The absolute number of occupied populations and the number of colonized, extinct, and persisting populations for each year. (**C**) Frequency distribution of occupancy according to hierarchic level: metapopulation, regional districts, semi-independent networks, and populations.

Fig. 2. Spatial distribution of host and pathogen populations. (**A**) Distribution of the host populations. (**B** to **D**) Colonized (yellow circles), extinct (red triangles), and persistent (blue squares) pathogen populations for a year with low, average, and high occupancy, respectively. (**E**) Fine-scale extinction-colonization dynamics in the regional district of Lemland in 2012, where black dots represent uncolonized host populations.

A

Average occupancy
(2001-2012)
- 1.80 to 1.96 years
- 1.00 to 1.16 years
- 0.84 to 0.99 years
- 0.68 to 0.83 years
- 0.52 to 0.67 years
- 0.36 to 0.51 years

10 km

B

Average occupancy
(2001-2012)
- 1.94 to 2.06 years
- 1.19 to 1.34 years
- 1.04 to 1.18 years
- 0.89 to 1.03 years
- 0.74 to 0.88 years
- 0.59 to 0.73 years

3 km

C

Years occupied
(2001-2012)
- 6
- 5
- 3
- 2
- 1
- 0

1 km
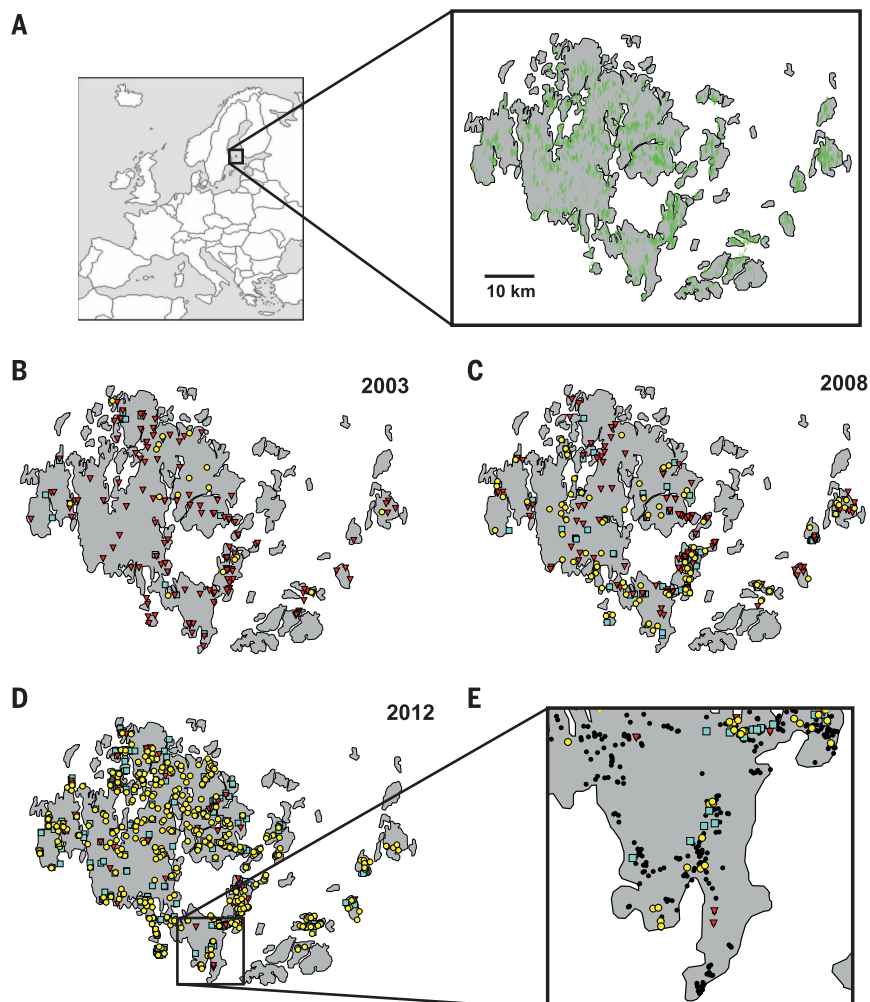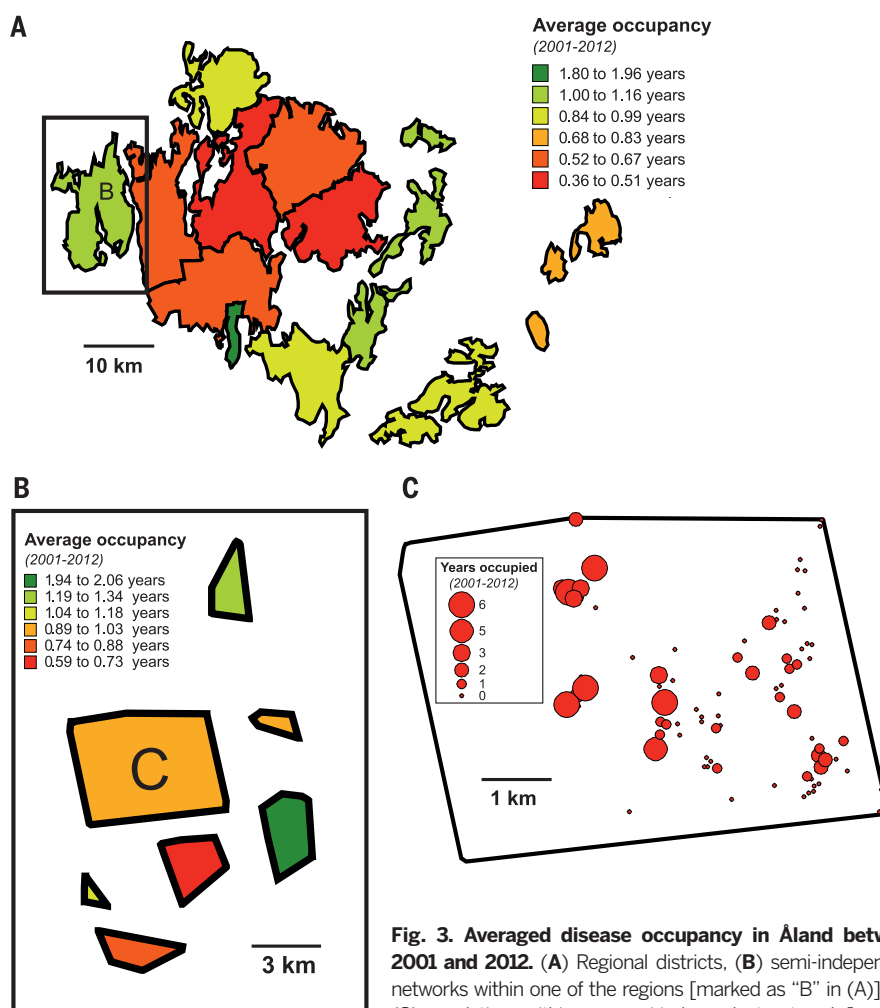
Fig. 3. Averaged disease occupancy in Åland between 2001 and 2012. (**A**) Regional districts, (**B**) semi-independent networks within one of the regions [marked as "B" in (A)], and (**C**) populations within one semi-independent network [marked as "C" in (B)].

Table 1. Environmental and spatial drivers of pathogen metapopulation dynamics and within population abundance analyzed with a Bayesian spatiotemporal logistic regression model. Mean parameter posterior estimate and its posterior standard error are shown for each factor for models analyzing the presence-absence of the pathogen (PA), the colonization of unoccupied (COL), and the extinction of occupied (EXT) plant populations. Population abundance (ABUND) was modeled by using categorical abundance data collected in autumn 2012. Nonsignificant (NS) parameters were removed from the models based on deviance information criterion. Patch shadow has three ordered levels, ranging from full exposure to sunlight (1) to shadow during most of the day (3).

Parameter	PA	COL	EXT	ABUND
Intercept	−4.2 ± 0.9	−3.9 ± 0.9	0.48 ± 0.36	2.02 ± 0.09
		Metapopulation configuration		
		Spatial factors		
Host population area (A)	0.30 ± 0.05	0.26 ± 0.06	−0.27 ± 0.19	NS
S^H	−1.2 ± 0.4	−0.82 ± 0.29	0.38 ± 0.16	−0.16 ± 0.13
Road presence	1.2 ± 0.1	1.2 ± 0.1	−0.43 ± 0.11	0.13 ± 0.07
S^M	not fitted	NS	−1.0 ± 0.2	not fitted
S^P	not fitted	0.08 ± 0.10	NS	not fitted
		Environmental factors		
Plant dryness	NS	NS	NS	0.18 ± 0.07
Habitat boundary type	NS	NS	NS	0.47 ± 0.31
Patch shadow, linear	−0.77 ± 0.14	−0.89 ± 0.19	NS	NS
Patch shadow, quadratic	−0.05 ± 0.09	−0.14 ± 0.11	NS	NS
July rainfall	NS	NS	NS	0.11 ± 0.10
August rainfall	−0.25 ± 0.09	−0.13 ± 0.11	NS	NS

succession, and, as a consequence, infection spreads within (*22*) and between host populations (*5*). By September, weather conditions turn unfavorable to pathogen transmission, and the epidemic spread ceases. At this time, since 2001, the entire host population network is surveyed for the presence-absence of the pathogen. This host-pathogen interaction is highly amenable to large-scale ecological studies, because infection is visually conspicuous (fig. S1) and the disease cycle lacks extended latency periods. Hence, field surveys at the end of epidemics yield a direct measure of host population size and disease prevalence within host populations, and the extinction-colonization dynamics of the pathogen can be inferred from these snapshot data by comparing the infection status of a given host population to its infection status in the previous year.

We complemented our analysis of spatial disease dynamics by assessing pathogen resistance through inoculation trials where host plants are challenged with fungal strains. The experiment consisted of 22 host populations representing different degrees of S^H. In the interaction between *P. lanceolata* and *P. plantaginis*, disease resistance is strain-specific, with the same host genotype blocking infection by some strains of the pathogen while being susceptible to others (*23*), and there is considerable variation among host individuals and populations in their degree of resistance (*24*). The obligate pathogen can only establish on susceptible hosts, and hence variation in resistance is expected to play a fundamental role in determining disease dynamics (*24*). Resistance should be favored because infected hosts suffer increased mortality in wild populations (*24, 25*). Ongoing studies suggest that resistance is costly for *P. lanceolata*, potentially explaining why this trait is so variable in Åland.

Our results show that in any one year the pathogen occupies only a small fraction of the available plant populations, generally ranging from 1.1 to 6.7%, with an exceptionally high occupancy in 2012 (16.9%; Fig. 1, A and B). The turnover of the pathogen populations is extremely high, because a large fraction of the pathogen populations (23.3 to 90.9%) go extinct every year (Figs. 1 and 2, B to E), predominantly during the critical overwintering stage (*21*). As a consequence, there is a major role for colonizations in maintaining the pathogen population regionally by counterbalancing the high extinction rates (Fig. 1, A and B). How these rates vary determines the annual disease occupancy levels (Figs. 1, A and B, and 2, B to E). Overall, our data show strong variation in the prevalence of the pathogen across several spatial scales (Figs. 1C and 3). Notably few host populations have been occupied for multiple years, and only a single host population has been occupied during the entire 12-year survey period (Fig. 1C). Instead, pathogen persistence takes place at regional district (Figs. 1C and 3A) and network levels (Figs. 1C and 3B).

The spatiotemporal Bayesian logistic regression model revealed the importance of spatial factors on disease dynamics at the metapopulation level as well as on local abundance (Table 1;
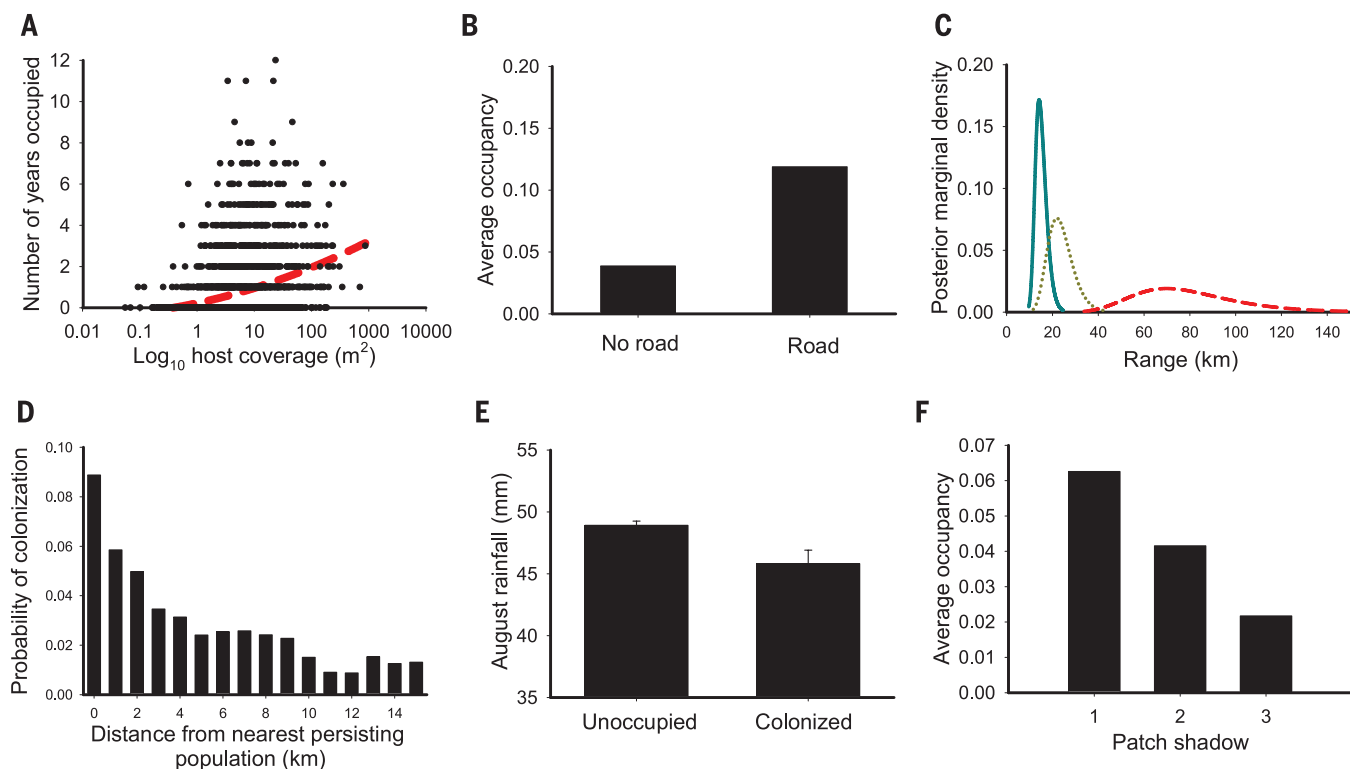
Fig. 4. The impact of environmental and spatial variables on disease dynamics. (**A**) Plant population area (correlates with number of individuals) and (**B**) the presence of a road had a positive effect on population occupancy by the pathogen. In (A), a dashed red trend line is shown. (**C**) Posterior Bayesian estimates for the range parameter for occupancy (solid blue line), colonizations (yellow dotted line), and extinctions (dashed red line). (**D**) Probability of colonization of an unoccupied host population as a function of the distance to the nearest persisting pathogen population (up to 15 km) for all years. (**E**) Average August rainfall was higher in populations that remained unoccupied (uncolonized) than in populations that became colonized by the pathogen during 2011. Error bars indicate SEM. (**F**) Increasing patch shadow reduces the pathogen average occupancy.

Fig. 5. The impact of host network configuration on metapopulation dynamics and disease resistance. (**A**) Extinction probability of the pathogen populations in year 2012 from plant populations that were colonized in the previous year increased as a function of host population connectivity (generalized linear model: S^H; $F_{1,183} = 7.17$, $P = 0.008$). (**B**) Host resistance was higher in well-connected than in isolated populations (linear model: $F_{1,21} = 4.63$, $P = 0.0437$).
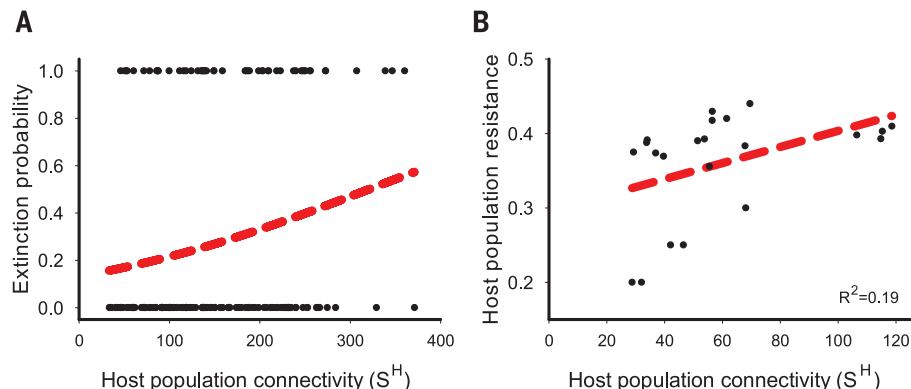
Fig. 4, A and B; tables S1 to S4; and figs. S2 to S5). As predicted by metapopulation theory (*12*), large host populations sustained higher colonization rates and lower extinction rates by the pathogen than did small host populations (Table 1 and Fig. 4A). Limited availability of large host populations is likely to be a key restricting factor of infection in this system. The colonization rate of the pathogen in any particular host population was affected by its spatial connectivity to other pathogen populations that had survived the winter that same year (S^P, Table 1), providing strong evidence for dispersal limitation with the majority of colonizations occurring within a ra-

dius of several kilometers from an occupied population (Fig. 4, C and D). Colonization happens predominantly over short distances, but there are infrequent longer-distance dispersal events that enable colonization of unoccupied areas (Fig. 4D) and are recognized to be important for large-scale dynamics of aerially transmitted fungi (*26*). Host populations that occurred along roadsides had significantly higher occupancy, abundance, and colonization probabilities and lower extinction probabilities of the pathogen (Table 1 and Fig. 4B). These results suggest that host plants growing in low densities along roadsides increase pathogen dispersal into the host populations

(*5*), thus enhancing colonization and preventing extinction (Table 1).

The drivers of extinction and colonization dynamics appear to be distinct because none of the measured environmental factors explained extinction dynamics, whereas the transmission and establishment phase of infection were sensitive to abiotic variation (Table 1 and table S4). Rainfall in August, a critical transmission period for the pathogen, may have impeded transmission by removing spores from the air and from host leaves (Table 1 and Fig. 4E) (*27*). How large a proportion of the host population is shaded by surrounding forest had a negative effect on

occupancy and colonization rates, which may be explained by light, humidity, and temperature conditions (Table 1 and Fig. 4F). In contrast to colonizations, extinctions were characterized by a much higher range of spatial autocorrelation (Fig. 4C), suggesting the importance of large-scale (unmeasured) environmental variation in weather, such as snow cover, for this phase.

Last, our study confirms that host resistance increases as host population density increases, with direct effects on epidemiological dynamics. In contrast to predictions from metapopulation theory (12), isolated host populations were more frequently infected by the pathogen than host populations in dense networks (S^H, Table 1). Moreover, pathogen colonization rate decreased and extinction rate increased as a function of host population connectivity (Table 1). This trend was particularly pronounced for the extinction probability of the pathogen in recently colonized host populations (Fig. 5A). Our experimental challenge of 22 *P. lanceolata* populations with the pathogen confirmed that resistance was significantly higher in the highly connected than in the isolated host populations (Fig. 5B). *Plantago* populations rapidly evolve resistance against the powdery mildew (25), and potentially higher gene flow into well-connected host populations may result in increased evolutionary potential and hence may reduce the probability of pathogen establishment and persistence. In addition to putatively higher rates of gene flow between *P. lanceolata* populations in high-density networks, an alternative, although not mutually exclusive, hypothesis is that areas supporting high-density host networks may also represent high-productivity environments for the host. In high-quality environments, hosts may be able to invest more toward disease resistance than hosts growing in areas that are nutritionally limited (28), resulting in resistance aggregating in high-connectivity areas of the landscape. The study of how spatial connectivity affects host-pathogen coevolution has been predominantly theoretical (16, 17), but, as shown in our work, variation in disease resistance among host populations may act as a powerful barrier against disease establishment in natural systems (24, 29) and may stabilize co-evolutionary dynamics (30).

Our study provides direct evidence of spatial structure having a profound effect on the ecology and evolution of disease dynamics in natural populations. Although comparable long-term and spatially explicit ecological and evolutionary data are lacking, we suggest that metapopulation structure and spatial heterogeneity at larger geographic scales may similarly drive antagonistic and mutualistic species interactions and their coevolutionary trajectories in a range of species interactions (31). Many factors explaining spatiotemporal variation in disease incidence across the metapopulation also explained within host population variation in disease abundance (Table 1), suggesting that incidence data at the metapopulation scale may be sufficient for understanding disease dynamics at finer spatial scales. With the majority of epidemiological studies targeted toward understanding the growth phase of epidemics, we argue

that more research should be directed toward understanding the drivers behind disease persistence at low endemic levels and the between-epidemic phase. Spatial and environmental heterogeneity, combined with the evolution of increased host resistance in high density host networks, are likely to be important components that prevent runaway dynamics of infection in nature. These insights may provide a natural blueprint for managing emerging diseases, as well as managing outbreaks that threaten sustainable agriculture and human health.

REFERENCES AND NOTES

1. D. J. Bradley, G. S. Gilbert, J. B. H. Martiny, *Ecol. Lett.* **11**, 461–469 (2008).
2. E. T. Borer, E. W. Seabloom, C. E. Mitchell, A. G. Power, *Ecol. Lett.* **13**, 810–818 (2010).
3. D. L. Smith, L. Ericson, J. J. Burdon, *J. Ecol.* **99**, 634 (2011).
4. J. Antonovics, in *Ecology, Genetics, and Evolution of Metapopulations*, I. Hanski, O. Gaggiotti, Eds. (Academic Press, San Diego, CA, 2004), pp. 471–488.
5. A.-L. Laine, I. Hanski, *J. Ecol.* **94**, 217–226 (2006).
6. D. Cressey, *Nature* **493**, 587 (2013).
7. B. A. McDonald, C. Linde, *Annu. Rev. Phytopathol.* **40**, 349–379 (2002).
8. P. K. Anderson et al., *Trends Ecol. Evol.* **19**, 535–544 (2004).
9. C. Viboud et al., *Science* **312**, 447–451 (2006).
10. J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, *Nature* **438**, 355–359 (2005).
11. A. S. Siraj et al., *Science* **343**, 1154–1158 (2014).
12. I. Hanski, *Metapopulation Ecology* (Oxford Univ. Press, Oxford, 1999).
13. B. Grenfell, J. Harwood, *Trends Ecol. Evol.* **12**, 395–399 (1997).
14. P. H. Thrall et al., *Ecol. Lett.* **15**, 425–435 (2012).
15. S. Gandon, Y. Michalakis, *J. Evol. Biol.* **15**, 451–462 (2002).
16. S. Gandon, S. L. Nuismer, *Am. Nat.* **173**, 212–224 (2009).
17. U. Carlsson-Granér, P. H. Thrall, *Oikos* **97**, 97–110 (2002).
18. P. H. Thrall, J. J. Burdon, *Plant Pathol.* **51**, 169–184 (2002).
19. J. D. Hoeksema, S. E. Forde, *Am. Nat.* **171**, 275–290 (2008).
20. C. Tollenaere, A.-L. Laine, *J. Evol. Biol.* **26**, 1716–1726 (2013).
21. A. J. M. Tack, A.-L. Laine, *New Phytol.* **202**, 297–308 (2014).
22. O. Ovaskainen, A.-L. Laine, *Ecology* **87**, 880–891 (2006).
23. A.-L. Laine, *J. Evol. Biol.* **20**, 1665–1673 (2007).
24. A.-L. Laine, *J. Ecol.* **92**, 990–1000 (2004).
25. A.-L. Laine, *Proc. Biol. Sci.* **273**, 267–273 (2006).
26. J. K. M. Brown, M. S. Hovmøller, *Science* **297**, 537–541 (2002).
27. D. E. Aylor, *Annu. Rev. Phytopathol.* **28**, 73–92 (1990).
28. M. A. Duffy et al., *Science* **335**, 1636–1638 (2012).
29. P. H. Thrall, J. Antonovics, *Can. J. Bot.* **73**, S1249 (1995).
30. E. Decaestecker, H. De Gersem, Y. Michalakis, J. A. M. Raeymaekers, *Ecol. Lett.* **16**, 1455–1462 (2013).
31. J. N. Thompson, *The Geographic Mosaic of Coevolution* (Univ. of Chicago Press, Chicago, 2005).

COMPARATIVE BEHAVIOR

Anxiety-like behavior in crayfish is controlled by serotonin

Pascal Fossat,[1,2] Julien Bacqué-Cazenave,[1,2] Philippe De Deurwaerdère,[1,3] Jean-Paul Delbecque,[1,2]* Daniel Cattaert[1,2]*†

Anxiety, a behavioral consequence of stress, has been characterized in humans and some vertebrates, but not invertebrates. Here, we demonstrate that after exposure to stress, crayfish sustainably avoided the aversive illuminated arms of an aquatic plus-maze. This behavior was correlated with an increase in brain serotonin and was abolished by the injection of the benzodiazepine anxiolytic chlordiazepoxide. Serotonin injection into unstressed crayfish induced avoidance; again, this effect was reversed by injection with chlordiazepoxide. Our results demonstrate that crayfish exhibit a form of anxiety similar to that described in vertebrates, suggesting the conservation of several underlying mechanisms during evolution. Analyses of this ancestral behavior in a simple model reveal a new route to understanding anxiety and may alter our conceptions of the emotional status of invertebrates.

Sources of stress or danger (called stressors) provoke fear, a basic emotion, and generate immediate responses, such as escape, freezing, or aggression. Stress can also lead to anxiety, a more complex state that is

considered a secondary emotion because it occurs when the stressor is absent or not clearly identified (1–3). In humans and rodents, anxiety is experienced as an anticipatory fear that facilitates coping with unexpected situations and

is revealed by a long-lasting behavioral adaptation intended to minimize threats, even in a different context and without the stressor (1–3). After the discovery of anxiolytics that act either on the serotonergic pathway or on the benzodiazepine site of γ-aminobutyric acid (GABA) receptors, anxiety has been intensively studied in humans and rodents (4, 5). However, most animals are capable of perceiving danger and exhibit varying degrees of behavioral adaptation to stress (1). Studies on zebrafish (6) have extended the concept of anxiety to all vertebrates. However, although studies have described fear response after aversive conditioning in *Aplysia* or pessimistic bias after stress in bees (7, 8), the characteristics of anxiety have not been fully observed in a single invertebrate model. In this study, we demonstrate that stressed crayfish (*Procambarus clarkii*) express context-independent anxiety-like behavior that can be promoted by 5HT and abolished by a benzodiazepine.

Crayfish naturally explore new environments but generally display a preference for dark places

[1]Department of Life Science and Health, Université de Bordeaux, 33 076 Bordeaux Cedex, France. [2]Centre National de la Recherche Scientifique (CNRS), UMR 5287, Institut des Neurosciences Cognitives et Intégratives d'Aquitaine, Avenue des Facultés, F-33405 Talence Cedex, France. [3]CNRS, UMR 5293, Institut des Maladies Neurodégénératives, Rue Léo Saignat, F-33076 Bordeaux, France.
*These authors contributed equally to this work. †Corresponding author. E-mail: daniel.cattaert@u-bordeaux.fr

(9). To measure possible variations in light aversion, we designed an aquatic dark/light plus-shaped maze (D/L+maze) that was divided into two shaded ("dark") arms, two exposed ("light") arms and a starting ("middle") zone (Fig. 1, A and B, and movie S1). During a 10-min testing period, 10 variables were measured to characterize the distance walked and the number/duration of visits in the compartments (table S1). Unstressed crayfish generally explored the entire maze (Fig. 1C). Dark preference was confirmed by a higher time spent in dark zones ($P < 0.05$) (Fig. 1D), a shorter latency of the first entry ($P < 0.05$) (Fig. 1E, left), and a greater number of entries ($P < 0.001$) (Fig. 1E, right) into dark arms. Crayfish often ceased locomotion before entering a light arm and frequently retreated (movie S2), a behavior quantified by calculating the retreat ratio (RR) (table S1).

We stressed other crayfish (Fig. 2A) by exposure to repetitive electrical fields that triggered tail-flips (rapid backward movements), which were considered aversive responses (10, 11). Stress induced a significant increase in blood glucose, which is a stress biomarker (fig. S1). Behavioral consequences were analyzed in the D/L+maze, where in contrast to unstressed crayfish, the stressed animals rarely explored or rapidly abandoned the light arms ($P < 0.001$, Mann-Whitney) (Fig. 2, B and C; table S2; and movie S3). In addition, the number of entries, mean duration per visit, laten-

cy to first entry into a light arm, and RR were significantly altered ($P < 0.001$, Mann-Whitney) (table S2). These changes were not due to a conditioned reflex but, in agreement with the anxiety criteria, were displayed in the absence of the stressor and in a new context.

Increasing stress duration up to 30 min decreased the time spent in the light arms, but longer exposure had no additional effect (Fig. 2D). The other measured variables also remained stable after 30 min of stress (fig. S2). Enhanced light avoidance persisted for at least 30 min and regularly returned to normal levels within 90 min (Fig. 2E and fig. S3). Therefore, the stress-induced behavioral adaptation of crayfish was sustainable, which is another criterion of anxiety.

Stress generally involves rapid mobilization of neuroendocrine centers that are considered homologous to the hypothalamo-pituitary axis of vertebrates (12). In crustaceans, such a center (X-organ/sinus gland) secretes hyperglycemic/steroidostatic hormones (CHH/MIH) under the control of serotonin (13–15). In crayfish, endogenous 5HT concentrations measured by means of high-performance liquid chromatography (HPLC) (16) were significantly elevated in the brains of stressed animals ($P < 0.01$, Mann-Whitney) (Fig. 2F) but not in their ventral cords (thoracic and abdominal ganglia), possibly because of the lower number of 5HT cells in cords than in brains (17, 18). To examine the functional
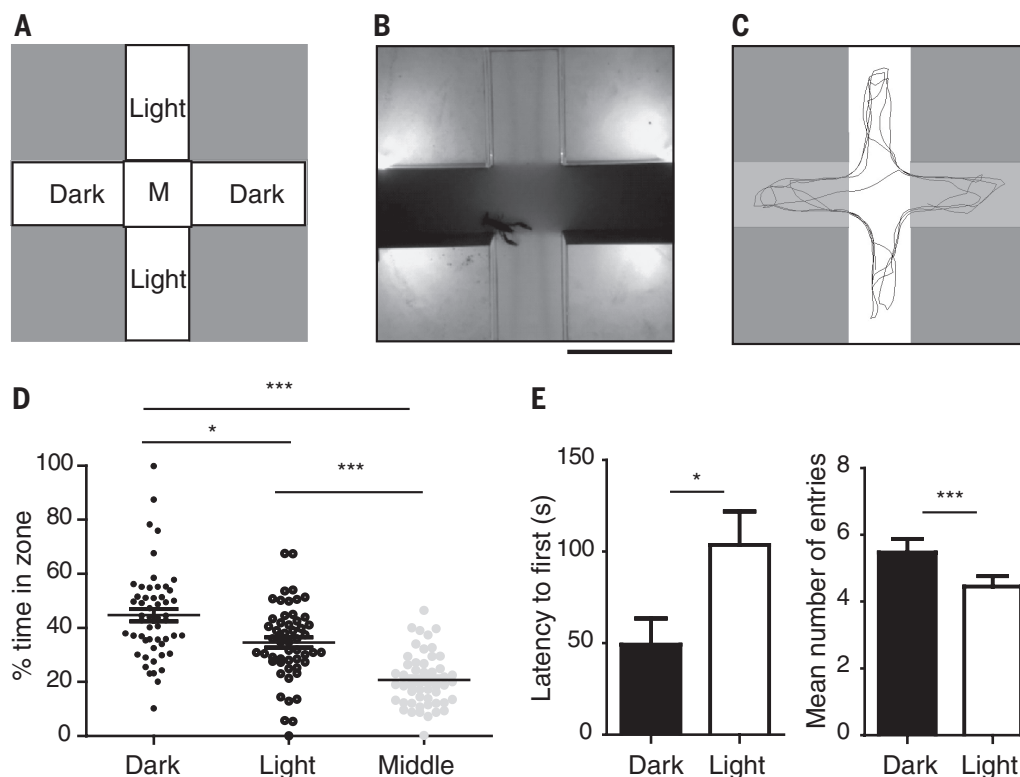
Fig. 1. Behavior of unstressed crayfish in the D/L+maze. (**A**) Illustration and (**B**) image of the D/L+maze comprising two illuminated (light) arms, two dark arms, and a middle (M) area from which the spontaneous behavior began. Scale bar, 25 cm. (**C**) An example maze route of an unstressed crayfish during the total experimental time (10 min). (**D**) Crayfish ($n = 55$) spent more time in the dark arms (43 ± 2% of total time) than in the light arms (35.7 ± 1.7%, $P < 0.05$). (**E**) (Left) The latency to the first entry into a dark arm (49 ± 14 s) was significantly shorter ($P < 0.05$) than that into a light arm (92 ± 15 s). (Right) The number of entries into dark arms (5.5 ± 0.4) was significantly greater ($P < 0.001$) than those into light arms (4.7 ± 0.3).

importance of these elevated 5HT levels, unstressed crayfish were injected with 5HT. As in other crustaceans (15), this bioamine was able to penetrate into the nervous structures of crayfish (supplementary materials, drug treatments) and induced a rapid increase in blood glucose (fig. S4). When transferred to the D/L+maze, in addition to reduced locomotion ($P < 0.05$) (table S2) the 5HT-injected crayfish displayed greater light avoidance as compared with that of saline-injected unstressed animals ($P < 0.05$, regarding percent of time in light arms, mean number of entries, latency to first entry in a light arm, and RR) (Fig. 2, G and H), similar to electrically

stressed animals ($P > 0.05$ for all variables; compare with Fig. 2, B and C). In contrast, a mixture of mianserin and methysergide, two broad-spectrum serotonin antagonists (14, 19, 20), prevented avoidance behavior when administered before electric field exposure [$P > 0.05$, compare with unstressed+saline for all variables (table S2)].

We next examined whether a well-characterized anxiolytic drug could abolish the behavioral changes induced by electric fields or 5HT injections. Chlordiazepoxide (CDZ), a potent benzodiazepine (5), prevented changes in exploratory behavior in animals previously submitted to ei-

ther electric fields (Fig. 3A, compare with Fig. 2B; and movie S4) or 5HT injection (Fig. 3B, compare with Fig. 2G) and restored the time spent in the light arms ($P < 0.05$, Dunn's test, compared with stressed+saline or unstressed+5HT) (Fig. 3, C and D), the latency to the first entry into a light arm, and RR [$P < 0.05$, Dunn's test, compared with stressed+saline or unstressed+5HT (table S2)]. This effect of CDZ was dose-dependent (fig. S5) and specific to avoidance behavior because it did not prevent the increase in glucose levels (fig. S6). When administered to unstressed crayfish, CDZ had no significant effect on time spent in light arms or on locomotion performance (table S2).
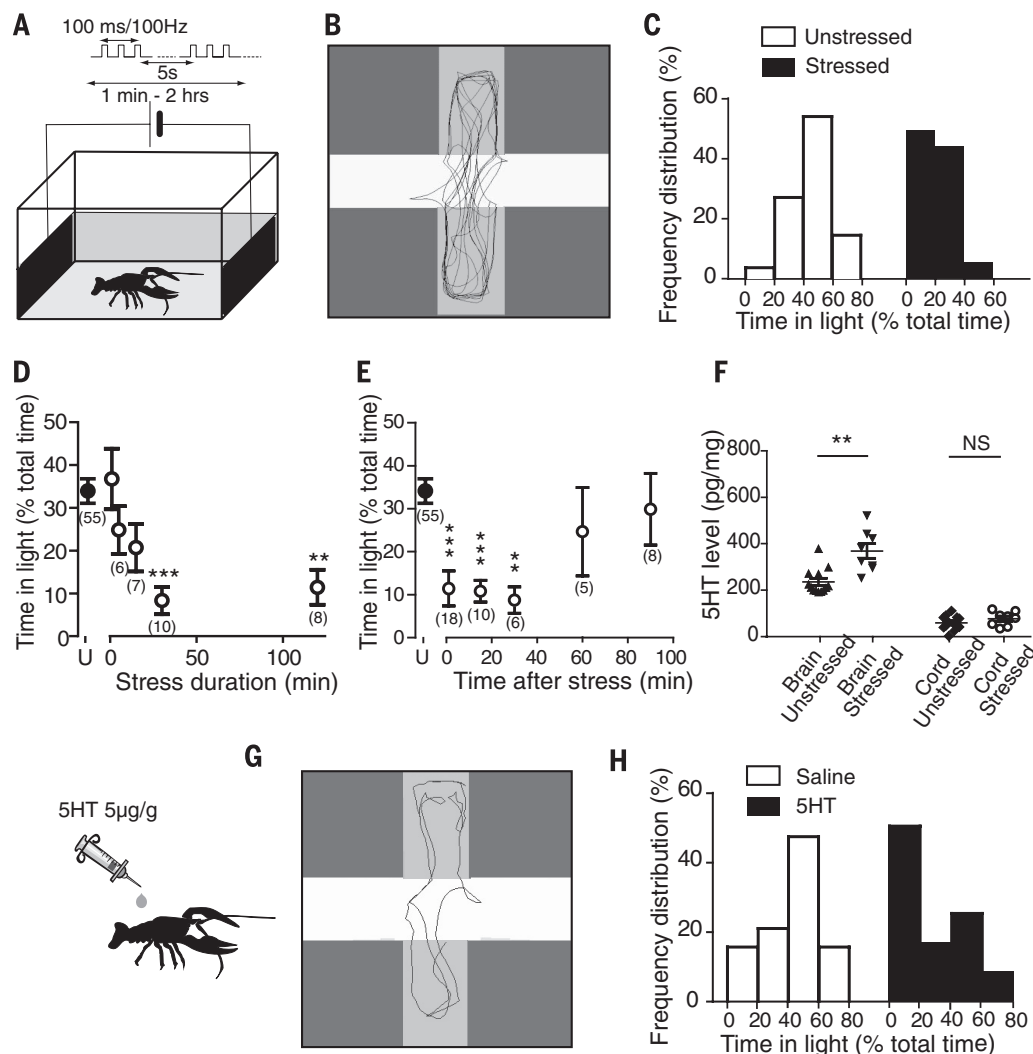
Fig. 2. Crayfish develop 5HT-dependent light avoidance after exposure to stress. (**A**) Experimental procedure for stress induction in crayfish (supplementary materials). (**B**) An example crayfish route after a 30-min exposure to an electric field. Walking occurred nearly exclusively in the dark arms. (**C**) Frequency distribution histograms of the percent time spent in light arms by stressed and unstressed crayfish. (**D**) Effect of stress duration on time spent in light arms (U = unstressed; P versus unstressed < 0.001 after 30 min and P versus unstressed < 0.01, after 2 hours of stress, Dunn's test). (**E**) Time course of behavioral changes (as measured by time spent in light arms) after exposure to a 30-min stressful experience. Crayfish

recovered "normal" behavior after 90 min (U = unstressed; P versus unstressed > 0.05, Dunn's test). The number of animals (n) is in parentheses in (D) and (E). (**F**) Serotonin concentrations (in picograms per milligrams of fresh weight) measured by means of HPLC in the brain and ventral cord of unstressed and stressed crayfish. Brain concentrations of 5HT were significantly higher in stressed than in unstressed animals. (**G**) After injection of 5 μg/g 5HT into the hemolymph, the crayfish route was similar to that of stressed crayfish. (**H**) The frequency distribution histograms of percent time in light arms for saline- and 5HT-injected crayfish were similar to those for (C) unstressed and stressed crayfish, respectively.

Fig. 3. Injection of benzodiazepine suppresses stress- and 5HT-induced light avoidance. (**A** and **B**) Typical crayfish routes either (A) after 30 min of stress followed by CDZ injection or (B) after 5HT and CDZ injection. In these examples, the crayfish entered all arms, similar to unstressed animals. (**C**) Frequency distribution histograms of percent time in the light arms for stressed crayfish injected with either saline (black bars) or CDZ (white bars). (**D**) Frequency distribution histograms of percent time in the light arms for the 5HT-injected crayfish (black bars) and the crayfish injected with 5HT and CDZ (white bars).
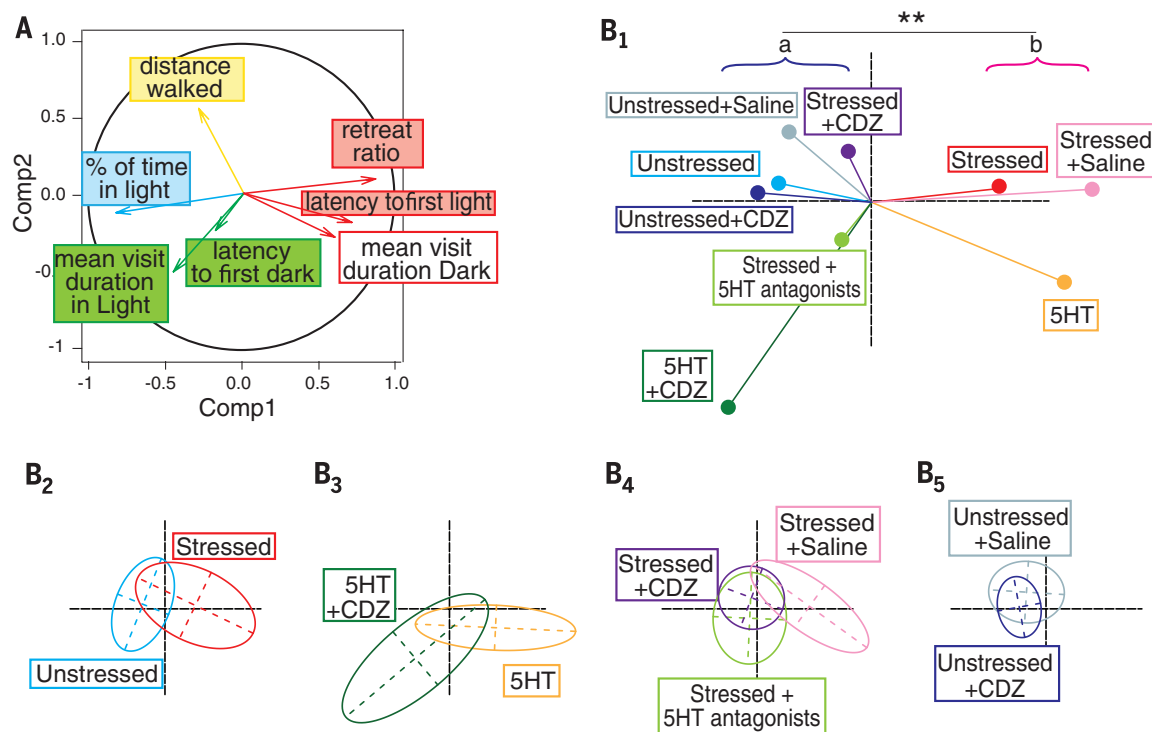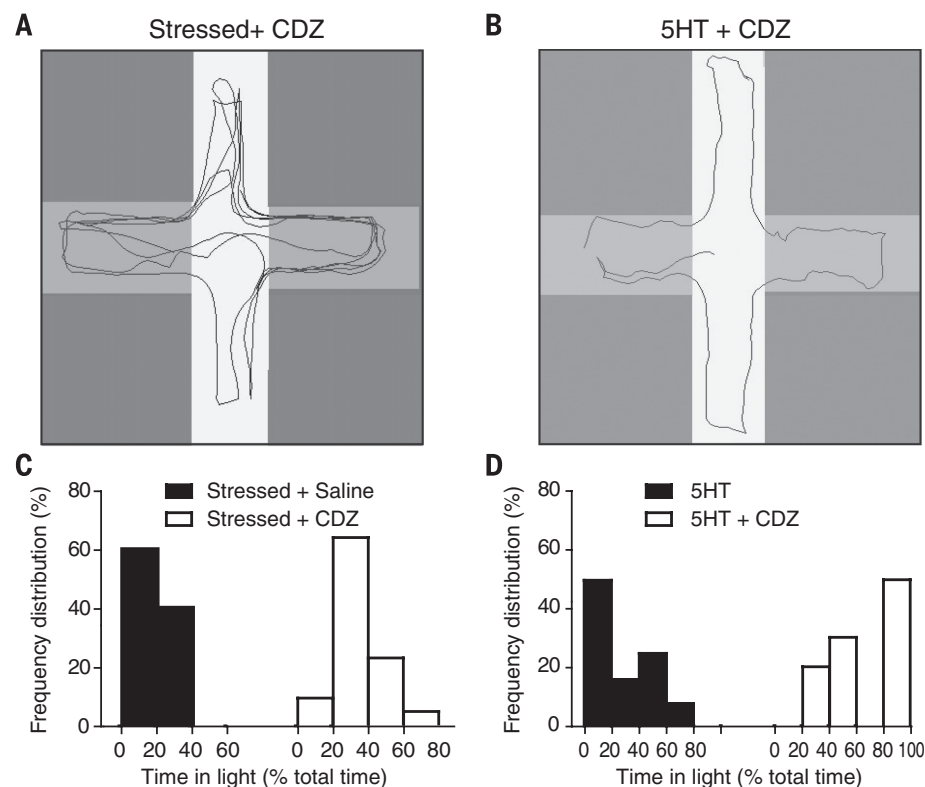
Fig. 4. PCA of the seven behavioral variables measured in experimental groups. (**A**) Contribution of the uncorrelated variables to the variances of the first and second components. (**B1**) Projection of group barycenters into the plane of the first and second components. (**B2** to **B5**) The envelopes of selected groups from the same PCA are presented separately in order to facilitate comparisons. (B2) Stressed versus unstressed crayfish. (B3) Crayfish injected with 5HT alone or in combination with CDZ. (B4) Stressed crayfish injected with saline, CDZ, or 5HT antagonists. (B5) Unstressed crayfish injected with saline or CDZ.

To compare all experimental groups, we performed principal component analysis (PCA) based on the seven uncorrelated variables (table S1, asterisk). The PCA enabled the representation of the variables (Fig. 4A) and experimental groups (Fig. 4B) on two orthogonal axes (supplementary materials, PCA). The first component (Fig. 4, x axis) can be interpreted as the anxiety level. Three groups (the stressed, stressed injected with saline, and 5HT-injected unstressed animals) exhibited substantially greater levels of anxiety (Fig. 4B1, positive x values of barycenters) and were significantly different from the other groups (Fig. 4B1, a/b groups, and table S2, a/b groups in PCA column). The stressed groups injected with either CDZ or 5HT antagonists were not significantly different from the unstressed group (and the other "a" groups). The groups injected with either 5HT or 5HT and CDZ had negative y values and were therefore separated from the other groups in the second component, mainly because of the inhibitory effects of 5HT on crustacean locomotion (21). Thus, 5HT had a CDZ-sensitive anxiogenic effect that was distinct from its inhibitory (and CDZ-independent) effect on locomotion.

The first conclusion of this study is that the stress-induced avoidance behavior in crayfish exhibits striking homologies with vertebrate anxiety. In the elevated plus-maze, a paradigm similar to the D/L+maze, the behavior of rodents after exposure to electric shocks (22) resembles that of crayfish—a context-independent and durable avoidance behavior. Similar to the risk assessment behavior described in rodents (23), the stressed crayfish also stopped before entering the aversive arms and frequently decided not to enter. Furthermore, as in rodents, crayfish anxiety-like behavior was sensitive to systemically applied CDZ, a potent benzodiazepine anxiolytic that modulates vertebrate GABA type A receptors (5). In crustaceans as in other animals, GABA is a fundamental inhibitory neurotransmitter (24) whose receptors are also sensitive to benzodiazepines

(25). Our results suggest that GABA is involved in the regulation of crayfish anxiety-like behavior.

The second conclusion of this study concerns the role played by 5HT. Serotonin, already involved in the control of stress response and arousal in crustaceans, triggers the rapid secretion of CHH and the mobilization of carbohydrate reserves (14, 15). Moreover, 5HT is frequently involved in the control of aggression (26), in which CHH also participates (27), implying a link between stress response and aggression. Beside the metabolic aspect of stress response, the present study shows that 5HT can also induce anxiety-like behavior in crayfish, interacting with GABA signaling (fig. S7, schematic diagram). In vertebrates, the relationships between 5HT and anxiety appear to be more complex; a classic hypothesis considering that 5HT-promoted anxiety has been contradicted by many studies (4). In this context, the crayfish represents a new model that might provide insights into the mechanisms underlying anxiety that have been conserved during evolution. Our results also emphasize the ability of an invertebrate to exhibit a state that is similar to a mammalian emotion but which likely arose early during the evolution of metazoans.

REFERENCES AND NOTES

1. C. Belzung, P. Philippot, *Neural Plast.* **2007**, 59676 (2007).
2. D. C. Blanchard, R. J. Blanchard, in *Handbook of Anxiety and Fear*, R. J. Blanchard, D. C. Blanchard, G. Griebel, D. Nutt, Eds. (Academic Press, New York, 2008).
3. T. Steimer, *Dialogues Clin. Neurosci.* **13**, 495–506 (2011).
4. S. L. Handley, J. W. McBlane, *Psychopharmacol. (Berl.)* **112**, 13–20 (1993).
5. F. López-Muñoz, C. Alamo, P. García-García, *J. Anxiety Disord.* **25**, 554–562 (2011).
6. R. J. Egan et al., *Behav. Brain Res.* **205**, 38–44 (2009).
7. E. T. Walters, T. J. Carew, E. R. Kandel, *Science* **211**, 504–506 (1981).
8. M. Bateson, S. Desire, S. E. Gartside, G. A. Wright, *Curr. Biol.* **21**, 1070–1073 (2011).
9. S. Yamane, M. Takahata, *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **188**, 13–23 (2002).
10. D. H. Edwards, W. J. Heitler, F. B. Krasne, *Trends Neurosci.* **22**, 153–161 (1999).
11. N. Kawai, R. Kono, S. Sugimoto, *Behav. Brain Res.* **150**, 229–237 (2004).
12. V. Hartenstein, *J. Endocrinol.* **190**, 555–570 (2006).
13. M. P. Mattson, E. Spaziani, *Gen. Comp. Endocrinol.* **62**, 419–427 (1986).
14. C. Y. Lee, P. F. Yang, H. S. Zou, *Physiol. Biochem. Zool.* **74**, 376–382 (2001).
15. S. Lorenzon, P. Edomi, P. G. Giulianini, R. Mettulio, E. A. Ferrero, *J. Exp. Biol.* **208**, 3341–3347 (2005).
16. P. De Deurwaerdère, L. Stinus, U. Spampinato, *J. Neurosci.* **18**, 6528–6538 (1998).
17. K. Langworthy, S. Helluy, J. Benton, B. Beltz, *Cell Tissue Res.* **288**, 191–206 (1997).
18. L. Rodriguez-Sosa, A. Picones, G. Rosete, S. Arechigay, *J. Exp. Biol.* **200**, 3067–3077 (1997).
19. L. Rodriguez-Sosa et al., *Comp. Biochem. Physiol. C* **142**, 220–230 (2006).
20. Y. Momohara, A. Kanai, T. Nagayama, *PLOS ONE* **8**, e74489 (2013).
21. H. V. Peeke, G. S. Blank, M. H. Figler, E. S. Chang, *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **186**, 575–582 (2000).
22. S. Pellow, P. Chopin, S. E. File, M. Briley, *J. Neurosci. Methods* **14**, 149–167 (1985).
23. R. J. Rodgers, A. Dalvi, *Neurosci. Biobehav. Rev.* **21**, 801–810 (1997).
24. E. Florey, *Can. J. Physiol. Pharmacol.* **69**, 1049–1056 (1991).
25. M. J. Snyder, S. Watson, H. V. S. Peeke, *Mar. Freshwat. Behav. Physiol.* **34**, 37–51 (2000).
26. R. Huber, K. Smith, A. Delago, K. Isaksson, E. A. Kravitz, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5939–5942 (1997).
27. L. Aquiloni et al., *PLOS ONE* **7**, e50047 (2012).

Scientifica

The Scientific World Journal

MEDIATORS of INFLAMMATION

Neural Plasticity

Journal of Diabetes Research

BioMed Research International

International Journal of Genomics

Journal of Nucleic Acids

Hindawi

Submit your manuscripts at
http://www.hindawi.com

International Journal of Polymer Science

Journal of Obesity

Journal of Immunology Research

PPAR Research

Biochemistry Research International

Journal of Chemistry

Bioinorganic Chemistry and Applications

International Journal of Photoenergy

Advances in Condensed Matter Physics

Advances in Materials Science and Engineering

Journal of Spectroscopy

BD Horizon Brilliant™ Blue 515

A brighter blue laser dye with less spillover and better resolution.
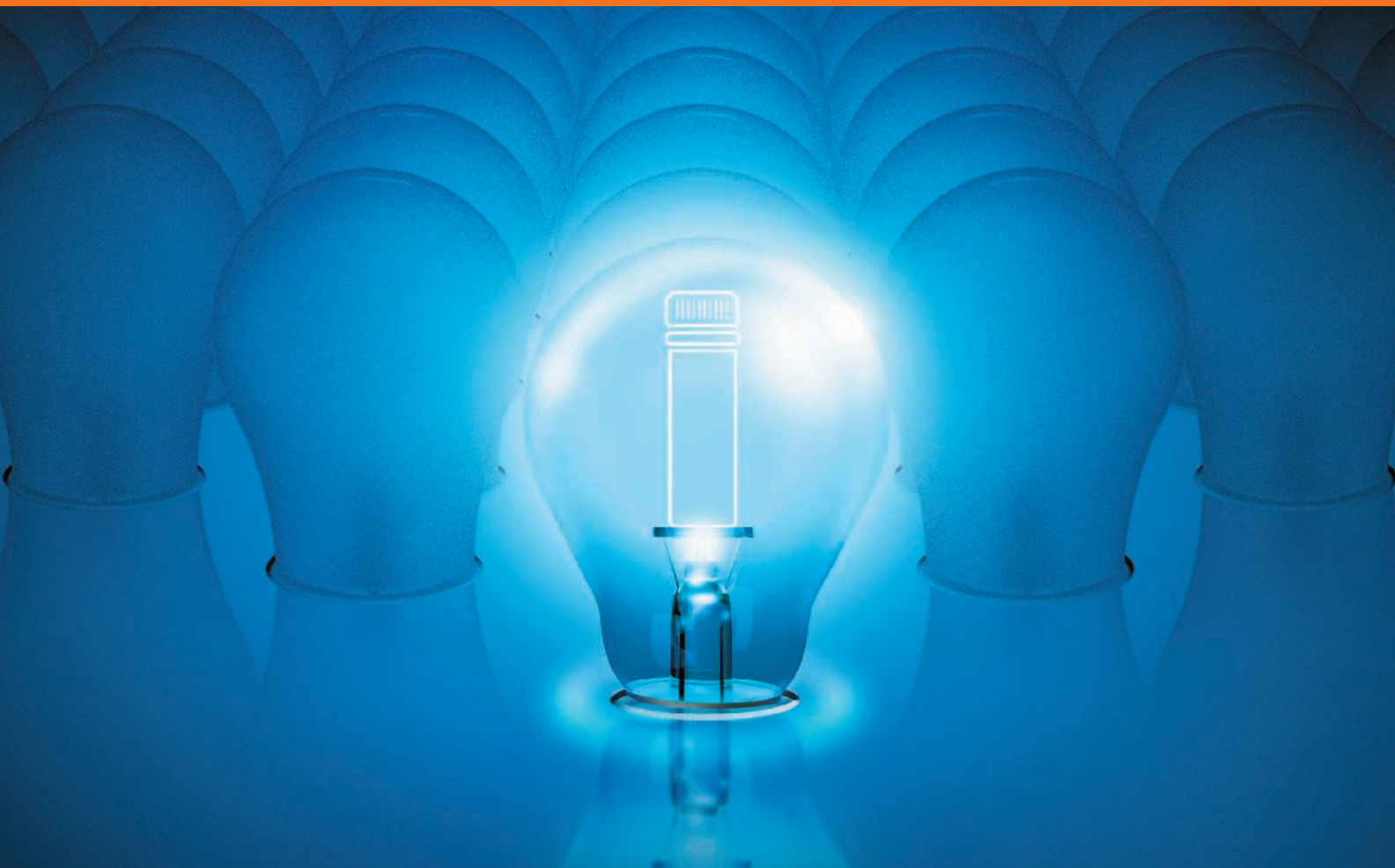
Makes working with FITC seem rather dull.

BD Helping all people live healthy lives

BD Horizon Brilliant™ Blue 515 brings innovation to the blue laser in the form of a novel format that is resolved in the FITC channel. Developed using technology from Sirigen, BD Horizon Brilliant Blue 515 is up to seven times brighter than FITC and offers significantly better resolution, with less spillover into the PE channel.

BD Horizon Brilliant Blue conjugated antibodies give researchers new avenues to resolve dim cell populations, take

advantage of better panel design, and have more confidence in their results.

This unique blue laser dye is just another way BD Biosciences continues to innovate and deliver improvements to the world of multicolor flow cytometry so that you can get more out of your research. To find out about brighter ways to use blue and obtain free samples, visit **bdbiosciences.com/go/newblue.**

BD Biosciences
2350 Qume Drive
San Jose, CA 95131
bdbiosciences.com

Always been ready for change?

MAKE GREAT THINGS HAPPEN

Opportunities for graduates and young professionals: If you are in the process of laying out your future career, why not start with us? To position ourselves for the years to come, we focus on continuous improvements in all our businesses and operations. For that, we need the best and brightest people – fresh talents willing to help us set the course for Merck KGaA, Darmstadt, Germany with courage, creativity and competence. And ready to seize an excellent opportunity to advance your own career. We are ready for change – are you?

Merck KGaA, Darmstadt, Germany – with more than 300 years of progress and about 38,000 employees in over 60 countries, we are leading in chemicals and pharma.
Our subsidiaries in Canada and the United States operate under the umbrella brand EMD. With passion, dedication and innovative ideas, we pursue one global goal: to improve people's quality of life.
Like to join in? Welcome to the team!

come2emd.com

EMD

the old world

the new world

LIFE SCIENCE TECHNOLOGIES
BIG DATA

Big Biological Impacts from Big Data

In the life sciences, data can come in many forms, including information about genomic sequences, molecular pathways, and different populations of people. Those data create a potential bonanza, if scientists can overcome one stumbling block: how to handle the complexity of information. Tools and techniques for analyzing big data promise to mold massive mounds of information into a better understanding of the basic biological mechanisms and how the results can be applied in, for example, health care. **See the full story on page 1298.**

Upcoming Features

Digital Lab Management—July 25
Metabolomics—September 19
Genomics—October 3

FOCUS ON CAREERS
BIOTECH AND PHARMA

An Explosion of Bioinformatics

Big data is everywhere, and its influence and practical omnipresence across multiple industries will just continue to grow. For life scientists with expertise and an interest in bioinformatics, computer science, statistics, and related skill sets, the job outlook couldn't be rosier. Big pharma, biotech, and software companies are clamoring to hire professionals with experience in bioinformatics and the identification, compilation, analysis, and visualization of huge amounts of biological and health care information. With the rapid development of new tools to make sense of life science research and outcomes, spurred by innovative research in bioinformatics itself, scientists who are entranced by data can pursue more career options than ever before. **See the full story on page 1303.**

Upcoming Features

Postdocs: Broadening Your Skills —August 22
Faculty: Professionalism 1—September 12
Faculty: Professionalism 2—October 3

AAAS | 2015
ANNUAL MEETING
12–16 FEBRUARY • SAN JOSE, CA

INNOVATIONS, INFORMATION, AND IMAGING

The meeting theme reflects on the information transformation happening in science and technology brought about by recent advances in organizing, visualizing, and analyzing data.
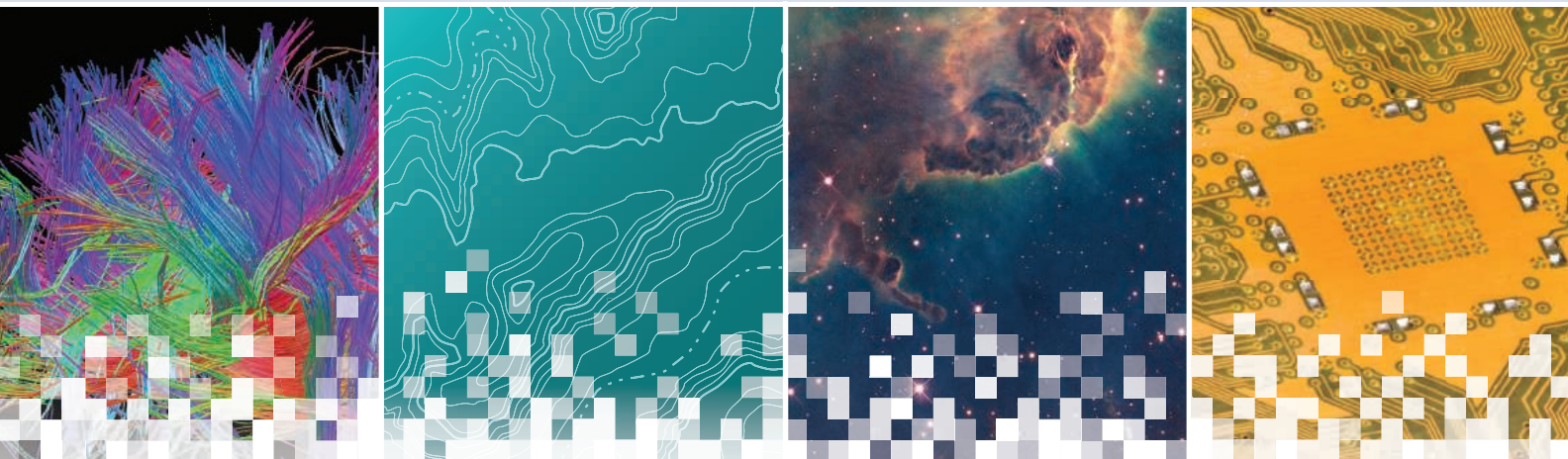
Save the Date: 12–16 February 2015

The AAAS Annual Meeting is interdisciplinary and inclusive. Each year, thousands of leading scientists, engineers, educators, policymakers, and journalists from around the world gather to discuss recent developments in science and technology.

Registration opens in early August.

www.aaas.org/meetings

Big Biological Impacts from Big Data

In the life sciences, data can come in many forms, including information about genomic sequences, molecular pathways, and different populations of people. Those data create a potential bonanza, if scientists can overcome one stumbling block: how to handle the complexity of information. Tools and techniques for analyzing big data promise to mold massive mounds of information into a better understanding of the basic biological mechanisms and how the results can be applied in, for example, health care. **By Mike May**

"Big data" is one of today's hottest concepts, but it can be misleading. The name itself suggests mountains of data, but that's just the start. Overall, big data consists of three *v*'s: volume of data, velocity of processing the data, and variability of data sources. These are the key features of information that require big-data tools.

Although biologists have spent decades struggling to collect enough data, says Keith Crandall, director of the Computational Biology Institute at **George Washington University** in Ashburn, Virginia, "the new bottlenecks in biology are big-data issues." As an example, he points out that the first human genome sequence, announced in April 2002, utilized the expertise, infrastructure, and people from 20 institutions and took 13 years of work and about $3 billion to determine the order of approximately three billion nucleotides. Now, says Crandall, "We can

sequence a human genome for $1,000, and we can generate more than 320 genomes per week!"

As life scientists explore more ways to deal with big data's volume, velocity, and variability, they are starting to develop new approaches to analyzing information.

Ever-Expanding Volume

When it comes to collecting large volumes of information about human biology, the pharmaceutical industry started battling large data sets decades ago. As Jason Johnson, associate vice president for scientific informatics at **Merck Research Labs** in Boston, Massachusetts says, "Merck has for many years had clinical trials with thousands of patients, and the ability to query millions of de-identified patient records, and now we have next generation genomic sequencing that can create a terabyte of data per sample."

To deal with so much data, even large pharmaceutical companies need help. For example, Bryn Roberts, global head of R&D operations at **Roche** in Basel, Switzerland, says, "A century's worth of Roche R&D data were more than doubled in 2011–2012 in a single large-scale experiment to sequence hundreds of cancer cell lines." Roberts and his colleagues want to derive more value from these data sets and others collected years ago. So they are collaborating with PointCross in Foster City, California, to create a data platform that allows flexible searching of data from the past 25 years of Roche studies, including those outsourced to contract research organizations. Those data, along with information about thousands of compounds, will be mined to use the existing knowledge to develop new drugs.

To generate large volumes of data, though, a biologist does not need the infrastructure of a large pharmaceutical company. For example, consider the specifications of the Ion Personal Genome Machine (PGM) System from **Life Technologies** in Carlsbad, California (now a part of **Thermo Fisher Scientific**). This next generation device can sequence up to two gigabases in less than eight hours, and this is marketed as a "personal genome machine" that can go on a scientist's benchtop. Life Technologies' larger Ion Proton machine pumps out up to 10 gigabases in four hours or less.

In general, for academic and industrial life scientists, next generation sequencing supplies a bonanza and a bottleneck. As Crandall explains, "We cannot effectively study this volume of genomes until our computational software scales up to these big data needs." So his team is working with W. Evan Johnson, an assistant professor of medicine at **Boston University School of Medicine,** to develop software, PathoScope, that can handle the data from today's next generation sequencing (NGS) platforms, which turn information on gigabases of DNA into gigabytes of computer data—the exact ratio tends to be about linear, depending on the NGS platform being used. This software compares DNA samples to reference genomes in an effort to identify a pathogen. Crandall says, "Our data sets can

Upcoming Features

Digital Lab Management—July 25 ■ *Metabolomics—September 19* ■ *Genomics—October 3*

be 20 gigabytes of data per sample for hundreds of samples with downstream analyses generating 100's of gigabytes of data per sample."

Such large volumes of data can be especially useful in health care, where pharmaceutical scientists must take into account the variability among people when designing their experiments. "You can't draw any reasonable conclusion by studying only 10 people, but you can come up with important lessons by studying say 500,000," says Chas Bountra, professor of translational medicine at the **University of Oxford,** United Kingdom. Large-scale studies might reveal genetic contributions to disease, whether a drug could help a subset of patients, or which individuals are likely to develop a particular disorder.

Other experts also expect to see a growing impact of genetic data on health care. "Genetics gives us a huge, powerful foothold into understanding how people get sick and what you can do about it," says Gil McVean, professor of statistical genetics at the **Wellcome Trust Centre for Human Genetics** in Oxford, United Kingdom. For example, genetic information might reveal biomarkers, or indicators of a specific kind of disease, like a molecule involved in a particular form of cancer. As McVean explains, "Genetics can tell you if a biomarker associated with a disease is worth going after as a [therapeutic] target." For instance, a molecule that drives a particular form of cancer could make a good target for treating the disease.

To apply this thinking, McVean and his colleagues are creating the Li Ka Shing Centre for Health Information and Discovery at Oxford University, which was launched on a huge contribution—about $33 million—from Chinese billionaire Li Ka-shing. This center will include a big data institute, which is currently under development. Overall, says McVean, the center "will bring together analytical data processing and genetics in one institute, so we can tackle some of the thorny but fascinating questions about collecting and analyzing big data sets."

Seeking High Velocity

The second *v*, velocity, depicts the speed of processing and analyzing the data. Scientists need high-speed processes to analyze the growing volumes of data.

In the past, analyzing gene-related data created a bottleneck. "Traditionally, these analysis platforms have posed productivity limitations on researchers," says Alan Taffel, president of **BioDatomics** in Bethesda, Maryland. "They've been difficult to use, often requiring bioinformatician support, and they've also been very slow in executing workflows." In fact, he says, it could take days or weeks to turn around a large DNA analysis. So BioDatomics developed its BioDT software, which provides more than 400 tools for analyzing genomic data. It integrates these tools into one software package to make them easier to use, and it can outrun any desktop computer.

BioDT runs on a computer cluster, which consists of machines, called nodes, inter-connected and working as one. "You need at least four nodes," says Maxim Mikheev, chief technology officer at BioDatomics. Yet BioDT can run on many nodes to process data even faster. "Scalability is theoretically unlimited,"

says Mikheev. "There are clusters that are using over 40,000 nodes." For users not inclined toward building a computer cluster, BioDT can also be accessed through the cloud.

Overall, says Taffel, BioDT "can execute workflows up to 100 times faster than traditional systems. Days or weeks become just minutes or hours."

Other experts also see the need for new tools driven by sequencing. According to Jaroslaw Zola, associate research professor in the Department of Electrical and Computer Engineering at **Rutgers University** in Piscataway, New Jersey, "The almost ubiquitous adoption of the next generation sequencing technologies requires new computational strategies to handle the data all the way from how it is stored, to how it is transferred, to how it is analyzed." This means that biologists must learn to use cutting-edge computational technologies, but, as Zola says, that "puts a pressure on information technology experts to deliver efficient solutions that are easy to adopt by the domain experts, yet hide the complexity of the underlying algorithm, software, and hardware architecture without sacrificing the efficiency." That, says Zola, requires novel algorithms, which he's working on now.

> **"Our data sets can be 20 gigabytes of data per sample for hundreds of samples with downstream analyses generating 100's of gigabytes of data per sample."**

Versions of Variability

The third *v*, variability, creates a big challenge for biologists. As Bountra says, "We are now bringing together people from lots of different areas with lots of different data sets."

For one thing, a biology lab includes a variety of instruments, and they often collect data in specific file formats. So **ACD/Labs,** headquartered in Toronto, Canada, developed computing systems that integrate a wide range of data types when working with big data. As Ryan Sasaki, director of global strategy at ACD/Labs, explains, "We support more than 150 file formats from various instruments." He adds, "This lets us bring data into one environment, our Spectrus database, which can be made available through desktop-client software or accessed through the web as well as via other laboratory informatics systems."

Biology's big data also include new kinds of variability. Scientists at **Definiens** in Munich, Germany, for example, analyze what the company calls tissue phenomics, or information about a tissue sample's make-up, such as cell size, shape, the stains that they absorb, and which cells contact each other. This technique can be applied to a variety of studies, such as those designed to

track the characteristic changes cells undergo during development, measure the impact that environmental factors have on an organism, or that quantify the cellular effects a medication may have on specific tissues.

Structured data, such as tables of numbers, do not reveal everything that is known about a medication or biological process. Much of what we know about living organisms exists in unstructured formats, like journal article text. As Johnson of Merck says, "There are thousands of ways to describe biological processes," and it is difficult to extract data from the literature.

At **IBM**'s Almaden Research Center in San Jose, California, analytics expert and research staff member Ying Chen and her colleagues have worked for years on creating technologies for mining text, which they now use for their "accelerated drug discovery solution." Their platform aggregates patents, scientific literature, basic chemistry and biology knowledge (such as how chemicals and molecules interact), more than 16 million unique chemical structures, and information about nearly 7,000 diseases. Using this system, researchers can search for compounds that might be useful for treating specific diseases.

Other companies also hope to mine existing resources to learn more about the biology of diseases and how to treat them. **NuMedii,** a big-data company in Silicon Valley, and Thomson Reuters, a provider of intelligent scientific information in New York, have teamed up to find new uses for existing drugs, known as drug repurposing. "Using genomic databases, integrated knowledge sources, and bioinformatic approaches, we can quickly discover novel uses for drugs," says Craig Webb, NuMedii's chief scientific officer. "We then leverage the safety data for the drugs in their original use to get to clinical trials faster and cheaper." NuMedii is contributing databases and analytics to the project, while Thomas Reuters is supplying in-depth knowledge on diseases and drugs.

One such project, Webb says, has researchers compiling gene expression data from more than 2,500 ovarian tumor samples and using several computer algorithms to predict whether any existing drugs could potentially treat ovarian cancer broadly or treat specific molecular subtypes. "Big data allows us to cast a wide net initially to identify leads, while 'big knowledge' allows us to quickly select viable compounds to test," Webb says.

Featured Participants

ACD/Labs
www.acdlabs.com

BioDatomics
www.biodatomics.com

Boston University School of Medicine
www.bumc.bu.edu/busm

Definiens
www.definiens.com

George Washington University
www.gwu.edu

GNS Healthcare
www.gnshealthcare.com

IBM
www.ibm.com

Life Technologies
www.lifetechnologies.com

Merck
www.merck.com

Novartis Institutes for BioMedical Research
www.nibr.com

NuMedii
www.numedii.com

Roche
www.roche.com/index.htm

Rutgers, The State University of New Jersey
www.rutgers.edu

Thermo Fisher Scientific
www.thermofisher.com

University of Oxford
www.ox.ac.uk

Wellcome Trust Centre for Human Genetics
www.well.ox.ac.uk

Complexity from Combinations

To the three *v*'s of big data, Stephen Cleaver, executive director of informatics systems at the **Novartis Institutes for BioMedical Research** (NIBR) in Cambridge, Massachusetts adds complexity. He says that scientists in the pharmaceutical industry analyze the data by "patients individually and then as a group, and then we integrate everything we have." That gets complex.

In health care, the complexity of big data analysis also arises from combining different types of information, such as data from genomics, proteomics, cellular signaling, clinical research, and even environmental studies. The results could reveal entirely new approaches to treating diseases. But Iya Khalil, cofounder of **GNS Healthcare** in Cambridge, Massachusetts, asks: "How do you make sense of those data and get insights from those data that will advance our understanding of the disease mechanism?" For Khalil and her teammates, the answer comes from machine learning, mathematics, computational algorithms, and supercomputers—all combined to explore the underlying pathways of disease and to follow a patient's likely response to a particular treatment.

At GNS Healthcare, such big-data analysis depends on a computational platform called REFS, which stands for reverse engineering and forward simulation. In short, the software analyzes data to construct possible molecular networks underlying a specific disease—that's the reverse part—and then it uses that information to simulate the impact a particular compound would have upon the pathway—the forward aspect of the process.

In addition to health care, REFS can be applied to basic biology. For example, Khalil and her colleagues have used this technology to make a molecular model of part of the cell replication cycle.

For Khalil and other scientists, the key is using big data in ways that move science forward. At NIBR, for instance, Cleaver and his colleagues want to make sure that the data is informative, first and foremost. "It's great to run advanced data-mining methods, but it must suggest the next scientific hypothesis," he says. That way, today's big data will change tomorrow's biology and medicine.

Digital Gene Expression Assay

The nCounter PanCancer Pathways Panel is a highly multiplexed, digital gene expression assay that offers a unique way to investigate cancer biology across all major cancer pathways. From a single tube, the PanCancer Pathways Panel simultaneously analyzes all of the key cancer pathways: PI3K, STAT, MAPK, TGFβ, Notch, Hedgehog, Wnt, Apoptosis, Cell Cycle, RAS, Chromatin Modification, Transcriptional Regulation, and DNA Damage Control. The 770 genes included in the PanCancer Pathways Panel were selected using a biology-guided, data-driven methodology that provides high confidence that the gene variance within each cancer pathway was captured. Each of the canonical cancer pathways was mapped to publicly available data sources such as Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and the Gene Ontology (GO) databases. Then, leveraging data generated through The Cancer Genome Atlas (TCGA) and the collective work of decades of cancer research, each pathway gene and driver gene was scored and ranked based on its biological relevance to cancer.

NanoString Technologies
For info: 888-358-6266
www.nanostring.com/pancancer

Compound Screening Software

The new version of the StarDrop software platform introduces MPO Explorer, which provides innovative new methods to guide strategic decisions in research programs by helping to identify the key properties and selection criteria with which to select compounds with a high chance of success. MPO Explorer introduces unique methods that guide the development and validation of multi-parameter scoring profiles, tailored to a discovery project's specific objectives. The resulting scoring profiles can be applied with StarDrop's Probabilistic Scoring approach for multi-parameter optimization (MPO), to effectively target novel compounds with a high chance of success. The most important property criteria are also highlighted, focusing experimental resources to generate the key data with which to identify successful compounds. To promote objective decision-making, the new module also tests if a specific property criteria in a scoring profile may be artificially distorting the selection of compounds to pursue.

Optibrium
For info: +44-(0)-1223-815900
www.optibrium.com

DNA Copy Number Variation Array

The range of research-validated CytoSure Molecular Arrays has been expanded to investigate DNA copy number variation (CNV) underlying a variety of genetic disorders. The arrays are the ideal complement to DNA sequencing, providing a particularly powerful tool for investigating the variety of aberrations underlying genetic disorders. Comparative genomic hybridization arrays (aCGH) are the gold standard for CNV detection and the 60-mer oligonucleotide probes utilized by OGT's aCGH platform have been shown to deliver superior CNV detection than alternative platforms. The expanded CytoSure Molecular Array portfolio now enables detection of CNV in genes associated with over 20 genetic disorders, including cardiovascular, inherited eye, intellectual disability, and neuromuscular disorders, as well as a range of inherited cancers. In addition, genes covering each disorder can be combined to create bespoke custom arrays, or further customized by the addition of novel content to suit each individual research project. All CytoSure Molecular Arrays are supplied with OGT's class-leading CytoSure Interpret Software.

Oxford Gene Technology
For info: +44-(0)-1865-856800
www.ogt.com/cytosure

Spectral Analysis Software

KnowItAll ATR/IR ID Expert and Raman ID Expert are new technologies that provide the fastest, most accurate answers possible to scientists identifying unknown infrared and Raman spectra. The spectral intelligence built into KnowItAll ATR/IR ID Expert combined with the world's largest spectral reference collection provides the highest level of expertise to any scientist, whether a novice or power user. Scientists often employ multiple software tools to identify a spectrum depending on its chemical composition and the availability of reference spectra, and even the most experienced researchers can misinterpret results. With ATR/IR ID Expert and Raman ID Expert, the user simply opens an unknown spectrum and the software automatically performs single and multiple component searches as well as functional group analyses simultaneously and summarizes the results on a single screen. If there are problems with the user's query spectrum, ID Expert has the spectral intelligence to identify these issues and suggest ways to fix them.

Bio-Rad
For info: 888-524-6723
www.knowitall.com

High-Content Screening Platform

A new high-content screening software platform enables cancer researchers to uncover more about basic and advanced cellular function to accelerate cell-based research. The Thermo Scientific HCS Studio 2.0 is a high-content quantitative imaging and analysis software platform for a range of cancer research applications, including angiogenesis and migration/invasion, as well as the development of three-dimensional models of cancer stem cells and spheroid tumors. The new Thermo Scientific HCS Studio 2.0, coupled with Thermo Scientific high-content analysis platforms, are increasingly being used in cancer biology as they are well suited for large-scale, high throughput biology, including complex morphological analyses of cells, cell structure, and aggregations of cells in colonies, yet simple enough to be applied to everyday assays such as viability and proliferation. The challenges posed by complex cancer applications and cell models demand new types of quantitative analysis techniques, and the new software provides innovative assay techniques and functionality.

Thermo Fisher Scientific
For info: 866-984-3766
www.thermoscientific.com/highcontent

Science
Translational
Medicine

Integrating Science, Engineering, and Medicine

Publishing results that harness the basic sciences to advance human health in all areas of medicine

Submit your research
ScienceTranslationalMedicine.org

Recommend to your library
ScienceOnline.org/recommend

Join the ranks of high-profile papers published in *Science Translational Medicine*:

NEUROLOGY
A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior
A. G. Casali *et al.* (M. Massimini), *Sci. Transl. Med.* **5** 198ra105 (2013)

IMMUNOLOGY
TGFβ Receptor Mutations Impose a Strong Predisposition for Human Allergic Disease
P.A. Frischmeyer-Guerrerio *et al.* (H. Dietz), *Sci. Transl. Med.* **5** 195ra94 (2013)

BIOENGINEERING
A Human Disease Model of Drug Toxicity-Induced Pulmonary Edema in a Lung-on-a-Chip Microdevice
D. Huh *et al.* (D. Ingber), *Sci. Transl. Med.* **5** 159ra147 (2012)

CANCER
Genome-Wide Mutational Signatures of Aristolochic Acid and its Application as a Screening Tool
S.L. Poon *et al.* (B. T. Teh), *Sci. Transl. Med.* **5** 197ra101 (2013)

TISSUE ENGINEERING
Human Cartilage Repair with a Photoreactive Adhesive-Hydrogel Composite
B. Sharma *et al.* (J.J. Elisseeff), *Sci. Transl. Med.* **5** 167 ra6 (2013)

Chief Scientific Advisors

Elazer Edelman, M.D., Ph.D.
Massachusetts Institute of Technology

Garret FitzGerald, M.D.
University of Pennsylvania

Editor

Katrina L. Kelner, Ph.D.
AAAS, Washington, DC

scitranslmededitors@aaas.org

coming soon

bio-techne.com/launch

POSITIONS OPEN

CHEMISTRY FACULTY POSITION
Drug Discovery Department

Moffitt Cancer Center, an NCI-designated Comprehensive Cancer Center, is seeking applications to join the faculty of the Drug Discovery Department at the **ASSISTANT, ASSOCIATE,** or **FULL PROFESSOR/** Member level. Structure-based drug design and focused combinatorial chemistry library approaches are strongly supported by current and developing research infrastructure. Moffitt Cancer Center has outstanding core laboratory facilities and services available including: NMR and mass spectrometry, synthetic and parallel chemistry, structural biology, molecular modeling, high throughput screening, proteomics, and genomics shared facilities. The chemistry laboratories are newly renovated, providing a state-of-the-art environment for synthetic and medicinal chemistry research programs.

Synthetic organic chemists looking for strong collaborations with faculty in cancer biology, targeted therapeutics, tumor metabolism, imaging, tumor virology, and/or immunology are encouraged to apply. Successful candidates must possess a Ph.D. in synthetic organic chemistry or related areas with preferred experience in molecular recognition and/or bioorganic/synthetic/medicinal chemistry. Experience in the design of small molecules that disrupt protein-protein interactions and/or that target non-coding RNAs is desirable but not required.

Candidates for the rank of Assistant Professor/Member must have at least two years postdoctoral experience. Candidates for the rank of Associate Professor/Member must have a proven track record of independent federal funding and research and at least four years experience at the Assistant Professor or equivalent level. Candidates for Full Professor/Senior Member must have a proven track record of independent research and a national/international reputation in their field and at least five years experience at the Associate Professor or equivalent level. Academic rank and salary will be commensurate with experience and qualifications. The Moffitt Comprehensive Cancer Center is affiliated with the University of South Florida and secondary University appointments are available.

Qualified applicants should send curriculum vitae, a synopsis of current and future research programs, and the names and addresses of three or more references to: **Dr. Said Sebti, Professor and Chair, Drug Discovery Department, Moffitt Cancer Center, 12902 Magnolia Drive, MRC-DRDIS, Tampa, FL, 33612.** Electronic documents (in PDF format) are preferred and should be sent to **e-mail: said.sebti@moffitt.org.** The selection committee will begin reviewing applications July 1, 2014.

To apply, visit our **website: http://www.MOFFITT. org/Careers. REQ: 11928** *Moffitt Cancer Center provides a tobacco-free work environment. It is an Equal Opportunity/Affirmative Action Employer and a drug free workplace.*

BIOMEDICAL FACULTY POSITION
University of Wisconsin-Madison

The Department of Comparative Biosciences, School of Veterinary Medicine invites applications for tenure-track faculty position (**ASSISTANT/ASSOCIATE PROFESSOR**). Qualifications include Ph.D. (or DVM/M.D.), postdoctoral experience, ability to develop an extramurally funded research program, and commitment to teaching excellence. Research area is open, but preference will be given to those complementing department strengths, including neuromuscular physiology, stem cells, cancer, and development. Teaching responsibilities based on expertise. Send curriculum vitae, brief statements of research interests and teaching philosophies, and three letters of reference in response to PVL 79859: **John Svaren, Chair of Search Committee, Department of Comparative Biosciences, University of Wisconsin, 2015 Linden Drive, Madison, WI, 53706.** Apply by August 15, 2014. For additional information, see **website: http://www.vetmed.wisc. edu/current-searches/.** *Equal Opportunity/Affirmative Action Employer.*

POSITIONS OPEN

TENURE AND TENURE-TRACK FACULTY POSITIONS

The School of Medicine at the University of Louisville invites applications for faculty position(s) in the Department of Microbiology and Immunology at the **ASSISTANT, ASSOCIATE,** or **FULL PROFESSOR** level. Scientists interested in all areas of microbiology or immunology are encouraged to apply. Preferential consideration will be given to applicants with a research program that emphasizes the study of host-microbe interactions including studies on microbiota and pathogenic microorganisms. For more details on these positions as well as faculty and research programs, please visit **website: http://louisville.edu/medicine/departments/ microbiology.** Applicants should submit: (1) a cover letter, (2) a summary of present and future research plans, (3) curriculum vitae, and (4) the names/e-mail addresses of three references at **website: https://highereddecisions.com/ uofl/emp_apply_login.asp** (**Job Id: UL157**).

CAREER OPPORTUNITY

Doctor of Optometry (O.D.) degree in 27 months for Ph.D.s in science and M.D.s. Excellent career opportunities for O.D./Ph.D.s and O.D./M.D.s in research, education, industry, and clinical practice. This unique program starts in March of each year, features small classes, and 12 months devoted to clinical care.

Contact the Admissions Office, **telephone: 800-824-5526** at the **New England College of Optometry, 424 Beacon Street, Boston, MA 02115.** Additional information at website: **http://www. neco.edu** or at **e-mail: admissions@neco.edu.**

MARKETPLACE

An Explosion of Bioinformatics Careers

Big data is everywhere, and its influence and practical omnipresence across multiple industries will just continue to grow. For life scientists with expertise and an interest in bioinformatics, computer science, statistics, and related skill sets, the job outlook couldn't be rosier. Big pharma, biotech, and software companies are clamoring to hire professionals with experience in bioinformatics and the identification, compilation, analysis, and visualization of huge amounts of biological and health care information. With the rapid development of new tools to make sense of life science research and outcomes, spurred by innovative research in bioinformatics itself, scientists who are entranced by data can pursue more career options than ever before. **By Alaina G. Levine**

Today's bioinformaticists are in for a real treat. With a seemingly endless stream of biological data being generated across sectors, there is high demand for talented, experienced professionals at the crossroads of biology, statistics, and computer science. Scientists who can analyze large amounts of information and present it in a clear manner to decision makers are finding the sky is the limit in terms of jobs and career pathways, especially in the big pharma and biotech sectors.

"It's a fun place to be and an exciting time to be in big data," remarks **Sriram Mohan,** professor of computer and software

engineering at Rose-Hulman Institute of Technology, who is spending his sabbatical developing bioinformatics software for Avalon Consulting, a data management firm.

And what an immense amount of data it is, due in part to a paradigm shift in the field, from data generation to data analysis, says **W. Jim Zheng,** associate professor in the School of Biomedical Informatics at The University of Texas Health Science Center at Houston. Now, with so much data being produced because of easier and more cost-effective tools, there is an even greater need for specialists who can make sense of the mountains of information in such a way that is meaningful for scientists and clinicians, and ultimately beneficial to customers and patients.

The increase in job opportunities is also being driven by a change in how bioinformatics is perceived in industry and academia. Previously, "scientists and companies used to look at bioinformatics as a tool," says **Wim Van Criekinge,** a professor of bioinformatics at Ghent University in Belgium and chief scientific officer at MDxHealth, a company developing epigenetics-based cancer diagnostics. Bioinformaticists would be called upon to answer a question about data; their role was to run an algorithm on a database that provided that answer. "But the subject has evolved from a service, like histology, to its own research arena…. Bioinformaticists are now the motor of the innovation," he adds. They not only answer the data inquiries, but also, more importantly, determine what questions need to be asked in the first place.

As a result, "there are many opportunities for scientists to pursue a bioinformatics/big data career in the biotech/big pharma industry at the moment," notes **Jared Kaleck,** senior director of computational chemistry/biology and formulation development at executive search firm Klein Hersh International.

Discovering Where the Careers Are
To begin to recognize where the jobs are and how the career tracks are accessed, it helps to understand how bioinformaticist positions are organized within different firms. In pharmaceutical and larger biotech companies, big data scientists may find themselves working in one of three different types of organizational structures. In one, all of the big data scientists and bioinformaticists work out of a central core. This large team could be concentrated in research and development (R&D) or information technology (IT) departments, and the scientists work almost like consultants on projects throughout the company, and are lent out as needed.

In a second model, bioinformaticist positions are decentralized, and located within different therapeutic areas. For example, at Johnson & Johnson (J&J), **Patrick Ryan** leads the epidemiology analytics group. As a clinical informaticist, he develops statistical methods to analyze "observational databases," such as electronic health records, to map disease patterns in order to better understand "the real-world effect of our medicines, and to develop safety protocols **continued>**

Upcoming Features

Postdocs: Broadening Your Skills —Aug. 22 ■ *Faculty: Professionalism 1—Sept. 12* ■ *Faculty: Professionalism 2—Oct. 3*

> "The deeper you understand the biology, the better you do your job in this area."
>
> —W. Jim Zheng

and mitigate risk for the patients," he explains. His team is part of an overall epidemiology department which reports to the chief medical officer of J&J. But he notes that the company also has a robust informatics and IT division, whose mission is to "provide technical perspective on how to manage and analyze data."

The third kind of organizational structure found in big pharma is a hybrid of the other two. **Christian Reich,** global head of discovery informatics at AstraZenica, shares that his company currently follows this model, although he notes that the trend sees enterprises restructuring themselves to follow one of the other paradigms every few years. His job entails overseeing a principal group of 25 specialists, but other informaticists are sprinkled throughout the company. Similarly at Pfizer, bioinformaticists are embedded in therapeutic units as well as core centers of excellence, says **Susan Stephens,** senior director of research and development business technology at Pfizer.

Genentech follows a similar mixed model, explains senior director of bioinformatics, **Robert Gentleman** (who is also a co-developer of the statistical computing and graphics programming language known as R). Bioinformaticists are organized in a core center, but they "integrate with different functional areas," he says. "They are in one department, but day-to-day they work directly with disease area specialists."

Depending upon where they are housed in their company, big data scientists can expect to have varying tasks. In R&D, bioinformaticists conduct research on new approaches to analyzing data and help design and possibly even build the analysis tools utilized by scientists throughout the company, says Reich. Here, the idea is to examine existing open-source algorithms to apply them in novel ways, or to create entirely new algorithms that rely heavily on mathematical and statistical expertise. "The goal is to put together a platform so the data analysis can be done easily, and return high-quality results," he adds.

Bryn Roberts, global head of operations, which includes informatics for Pharma Research and Early Development at Roche, notes that in his company's hybrid structure, informaticists and data scientists are involved in a wide range of activities. They develop and support software systems; they procure and make external scientific content available and actionable by scientists throughout the company; they implement and maintain workflow systems, such as e-lab notebooks, for both drug discovery and regulated functions; and they support and perform data, image, and text mining and analysis to support scientific decision making.

At Genentech, bioinformatics scientists participate in all levels of the investigative process, from helping to design experiments that will find genetic markers for disease, to leveraging their bioinformatics skills to help find biomarkers that will aid in patient selection.

Elsewhere in big pharma, big data scientists may be tasked to investigate trends in diseases and drug development and discovery, which can involve collaboration with the marketing team. They may also provide quantitative support for business decisions, such as in which therapies firms should invest, says **George Telthorst,** director of the Center for the Business of Life Sciences at Indiana University.

Contract Research Organizations (CROs) also offer much for those interested in the big data profession. As **Dimitris Agrafiotis,** vice president of informatics and chief data scientist of Covance, one of the world's largest CROs, attests, "CROs are becoming the R&D engine of the pharmaceutical industry." Covance data scientists can expect to be involved in myriad projects across the entire drug development continuum, from biomarker discovery to preclinical development, clinical trials, health economics and outcomes research, and even marketing.

Beyond working in big pharma or biotech, there are also opportunities for data scientists in industry support companies, such as those that produce bioinformatics software and other data analysis tools. Furthermore, bioinformaticists are recruited by health insurance corporations and hospital management organizations.

Even academia has seen an uptick in bioinformatics career opportunities, as the discipline itself is expanding. Zheng recalls a time in the early days of genomics when doing big data research meant scientists had to leave the university lab and head to industry, but the tide has changed. Now, programs like the National Institutes of Health's Big Data to Knowledge funds academic research in bioinformatics.

Pursuing Big Data Skills

Experts agree that the most successful bioinformaticists (and the ones who land the jobs) are those who have a multitude of skills. But the starting point is always knowledge of life sciences, also referred to as "domain expertise" in the industry. In fact, "the deeper you understand the biology, the better you do your job in this area," says Zheng. Hiring managers specifically seek out scientists who have doctorates in various areas of life sciences, including molecular and cellular biology, chemistry, genetics, immunology, and epidemiology. At Genentech, Gentleman looks for candidates who possess expertise in the biology of a particular disease.

Additional critical skills are required for big data careers in industry, such as text mining, ontology, data integration, machine learning, and information architecture. A superior "quantitative ability," as Gentleman calls it, which covers a range of statistical capabilities, is a must, as are overarching computing skills. These include core programming abilities, such as coding in C++ or Java, or scripting

Scientist's choice – University vs. Industry

There are a few key reasons why outstanding scientists should consider joining big pharma like Roche. Dr. Andrew Thomas, Head of Medicinal Chemistry, Neuroscience at Roche in Basel, explains why he joined one of the leading pharma companies a few years ago.

Dr. Andrew Thomas is Section Head Medicinal Chemistry at Roche in Basel.

What is your current role?

I am a member of our Chemistry Leadership Team, Extended Small Molecule Research Leadership Team, global Non Clinical Drug Safety Committee and I'm directly involved in the discovery and development of new medicines for the therapeutic areas of neuroscience ophthalmology and rare diseases. The goal of my team is design and discover molecules that will one day become medicines. One of the main responsibilities of my current role is to ensure that our department retain a reputation as being an employer of choice for the best scientists in our domain by attracting world class talent to join us.

Why did you decide to join Roche?

There are a few key reasons why I joined Roche. First and foremost – I had an immediate strong and authentic affinity to the outstanding scientists I met during my interviews – which continues to this day. Before I interviewed with Roche I was thinking to become a University Professor – but I was convinced by joining Roche I would be able to get the best of both worlds by generating new knowledge, teaching and mentoring students and more important being to directly impact healthcare through the implementation of medicinal chemistry knowledge. These have all been realised – which is why I have decided to stay with Roche.

What can Roche offer to young scientists?

We are able to offer a unique experience – where we provide the future leading scientists a stimulating and truly enriching environment interacting within very experienced teams where our young scientists gain an authentic insight into the leading capabilities of our organisation. Because we attract leading talents they are regarded right from the beginning as competent researchers and our skill is to support them advancing their own hypotheses and trying out new approaches to their research. The development of novel medicines and diagnostics gives them the opportunity to use their skills in complex areas of science, where there is still so much to discover. Industry has the advantage over academia that we focus very efficiently on the direct application of our research results rather than only generating them for others to use which is often the case in academia. We are able to drive this through our focus on creating new knowledge and value through inventions followed by timely publications. In addition the core of our industry is composed of a team of multi-disciplinary and recognized world experts across all key areas. Having access to experts internally is an asset unique to industry as well as our efficient infrastructure that ensures our company runs smoothly so we can focus on our research.

What would you advise job seekers who want to join Roche?

Make personal contact to our recruiters and departmental representatives who are out in the field at career fairs and/or giving scientific seminars around the world. If you are someone who has the passion to be a world class contributor to the discovery and development of new medicines demonstrate your potential early in your career during your formal education which will catalyse you being recognized by Roche employers.

To apply online for positions visit www.careers.roche.ch
Roche is an Equal Opportunity Employer.
The next step is yours.

in PERL or Python, says Van Criekinge. It is vital to be able to navigate operating systems like UNIX and Linux as well as have knowledge of common tools such as Hadoop and NoSQL databases, adds Mohan. Experience in data visualization and building effective user interfaces, as well as familiarity with hardware, buttresses your marketability.

In addition to scientific problem-solving skills, bioinformaticists must have business proficiency. "Bioinformatics is a team sport," says **Stephen Ruberg,** distinguished research fellow, advanced analytics of Eli Lilly and Company, and thus project management, teambuilding, and communications experience is a requirement. In fact, "being able to communicate with the other scientists is really the most important skill we look for," says Gentleman.

Nimbleness and the ability to adapt quickly are also fundamental. "It's a fast-paced environment," says Van Criekinge. "You have to have a mindset of constantly using new tools, or you will become obsolete in two years."

Landing Big Data Jobs

It would be ideal if companies could find candidates who have all of the above skill sets, but sources indicate that that is wishful thinking. More often than not, hiring decisions are made based on the immediate needs of the team, especially given their interdisciplinary nature. "We look for people whose expertise complements the existing group's skills," says Roberts. However, just because you lack a specified talent or interest area as noted in a recruitment ad, doesn't mean you shouldn't apply anyway. "We share CVs internally all the time," says Stephens. So even if she can't bring you in to her group at Pfizer, she may be able to find another team at the company for which you would be a good fit.

Vincent Fusaro

In some cases, companies are growing their own talent, as a result of the lack of large numbers of qualified, multi-skilled candidates. At Roche, "we offer continuous training in various areas and encourage our staff to attend conferences, publish, or pursue higher degrees," says Roberts. Pfizer data scientists have myriad chances to pursue professional development, and are also granted time to try out new techniques, says Stephens, something she refers to as "sandbox opportunities."

Experience plays a major role in gaining access to jobs. Kaleck highly recommends doing an industrial postdoc or internship, but in absence of these, scientists might consider "bridge" programs, like the Insight Data Science Fellows Program. This fully supported, six-week training opportunity offers postdoctoral fellows the chance to work on real-world problems for the likes of Facebook and Microsoft. This appealed to **Vincent Fusaro,** whose Ph.D. is in bioinformatics. As a fellow, he gained expertise in databases, Python, machine learning, and data visualization, which helped him land a position as a self-described "data ninja" for Invitae, a genetic information company. Today he is responsible for software engineering, data analysis, and pipeline and product development, among other tasks.

The Expanding Big Data Universe

Data scientists can expect the field to change and evolve in novel ways in the near future. But the bottom line is that "companies are growing their bioinformatics," says Kaleck. "There are 100% more job opportunities opening up in bioinformatics than ever before," much of which is driven by an increase in venture capital investment.

Given that big data "is the hottest field on the planet," says Agrafiotis, those who have the requisite skills and expertise often have their pick of opportunities. "I have to fight Google, Amazon, LinkedIn, and hedge funds to hire the top people. They are valuable in any industry."

In particular, the future of big data in big pharma and biotech sectors is bright and exciting. "Bring your expertise to health care," says Telthorst, "and you'll know you're going to make a difference, at the patient level and at the societal level."

Alaina G. Levine is a science writer based in Tucson, AZ.

Featured Participants

AstraZenica
www.astrazenica.com

Avalon Consulting, LLC
www.avalonconsult.com

Covance
www.covance.com

Eli Lilly and Company
www.lilly.com

Genentech
www.genentech.com

Ghent University
www.ugent.be/en

Indiana University
www.indiana.edu

Insight Data Science Fellows Program
www.insightdatascience.com

Invitae
www.invitae.com/en

Johnson & Johnson
www.jnj.com

Klein Hersh International
www.kleinhersh.com

MDxHealth
www.mdxhealth.com

Pfizer
www.pfizer.com

Roche
www.roche.com

Rose-Hulman Institute of Technology
www.rose-hulman.edu

University of Texas Health Sciences Center at Houston
www.uth.edu

Bioinformatics opportunities at the Wellcome Trust Sanger Institute

The Wellcome Trust Sanger Institute (WTSI), based south of Cambridge (http://www.sanger.ac.uk/), is an internationally renowned genomic research centre, funded primarily by the Wellcome Trust. Our mission is to use genome sequences to advance understanding of the biology of humans and pathogens in order to improve human health. The Institute is located on the 100 acre Genome Campus with the European Bioinformatics Institute (EMBL-EBI), the Wellcome Trust Conference Centre and its associated advanced courses and conferences programme; we share a broad vision to develop the Campus as a hub of science including informatics, business, advanced scientific training and cultural activities in the area of genetics and genomics.

Informatics is central to genome science at the WTSI. Our Bioinformatics programme develops and applies methods to process, store and analyse data generated by high-throughput projects. Its principal aims are to infer genomic knowledge through computational analysis and integration of data and to generate resources of lasting value to biomedical research.

We offer a wide range of Bioinformatics opportunities across a broad spectrum of Faculty-led research projects, such as:

- Computational genome biology
- Genome informatics
- Genomics of gene regulation
- Population genomics of molecular phenotypes
- Using outbred genetic variation to understand basic biology
- Vertebrate genome analysis

Working collaboratively in Bioinformatics is important to us at the WTSI. A major collaboration is centred on capacity building in bioinformatics and large-scale data analysis in Africa, including supporting the development of a new data centre in Uganda. More details on these projects and collaborations can be found on our website at: http://www.sanger.ac.uk/research/areas/bioinformatics/#proj

For more details on all of our bioinformatics positions and to submit your CV/apply online please go to https://jobs.sanger.ac.uk.

For more information on working at the Institute and the benefits and facilities available to staff visit http://www.sanger.ac.uk/workstudy/.

www.sanger.ac.uk
Working towards diversity through equality

Genome Research Limited is a Registered Charity No.1021457

Hunting for Talents
NUAA, Jiangsu, China

Nanjing University of Aeronautics and Astronautics(NUAA) is a research-oriented national key university of "211 Project". It also enjoys a well-balanced development of multiple disciplines in engineering, technology, natural sciences, economy, management and social sciences with the characteristics of aeronautics, astronautics and civil aviation. NUAA is qualified to be "Dominant Discipline Innovation Platform of 985 Project" and to independently recruit and receive international students who are granted the Chinese Government Scholarship. Now NUAA consists of 16 colleges with more than 3,000 staff members and approximately 26,000 degree students.

Academia and education at NUAA represent strong capacity among all the universities in China. It has acquired national status through the quality of its excellence research work, especially in the areas of Aerospace Engineering, Mechanics,Electromechanics, Economy and Management, etc.

NUAA gives a warm welcome to excellent experts, scholars and young students from both home and abroad, who are willing to serve the country, dedicate themselves to the development of aerospace science and make contributions to the industrialization, information technology of China. NUAA will provide teachers and researchers with a good academic environment, satisfactory working and living conditions and a stage on which they can put their talents to good use.

Contacts

Ms. Zhao Haiyan, Mr. Cao Yunxing
Personnel Division, NUAA
Address: 29# Yudao St. Nanjing, Jiangsu Province, Postcode: 210016
Tel: +86-25-84892461
Fax: +86-25-84895923
Email: zhaohaiyan@nuaa.edu.cn
Web:http://www.nuaa.edu.cn/nuaanew   http://rsc.nuaa.edu.cn

Opening of the Director's Position and Faculty Positions in the New Institute of Health Informatics at Peking University

DIRECTOR'S POSITION AND FACULTY POSITIONS

Peking University is establishing a new Institute of Health Informatics that aims to advance research and teaching in all areas of health informatics. Director's position and several faculty positions at all academic levels are open for application.

Faculty in the new institute can have joint appointments at Peking University's Medical Center and the School of Interdisciplinary Research. Faculty are expected to develop new methodologies and technologies to manage, query, and analyze clinical data and/or perform innovative analysis to find new patterns and trends that have significant impact on clinical practice or health policy. Faculty are especially encouraged to interact with the university's nine affiliated hospitals and other health care organizations in China to take advantage of China's vast clinical resource. Faculty are expected to teach core courses on the curriculum.

The Director, in addition to his/her role as a faculty member,is also expected to lead the institute in world-class research, develop a comprehensive curriculum, recruit faculty members, and establish collaborations with the university's nine affiliated hospitals and other health care organizations in China.

Candidates should have a doctorate degree and strong research background in a related field. Hiring levels will be commensurate with experience and achievements. Interested applicants please send CV, research statement, teaching statement, five representative publications, and contact information of three references to Ms Siyuan Gong *(gongsy@pku.edu.cn)*, and include a cover letter indicating the position applying to.

Peking University provides competitive salary and funding support for research and teaching. Successful candidates may also be supported by the Peking-Tsinghua Center for Life Sciences. The recruitment continues until all the positions are filled.

Faculty Positions available at Hohai University, Nanjing, China

Hohai University invites applications for faculty positions at the assistant, associate, or full professor level in the area of engineering, science, economics, management, liberal arts, and law. Applicants should have a doctoral degree from a prestigious university. For the complete job announcements and directions on how to apply, visit: rsc.hhu.edu.cn or contact the Department of human resource at 86-25-83786205.

Hohai University, founded in 1915, wins its worldwide reputation on the research of Water Science& Civil Engineering&Environment Engineering. It is a National key university of China, and among the universities of the National "*211 Project*" and Innovation Bases of the National "*985 Project*". Hohai University aims to be a research oriented university.

Chief Editor
Nature Methods

The Nature Publishing Group is looking for a Chief Editor for *Nature Methods*, the leading biomedical research journal devoted to publishing the latest and most exciting advances in life science research methods. The journal covers a wide range of fields including biophysics, structural biology, genomics, proteomics, cell biology, systems biology and neuroscience and publishes technologies in areas such as sequencing, microscopy, mass spectrometry and computational tools.

Nature Methods is an interdisciplinary research journal that publishes novel techniques and tools for basic laboratory research in the life sciences, as well as significant improvements to tried-and-tested methods. The journal combines practical, technique-driven subject matter with rigorous peer review standards to ensure that readers are consistently presented with only the most valuable and highest quality methodological research. It provides laboratory researchers with new tools to conduct their research and places a strong emphasis on the immediate practical relevance of the work presented.

The position of Chief Editor is a senior appointment, reporting to the Executive Editor, and is backed by a strong commitment to excellence and investment in resources. This is an exciting opportunity to join the world's leading scientific publishing company and to be involved with the development of a prestigious journal. The Chief Editor is responsible for leading a skilled and enthusiastic editorial team and for developing the editorial content of the journal, both in print and online.

To meet these challenging tasks, the ideal candidate will have intellectual vision, strong leadership qualities, and be able to demonstrate commercial awareness appropriate for a controlled circulation journal with substantial print and online advertising. She or he should have a PhD in a relevant field, have excellent scientific judgment, and a broad knowledge and understanding of basic research in the life sciences.

The ideal candidate will have a strong research background and publication record while previous editorial and/or managerial experience is an additional advantage. A key aspect of the job is interacting with the scientific community and attending international conferences. The successful candidate must, therefore, be dynamic and outgoing, be prepared to travel, and have excellent interpersonal skills.

The position will be based in Nature Publishing Group's office in New York, and accordingly candidates must be able to demonstrate the right to live and work in this location. The terms and conditions are highly competitive, reflecting the importance and responsibilities of the role.

For further information about *Nature Methods*, please visit **http://www.nature.com/nmeth**

To apply, please go to **https://home.eease.adp.com/recruit/?id=9631521** and submit a cover letter stating your suitability for this post, salary expectations, a current CV, potential starting date, and a statement that encapsulates your vision for Nature Methods.

All applicants will be reviewed upon receipt with a close date of **June 27, 2014**.

Please note that our system only allows a single file upload. If you cannot create a single document to include all of the requirements for posting, please reply to the thank you message you will receive at the conclusion of your application with any additional details.

To learn more about our company, please visit our web site at www.nature.com.

The following links cover our policies on Privacy, Use of Cookies, Legal Notice, and Terms:
http://www.nature.com/info/privacy.html
http://www.nature.com/info/cookies.html
http://www.nature.com/info/legal_notice.html
http://www.nature.com/info/tandc.html

Nature Publishing Group is an Equal Opportunity Employer.

DEPARTMENT OF BIOTECHNOLOGY MINISTRY OF SCIENCE & TECHNOLOGY GOVERNMENT OF INDIA

NOMINATIONS/APPLICATIONS FOR THE POST OF DIRECTOR, IN-STEM, BANGALORE

Nominations/Applications are invited for the post of Director on deputation (including short term contract), Institute for Stem Cell Biology and Regenerative Medicine (**inStem**) [**www.instem.res.in**], Bangalore, an autonomous Institute under the Department of Biotechnology, Government of India (**www.dbtindia.nic.in**).

inStem is a state-of-the-art research institute in Bangalore, dedicated to the study of stem cell and regenerative biology. In-Stem emphasizes collaborative research in stem cell biology. inStem's mandate to allow this cross-disciplinary, multi-pronged approach to research, straddles the divide between clinical and laboratory research in stem cell biology. In trying to answer intractable and challenging questions that face the field, inStem seeks to rewrite the paradigm of the research institute: without barriers and across disciplines.

The applicants should be below **56 years of age**. The post carries Pay Scale of **Rs.80000/-** (Fixed) with usual allowances **as per the Govt. of India Rules**.

Those applicants from Central/State Governments, **research institutes, autonomous bodies** etc. are typically expected to be:
1. Holding analogous post or post in HAG + Scale on regular basis in the parent Department.
2. 3 Years service in PB-IV with Grade Pay of Rs.10,000/- ; or
3. 5 years service in PB-IV with Grade Pay of Rs.8900/- ; or
4. 7 years service in PB-IV with Grade Pay of Rs.8700/-

Essential
1. Ph.D. in Life Sciences or other Allied Sciences/Engineering or M.D. or equivalent
2. Relevant experience in the area of organismal, cellular, molecular, systems biology, or biotechnology, and
3. Excellent track record, proven expertise of guiding research activities in the area of specialization.
Experience in Science – administration will be an asset.

The nomination/application along with detailed curriculum vitae including the date of birth, address for correspondence including telephone no., fax and e-mail address, qualifications acquired, professional and research experience, present position and scale of pay with total emoluments, publication details and a 500 words write-up on the candidate's vision of inStem, Bangalore for the next ten years may be sent to **Smt. Kusum Lata Sharma, Deputy Secretary, Department of Biotechnology, Block-2, CGO Complex, Lodi Road, New Delhi-110 003**, superscribing the cover **"Nomination/Application for the Post of Director, inStem"** so as to reach her within 30 days from the date of publication of this advertisement. The detailed advertisement and format of application is available at DBT website **www.dbtindia.nic.in** and at the inStem website **www.instem.res.in**

The Department reserves the right to relax any of the requirements prescribed above.

Cathy Walker

The fun of science

Recently, I was reminded that my job as a graduate student in a science lab is actually "really cool." I was in need of a reminder. I was nearing the end of my fourth year, and I was in a research rut I couldn't seem to escape. A group of friends from my undergraduate days invited me to spend a late-spring weekend at a lakefront cabin. Our plans to enjoy the lake and the lengthening daylight were dashed by a cold front that came through, bringing 2 days of rain. We still managed to enjoy ourselves, though, by playing card games, making ridiculously good barbecued pork, and reminiscing.

The weather changed eventually, and on our last night at the cabin, we were able to spend some time outside. We built a bonfire. One of the guys went to look for wood and came back with a butane torch. He discovered that he could melt holes in empty aluminum beverage cans. (I quickly located my cellphone in case I needed to call 911; even on weekend trips, apparently, I cannot escape my role as safety officer.) One of the guys asked if I wanted a turn, but I declined. "No thanks," I said. "I get to use one of those at work."

After a short, puzzled look, my friend replied, "Really? Why?"

Me: Usually to make sure our glassware and sample pans are clean and dry.

My Friend: Cool. And nothing melts?

Me: No, it's either glass or platinum. The metal gets all glowy. (Clearly my public-outreach vocabulary needs work.)

My Friend: What else do you get to use?

The conversation continued along those lines. I learned that my work—with torches, liquid nitrogen, reagents, and a bunch of solvents and, especially, the fact that I get to make things—is impressive to people who spend 40 hours a week in an office. True, they are impressed by different things than I am: To me, fluorine-19 NMR is really cool, but that one didn't rate so high with my nonscientist friends. They preferred the fact that we occasionally use liquid nitrogen to make ice cream.

Refreshed from the trip, I returned to the lab and tried to convince my lab mates that, according to some people, doing science is cool because we get to use butane torches and do other interesting things. I'm not sure they were convinced, but for me the message really hit home a few weeks later, as I was reading a post at the blog Tenure, She Wrote. The post I was reading (which I highly recommend) was about the reasons why it's good to be a woman in science (http://bit.ly/1kDJ5zT). Number one is, "We get to do science." The writer even describes science as "kick-arse."

I hadn't recognized it until that reminder from my friends, but it's true: We get to do cool things, and it's not just playing with torches. Whenever I decide that my work is getting boring, I can go and learn a new technique. I'm encouraged to argue with my lab mates and my adviser, and I win praise for making things up—that is, for coming up with ideas and solving problems in creative ways. We make things. We play. And the people we work with are great.

We make things. We play. And the people we work with are great.

Shortly after my return from the cabin, I started working with two undergraduate students who were participating in a summer research experience. Having their fresh eyes and energy in the lab served as a further reminder that I had chosen an interesting and exciting career. Furthermore, the added responsibility forced me to organize my work and manage my time more effectively; I had to be sure that the undergrads' project stayed on schedule and that they always had something to do. Being a mentor while completing my own assignments raised my confidence. I escaped my rut.

It's true: Being a scientist really is "kick-arse." If you ever forget that, go spend time with friends whose jobs don't allow them to make things and play with cool toys. They surely will remind you.

Cathy Walker is completing her graduate studies in polymer science and engineering at the University of Massachusetts, Amherst. For more on life and career issues, see http://www.sciencecareers.org.

ILLUSTRATION: MARC ROSENTHAL