Science

▲AAAS

High flyers

Migrating geese hug land to cross the Himalayas *p. 250*

Published by AAAS

Toward designing safer chemicals

One year ago, an industrial coal-processing liquid contaminated the Elk River in West Virginia and affected the tap water of 15% of the state's population. The spill was declared a federal disaster, and ongoing investigations remain. Last month, a report assessing the water and health impacts of the Elk River spill pointed to the lack of a sound scientific approach for responding to and recovering from such incidents.* This year also marks 5 years since the Deepwater Horizon oil spill in the Gulf of Mexico, and last month brought the 30-year anniversary of the Bhopal gas tragedy that killed thousands, considered the world's worst industrial disaster. Despite our best efforts and intentions, human-made chemicals continue to be released into the environment, often with unquantified and potentially unquantifiable deleterious consequences. The questions posed to science are how to better understand the nature of synthetic substances in order to predict their potential adverse impacts on humans and the biosphere, and how do we design future substances to eliminate the need for engineered control systems.

> *"...how do we design future substances to eliminate the need for engineered control systems."*

Until recently, descriptive toxicology characterized the impact of toxic substances on living organisms and ecosystems. Today, the emerging fields of mechanistic and molecular toxicology are evolving our understanding of how toxic exposure happens. We also now know that many of the physical and chemical properties that we impart to molecules to gain function and performance are linked to adverse consequences. Furthermore, improved knowledge about the human body's absorption, distribution, metabolism, and excretion of chemicals suggests a path toward reducing hazards through molecular design. These four criteria are enabling predictive modeling that uses the physicochemical properties and structural motifs of a chemical to provide insights into the transport and fate of chemicals in the environment, their metabolism and (bio)degradation, and their epidemiology.

However, associating the physical and chemical properties of a chemical with the mechanism of an adverse outcome is only a beginning. These associations require knowledge about large numbers of combinations of physicochemical properties. Although it is necessary to investigate chemicals individually, this is never the real-world situation. Chemicals interact in ways that can magnify or mitigate their effects, sometimes dramatically. All toxicological data to date must be considered with this limitation in mind. One of the greatest future challenges in safer chemical design is recognizing that the timing of dosing—not just the dose alone—is a critical factor in toxicity. Indeed, the adverse effects of chemical exposure on genes in one generation can be caused by exposures (doses) of prior generations, including during fetal and neonatal development.

Do more data mean more knowledge? With the advent of high-throughput screening in efforts such as ToxCast by the U.S. Environmental Protection Agency and Tox21, a consortium of government agencies, private organizations, and university partners, there is an urgent need for new ways to mine, curate, store, and manage data in ways that can ultimately be used to design more benign products, processes, and systems.

To begin to address these challenges, however, one area perhaps requires primary focus: transdisciplinarity. Research institutes, universities, industry, and funding and regulatory agencies (among other stakeholders) must cultivate a research ecosystem in which efforts are collaborative and knowledge is shared across disciplines, including pharmacology, (eco)toxicology, chemistry, modeling, and biostatistics.

If traditional analyses can be coupled with integrated systems approaches, then the knowledge gained about the nature of complex systems may well lead to the design of chemicals that are compatible with life.

– Julie B. Zimmerman and Paul T. Anastas

Julie B. Zimmerman is a professor in the Department of Chemical and Environmental Engineering in School of Engineering and Applied Sciences, Yale University, New Haven, CT, 06520. E-mail: Julie.zimmerman@yale.edu

Paul T. Anastas is a professor in the School of Forestry and Environmental Studies, Yale University, New Haven, CT, 06520.

PHOTOS: (INSET) CARLOS BARRIA/REUTERS; (TOP RIGHT) COURTESY OF JULIE ZIMMERMAN; (BOTTOM RIGHT) JOANNE WILCOX

*http://pubs.acs.org/doi/abs/10.1021/es5040969.

10.1126/science.aaa6736

IN BRIEF

An expanded range
and new protections
may help Mexican
gray wolves thrive.

More room for Mexican wolves

In an effort to improve the recovery of Mexican gray wolves, the rarest subspecies of the gray wolf (*Canis lupus*), the U.S. Fish and Wildlife Service is expanding the animals' range and placing them under the protection of the Endangered Species Act. The wolves can now roam from the Mexican border across much of New Mexico and Arizona south of the Grand Canyon. Since 1998, when the first captive-bred Mexican wolves were released in Arizona's Blue Range, the animals have struggled. In contrast with the successful 1995 release of Rocky Mountain gray wolves in Yellowstone National Park (which today have spread into several states and number over 1000), the Mexican wolf numbers only 83, and the population is inbred. Under the new ruling, more captive-bred wolves will be freed in national forests in the two states and allowed to disperse west and east to the California and Texas borders. Wolves moving beyond the new range will be captured and returned to the core area. http://scim.ag/mexwolf

AROUND THE WORLD

Slew of deals for 23andMe

MOUNTAIN VIEW, CALIFORNIA | Direct-to-consumer genetic analysis company 23andMe has announced the first two of a string of anticipated deals to share its data with drug developers. Through the sale of $99 saliva testing kits, the company has built a coveted genetic database of 800,000 people, more than 80% of whom have agreed to participate in research. Last week, 23andMe announced that it will share 3000 whole genomes with biotech company Genentech to help identify new drug targets for Parkinson's disease—in return for an upfront $10 million, plus more to come, depending on the project's progress. And on 12 January, it revealed another agreement with pharma giant Pfizer, for a lupus research effort that will involve enrolling 5000 patients in a longitudinal study. The company has said that it plans to roll out 10 such deals this year.

Scheme reveals anti-HIV stigma

NANYANG, CHINA | A local government in Henan province reprimanded four housing officials after a developer hired six people with HIV to harass residents who resisted demolition of their homes, according to Chinese state news agency Xinhua. The patients brandished certificates showing they had HIV and needles filled with their blood, according to other Chinese press reports. The chance of transmission through needle pricks is low, especially when individuals are on antiretroviral therapy, says Joseph Tucker, a medical researcher at the University of North Carolina, Chapel Hill, who directs a project focused on sexually transmitted diseases in China. "This thuggery, underpinned by a fear of HIV-infected individuals, shows that stigma still powerfully shapes the lives" of many Chinese with the virus, he says.

Flight of fancy

MUMBAI, INDIA | A talk on ancient flying contraptions at the annual Indian Science Congress this month has caused an uproar in India's scientific community.

Kepler-186f
WHERE THE GRASS IS ALWAYS **REDDER** ON THE OTHER SIDE

RELAX ON KEPLER-16b
THE LAND OF TWO SUNS
WHERE YOUR SHADOW ALWAYS HAS COMPANY

EXPERIENCE THE GRAVITY OF HD 40307g A SUPER EARTH
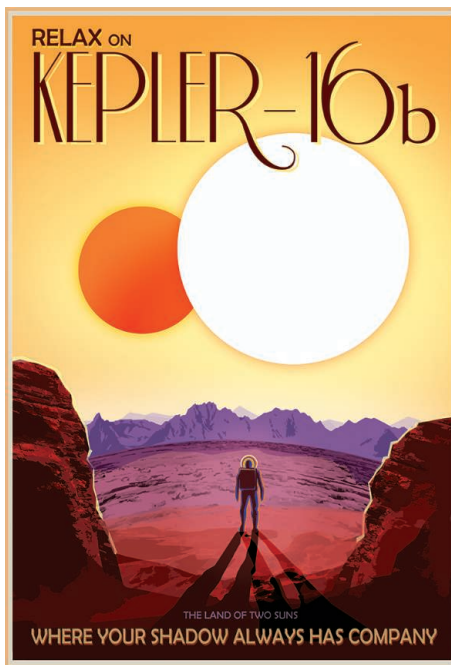
Visit scenic Kepler-186f: NASA invites you to try a holiday in space

NASA scientists announced earlier this month at a meeting of the American Astronomical Society in Seattle that the catalog of possible planets around other stars detected by the Kepler spacecraft has now topped 4000; more than 1000 of those had been confirmed as exoplanets by other methods. Fergal Mullally of the Kepler Science Office says that the new catalog contains "more Earth-like candidates than ever before." One is "the closest analog to Earth found to date," he says. To celebrate Kepler's achievements—cut short in 2013 when its pointing mechanism failed—NASA's Jet Propulsion Laboratory has produced a series of retro-style travel posters for potentially habitable exoplanets, extolling the virtues of their strong gravity, red grass, and binary suns. Throughout last year, NASA researchers worked to develop a new mission for Kepler, using the pressure of sunlight to help steady it. The so-called K2 mission announced its first discovery of a new exoplanet, HIP 116454b, last month. As well as exoplanets, K2 will study planet formation, stellar evolution, and other, more distant objects. **http://scim.ag/_Earthlike**

Retired pilot Anand Bodas drew ire when he spoke of the existence of 7000-year-old airplanes capable of both terrestrial and interplanetary flight, as described in an early 20th century Indian text. One example, the Shakuna Vimana (pictured), was supposedly constructed of ammonium chloride, chickpeas, and mercury. Naresh Chandra, a secretary of the Indian Science Congress, says critics of Bodas's talk miss the point; ancient Indian science was very advanced, he says, and "anybody doing research would like to learn from history." That argument may fly for a subject like trigonometry, but "planes and missiles are ludicrous," Abhay Ashtekar, a theoretical physicist at Pennsylvania State University,
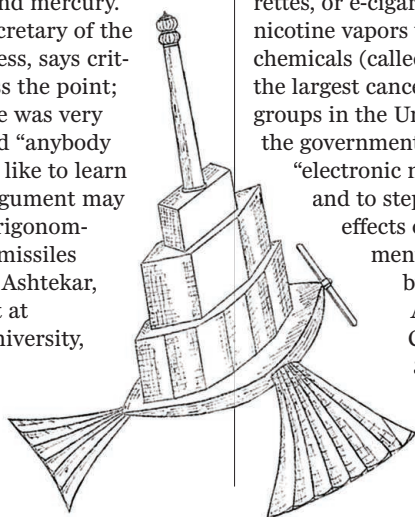
The "ancient" Shakuna Vimana, a spacecraft described in a modern text based on Sanskrit writings.

University Park, told India's *Economic Times*. "Let's not insult the great minds of our antiquity with stuff like this."

Call for e-cigarette regulation

ALEXANDRIA, VIRGINIA | Electronic cigarettes, or e-cigarettes, allow users to inhale nicotine vapors without other harmful chemicals (called "vaping"). Now, two of the largest cancer science and treatment groups in the United States have called on the government to start regulating such "electronic nicotine delivery systems" and to step up research on the health effects of vaping. A joint statement released on 8 January by the 35,000-member American Society of Clinical Oncology and the 33,000-member American Association for Cancer Research noted that "e-cigarettes may reduce smoking rates

and attendant adverse health risks" but that their actual effects are unclear in the absence of research and regulation. The statement also expressed concern that "e-cigarettes may encourage nonsmokers, particularly children, to start smoking and develop nicotine addiction." **http://scim.ag/ecigreg**

FASEB on NIH funding crunch

BETHESDA, MARYLAND | The Federation of American Societies for Experimental Biology (FASEB), which represents more than 120,000 scientists, last week weighed in on how the National Institutes of Health (NIH) can stretch its flat budget to cover more grants. In a new report, the group suggests cutting regulatory costs, encouraging sharing of large instruments, and awarding longer grants based on an investigator's track record rather than a specific project. FASEB also urges NIH to give extra scrutiny to proposals from well-funded labs "to ensure a global

distribution of research funding." Graduate programs should consider lowering admissions, and labs should rely more on staff scientists and less on trainees. http://scim.ag/FASEBreport

Brazil's climate-skeptic minister

On 1 January, a reputed climate change denier took the reins of Brazil's science ministry, causing some scientists to worry about the country's environmental future. **Aldo Rebelo**, a hard-line communist whose party supported the October reelection of President Dilma Rousseff, has called climate change an "environmental scam" and dubbed the movement to curb greenhouse gas emissions "the bridgehead of imperialism." As a member of Brazil's Congress, Rebelo also worked on a 2012 revision of the country's forest code that many experts believe will increase deforestation. "His positions on climate change are completely out of phase with the Brazilian scientific community," says Paulo Artaxo, an atmospheric physicist at the University of São Paulo who studies climate change in the Amazon. http://scim.ag/Rebeloscience

Three Q's

Eric Horvitz, managing director of Microsoft Research, launched the One Hundred Year Study on Artificial Intelligence (AI100) late last year to track the impact of artificial intelligence (AI) on aspects of life from national security to public psychology to privacy. *Science* caught up with Horvitz to discuss the ambitious project. http://scim.ag/Horvitz100yr

Q: AI100 is a continuation of a 2008 to 2009 study. Why extend it to 100 years?
A: Machine intelligence will have deep effects on people and society. A lot can happen in 100 years of technology. ... The goal is to set up an initial standing committee that will do a great job at self-sustaining and continuing a chain of standing committees and study panels over 100 years. I think it will be interesting for these panels to look back at what [earlier panels] had forecast.

Q: Do you hope to change public opinion?
A: It's not clear what public attitudes are on machine intelligence. Many people enjoy the fruits of systems like search engines without thinking that they are AI. As scientists, we need to [make] sure that concerns are addressed ... by asking questions scientifically: Are the outcomes that some people fear possible? And, if possible, how can we make sure they don't happen?

Q: What's your vision for the future of AI?
A: My view is that AI will be incredibly empowering to humanity. It will help solve problems, it will help us do better science, it promises to really help with challenges in education, health care, and hunger.
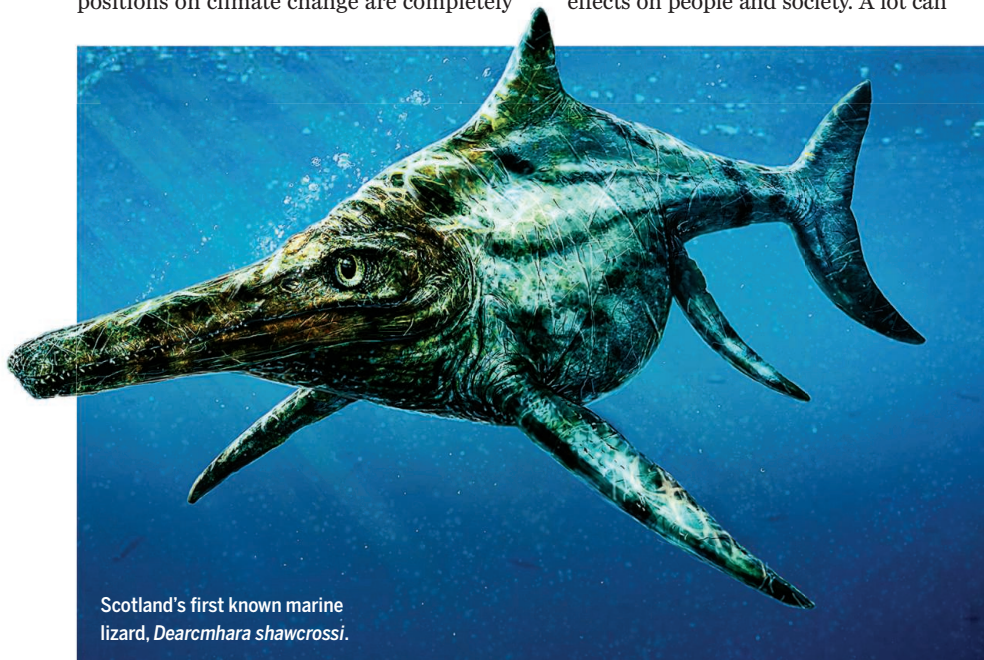
Scotland's first known marine lizard, *Dearcmhara shawcrossi*.

Scotland's first ichthyosaur

Forget the Loch Ness monster: Scotland was once home to a swimming reptile the size of a motorboat. Scientists have discovered the country's first known ichthyosaur, a large marine creature that lived during the Middle Jurassic period about 170 million years ago. The fragmentary specimen—dubbed *Dearcmhara shawcrossi* by researchers who described it online this week in the *Scottish Journal of Geology*—is named after amateur collector Brian Shawcross, who found the fossils on the shores of Scotland's rugged Isle of Skye. (*Dearcmhara*, pronounced "jark vara," is Scottish Gaelic for "marine lizard.") The ichthyosaur, pictured here in an artist's reconstruction, was about 4 meters long and hunted fish and smaller reptiles in the then-warm seas around Skye, which has some of the world's best preserved Middle Jurassic sediments.

BY THE NUMBERS

800

Number of migrating blacktip sharks within a square kilometer of ocean 500 meters off South Florida, scientists reported last week at the Society for Integrative and Comparative Biology meeting.

12

Weeks it took to clear people of hepatitis C virus (HCV) infection with a combination of three anti-HCV drugs, according to a study in *The Lancet*. Previous regimens took twice as long.

7534

Record number of patents IBM received in 2014. The company has topped the annual list of U.S. patent recipients for 22 years.

A health worker at the Ebola treatment center run by the French Red Cross society in Macenta, Guinea.

THE EBOLA EPIDEMIC

High hopes for Guinean vaccine trial

In nation with unrelenting epidemic, an unusual trial may provide key answers

By **Martin Enserink**

When Western researchers began laying the groundwork for real-world tests of Ebola vaccines late last summer, Guinea was last on their list of likely locations. It had fewer weekly cases than neighboring Sierra Leone and Liberia, and patients were scattered over a country the size of the United Kingdom with a very weak infrastructure. At an expert meeting in Geneva, Switzerland, on 23 October, the U.S. National Institutes of Health (NIH) presented a trial plan for Liberia, while the U.S. Centers for Disease Control and Prevention (CDC) offered proposals for Sierra Leone—but there was nothing for Guinea. "What about us?" the Guinean delegates asked, says biostatistician Ira Longini from the University of Florida in Gainesville, who attended the gathering. "You've left us out."

Now, in an ironic twist, many scientists say that an unusual Ebola vaccine trial set to begin next month in Guinea may produce the earliest—and perhaps the only—answers about vaccine efficacy. New cases of Ebola have dropped precipitously in Liberia, and begun to ebb in Sierra Leone, making it much harder to demonstrate whether a vaccine works—perhaps even impossible. In Guinea, however, a less intense epidemic has not relented, possibly making it "the only country where the trial will have enough power to give an answer," says Adrian Hill, a vaccine researcher at the University of Oxford in the United Kingdom.

The push to test Ebola vaccines is accelerating. Last fall, an international ethical panel convened by the World Health Organization (WHO) judged that the unprecedented threat made it acceptable to take Ebola vaccines straight from small, phase I safety studies to large phase III efficacy studies, while simultaneously conducting phase II studies. Two vaccines are now almost ready to make the leap. One, made by GlaxoSmithKline (GSK), contains a chimp adenovirus equipped with a gene that codes for the Ebola surface protein. The other, developed by NewLink Genetics but now manufactured by Merck, uses a livestock virus called VSV to carry an Ebola gene.

How to design the studies has been the subject of intense ethical and methodological debates. In the end, researchers settled on very different designs for each country. In Liberia, the NIH-led study will use a classic randomized controlled trial with three arms, one for each of the two vaccines and a third in which people receive neither. It will target the general population in the capital, Monrovia. CDC's plan for Sierra Leone is to recruit workers at Ebola treatment units; everyone will receive the vaccine, but some earlier than others. The trial has yet to select the vaccine it will test.

The welcome decline in cases in those two nations means that the trials need to include large numbers of people to determine efficacy. Although NIH now plans to enroll nearly 10,000 people in each arm of the Liberian trial, current infection rates mean it will take about 6 months to reach a statistically significant answer if the vaccine protects 50% of those who get it, says Ripley Ballou, who heads GSK's Ebola vaccine project. (An answer would come faster if protection rates are better.)

But Liberia's epidemic could end by mid-2015, suggests a model published this week in *PLOS Biology*. "There is a real risk that we won't have an answer at all at the

Planned Ebola vaccine trials

1. LIBERIA
Led by: NIH
Participants: 30,000
Design: Randomized trial with control arm in general population
Vaccines: GSK, Merck

2. GUINEA
Led by: International consortium
Participants: 9000
Design: 1. Ring vaccination trial;
2. Observational study in Ebola workers
Vaccines: To be determined

3. SIERRA LEONE
Led by: CDC
Participants: 6000
Design: Stepped-wedge trial in Ebola workers
Vaccines: To be determined
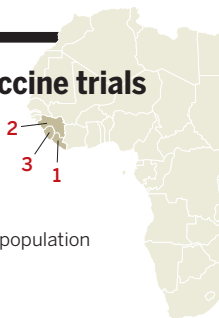
end of the day," Ballou says. That's why, in addition to counting cases and deaths, researchers should carefully study immune responses in vaccine recipients to learn as much as possible from each trial, says vaccinologist Claire-Anne Siegrist of the University of Geneva.

The Guinea study will feature an old vaccine strategy, never used before in a clinical trial, which may yield results more quickly. Inspired by the global campaign that eradicated smallpox in the 1960s, it rests on a concept called ring vaccination. Researchers will search for people newly infected with Ebola, then attempt to vaccinate a "ring" of people living around them—most likely their village or neighborhood. In half of the rings, researchers will administer the vaccine immediately to an average of 50 people; they will compare their outcomes to other rings where researchers wait between 4 and 8 weeks to administer the vaccine, says John-Arne Røttingen of the Norwegian Institute of Public Health in Oslo, who leads the study's steering group. (It was designed by a coalition including WHO, Doctors Without Borders, Longini, and several other academic partners.) Front-line health workers will also receive the vaccine, without a control group.

The vaccine for the Guinea trial has yet to be chosen, but "we will start in February, even if it is the 28th of February at 11:30 p.m.," WHO's Ana Maria Henao-Restrepo told an Ebola vaccine meeting at the Mérieux Foundation in Veyrier-du-Lac, France, this week. The study could have answers after 3 months, Røttingen says.

If one of the vaccines works and case numbers keep dwindling, a targeted ring vaccination strategy to mop up remaining virus pockets may become the best way to battle the epidemic, Røttingen says.

"It's a very interesting design from a group of very smart people," says Ballou, who adds that the logistical challenges in Guinea remain daunting. The study will have to chase cases over a vast area, and it doesn't have the operational and financial support of major players like NIH or CDC.

Meanwhile, a third vaccine candidate has entered the fray: a pair of new vaccines used in tandem. Last week, Johnson & Johnson announced that it had launched a phase I test; it is also looking for phase III study sites. Whether one of the three countries can support a second big study—and whether there will be enough new infections once the vaccines are ready for their big leap—remains to be seen. ∎

With reporting by Kai Kupferschmidt and Jon Cohen.

All in the (bigger) family

Revised arthropod tree marries crustacean and insect fields

By **Elizabeth Pennisi**, *in West Palm Beach, Florida*

The next time you are about to dig into a freshly steamed lobster for dinner, think "cockroach," or better yet, "dragonfly." A decade of genetic data and other evidence has persuaded most researchers that insects and crustaceans, long considered widely separated branches of the arthropod family tree, actually belong together. Now they are exploring the consequences of the revision, which traces insect ancestry to certain crustaceans. "When I think about traits in insects, I now have a context for where they came from," says Jon Harrison, an evolutionary physiologist at Arizona State University, Tempe, who has spent 25 years investigating insect respiration. "It's a total change."

Last week, at a special symposium of the annual meeting of the Society for Integrative and Comparative Biology (SICB) here, Harrison and many others reported new parallels between these two very successful groups of animals and new insights about what it took for an ancient aquatic crustacean to give rise to insects, which have thrived on land. Insects and crustaceans "are the most divergent organisms on the Earth, and as biologists we'd really like to understand how that came to be," says Jonathon Stillman, a marine environmental physiologist at San Francisco State University's environmental research center in Tiburon.

Because of their importance to agriculture and the use of the fruit fly *Drosophila* as a model organism, much more is known about insects than about crustaceans, such as shrimp, barnacles, and crabs. That knowledge, and the many research techniques and tools developed by insect scientists, have now become relevant to crustacean researchers. Likewise, because insects come from crustaceans, "all of a sudden it becomes quite interesting to look at specific groups of crustaceans," says Miriam Henze, a physiologist at Lund University in Sweden.

In 1926, noting similarities in brain structure, a Swedish anatomist first suggested that lobsters and cockroaches were closely related. Yet throughout the 20th century, insects were seen as closer kin to centipedes and millipedes, having many shared features in their respiratory systems, limbs, and elsewhere. A comparison of arthropod DNA in 1995 raised questions about this kinship, but most biologists were swayed only as molecular studies piled up more recently. "It's not unanimous that crustaceans and insects have won the day, but it's now the more strongly supported point of view," says Greg Edgecombe, a systematist at the Natural History Museum in London.

In the new arthropod tree, hexapods—six-limbed creatures that include insects, springtails, and silverfish—are closer kin to crabs, lobster, shrimp, and crayfish than those "standard" crustaceans are to others such as seed shrimp (ostracods). Many researchers also now put hexapods next to remipedes, shrimplike crustaceans discovered in 1979 that live in land-bound caves infiltrated with seawater. (The all-inclusive group of hexapods and crustaceans has been dubbed pancrustacea.)

The old view of insect evolution may explain why paleontologists have had so much trouble identifying the ancestral insect: They were looking for something that had insectlike and millipedelike traits, rather than a crustacean. "We've been looking from the wrong perspective," says Thorsten Burmester, a comparative physiologist at the University of Hamburg in Germany.

Even when insect and crustacean scientists have studied similar problems, their approaches differed. Some members of both groups of animals are adept at regenerating

Compound eyes from crustaceans such as mantis shrimp (*left*) and insects such as horseflies (*right*) have a common origin.

missing limbs. But insect specialists concentrated on understanding the genetic basis of restoring legs, while crustacean researchers, lacking genetic tools, have focused on its hormonal controls. "They are not talking a common language," Edgecombe says.

Now they are starting to do so, as the SICB symposium showed. And it is paying off in new insights into the water-to-land transition that a crustacean ancestor must have made on the way to becoming an insect. One big change involved respiration; another, vision.

Both insects and traditional crustaceans have compound eyes, thought to be an example of convergent evolution in the two groups. But Todd Oakley, an evolutionary biologist at the University of California, Santa Barbara, reported that he has compared genetic data on eye-related proteins and other molecules in species of insects and in ostracods. The results indicate that compound eyes evolved just once, in the common ancestor of pancrustacea.

That means compound eyes must have originated in water. There they are well adapted for the existing environment, where light scattering makes long-distance vision impossible. But "on land, you have a problem," Henze says. Indeed, insects can't distinguish much far away, but have instead evolved to efficiently spot motion—

and to use their fast reflexes if they need to evade something.

Breathing, too, had to change in the move to land. Shrimp and lobsters extract oxygen from water by actively circulating it over gills on their legs. Most terrestrial arthropods such as insects instead take in air through holes in their tough outer skin, or cuticle. As muscles connecting the segments of the insect's body contract, air is pumped through tubes called tracheae to the deepest recesses of the body.

Comparing more than 300 insect species and 300 crustaceans, Harrison has found that crustaceans have a much higher metabolic rate at rest. Living on land, with more plentiful oxygen and warmer temperatures, releases insects from the need to work so hard just to breathe, he concluded. But because of their easier access to oxygen, insects can increase their metabolic rate

10-fold when running or flying, whereas most crustaceans are limited to a threefold increase, he reported at the meeting.

A flurry of other presentations compared different classes of molecules—respiratory proteins, microRNAs, and the heat shock proteins produced in response to stress—in insects and standard crustaceans. For example, hemocyanins, oxygen-carrying molecules once known to exist only in shrimp, lobsters, and other so-called malacostracans, are also made by some insects, including the embryos of cockroaches and grasshoppers. The hemocyanin in insect embryos closely resembles the version Burmester discovered in remipedes, further strengthening the close tie between that group and insects. As land insects evolved more efficient breathing, hemocyanin became superfluous and disappeared in many groups, he hypothesizes.

The study of key developmental processes also stands to benefit from the new family tree. In insects, falling levels of so-called juvenile hormone stimulate the transition to adulthood; the analogous hormone in crustaceans is methyl farnesoate, which spurs growth and molting. Juvenile hormone is thought to be specific to insects, whereas methyl farnesoate is inactive in that group. But researchers have learned that production of both hormones depends on the same rate-limiting enzymes. And Jerome Hui of the Chinese University of Hong Kong found that in both insects and crustaceans, the same set of microRNAs control expression of the genes for those enzymes.

Even as uncertainty lingers about certain branches of the arthropod family tree, some researchers are delighted with the new ties between insects and crustaceans. Stillman predicts studying *Drosophila* may help him answer questions about heat shock proteins in the crustaceans he typically analyzes. And anyone who loves lobsters might start thinking more broadly about dinner. "We should get comfortable eating crickets," says symposium co-organizer Sherry Tamone, an endocrinologist at the University of Alaska Southeast, Juneau. "It's all one big group." ∎

Swapping branches

The current arthropod tree links insects with crustaceans, not with centipedes and millipedes (myriapods), as was once thought.

| Myriapods | Trilobites | Chelicerates | Crustaceans | Insects (and other hexapods) |

Pancrustacea

CREDITS: (PHOTO) O_LUKYANOV/ISTOCKPHOTO.COM; (ILLUSTRATIONS, FROM LEFT TO RIGHT) PHYLOPIC.COM; T. MICHAEL KEESEY (3); FABIO MERONI/MISIRLOU

BIOMEDICINE

Google[x] searches for ways to boost cancer immunotherapy

At gathering, researchers brainstorm "10x" ideas

By **Jon Cohen**, *in Mountain View, California*

Google[x], a semisecretive branch of the company that's famous for making information as accessible as possible, thinks big—and wants others to do the same. When it comes to human disease, there are few bigger challenges than cancer. So last week, the corporate headquarters known as Googleplex hosted a most unusual science gathering: nine people from Google[x], two dozen outside researchers, and others such as Sean Parker, co-founder of Napster and an infamous Silicon Valley hellion-turned-philanthropist. This diverse group's mission: determining how Google[x] can boost the suddenly hot field of cancer immunotherapy, which enlists a body's own immune system to fight cancer.

The meeting, co-sponsored by the charity Stand Up To Cancer, brought oncologists and immunologists together with specialists in imaging, engineering, and bioinformatics. All were urged to improve the field by "10x, not 10%"—the motto of Google[x], whose work has already spawned computer-equipped eyeglasses, driverless cars, and balloon-borne Internet beacons. The goal was to come up with two ideas that Google[x] could support with abundant cash, its extraordinary computing power, and its in-house expertise. "I would like to ask everybody to park their own self-interests and ask the questions that are the pressing questions for the field," said Arnold Levine, a cancer researcher at Princeton University who co-chaired the gathering.

Andy Conrad, who heads the nascent and hugely ambitious life sciences program at Google[x], did not commit any money but made it clear that "this space" greatly interests Google[x]. He described the kind of ideas his team of 100 scientists sought: "You've been dreaming about this since you're 11, and it's the right thing and it's a crazy wonderful experiment to try." From the outset, Conrad made clear that this was no stuffy advisory meeting at the U.S. National Institutes of Health (NIH). A cell biologist who previously worked at LabCorp in Los Angeles and co-

founded that company's National Genetics Institute for blood screening and clinical research, he asked participants to think "very un-NIH–like."

They were happy to do so. NIH's National Cancer Institute "hasn't recognized this area of science nearly enough," meeting co-chair Phillip Sharp, a Nobel laureate based at the Massachusetts Institute of Technology in Cambridge (and the board chair of AAAS, which publishes *Science*), said in an inter-

Google[x]'s Andy Conrad pushed his guests to think exponentially.

view. "If they don't change, there's going to be a credibility issue."

Until recently, cancer immunotherapy had a reputation for overpromising and underdelivering. Two developments have resulted in a dramatic about-face. Tumors can elude the immune system by deploying proteins that bind to particular surface receptors on immune cells, shutting down the their response. Recently developed antibodies to these receptors take off the brakes—technically referred to as a checkpoint blockade—and have helped some people with cancers such as melanoma and renal cell carcinoma whom traditional therapies have failed. Meanwhile, a strategy of treating blood cancers by removing T cells from patients, tweaking them in the lab, and then returning them has had remarkable success in several clinical trials. The engineered T cells feature chimeric antigen receptors (CARs) that destroy tumor cells by

targeting specific proteins on their surfaces.

But checkpoint blockade inhibitors and CARs have worked on only a few cancers so far and still fail in many patients. Seeking ideas to overcome those limitations, the meeting organizers invited the visiting scientists to give 15-minute talks about their own work. The research spanned everything from novel clinical trials that combine different immunotherapies to a state-of-the-art technique that borrowed a machine from geologists to do what's called multiplexed ion beam imaging, which can probe single cells and reveal huge numbers of cancer-related proteins simultaneously.

The group ultimately decided that the field's most pressing need was to ramp up the development of better tools to distinguish patients who respond to a cancer immunotherapy from those who do not. Antoni Ribas, a surgical oncologist at the University of California, Los Angeles, said that "incredible technologies," such as multiplexed ion beam imaging, could one day create movies of the molecular and cellular mechanisms that lead tumors to shrink in some patients who receive immunotherapies. "Every time that happens, if we don't study everything possible, we're missing opportunities," Ribas said.

The second chosen priority was to help medical centers more aggressively and routinely collect tumors from those living with the disease and from cancer patients who die. "Nothing would change the field faster" than having better access to such tumors, said Irving Weissman, a leading stem cell biologist at Stanford University in Palo Alto, California. "We can begin using these wonderful microscopic tools to say what is in the tumor early, mid and late, and what happens when you do the perturbation of your clinical trial."

Not everyone at the meeting was convinced that these were 10x ideas. "A lot of what I've heard, and even what we're prioritizing now, is still not the big, bold thinking of what can happen a decade out or 2 decades out to solve these problems," said Sanjiv Sam Gambhir, a Stanford materials scientist who already collaborates with Google[x]. Conrad agreed with Gambhir that getting more tumor samples was not a high-risk proposition, but he had a different takeaway. "We don't need to get a new principle of physics, we don't have to invent, we just have to make people do the right thing," he said. "And it may take money to make people do the right thing, and I wonder if that doesn't supply the material for the big dream." ∎

PHOTO: GOOGLE

Members of Sahar Al Bayatti's team collecting data for a livestock genetic diversity study.

IRAQ

Throwing a lifeline to a one-time Arab science power

Iraqi scientists gain expertise from U.S. fellowships

By **Richard Stone**

In Baghdad, scarred from a decade of war and anxious about the advance of Islamic State forces in northern Iraq, a modest symbol of hope is taking shape: a $5 million facility for storing sperm and other tissue from the country's livestock breeds. Iraq's government, which deems the livestock gene bank vital to food security, is providing the funding. But the expertise behind the project owes much to a small U.S. State Department effort called the Iraq Science Fellowship Program (ISFP), which enabled Sahar Al Bayatti, Iraq's national coordinator for animal genetic resources, to spend 6 months in 2012 learning the ropes at the U.S. National Animal Germplasm Program in Fort Collins, Colorado. Al Bayatti, who is overseeing the gene bank's creation, calls her chance for overseas training "a miracle."

Now the U.S. government is stepping up such efforts, which aim to restore the intellectual capital that Iraq will need to rebuild. The State Department program, launched in 2008, has already trained 52 Iraqi scientists in various specialties. This month it will be joined by a second fellowship program, funded by the U.S. Defense Threat Reduction Agency (DTRA), to bring several scientists a year to the United States to train in top labs. The DTRA program will mainly target biologists in dual-use areas—those whose research could be used to create new toxins or biological weapons—and will emphasize biosecurity and threat reduction. The State Department is now planning to focus ISFP on chemistry and other fields not covered by the DTRA program.

"This kind of engagement is so powerful. I totally support it," says David Franz, a veterinary scientist and former commander of the U.S. Army Medical Research Institute for Infectious Diseases who has been to Iraq on several occasions, most recently for a biotech conference in Baghdad in 2013.

Iraq was a rising science power in the Arab world until 1980, when war broke out with Iran and research became tightly yoked to military aims. The years of sanctions and isolation that followed the 1990 to 1991 Gulf War took a heavy toll, spurring a massive exodus of scientific talent, and the research infrastructure was hammered anew during the U.S.-led invasion to topple Saddam Hussein in 2003. Iraq's science "truly collapsed almost to the zero line," Sandia National Laboratories and the Arab Science and Technology Foundation concluded in a 2004 report.

Years of sectarian violence, when hundreds of scientists, professors, and doctors were murdered or kidnapped, only deepened the plight of researchers. At first, the United States and its allies focused on lining up peaceful work for former weapons scientists, but later the coalition partners looked for ways to throw lifelines to the broader scientific community. "You have a lot of researchers who want to reengage with the rest of the world," says Galaal Elsamadicy, a project manager at CRDF Global, a nonprofit in Arlington, Virginia, that administers the fellowship schemes and set up a virtual science library that gives 80,000 Iraqi scientists and graduate students free access to thousands of journals.

Fellows give the ISFP program high marks. As head of molecular diagnostics at Iraq's Central Veterinary Diagnostic and Research Laboratory in Baghdad, Layth Abdulrasool says one of his main tasks is keeping tabs on pestiviruses, a genus of viruses that attack livestock, causing swine fever, bovine viral diarrhea, and border disease virus. "We don't have much information on the real percentage of livestock infected with these viruses in Iraq, and we're missing some techniques that would help us," he says. Last month Abdulrasool wrapped up an ISFP stint at the U.S. Agricultural Research Service's National Animal Disease Center in Ames, Iowa, where he worked with microbiologist Julia Ridpath to design a pestivirus surveillance program for Iraq.

"The fellowship programs are attracting the best of the best," says Gavin Macgregor-Skinner, a veterinary scientist at Pennsylvania State University, Hershey, who consults for the U.S. government on overseas biological threats and is advising DTRA on its new fellowship scheme. Macgregor-Skinner says he has no doubt that Abdulrasool and other fellows soak up a lot of knowledge in U.S. labs. But he also feels the U.S. government should think more ambitiously and dispatch U.S. researchers to Iraq, where they can tailor training to local conditions. He notes that some skill sets simply don't transfer back to Iraq, where easy access to reagents and instruments is a pipe dream. "We should mentor them in-country and bring them to the United States for conferences," he says. That way, instead of helping a handful of Iraqi scientists a year tool up, the U.S. government could train hundreds, Macgregor-Skinner explains. Franz agrees, though he notes it would be a challenge to persuade U.S. researchers to venture to Iraq while sectarian violence and the Islamic State group are potent threats. "The major issue obviously is security. Imagine if you had to change your route to work every day?"

Even Al Bayatti, who says she has grown inured to the risks, concedes that "the unrest has slowed our work." The Islamic State group advance in northern Iraq has crippled efforts to collect livestock germplasm. But she vows that the gene bank will be up and running next year. ∎

SCIENCE COMMUNICATION

Backlash greets 'bad luck' cancer study and coverage

How subtleties got lost in the telling

By **Jennifer Couzin-Frankel**

On New Year's Day, some striking news broke in *Science*: a piece by this reporter called "The bad luck of cancer" and the research study it was based on, published simultaneously in the journal. As *Science*'s story and many other news accounts of the paper explained, the study authors concluded that simple "bad luck"—their description of random mutations accumulating in healthy stem cells—could explain about two-thirds of cancers. That would exceed the risk of environmental and genetic factors combined.

Readers wasted little time in skewering the authors, mathematician Cristian Tomasetti and cancer geneticist Bert Vogelstein of Johns Hopkins University in Baltimore, Maryland. "Seems some cancer researchers are simply running out of intelligent questions or arguments," read one of the 210 online comments on *Science*'s news piece. Reporters, including this one, fared worse. "Please, journalists, get a clue before you write about science," pleaded an irate column in *The Guardian* by an evolutionary biologist and a statistician. Critics charged that media stories misinterpreted the study, seeking a streamlined message that downplayed the value of cancer prevention.

The furor exposed the challenges that come with communicating the science of risk, especially in charged areas such as cancer, and the desire by both scientist-authors and reporters to simplify a story. Many news accounts, including *Science*'s, glossed over the fact that the study didn't include all cancers. In fact, it omitted two of the most common, prostate and breast, because data needed to include them were lacking.

Still, was the "two-thirds" figure actually referring to a fraction of cancer cases? Tomasetti had explained to *Science* that "if you go to the American Cancer Society website and you check what are the causes of cancer, you will find a list of either inherited or environmental things. We are saying two-thirds is neither of them." He also confirmed the news story's language describing the study before it was published.

In a follow-up interview after the uproar, Tomasetti clarified that the study argued that bad luck explained two-thirds of the *variation* in cancer rates in different tissues—a subtly different claim. Some tissues are overtaken by cancer more readily than others, and mutations in stem cells explain two-thirds of that variability, Tomasetti and Vogelstein concluded.

Despite the confusion among reporters, Tomasetti did not feel they had been careless—quite the contrary. "Overall, the reporters who interacted with us made a very honest and sincere effort to be as accurate as possible," he says. And, he believes, he did his best to convey his findings to nonexperts. "If given enough time, or space, I can explain the subtleties of any given scientific result to anyone really," but there were only so many hours he could spend speaking with reporters on deadline. The material is complicated even for statistical gurus, he believes. He has been busy preparing a technical report with

additional details, and Johns Hopkins also sent a follow-up explainer to journalists and posted it online.

"It's too easy to blame the media," says David Spiegelhalter, a biostatistician at the University of Cambridge in the United Kingdom who blogs at Understanding Uncertainty. In this case, he felt, "the gist of the coverage is very reasonable—most cases of cancer are due to chance."

Many scientists felt the paper's authors were also guilty of oversimplifying. The paper included a striking figure splitting cancers into green and blue categories. The green were cancers "mainly due" to random mutations—suggesting, the authors wrote, that they were less likely to bow to prevention. However, that category included esophageal cancer and melanoma, both thought to have strong environmental drivers—heavy alcohol consumption for esophageal cancer, for example, and sun exposure for melanoma. Melanoma sat just inside the border of green country—but that was enough to incense some readers. "They've ignored some of the fundamental lifestyle factors," said Graham Colditz, a cancer prevention specialist at Washington University in St. Louis. Vogelstein says his paper doesn't dispute that the environment contributes to melanoma. "This is a mathematical theory," he says, and it doesn't explain every facet of every cancer it includes.

Anne McTiernan, a physician and epidemiologist at the Fred Hutchinson Cancer Research Center in Seattle, Washington, believes the authors had "good intentions," but she criticizes their assumption that a correlation between stem cell divisions and cancer risk means that one causes the other, something their data didn't prove. Tomasetti agrees, but he notes that "all the biology we have on this topic supports causality."

A key unanswered question is whether the furor will dampen cancer prevention efforts. "The message shouldn't be, 'Oh, it's all chance, there's nothing we can do about it,' " says Timothy Rebbeck, a cancer prevention specialist at the University of Pennsylvania. "There should be something we can do about it," because risk varies so much among individuals. If anything, he says, the study points to the value of prevention. It shows huge risk gaps between cancers driven by the environment or genetics—such as lung cancer in smokers—and cancers at the same site without a clear cause.

Spiegelhalter isn't surprised that coverage of the paper had its flaws. "It's one of those things that's so superficially simple," he says, "and yet the superficial simplicity is not correct." ∎

A recent study slighted cancer prevention efforts, such as wearing sunscreen—or so critics charged.

SOLAR CELLS

Devices team up to boost solar power

Hot new solar cell materials are poised to rev up the output of traditional photovoltaics

By **Robert F. Service**

The silicon solar cells spreading across rooftops and deserts are only middling when it comes to turning sunlight into electricity. Today's efficiency champions are complex devices called tandems: stacks of solar cells, each optimized to absorb a different part of the solar spectrum. That complexity—and the pricey gallium arsenide–based semiconductors from which tandems are made—makes them expensive, so they're used mostly in space, where the extra power is worth the premium. But now, solar cell researchers are beginning to apply the tandem strategy to the hottest and cheapest new solar cell material out there, crystals called perovskites, igniting new hopes for high-efficiency, low-cost cells.

"It's becoming a major undertaking," says Michael Grätzel, a chemist at the Swiss Federal Institute of Technology in Lausanne. For now, researchers are still struggling to marry different solar cell materials that have never been paired before. But recent successes have encouraged them. "I really believe high-efficiency tandems will be made in the next year or two," says Michael McGehee, a materials scientist at Stanford University in Palo Alto, California. But Grätzel adds a caveat: "Whether it ends up being worth it in cost is another question."

The wave of enthusiasm began last October, when researchers at IBM's Thomas J. Watson Research Center in Yorktown Heights, New York, showed how to grow a perovskite cell atop another cell made from a blend of copper, zinc, tin, sulfur, and selenium (CZTSSe). That tandem, described in *Applied Physics Letters*, was only 4.6% efficient—far below the best CZTSSe cell or the best perovskite cells by themselves. A big part of the problem was that a metal electrical contact at the top of the stack blocked most of the light from passing through to the lower layers.

Now, a group led by McGehee and Grätzel has largely solved that problem. In a paper posted online in *Energy & Environmental Science* on 23 December 2014, the researchers reported that by combining perovskite cells with either a copper-based cell abbreviated CIGS or a more standard silicon cell, they had created devices with an overall efficiency as high as 18.6%—beating both the

17% efficiency of their plain CIGS cells and the 11.7% of the perovskite cell used in the test. One key to its success was a top metal contact made of silver nanowires, which let most of the photons pass through to the light-absorbing layers beneath.

McGehee and others are confident they can do better. Last month, Nam-Gyu Park, a chemical engineer at Sungkyunkwan University in Suwon, South Korea, told a meeting of the Materials Research Society in Boston that he and colleagues had made a perovskite-silicon tandem solar cell with 28% efficiency. Instead of layering two cells one atop the other, as in a standard tandem cell, Park's team combined a high-efficiency silicon cell lying flat with a perovskite cell standing upright. They then used a device called an optical splitter to steer the higher energy bluish photons to the perovskite cell

while allowing the lower energy reddish photons to pass through the splitter to the silicon cell below. "It's not easy to apply" to making commercial solar cells, Grätzel says, but "it shows the potential if you split the spectrum."

That potential continues to grow, primarily because researchers continue to improve the recipe for perovskites. A group led by Sang Il Seok of the Korea Research Institute of Chemical Technology in Daejeon, South Korea, for example, reported online

last week in *Nature* that they've made a perovskite cell that's more than 18% efficient by itself. In unpublished research, independently verified at the National Renewable Energy Laboratory in Golden, Colorado, the group built a perovskite cell with an efficiency of 20.1%.

Grätzel says that by combining perovskites boasting that kind of efficiency with the top commercial silicon solar cells, researchers should be able to reach Park's 28% with the customary stacked tandem structure. That would beat the best silicon cells by 20%—and Grätzel says that "in this field, people get excited if you get a 0.1% improvement."

To maximize the overall power output of the tandems, researchers still need to increase the current generated by the perovskite cells to match the current from

the silicon cells. (A tandem's current is limited by the weaker of its component devices.) And they need to reduce electrical losses that occur at the interface of the two materials, says Yang Yang, a materials scientist at the University of California, Los Angeles. Decades of painstaking work have taught researchers how to do that with gallium arsenide cells used in space. Solving those same riddles with perovskites could bring tandems down to Earth. Says Yang: "We are confident we can do it." ∎

Two ways to make a tandem

In tandem solar cells, different layers absorb different parts of the solar spectrum. "Four-terminal" cells (left) stack two separate solar cells; "two-terminal" devices (right) layer the two absorbers into a single device. New materials may replace expensive gallium arsenide.

Four-terminal

Light

Glass
Transparent electrode
Top cell (perovskite)
Silver nanowire electrode
Transparent electrode (CIGS)
Si emitter + metal grid (Si)
Bottom cell (Si or CIGS)
Metal electrode

Two-terminal

Light

Transparent electrode
Top cell (perovskite)
Tunnel junction/
recombination layer
Bottom cell (Si or CIGS)
Metal electrode

THE CANCER STEM CELL
GAMBLE

Researchers are betting that a round of clinical trials will prove a controversial cancer theory and deliver new treatments

By Jocelyn Kaiser

Robert Weinberg is one of the world's best known cancer biologists, thanks largely to his pioneering work identifying genes that underlie tumor development. He has seen hopes for cancer treatments come and go. "I've been in this business for better or worse for 40 years. Many of the things that we've worked on have proved to be relatively useless in the clinic." But at 72, he is optimistic again. "This is really the first time where I'm positioned to help effect the development of an agent or agents that actually will benefit cancer patients," he says.

The Massachusetts Institute of Technology researcher is now staking part of his considerable reputation, and nearly $200 million that investors have given to a company he co-founded, on a bold theory that has divided the cancer field. Weinberg and others contend that tumors contain a small number of cells that are distinctive because they resemble the stem cells that give rise to normal tissues. They believe that these cancer seeds, able to resist chemotherapy and spring back months or years after treatment, may explain the tragic relapses people often experience. Target these cancer stem cells specifically, the thinking goes, and the disease can be kept under control.

Verastem Inc., Weinberg's company in Needham, Massachusetts, is one of several that are launching a new round of clinical trials to find out whether the theory actually works. Beyond the promise of changing cancer care, the financial stakes are huge. OncoMed Pharmaceuticals Inc., another leader in this area, could win $5 billion in additional funding from major drug companies if its trials succeed.

But as Weinberg and others in the field acknowledge, it may be difficult to draw definitive conclusions from these trials. Unlike traditional chemotherapy, the drugs undergoing testing are not expected to quickly shrink tumors, because they are designed to kill just the tiny subset of cells that seed and resupply the main tumor. So detecting whether the drugs are working
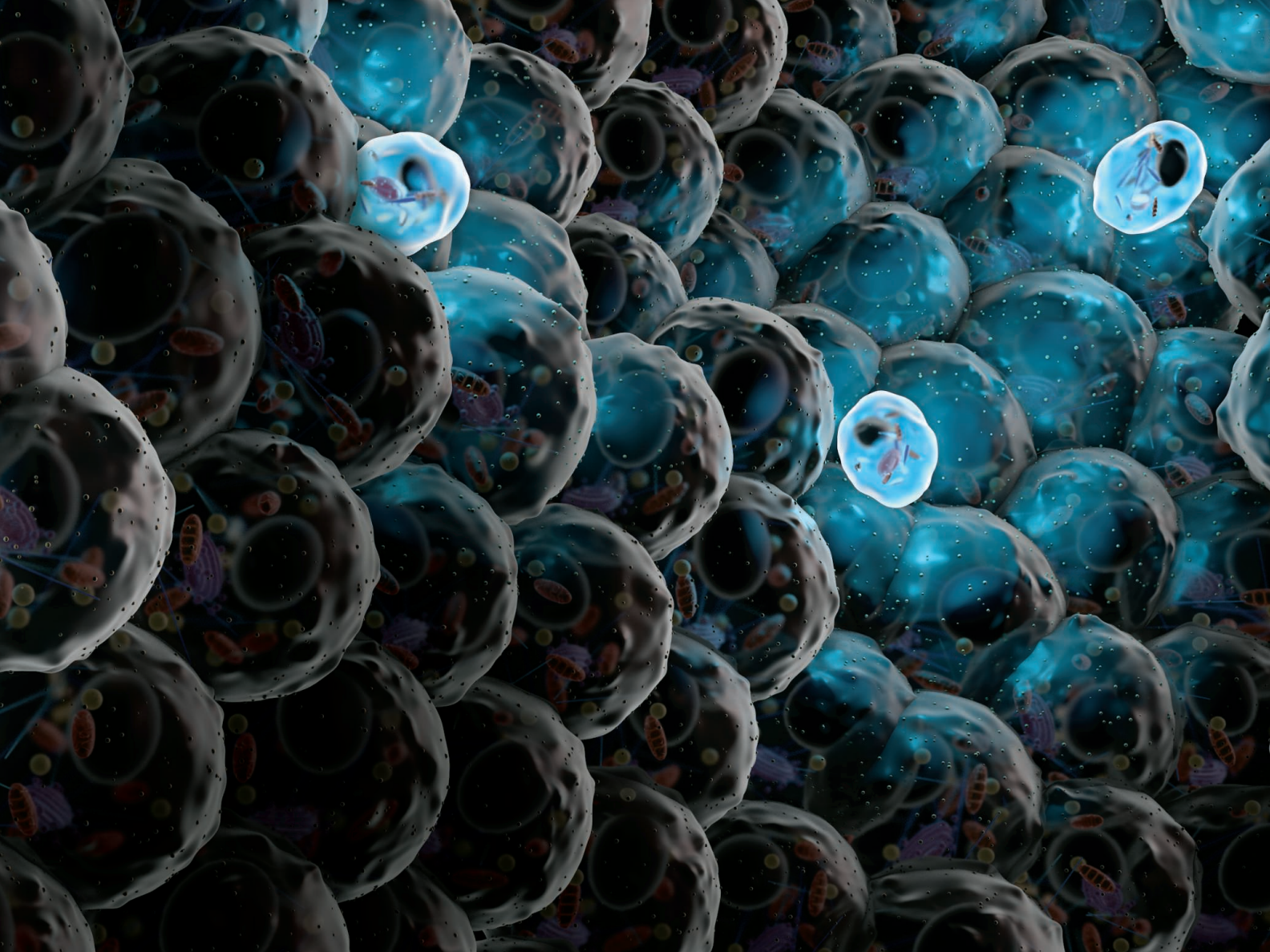
in the intended way is not straightforward. Indeed, for solid tumors, researchers lack simple, rigorous assays for measuring the number of cancer stem cells.

The efforts also face some fundamental skepticism: Many still don't believe cancer stem cells exist as a cell type distinct from other tumor cells, and some suggest that companies are hyping or at least over-simplifying the premise. A win in the clinic could resolve some of the controversy. "I think the onus is on all of us in the community that's developing cancer stem cell therapies to show beyond a doubt that these therapies really work," says Max Wicha of the University of Michigan, Ann Arbor.

THE CANCER STEM CELL model emerged in the mid-1990s, when stem cell biologist John Dick of the University of Toronto reported that his team had isolated rare cells in the blood of people with leukemia that seemed to play a key role in the cancer. Although such patients' blood teems with aberrant white blood cells, only a few of them were capable of growing into a new leukemia when injected into mice. Those cells appeared to be misguided versions of the normal adult blood stem cells that differentiate into mature blood cells. Like normal stem cells, the cancer stem cells carried distinctive surface proteins and were self-renewing: They could divide to produce both a regular cancer cell and a new stem cell.

Other teams subsequently reported finding apparent cancer stem cells in solid tumors such as breast, colon, brain, and pancreas. Dick and others suggested that these cells, making up perhaps 1% to 3% of most solid tumors, evade chemotherapy and radiation, partly because most treatments selectively kill rapidly dividing cells, and cancer stem cells grow more slowly than other malignant cells. After lying low, these cancer stem cells could later regenerate the original tumor or spawn metastases in other organs (*Science*, 24 August 2007, p. 1029).

There were problems with these studies, however. Scientists usually picked out the cancer stem cells within a solid tumor by isolating cells with certain surface proteins, thought to be markers of stem cells. But it turned out that not all tumor-generating cells carried these markers, and other cells making up the bulk of the tumor sometimes did as well. Relying on markers "will fool you," says oncologist and cancer stem cell researcher William Matsui of Johns Hopkins University in Baltimore, Maryland.

Still, excitement about cancer stem cells inspired the U.S. National Cancer Institute and major companies starting in the late 2000s to launch small-scale safety trials of drugs aimed at signaling pathways active in stem cells. The pathways targeted were those controlled by the gene called Sonic

>60
Ongoing or planned cancer stem cell clinical trials

200
million
Dollars invested so far in Verastem

hedgehog and genes belonging to the Notch family, best known for shaping embryonic development. This first wave of trials, now winding up, has proved disappointing. Often the drugs caused serious side effects, most likely because they harmed normal stem cells, such as those needed to regenerate the gut lining, Matsui says. Even when these side effects could be managed, for most cancers the drugs showed no signs of efficacy in larger studies.

One hedgehog inhibitor did make it onto the market: vismodegib (Erivedge). But it is approved only for basal cell carcinoma, the most common skin cancer, in which the hedgehog signaling pathway goes awry in all of the malignant cells. As a result, it is unclear if the drug works by actually homing in on rare cancer stem cells.

Despite the early failures, a few compounds have shown enough promise in initial testing that their developers are moving them into larger efficacy trials. Wicha, an OncoMed co-founder and consultant to many companies targeting cancer stem cells, counts more than 60 ongoing trials at universities and companies, testing either antibodies or small molecules aimed at cancer stem cells. Most are designed primarily to gauge safety, while also looking for signs of efficacy. But a few are more advanced phase II trials in which patients are randomly assigned to receive either the experimental drug or a conventional drug or placebo—the classic way to determine whether a new drug is effective.

OncoMed's tarextumab, a monoclonal antibody targeted at proteins in the Notch pathway, is among those furthest in the pipeline. In a safety study that combined tarextumab with two conventional drugs for pancreatic cancer—a disease in which traditional chemotherapy rarely helps—83% of 29 patients' tumors were stable or shrank over periods of from 8 weeks to about a year, the company reported in 2014 at meetings. OncoMed last year began larger, phase II trials for tarextumab in pancreatic and lung cancer. The company says tarextumab and other drugs it is developing seem to work not by killing cancer stem cells, but by nudging them to differentiate into bulk tumor cells that get wiped out by the chemotherapy.

> *"I think the onus is on all of us in the community that's developing cancer stem cell therapies to show beyond a doubt that these therapies really work."*
>
> **Max Wicha**, University of Michigan, Ann Arbor

Verastem's strategy is to screen approved drugs and other chemicals for their ability to block focal adhesion kinase (FAK), an enzyme that helps tumor cells stick to each other and also helps cancer stem cells survive. In the body, Weinberg believes, blocking FAK kills cancer stem cells directly and also makes it harder for these rare cells within a primary tumor to travel through the bloodstream and seed metastases.

The company's first candidate FAK inhibitor, a livestock antibiotic, did not pan out in the clinic. But according to before-and-after biopsies in 10 mesothelioma patients, who were given another FAK inhibitor called defactinib daily for nearly 2 weeks before having their tumor surgically removed, that drug seemed to knock down the portion of tumor cells carrying a specific stem cell marker. None of the patients' tumors grew in this safety study and in two cases they unexpectedly shrank, for reasons the company is still exploring, Verastem says. Defactinib is now being tested in phase II trials in people with lung cancer and in mesothelioma patients who responded to traditional chemo to prevent the cancer from recurring.

Reparixin, a drug initially developed by the Italian company Dompé to fight transplant rejection, seems to kill cancer stem cells by blocking a receptor that triggers their growth in response to inflammation, Wicha's team has reported. The receptor, which binds inflammatory molecules called cytokines, is found on the surfaces of the stem cells but not on other cells in the bulk tumor. After showing hints of efficacy against metastatic breast cancer in a safety trial, reparixin is now in a phase II trial in which it is given to women diagnosed with breast cancer for 3 weeks before surgery to see if it knocks down cancer stem cells in their tumors.

SOME ARE NOT OPTIMISTIC about the new drugs. Harvard University cancer biologist William Kaelin, a prominent skeptic of the cancer stem cell hypothesis, says that even if these cells exist as a small, distinct population in solid tumors—he's not convinced—tumors can resist chemotherapy in many ways. It's misleading to suggest that "if you kill the cancer stem cells, your work is done," he cautions.

Even if the trial results are encouraging, it won't be easy to tease out the effects of the experimental drug. For one thing, many trials combine a drug targeting can-

A cancer hypothesis on trial

Some efficacy trials of drugs aimed at cancer stem cells, often combined with conventional tumor treatments.

COMPANY	DRUG	TARGET	CANCER	STAGE	COMBINATION
OncoMed	Tarextumab	Notch 2,3 receptors	Pancreatic, lung	Phase II	Yes
	Demcizumab	DLL4 (Notch ligand)	Ovarian	Phase II, mid-2015	Yes
Verastem	VS-6063	Focal adhesion kinase	Mesothelioma, lung	Phase II	No
Boston Biomedical (Sumitomo Dainippon)	BBI608	STAT3, ß-catenin, Nanog	Colon	Phase III halted*	No
			Gastric, esophageal	Phase III	Yes
			Colon, other cancers	Phase II	Yes
	BBI503	Multiple kinases	Solid tumors	Phase II	No
Stemline Therapeutics	SL-401	Interleukin-3 receptor	Leukemia	Phase I/II	No
Dompé	Reparixin	Chemokine receptors 1 and 2	Breast	Phase II	No

*Failed to meet efficacy endpoint

cer stem cells with more conventional cancer therapy. The true test would be to give only a drug that kills the stem cells, says stem cell biologist Irving Weissman of Stanford University in Palo Alto, California. If cancer stem cells exist as proponents envision, then eliminating them should eventually wipe out a tumor because the bulk tumor cells cannot divide indefinitely and will eventually die, he explains.

But giving people only a cancer stem cell drug is not practical because it would be so slow to take effect—perhaps taking many weeks or months to wipe out a human tumor, says developmental biologist Michael Lewis of Baylor College of Medicine in Houston, Texas, who suggests this is why some early trials found no efficacy. What's more, ordinary cancer cells may sometimes "revert" to cancer stem cells, work in Weinberg's and other labs suggests. If so, conventional cancer drugs may also help kill cancer stem cells—and disentangling the effects of stem cell versus conventional treatments could be even more difficult.

A sign that the new drugs are working as advertised would be a drop in the fraction of stem cells within a person's tumor. But monitoring tumors' stem cell content isn't easy. Serial biopsies of the tumor are needed to track any change, which requires cooperating oncologists, consent from the patient to an invasive procedure, and a means of storing and shipping the tumor samples for analysis.

How to quantify the cancer stem cells in a biopsy is another puzzle. Just counting the number of cells with a particular stem cell marker, such as a surface protein, isn't enough for many in the field. The gold standard for proving that putative cancer stem cells are actually that is to inject them at various doses into immune-deficient mice and see if they form tumors. But these "lim-

Taking aim at the seeds of cancer

According to the cancer stem cell model, chemotherapy kills bulk tumor cells but leaves rare, stem cell–like tumor cells untouched (left). These cells then seed the tumor's regrowth. Killing the stem cells may lead to tumor regression over time (center), but combining a cancer stem cell drug with chemotherapy could be faster (right).

Conventional cancer therapy **Cancer stem cell–specific therapy** **Combination cancer therapy**

Tumor relapse **Slow regression** **Eradication**
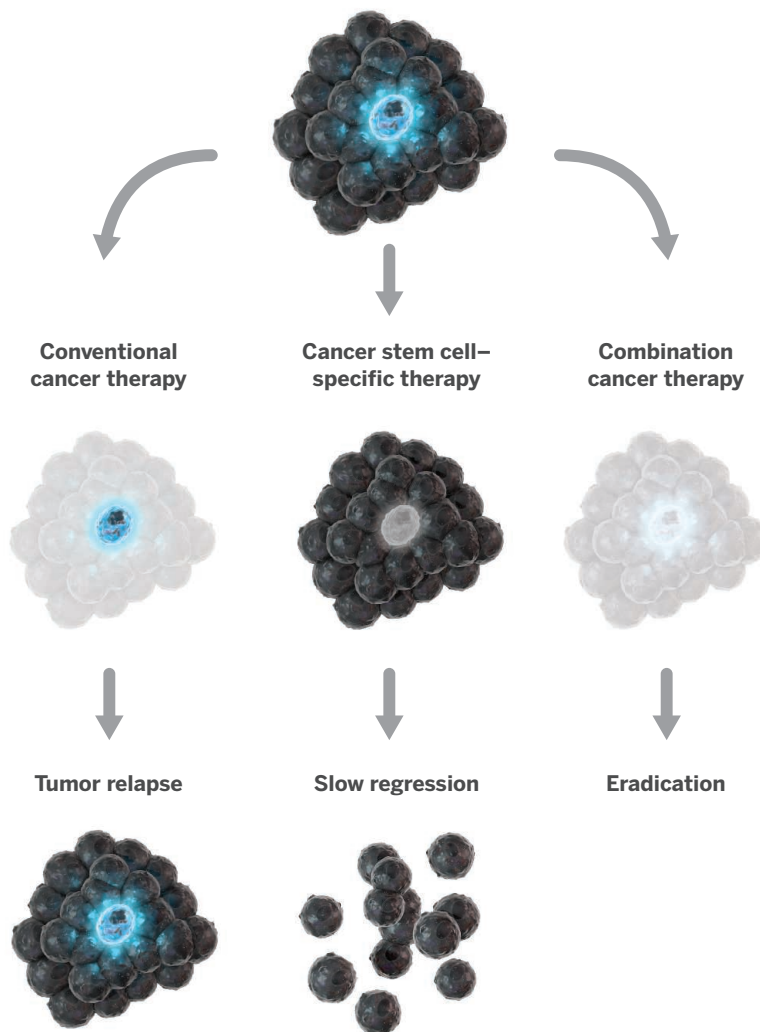
ited dilution" tests require four or five cell doses, several groups of mice, and at least a half-dozen mice per group. "It's time-consuming and expensive," notes stem cell researcher Mick Bhatia of McMaster University in Hamilton, Canada.

Joanna Horobin, chief medical officer for Verastem, says that for "pragmatic" reasons her company isn't trying to measure cancer stem cells with a battery of tests in its trials. Instead, its approach is to answer "different questions in different studies, then put the whole thing together." For example, in a study of mesothelioma, the company showed that chemotherapy alone correlated with an increase in stem cell markers in patients' remaining tumor tis-

sue, as would be expected if those cells resist treatment. Together with the drop in stem cells seen in mesothelioma patients who received a Verastem drug shortly before surgery, "it starts to create the case," Horobin says.

Timothy Hoey, senior vice president for cancer biology at OncoMed in Redwood City, California, notes that the company does have some limiting dilution data on patient samples suggesting that its drugs are hitting cancer stem cells. But the assay OncoMed relies on most is a gene expression signature that it believes indicates how much of a tumor sample is made up of stem cells. Looking at changes in this pattern in a tumor before and after a person receives treatment "connects the preclinical studies to the clinical studies," Hoey says. He notes that these studies have been presented at meetings but the results, including the so-far undisclosed gene signature, still need to be published in a research journal. "I think it's important for the field," he says.

Still in development are tests that will examine the cells released into the bloodstream by solid tumors. Wicha says his team is working closely with engineers and genomics experts on the University of Michigan campus to develop assays that can reliably identify circulating cells with stem cell–like properties.

For now, cancer patients, researchers and physicians, and investors in companies such as Verastem will anxiously wait for data to roll in from the clinical trials. For those with a stake in treatments, the results could bring hope. For researchers debating the reality of cancer stem cells, though, they may not bring resolution. Says Jeremy Rich of the Cleveland Clinic in Ohio, who is studying stem cells in brain cancer, "Even if we're wildly successful, which I don't think we will be, I don't think there will be a black-and-white answer." ∎

Precision of interfering atoms *p. 232*

How views of ability affect gender balance *p. 234*
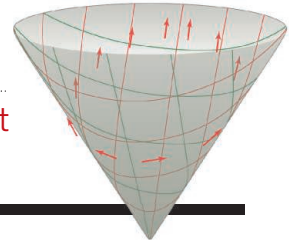
GEOLOGY

The birth of the geological map

Two hundred years ago, William Smith published the first geological map of a country

By **Tom Sharpe**

In 1815, William Smith, an English canal surveyor and land drainer, provided the young science of geology with the first true geological map of an entire country (see the first figure). Two hundred years on, Smith's map has become an icon of Earth science, and the basic principles he developed and applied are still used in interpreting rock sequences and making geological maps.

Smith's map is remarkable for many reasons. It was ambitious in its scale and scope, covering the whole of England, Wales, and southern Scotland, an area of more than 175,000 km² (larger than Washington state). It was constructed by the application of Smith's own discovery that the strata of southern Britain are arranged in a regular sequence and that each rock layer contains distinctive and diagnostic fossils. It was big, measuring about 2.6 m by 1.8 m. It was colorful, with each color carefully chosen for its similarity to that of the rocks it represented. And it was the work of a single individual with a limited rural education, the son of a village blacksmith, working independently outside of the developing structure of the recently formed (and gentlemanly) Geological Society of London.

Smith's was not the first map to show the distribution of different rock types; such "mineralogical" maps had been around for some time (*1*). But it was the first to show the rocks of a whole country in a way that indicated the sequence of the strata and with a key in an informative, stratigraphic

Lyme Regis Museum, Bridge Street, Lyme Regis, Dorset DT7 3QA, UK, and Department of Lifelong Learning, Cardiff University, Senghennydd Road, Cardiff CF24 4AG, UK. E-mail: sharpet@cardiff.ac.uk
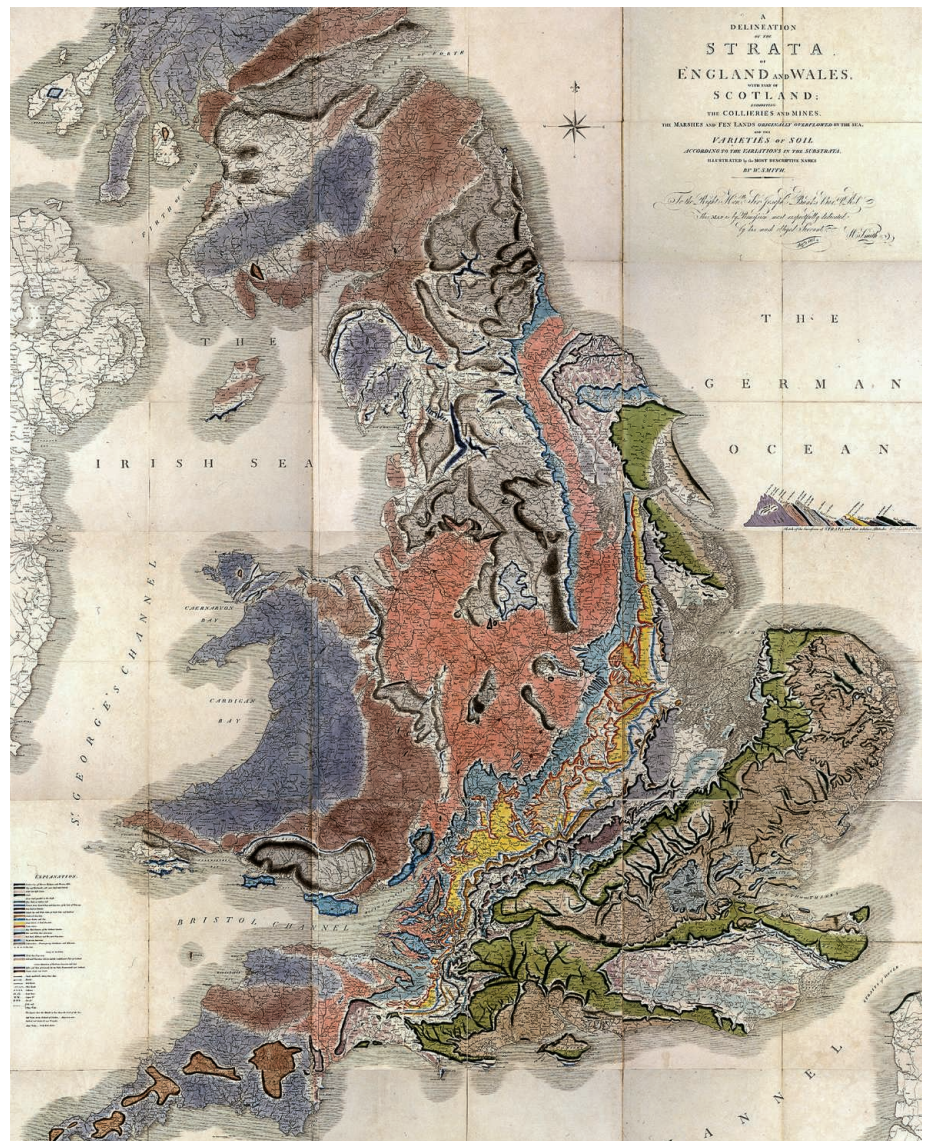
William Smith's 1815 map. The dark tones indicating the base of each stratum are clearly seen in southern and eastern England.

order, from oldest rocks to youngest. The real innovation, however, was the clever (although expensive) use of a darker watercolor tone to indicate the base of each stratum, fading out toward the top of the bed to give an impression of three dimensions. At a glance, not only is the sequence of tilted beds, layered one on top of the other, readily apparent, but the viewer can immediately envisage how the beds continue underground as they dip below overlying strata.

Smith's great map, *A Delineation of the Strata of England and Wales with Part of Scotland*, was first exhibited in London in early 1815 (*2*). Its subtitle—*Exhibiting the Collieries and Mines, the Marshes and Fen Lands Originally Overflowed by the Sea, and the Varieties of Soil According to the Variations in the Substrata*— reflected where Smith saw the value of his map: It was intended to be a practical tool for mineral exploration, land drainage, and agriculture.

Smith had begun his career as a trainee surveyor at a time of land enclosure to create large private estates, before being sent to survey coal mines in Somerset and working as surveyor for the route of the proposed Somerset Coal Canal near Bath in the west of England (*3*). During the course of this work, Smith noticed that the two branches of the canal in adjacent valleys passed through the same sequence of strata and that the different rocks were always in the same order and arrangement, tilted gently toward the southeast. In addition, he recognized that each bed of rock contained distinctive fossils. This discovery of the value of fossils in correlating strata was a fundamental breakthrough for the new science of geology. In continental Europe, unknown to Smith, others had been developing the same idea, mainly in small areas (*1*), but it was Smith who recognized that the principle of diagnostic fossils could be applied over a wide area.

Smith's discovery had immediate practical application. These were the early days of the Industrial Revolution in Britain, when coal was in huge demand. Land owners could make a fortune from coal on their properties, but many searching for it were being misled: Dark gray mudstones similar to those associated with coal rocks also occur far above and below the coal beds in much older and much younger strata. Fossils allowed Smith to distinguish between these different mudstone sequences and to say with confidence whether coal was likely to be found (*4*).

By 1799, Smith was in a position to describe a sequence of 23 strata around Bath and to draw up a rudimentary geological map of the area. Working as an itinerant surveyor and drainage engineer, Smith was able to validate his ideas on the regularity of strata and the use of fossils as he traveled the length and breadth of England and Wales. Within 2 years, he was in a position to draft outline geological maps, tracing some of the strata across England to the east coast. Much of the geology of southern and eastern England is a southeasterly tilted sequence of fossil-bearing sedimentary rocks, lacking the complications of igneous intrusion, severe folding, or metamorphism found in much of western and northern Britain. Nonetheless, the preparation of his great map was to take another 14 years.

Smith's map and his ideas on strata and the use of fossils did not receive universal acclaim from his contemporaries. George Bellas Greenough and the gentlemen of the Geological Society were unconvinced by Smith's approach (*5*) and, by the time Smith's map was published, had begun work on their own map of the geology of England and Wales. This collaboration by the Society's members around the country soon ran into difficulties, which were resolved—on the recommendation of John Farey, one of Smith's pupils and friends—by an application of Smithian principles. Smith's ideas worked.

Smith's map evolved during the years of its production and was modified as new

Geological map of Mars. The colors represent different rock units of the martian stratigraphy of the Noachian, Hesperian, and Amazonian Periods, based, in the absence fossils, on stratigraphical relationships and cratering density (*12*). The image shows the region around the Gale Crater, the landing site selected for Curiosity Rover because its geology was deemed promising for finding evidence for habitability. For the full map and color key see http://pubs.usgs.gov/sim/3292/pdf/sim3292_map.pdf.

Smith's map in the digital age. Part of Smith's map draped over a digital elevation model, allowing it to be viewed in three dimensions.

IMAGES: (TOP) UNITED STATES GEOLOGICAL SURVEY; (BOTTOM) PETER WIGLEY

information became available to him. As a result, at least five different versions are recognized today (*6*). In understanding Smith's contribution to geology, the 1815 map cannot be considered in isolation; it was accompanied by an explanatory memoir describing the rocks and soils (*7*), and the small cross section on the map (see the first figure) was soon supplemented by a series of eight large geological cross sections showing the strata below the surface (*8*). Smith also produced more detailed geological maps of 21 English counties, with the same geological coloring as his large map (*9*). He explained his discoveries on the use of fossils in a series of colored plates (*10*) and in a catalog of his fossil collection, purchased by the British Museum (*11*).

Making a geological map remains a first step to understanding regional geology and in the search for raw materials and hydrocarbons. The location and development of resources such as rare earth minerals in places like Greenland requires a geological map at the outset. In hydrocarbon exploration today, seismic surveys extend the mapping into the subsurface in three or even four dimensions. Fossils are as essential in correlating rocks as they were in Smith's time and are important in the identification of strata in the oil industry today.

Geological maps have, since the time of NASA's Apollo program, extended beyond the terrestrial, as exemplified by a new geological map of Mars (see the second figure) (*12*). The paper map is now being replaced by the digital, and the same technology allows us to view Smith's map differently by rendering it in three dimensions (see the third figure) (*13*). In his own lifetime, Smith was hailed the "Father of English Geology," but he can equally be regarded as the Father of Stratigraphy and of the modern geological map. ∎

REFERENCES

1. M. J. S. Rudwick, *Bursting the Limits of Time* (Univ. of Chicago Press, Chicago, 2005).
2. L. R. Cox, *Proc. Yorks. Geol. Soc.* **25**, 1 (1942).
3. H. S. Torrens, *Oxford Dictionary of National Biography* (Oxford Univ. Press, Oxford, 2004), vol. 51, p. 372.
4. H. S. Torrens, *Geol. Soc. Lond. Spec. Publ.* **190**, 61 (2001).
5. S. J. Knell, *The Culture of English Geology, 1815–1851* (Ashgate, Aldershot, UK, 2000).
6. V. A. Eyles, J. M. Eyles, *Ann. Sci.* **3**, 190 (1938).
7. W. Smith, *A Memoir to the Map and Delineation of the Strata of England and Wales* (Cary, London, 1815).
8. J. G. C. M. Fuller, *'Strata Smith' and His Stratigraphic Cross Sections, 1819* (American Association of Petroleum Geologists and Geological Society of London, 1995).
9. J. M. Eyles, *J. Soc. Biblphy. Nat. Hist.* **5**, 87 (1969).
10. W. Smith, *Strata Identified by Organized Fossils* (W. Smith, London, 1816–1819).
11. W. Smith, *Stratigraphical System of Organized Fossils* (Williams, London, 1817).
12. K. L. Tanaka *et al.*, *Geologic Map of Mars* (USGS Scientific Investigations Map 3292, 2014).
13. P. Wigley *et al.*, *'Strata' Smith: His Two Hundred Year Legacy* (Geological Society, London, 2007).

New SQUID on the Bloch

A cloud of cold atoms can operate as an ultrasensitive interferometer

By **Austen Lamacraft**

Schrödinger's famous cat showed us that the matter waves inside an atom could manifest on macroscopic scales, but the real workhorse of the quantum world is the SQUID, or superconducting quantum interference device. This piece of technology is the key component of some hospital MRI (magnetic resonance imaging) scanners and other medical equipment because of its ability to detect tiny magnetic fields. On page 288 of this issue, Duca *et al.* (*1*) show that identical experimental principles can be used to detect the phases of the waves describing the motion of atoms in a periodic potential, known as Bloch states. Electrons moving in a crystal lattice are described by the same states, but although they are central to solid-state physics, their properties are usually inferred

Department of Physics, University of Cambridge, Cambridge CB3 0HE, UK. E-mail: al200@cam.ac.uk

from other measurements—conductivity, for example. The study by Duca *et al.* provides a proof-of-principle demonstration of a new type of interferometry capable of providing much more detailed information than existing techniques.

In a SQUID, superconducting currents flowing in a ring at the heart of the device are sensitive to tiny changes in the magnetic flux through the ring; this sensitivity is due to the interference of quantum waves propagating around each half, as in the two arms of an interferometer. In contrast, the two interfering paths in the experiment of Duca *et al.* form a loop in momentum space. The corresponding phase that is measured is an example of the far-reaching generalization of magnetic flux discovered by Berry in 1984 (*2*). Berry realized that for any family of quantum states labeled by continuous parameters (the momentum of the particle in this case), it is in general not possible to compare the phase of the states at two different values of the parameters.

A close analogy is provided by vectors lying tangent to a curved surface, where it is similarly difficult to compare the directions of vectors attached to different points. A natural resolution is to consider vectors close to each other and to determine the angle by which one vector has to twist about the normal to the surface in order to be parallel to the other—a process called the parallel transport of a vector. Surprisingly, this local comparison cannot be extended over finite distances; instead, parallel transport around a loop causes a rotation of the vector by an angle
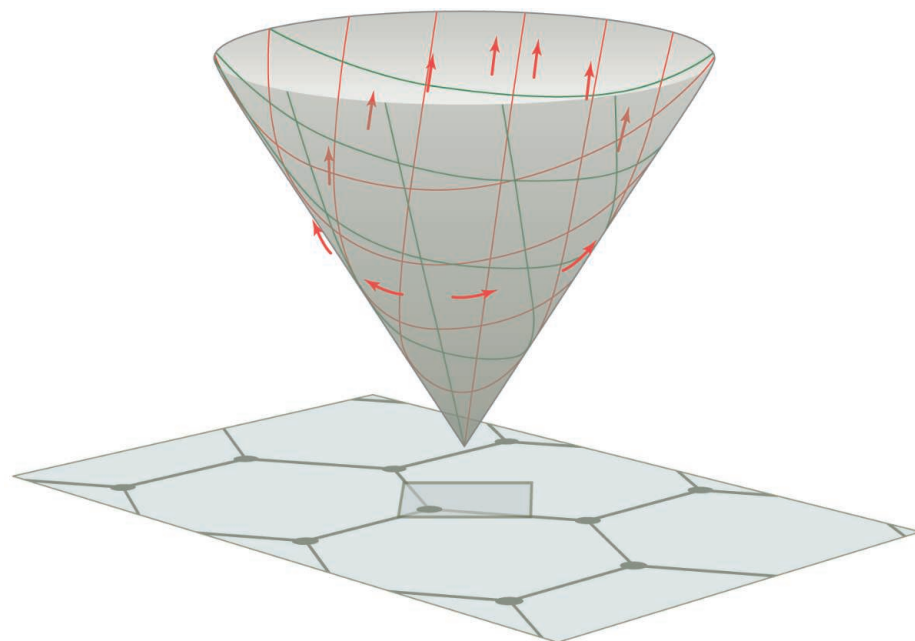
Interfering atoms. In the Aharonov-Bohm interferometer of Duca *et al.* (*1*), trajectories of wave packets in momentum space passing either side of a point where the Berry curvature is concentrated pick up a relative phase of 180°. The effect is related to the rotation of a vector parallel transported around a cone with a 60° aperture angle.

determined by the curvature of the surface inside. In the same way, Berry's phase carries information about an analogous property of the quantum states, often called Berry curvature. Berry's idea has found application in almost all areas of physics, most recently in the theory of topological insulators (3).

Instead of a superconductor, Duca *et al.* use a Bose-Einstein condensate of atoms to amplify the effects of single-atom interference. In addition, the tools of atomic physics allow precise control over the trajectory of the atoms in momentum space. They applied their technique to a hexagonal optical lattice formed by superimposing three laser beams, so that the Bloch states have a particularly simple distribution of Berry curvature concentrated at isolated values of momentum. Choosing trajectories to encircle these values reproduces the conditions of the Aharonov-Bohm effect, in which the interference of charged particles is affected by a magnetic flux in a region of space they do not enter. The Berry curvature matches that of a conical surface, which is concentrated at the tip of the cone. A trajectory that passes around the tip leads to a phase change of 180° (see the figure).

These effects are by now very familiar to condensed matter physicists, as they determine the conductivity of graphene in a magnetic field (4). Graphene consists of a two-dimensional sheet of carbon atoms arranged in a honeycomb lattice, so the Bloch states closely resemble those of the artificial lattice formed by laser light in the present experiment. An earlier experiment using an atomic Fermi gas was able to show that the energies of the Bloch states have the same conical features as those in graphene, but gave no information on the phases of the states (5). Although it was then clear that an interferometric measurement would be by far the cleanest and most convincing demonstration of the Berry phase, its achievement is an experimental tour de force.

The Berry phase measured by Duca *et al.* represents only the simplest example of a rich set of associated phenomena. The realization of more complex families of quantum states, involving the so-called non-Abelian Berry phases—where the quantum state rotates upon completing a circuit, rather than merely changing its phase—is a natural next milestone. ∎

REFERENCES

1. L. Duca *et al.*, *Science* **347**, 288 (2015).
2. M. V. Berry, *Proc. R. Soc. London Ser. A* **392**, 45 (1984).
3. M. Z. Hasan, C. L. Kane, *Rev. Mod. Phys.* **82**, 3045 (2010).
4. Y. Zhang, Y.-W. Tan, H. L. Stormer, P. Kim, *Nature* **438**, 201 (2005).
5. L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, T. Esslinger, *Nature* **483**, 302 (2012).
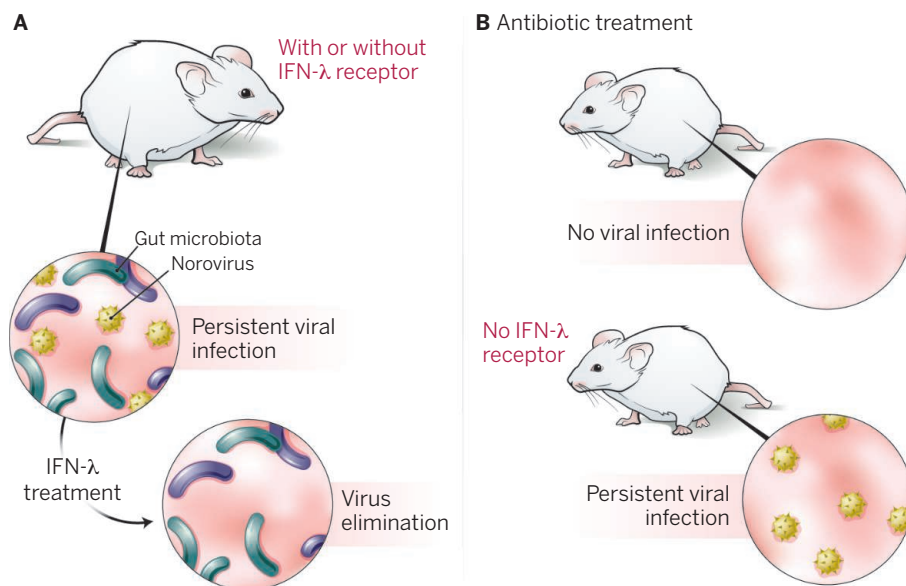
Controlling persistent infection. (**A**) Both wild-type mice and mice lacking the receptor for IFN-λ cannot clear murine norovirus (MNV) infection, but treatment with IFN-λ leads to virus elimination. (**B**) When the gut microbiota is eliminated by antibiotic treatment, MNV disappears from wild-type mice (and from mice that lack the adaptive immune system), but not from mice that lack IFN-λ signaling.

IMMUNOLOGY

Interfering with interferons

The gut microbiota helps norovirus to persist by controlling λ interferons

By **Jessica Wilks** *and* **Tatyana Golovkina**

Living organisms must resist viral infection. In mammals, both infected cells and innate immune cells release signals (cytokines) that program the infected cells for antiviral defense, as well as alert neighboring cells that trouble is afoot. These signals—exemplified by the type I (α and β), type II (γ), and type III (λ) interferons (IFNs)—control the mammalian response against the vast majority of viruses. The host's control of an enteric pathogen, rotavirus, requires type III IFNs (1, 2). On page 269 and 266 of this issue, Nice (3) and Baldridge (4), respectively, show that protection provided by λ IFNs is generalizable to another enteric pathogen, norovirus. Notably, this protection is independent from the adaptive immune response, which has long thought to be absolutely required for clearing viral infection.

The principal difference between type I and type III IFNs lies in their cognate receptor (5, 6). Whereas type I IFNs are recognized by the IFNαR1-R2 heterodimer (also called IFNαβR), type III IFNs are recognized by a distinct heterodimeric receptor composed of IFNλR1 and interleukin-10 receptor beta (IFNλR1-IL10Rβ1 or IFNλR). Even though the IFNαβR and IFNλR are independent, both receptors signal through a series of common adaptors leading to an indistinguishable antiviral immune response. Unlike IFNαβR, which is ubiquitously expressed in all nucleated cells, IFNλR is primarily expressed by the mucosal epithelium, with the highest expression in the small and large intestine.

A common cause of gastroenteritis, human norovirus is notoriously difficult to study in vitro and in vivo. For this reason, murine norovirus (MNV) has been an invaluable tool for analyzing the relationship between enteric viruses and the host immune system. To uncover the antiviral mechanism controlling persistent MNV infection, Nice *et al.* used a genetic approach to discover that type III IFN, but not type I IFN, is required to control persistent MNV, as virus shedding was substantially increased in IFNλR-deficient mice but not

Department of Microbiology, University of Chicago, Chicago, IL, USA. E-mail: tgolovki@bsd.uchicago.edu

ILLUSTRATION: V. ALTOUNIAN/*SCIENCE*

in IFNαβR-deficient mice. Administration of IFN-λ to mice with persistent norovirus infection reduced virus shedding below detectable amounts (see the figure). Moreover, injection of exogenous IFN-λ cleared the virus from mice devoid of an adaptive immune system (thus, eliminating the possibility that the animals invoked norovirus antigen–specific targeting by T and B cells of the adaptive immune system). Similarly, the administration of exogenous IFN-λ also ablated acute rotavirus infection in vivo (2). Consequently, this finding may have far-reaching implications regarding "sterilizing" innate immunity against enteric viral infections.

Wild-type mice should be able to induce IFN-λ, so why did they fail to clear persistent norovirus infection? Baldridge et al. provide a possible link between the host's microbiota and the antiviral response governed by λ IFNs. The authors found that ablation of the gut microbiota by antibiotic treatment results in clearance of persistent murine norovirus infection; restoring gut microbiota with that from untreated mice (through fecal transplant) rescued virus replication. This result has been separately confirmed in both antibiotic-treated (7)

> *"Exploring the interaction between viruses and the surrounding microbial community should reveal how commensals contribute to the transmission of an array of viruses..."*

and germ-free animals (8). By stark contrast, antibiotic-treated mice lacking IFN-λ signaling were unable to clear viral infection. Thus, IFN-λ is required to clear the virus, but its antiviral activity is diminished in the presence of the gut microbiota. This raises the possibility that the microbiota may directly or indirectly benefit the virus by inhibiting virally induced IFN-λ signaling. Indeed, another murine virus, mouse mammary tumor virus, exploits the host's gut microbiota by cloaking itself in bacterial lipopolysaccharide, a constituent of the outer membrane of Gram-negative bacteria (9). Virus-bound lipopolysaccharide triggers the pattern recognition receptor Toll-like receptor 4, which blocks the antivirus immune response by eliciting the production of IL-10, an immunosuppressive cytokine.

Enteric viruses, including reoviruses (10, 11), norovirus (4, 7, 8), and poliovirus (10, 12), are known now to require the gut microbiota for successful replication and transmission. The mechanisms through which the microbiota facilitates propagation of these viruses are not yet clear. In the case of murine norovirus, it may be that viral infection of B cells requires the presence of glycans, resembling histo-blood group antigens synthesized by specific types of enteric bacteria (7). Another possibility, suggested by the study of Baldridge et al., is that the gut microbiota may interfere with antiviral innate immunity, quenching IFN-λ signaling by an as-yet-undiscovered mechanism.

The findings of Nice et al. and Baldridge et al. prompt many questions. For example, it is unclear how MNV and other viruses elicit IFN-λ production, and what cell types sense this cytokine in the gut. Another question is why type I IFN production, which is triggered by MNV (8), does not contribute to the antivirus response in the gastrointestinal tract. One possibility is that the virus blocks the effects of both type I and type III IFNs with the help of the microbiota. It is also possible that a specific hierarchy exists through which IFNs control the virus, with IFN-λ providing superior protection in the gut whereas type I IFNs mainly protect systemic organs. That type III IFNs induce an antiviral response identical to that of type I IFNs and that their cognate receptor primarily lies within the intestinal epithelium can explain their essential role in the control of acute intestinal infections.

It appears that all orally transmitted viruses studied thus far exploit the gut microbiota for efficient transmission. However, many viruses enter the host through other surfaces, which also harbor commensal bacteria. Exploring the interaction between viruses and the surrounding microbial community should reveal how commensals contribute to the transmission of an array of viruses, not only enteric pathogens. ∎

REFERENCES

1. J. Angel, M. A. Franco, H. B. Greenberg, D. Bass, *J. Interferon Cytokine Res.* **19**, 655 (1999).
2. J. Pott et al., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7944 (2011).
3. T. J. Nice et al., *Science* **347**, 269 (2015).
4. M. T. Baldridge et al., *Science* **347**, 266 (2015).
5. S. V. Kotenko et al., *Nat. Immunol.* **4**, 69 (2003).
6. P. Sheppard et al., *Nat. Immunol.* **4**, 63 (2003).
7. M. K. Jones et al., *Science* **346**, 755 (2014).
8. E. Kernbauer, Y. Ding, K. Cadwell, *Nature* **516**, 94 (2014).
9. M. Kane et al., *Science* **334**, 245 (2011).
10. S. K. Kuss et al., *Science* **334**, 249 (2011).
11. R. Uchiyama, B. Chassaing, B. Zhang, A. T. Gewirtz, *J. Infect. Dis.* **210**, 171 (2014).
12. C. M. Robinson, P. R. Jesudhasan, J. K. Pfeiffer, *Cell Host Microbe* **15**, 36 (2014).

SOCIAL SCIENCE

Gender inequality in science

How should a better gender balance be achieved?

By **Andrew M. Penner**

W hy are women underrepresented in many areas of science, technology, engineering, and mathematics (STEM)? This is a question with no easy answers. In science, as in many areas of life, bias against women exists (1), but researchers disagree on how much bias matters: Some suggest that the effects of bias accumulate over time to shape careers (2), whereas others argue that gender differences in preferences are much more important (3). However, it is likely impossible to disentangle the effects of societal bias and individual preferences, because people's understanding of gender differences shape their preferences (4). Research suggests differences in innate ability are unlikely to play a major role (3), but one route to more equal representation across academic fields might be convincing both women and men that this is true. On page 262 of this issue, Leslie et al. (5) show that how ability is viewed within a field plays a key role in how well women are represented.

Two puzzles complicate typical explanations of women's underrepresentation in science. First, race and gender interact in ways that are problematic for one-size-fits-all approaches. In the United States, for example, although Asian women choose physical science majors at lower rates than Asian men, they do so at similar rates to white men, and at nearly twice the rate of white women. Of U.S. Asians who earned Bachelor's degrees in 2011, 1.9% of women and 2.4% of men majored in the physical sciences, compared to 2.1% of white men and 1.0% of white women (6). Second, gender representation varies considerably both within STEM and within non-STEM fields. As noted by Leslie et al., in 2011 women received 54% of U.S. Ph.D.'s in molecular biology, compared with 18% in physics, 72% in psychology, and 31% in philosophy.

Leslie et al. offer a novel framework for understanding this second puzzle by showing that how ability is viewed in different fields correlates with the degree to which

Is leaving science bad? German Chancellor Angela Merkel trained as a physical chemist but left research to enter politics. Many women leave STEM fields to work in other areas and make important contributions to society. Leslie *et al.* show that women are better represented in fields of study that view ability as related to effort than in those that view ability as innate.

PHOTO: HERBERT KNOSOWSKI/REUTERS/POOL

women are represented. In philosophy and physics, which are dominated by men, ability is considered to be innate. In molecular biology and psychology, in which women are well-represented, effort is viewed as important. This intriguing finding accounts for gender sorting into STEM versus non-STEM fields. It also explains why women are more represented in some STEM fields than others. Further, it avoids a problem plaguing many popular accounts for the underrepresentation of women in STEM, which fail to explain why women now pursue law degrees at similar rates to men, even though law school has a competitive culture, lawyers work long hours, and law firms are not yet as family friendly as one might hope.

Past studies of the underrepresentation of women in science have paid considerable attention to tracing the development of gender-related attitudes and stereotypes about science over students' careers. Building on this work, understanding how and when students become aware of the ability beliefs of specific fields will be instructive. Given how vital people think mathematical ability is for success in STEM fields, it will also be important to examine whether mathematical ability is viewed in particularly innate terms. If so, then we might expect fields in which mathematics is viewed

Department of Sociology, University of California, Irvine, Irvine, CA, USA. E-mail: andrew.penner@uci.edu

as more central to have particularly low participation of women.

A number of broader questions and challenges remain. Given that women have been graduating from college at higher rates than men for over three decades, the idea that men's curricular choices should be used as the baseline for women to emulate seems problematic. It is commonly suggested that more women are needed in science to meet U.S. workforce needs, but many students with STEM degrees do not end up employed in STEM fields (7). Some fields, like computer science, have experienced job growth and have a dearth of women, but it is unclear that women should be encouraged to enter fields solely because they are underrepresented. Rather, we might consider achieving gender parity in some fields by encouraging men to major in them at rates more similar to those of women.

It is also important to consider differences in the opportunities available to men and women outside of STEM, because women who excel in mathematics and science typically have stronger skills in other domains than do men (8). Angela Merkel (see the photo) and Margaret Thatcher were obviously exceptional in their achievements after leaving science, but we risk trivializing the contributions of women and men who choose to pursue other endeavors when we define success as becoming a STEM professor at a research university. This is especially true given that many women leave STEM to

go into fields such as education and health care. To be sure, there are substantial issues with how society devalues women's work, as pay declines in fields that come to be seen as women's work (9), but it is unclear that this constitutes failure on the part of the individuals who enter these fields, or that society does not benefit from their choices.

Like most researchers who study gender inequality in STEM fields, my inclination is to argue that we need these talented women in STEM fields. Yet, given the importance of having talented men and women in education, health care, and throughout the economy, it seems important to take a broader perspective on issues of gender equality. Perhaps it is time to ask a new question about gender representation in STEM: Would society be better off if men were more like women? ■

REFERENCES

1. D. Li, thesis, Massachusetts Institute of Technology (Cambridge, MA, 2012).
2. V. Valian, *Why So Slow?* (MIT Press, Cambridge, MA, 1998).
3. S. J. Ceci *et al.*, *Psychol. Sci. Public Interest* **15**, 75 (2014).
4. S. J. Correll, *Am. Sociol. Rev.* **69**, 93 (2004).
5. S. J. Leslie, A. Cimpian, M. Meyer, E. Freeland, *Science* **347**, 262 (2015).
6. T. D. Snyder, S. A. Dillow, *Digest of Education Statistics 2012* (National Center for Education Statistics, Washington, DC, 2013).
7. Y. Xie, A. A. Killewald, *Is American Science in Decline?* (Harvard Univ. Press, Cambridge, MA, 2012).
8. M. T. Wang, J. S. Eccles, S. Kenny, *Psychol. Sci.* **24**, 770 (2013).
9. P. England, *Comparable Worth* (Aldine de Gruyter, New York, 1992).

10.1126/science.aaa3781

SCIENCE sciencemag.org · 16 JANUARY 2015 • VOL 347 ISSUE 6219 **235**

Published by AAAS

INTELLECTUAL PROPERTY

Disclosing patents' secrets

Inventors prefer to disclose know-how before patent grant

By **Stuart Graham**[1] and **Deepak Hegde**[2]*

POLICY

The patent system is built on a grand bargain: To gain exclusive rights to practice their inventions, inventors must disclose their proprietary knowledge publicly. Economists have studied incentive benefits of exclusivity while implicitly assuming that disclosure of know-how in patent applications is costly for inventors. Yet, apart from facilitating diffusion of knowledge, disclosing know-how in a patent may privately benefit inventors by deterring rivals' duplicative research and development (R&D), preempting competitors' efforts to patent similar technology, and reducing informational asymmetries between patentees and potential investors [supplementary materials (SM)]. Understanding to what extent disclosure is viewed as a cost or a benefit by patenting inventors provides insights into our complex patent system and allows better policy-making to advance the diffusion of technical knowledge.

Whereas, historically, applications for U.S. patents were kept secret until grant, Congress in 1999 sought to harmonize the United States with virtually all other nations by requiring publication 18 months after first filing (1). This sparked disagreement, including from 26 Nobel laureates who warned that the legislation, which included the pregrant publication provision, "will prove very damaging to American small inventors and thereby discourage the flow of new inventions that have contributed so much to America's superior performance" (2). Congress adopted the change in the American Inventors Protection Act ("AIPA" or "Act"), but allowed applicants seeking no foreign patents on the same invention to opt out of publication, preserving secrecy until patent grant (3, 4) (SM).

This bifurcated choice—to disclose or not—presents a quasi-natural experiment. Inventors make choices to maximize private value. We investigate these "revealed preferences" using data on all 1.81 million applications filed with the U.S. Patent & Trademark Office (USPTO) in 1996–2005 and granted by mid-2012. Our focus on preferences, conditioned on the decision to patent, cannot measure benefits or costs of secrecy in general. Our findings are relevant to recent legislative proposals seeking to restrict pregrant publication (5) and to recommendations to eliminate the secrecy loophole altogether (6, 7). Disclosure provisions are a sticking point in international patent-system harmonization; thus, our findings inform ongoing negotiations [e.g., Trans-Pacific Partnership (8)].

U.S. patents and their disclosure status

USPTO patent applications (share, %)
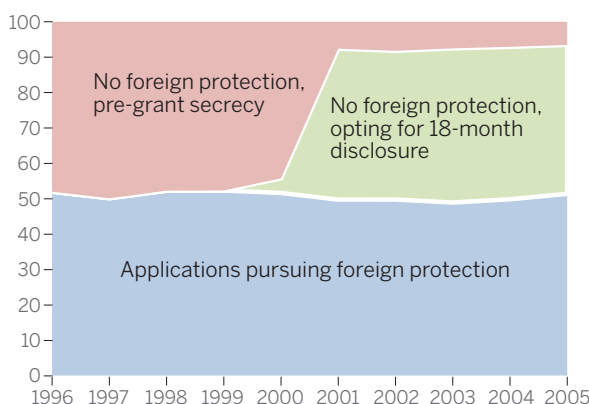
Disclosure status of all 1,809,932 patent applications filed at the USPTO between 1996 and 2005 for which patents were granted through mid-2012. Applicants shifted toward disclosing know-how after AIPA became effective 29 November 2000.

SMALL AND IMPORTANT. Before AIPA, the United States published only at patent grant, yet applications with a parallel foreign filing were published elsewhere. Beginning after 28 November 2000, AIPA requires the USPTO to publish all U.S. applications, with or without parallel foreign applications, 18 months from the first filing date. Applicants without a parallel foreign filing may opt out of 18-month publication at the time of application. Filing-to-grant lags averaged about 38 months in 2001, thus AIPA accelerated disclosure by about 20 months on average.

About 50% of U.S. granted patents, before and after AIPA, have parallel foreign

patenting. Conditional on choosing against (the 50% likely) foreign protection, about 85% of inventors choose pregrant disclosure (9) (SM) (see the first chart). In every technology field—drugs and medical, chemistry, mechanical, or electrical and electronics—a large majority of patentees choose disclosure over secrecy (SM). Opponents of pregrant disclosure argued that small U.S. inventors of important inventions would be particularly harmed by 18-month publication, and the opt-out was included to "protect" such inventors (10). We use USPTO data to classify applicants into four ownership and size types (11): (i) large U.S. (corporate; 34.2% of patents in our data); (ii) small U.S. (company and individual; 9.5%); (iii) foreign large (corporate; 38.2%); and (iv) others (18%) (12).

All inventor types are much more likely to choose pregrant disclosure over secrecy (SM). Conditional on U.S.-only patenting, small U.S. inventors prefer pregrant disclosure, and are no more likely than large U.S. entities to select secrecy (16.9% versus 16.4%; not significant at $P < 0.01$) (13). Secrecy is requested twice as often in "complex product" industries like "Computers and Communication" compared with "discrete product" industries like "Chemicals," but among large and small U.S. patentees alike, pregrant disclosure is preferred overwhelmingly in every technology sector (SM).

We compare choices made by applicants as a function of inventions' importance. As prior research has done, we use three different measures of "importance" (or "value" or "impact"): (i) the number of patent claims; (ii) periodic patent maintenance fee payments; and (iii) "forward" patent citation counts. Even if each measure has limitations, we find all three point in the same direction, which suggests that our results are not likely due to measurement errors.

Claims. Patent claims are the written descriptions of the invention indicating scope; more claims correlate with higher patent value and impact (14). Although our data show inconsistent results among other types of applicants (SM), average claim counts of patents from small U.S. inventors rank (from highest to lowest): those with foreign patent filings (24.4), U.S.-only patents published at 18 months (22.2), and patents selected into pregrant secrecy (20.6); all differences are significant at $P < 0.01$. These differences suggest that U.S. small inventors choosing secrecy produce patents having the narrowest scope and smallest economic impact.

[1]Scheller College of Business, Georgia Institute of Technology, Atlanta, GA 30308, USA. [2]Stern School of Business, New York University, New York, NY 10012, USA. *Corresponding author. E-mail: dhegde@stern.nyu.edu

Disclosure choices by citation's percentile, across patentee types

Patents granted after pre-grant secrecy (share, %)
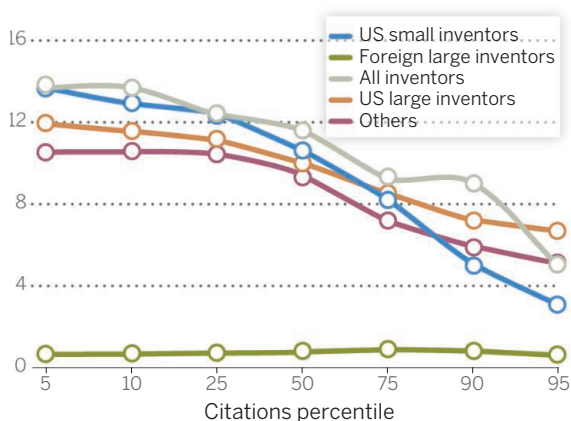
Share of patents filed in 2001 opting for pregrant secrecy. Scaled by citation percentiles for the different patent applicants in our data set. Higher percentiles indicate patents receiving higher numbers of citations. Larger invention impact is inversely correlated with opting for pregrant secrecy.

Maintenance fees. U.S. patentees must pay postgrant maintenance fees at 3.5, 7.5, and 11.5 years. These payments escalate over time, providing a proxy for the economic value of patents to their owners (*15*, *16*). We can observe 3.5- and 7.5-year renewal rates, finding them highest for U.S. patents with parallel foreign filings. Among large U.S. inventors, renewal rates are higher for patents issuing from pregrant secrecy. But among small U.S. inventors, both 3.5- and 7.5-year renewal rates are lowest for patents issuing from pregrant secrecy (significant at *P* < 0.01), which suggests that these patents are the least valuable and have the least impact particularly for U.S. small inventors (SM).

Citations. Increasing numbers of citations of a patent by other patents indicate more follow-on invention and private value (*17*). As 90% of citations arrive within 10 years (*18*), we minimize truncation bias by analyzing citations to post-AIPA patents filed in 2001, and accumulated over the 10 years from application date. Patents issuing from pregrant secrecy collect, on average, the least citations (7.8), with U.S.-only published patents (11.0) and foreign-filed patents (8.7) collecting more (significant at *P* < 0.01) (SM). Patents issuing from pregrant secrecy collected the least average number of citations for all inventor types (SM).

For robustness, we also measure citations from the disclosure date (18-month publication date versus grant date for "opt-out" patents). To avoid bias introduced by possible "strategic" citation behavior by industry applicants, we also restrict analysis to the ~35% of patent citations in our sample inserted only by patent examiners. With the exception of those held by non-U.S. large entities, we find that U.S. patents issuing from pregrant secrecy receive significantly fewer citations than those published at 18 months (SM).

Given the skew in patent value (*19*), we investigated patents in the most valuable tail of the citations distribution. We find an inverse correlation between citation percentile and probability of using pregrant secrecy (see the second chart) (Fig. 2). This pattern holds for all inventor types, barring foreign large patentees, and into the extreme tail: among the 198 small U.S. inventor patents in the 99th citations percentile, only 0.5%, or 1 of 198 patents, issued from pregrant secrecy.

When U.S. inventors make a disclosure choice for their patents with the greatest impact, they are much more likely to select early disclosure. Because the USPTO frames the choice between pregrant publication and secrecy saliently in patent applications and also because inventors likely monitor their valuable inventions most closely, our findings are at odds with any notion that the result is an artifact of 18-month publication (SM).

Maximum-likelihood logistic regressions reveal that U.S. small inventors' inventions with the most impact are systematically less likely than large inventors' to issue from pregrant secrecy in the overall population of patents, as well as in every technology field except "mechanical" and "other" (SM).

POLICY IMPLICATIONS. Although our analysis cannot unravel mechanisms, we speculate that patent disclosure confers private benefits to inventors, such as by credibly publicizing an invention's existence, quality, and scope to competitors, external investors, and licensees (*20*). Publication after 18 months offers patentees the provisional right to claim royalties from licensees from publication date. These benefits may be particularly salient for individuals and small firms, because they often depend on external observers to commercialize their important inventions.

Regardless of their motives, our finding suggests that patent publication may not only benefit society by accelerating the diffusion of scientific knowledge but also confers private benefits to patentees. Thus, recent arguments that pregrant disclosure should be limited to patent abstracts because they harm small U.S. inventors appear to have no empirical basis. Because small U.S. inventors are not choosing pregrant secrecy to protect their most important inventions, AIPA's "opt out" may be imposing avoidable costs on society by delaying cumulative innovation and encouraging duplicative R&D investments, while standing in the way of an internationally harmonized patent system. Our results challenge us to reconsider and focus more scientific inquiry on the benefits societies—and patentees themselves—are receiving from nondisclosure, even the 18 months now available as the worldwide standard for all types of patents, and patentees. ■

REFERENCES AND NOTES

1. U.S. applicants (applications) include USPTO filings from any country, unless otherwise stated.
2. F. Modigliani, An open letter to the U.S. Senate (1999); www.eagleforum.org/patent/nobel_letter.html.
3. P.L.106-113.
4. Other AIPA provisions were relatively minor changes.
5. H.R. 5980, 111th Congress 2nd Session (2010); https://www.govtrack.us/congress/bills/111/hr5980/text.
6. S. A. Merrill, R. C. Levin, M. B. Myers, Eds., *A Patent System for the 21st Century* (National Academy of Sciences, Washington, DC, 2004).
7. Federal Trade Commission, *The Evolving IP Marketplace: Aligning Patent Notice and Remedies with Competition* (FTC, Washington, DC, 2011).
8. R. Chen, *Japanese IP Practice II: IP Considerations with Japan's Entry into the Trans-Pacific Partnership* (American Bar Association, 2013); http://sugi.pat.co.jp/ABA_YLD_IP/JP_Trans-pacific_Partnership.pdf.
9. Separately, we obtained data on the 514,397 U.S. patent applications filed 11/29/2000–12/31/2005 that were abandoned prior to grant, finding that 34.1% filed for foreign protection, 56.6% chose U.S.-only publication, and only 9.3% chose pregrant secrecy.
10. E. Ergenzinger Jr., *Wake Forest Intell. Property Law J.* **7**, 146 (2002).
11. Based on disclosures to the USPTO at or before grant. USPTO defines "small entity" as fewer than 500 employees.
12. Governments, nonprofits, and foreign small entities and individuals.
13. The probability of secrecy conditional on not seeking foreign protection is $Ps/(1 - Pf)$ where Ps is the unconditional probability of pregrant secrecy and Pf is the unconditional probability of seeking foreign protection for a post-AIPA patent.
14. J. Lanjouw, M. Schankerman, *Rand J. Econ.* **32**, 129 (2001).
15. M. Schankerman, A. Pakes, *Econ. J.* **96**, 1052 (1986).
16. Current USPTO fees are $1130, $2850, and $4730 respectively for 3.5-, 7.5-, and 11.5-year renewal, discounted 50% for small firms and individuals.
17. M. Trajtenberg, *Rand J. Econ.* **21**, 172 (1990).
18. B. Hall, A. Jaffe, M. Trajtenberg, *The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools* (Working paper 8498, NBER, Cambridge, MA, 2001).
19. D. Harhoff, F. Narin, F. Scherer, K. Vopel, *Rev. Econ. Stat.* **81**, 511 (1999).
20. D. Hegde, H. Luo, *Imperfect Information, Patent Publication, and the Market for Ideas* (Working paper 14-019, Harvard Business School, Cambridge, MA, 2013).

Cheap and Clean
How Americans Think About Energy in the Age of Global Warming
Stephen Ansolabehere and David M. Konisky
MIT Press, 2014. 271 pp.

ENERGY

Fueling the future

Matthew E. Kahn considers an analysis of the public's perception of energy in the 21st century

Global energy demand is soaring. The major sources of power—including coal, oil, natural gas, and renewables such as wind and solar energy—differ with respect to the cost per unit and their local and global environmental impacts.

Consider coal. Today, China uses it to generate roughly 80% of its electricity (1). This "cheap" energy source is a major source of ambient particulate matter and greenhouse gas emissions that have serious consequences for our climate and our life expectancy.

In truth, there is no such thing as a "free lunch" in the energy sector, which means that we face tough choices when it comes to power. But how do we weigh our options?

In *Cheap and Clean: How Americans Think About Energy in the Age of Global Warming*, political scientists Stephen Ansolabehere and David M. Konisky present and analyze the results of a number of surveys on public attitudes about energy. The surveys, conducted between 2002 and 2013, sought to determine how well Americans understand the costs and benefits of different fuels, which fuels Americans prefer, how this relates to their understanding of these fuels, and whether global warming is changing the way people think about energy.

In terms of general energy preferences, the survey results indicate that Americans favor less reliance on coal and oil, are happy with the current level of natural gas and nuclear power generation, and favor the expansion of solar and wind power generation. Not surprisingly, most of those surveyed would prefer not to live close to coal-fired or nuclear power plants.

The authors are especially interested in determining how our perception of cost affects our energy preferences. Survey results indicate that Americans tend to view coal and natural gas as far less expensive than oil and nuclear power and tend to underestimate the cost of wind and solar power.

In chapter 6, the authors present results from a field experiment designed to examine the roots of our preferences. Participants received specific pieces of information intended to educate them about the true financial, health, and environmental costs of different energy sources. When they were given a summary of current energy prices and projected costs for each energy source, their attitudes toward coal did not change substantially; however, their support for renewable fuels, which are often considerably more expensive, diminished slightly. On the other hand, when participants learned about the social cost of coal—a measure that incorporates health and environmental cleanup costs into the cost estimate—they were less likely to support its use. They were also more likely to support the use of alternative fuels. Based on these results, the authors conclude that public attitudes are shaped by available information.
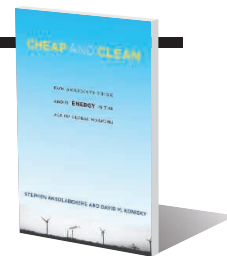
Environmental economists also seek to answer these types of questions; however, we tend to be weary of directly asking about energy priorities because what participants say in a survey is not always consistent with how they behave in the real world. An alternative research methodology is to study real estate prices in close proximity to pollution sources compared to neighborhoods with similar housing and demographic characteristics. A 2011 study found that neighborhoods within two miles of a fossil fuel power plant experienced 3 to 7% reductions in housing values and rents (2). On the other hand, in my own research, my coauthors and I have observed that California homes enjoy a 3.5% increase in resale value when they have solar roofs (3).

Our ability to have access to "cheap and clean" energy hinges on technological innovation, which the book briefly touches on in chapter 2. As we celebrate the increase in computer processing power enabled by advances in semiconductor technology, it is fair to ask when we can expect to see similar advances in renewable technology. Additionally, will free-market capitalism deliver the technological innovation required to make energy cheap and clean, or is the equivalent of a major federal investment—a green Manhattan Project—needed? Given the authors' focus on population preferences, I would like to have seen some discussion of the diverse population's willingness to support government investments in energy-related public goods.

The book highlights the important role that social science plays in evaluating the nation's emerging energy and climate change policy. Whether the preferences captured in these surveys ultimately translate into legislation such as a national carbon cap-and-trade policy remains an open question.

The reviewer is at the University of California–Los Angeles Institute of the Environment and Sustainability, La Kretz Hall, Suite 300, Los Angeles, CA 90095-1496, USA. E-mail: mkahn@ioe.ucla.edu

REFERENCES
1. International Energy Agency, 2012 Annual Report; www.iea.org/publications/freepublications/publication/IEA_Annual_Report_publicversion.pdf.
2. L. W. Davis, *Rev. Econ. Stat.* **93**, 1391 (2011).
3. S. R. Dastrup, J. G. Zivin, D. L. Costa, M. E. Kahn, *European Econ. Rev.* **56**, 961 (2012).

10.1126/science.aaa1579

The Myth of Mirror Neurons
The Real Neuroscience of Communication and Cognition
Gregory Hickok
Norton, 2014. 302 pp.

NEUROSCIENCE

The straw man in the brain

Pop culture claims about the brain make for easy targets, finds Christian Keysers

Two decades ago, a team led by the Italian neuroscientist Giacomo Rizzolatti noticed something peculiar in the premotor cortex of macaque monkeys. In addition to firing during the monkey's own actions, some neurons also fired when an experimenter performed comparable actions. These cells, which came to be known as "mirror neurons," generated tremendous scientific interest and the most-cited neuroscience paper of the past decade (1).

After the initial report, two mirror neuron "cultures" developed side by side. Most neuroscientists focused on basic questions in their scientific publications: where in the brain do mirror neurons exist; what do they respond to; do humans have them? At the same time, journalists, bloggers, and even some scientists, speculated enthusiastically about the function of these cells in popular culture outlets, implicating mirror neurons in everything from obesity to autism, despite the fact that many of these

The reviewer is at the Netherlands Institute for Neuroscience of the Royal Netherlands Academy of Arts and Sciences, Amsterdam, Netherlands. E-mail: c.keysers.nin.knaw.nl

claims so far remain untested.

The Myth of Mirror Neurons reflects author Gregory Hickok's frustration with the persistence of exaggerated claims about these cells. As a cognitive scientist specializing in language, Hickok has clearly been irritated by attempts to reduce the complexities of human language acquisition and comprehension into something that can be explained by mirror neurons alone. He is equally unimpressed with the claim that mirror neurons are all that is necessary for us to infer intent from the actions of others, an ability known as action understanding.

If mirror neurons in the motor circuitry by themselves explain all of language and action understanding, he argues, then patients with impaired motor circuitry should be unable to understand the language and the actions of others. Referencing a 2008 study (2), Hickok points out that while some patients with damage to the motor circuitry show subtle impairments in language perception and action understanding, others do not. In Chapter 4, he elaborates further, pointing out that speech perception is possible in patients who have suffered damage to motor speech centers in the brain and
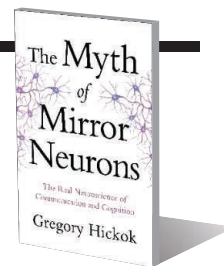
that we can understand actions that we cannot, ourselves, perform. His arguments are compelling refutations of the pop-culture claims that mirror neurons are necessary and sufficient for language and action understanding.

The problem is that Hickok fails to refute any actual scientific claims about mirror neurons. Rizzolatti and his colleagues have maintained all along that mirror neurons are not sufficient for language or action understanding (3, 4). They are thought, instead, to represent one component of a complex circuit that can enrich action perception. According to Hickok, scientists working in this arena also "minimize the importance of simulation in nonmotor systems." In reality, a number of groups, my own included, have published research showing that nonmotor circuitry, including the limbic and somatosensory systems, plays a critical role in social cognition (5, 6).

Unfortunately, these examples are not exceptions but are reflective of Hickok's methods throughout the book: He presents and then deconstructs dramatically oversimplified claims about mirror neurons, implying that these reflect the proposals of the scientists working in this field. His refutations of these claims are passionate and compelling but, ironically, lead down the very path of mystification he tries to denounce.

Although the book fails to address the current state of understanding about the function of mirror neurons, it does make it clear how urgent it has become for the field of neuroscience to ramp up efforts to determine what mirror neurons contribute to our mind and behavior. The truth is that a scarcity of evidence is not evidence for the scarcity of their contribution.

REFERENCES

1. G. Rizzolatti, L. Craighero, *Ann. Rev. Neurosci.* **27**, 169 (2004).
2. M. Pazzaglia *et al.*, J. Neurosci. **28**, 3030 (2008).
3. G. Rizzolatti, L. Fogassi, V. Gallese, *Nat. Rev. Neurosci.* **2**, 661 (2001).
4. K. Nelissen, E. Borra, M. Gerbella, S. Rozzi, G. Luppino, W. Vanduffel, G. Rizzolatti, G. A. Orban, *J. Neurosci.* **31**, 3743 (2011).
5. V. Gallese, C. Keysers, G. Rizzolatti, *Trends Cog. Sci.* **8**, 396 (2004).
6. C. Keysers, J. H. Kaas, V. Gazzola, *Nat. Rev. Neurosci.* **11**, 417 (2010).

10.1126/science.1259030

MEDICINE

What ails medical education?

Cynthia Whitehead welcomes a thorough appraisal of American medical residency training

Is medical education in a state of good health? A recent flurry of reports and commissions suggests that it is not (*1*, *2*). An article published in 2014 in the *New England Journal of Medicine* identifies a dearth of research as a major obstacle that must be overcome to improve medical education (*3*). According to the authors, in the absence of robust research, most ideas for medical education innovation are based on assumptions, traditions, and taken-for-granted notions of what should work. Given the vast resources spent on health care, the lack of investment in the study of medical education is puzzling and problematic.

Kenneth Ludmerer's new book, *Let Me Heal: The Opportunity to Preserve Excellence in American Medicine*, is a valuable contribution to our understanding of one of the most important components of medical education: the residency training program. Residency programs provide additional training after the completion of medical school, during which new doctors specialize in a particular area such as surgery, internal medicine, or family medicine. The book presents a history of residency education in the United States and elegantly deconstructs the assumptions and forces that have shaped the current state of American residency education.

Before the inception of U.S. residency programs, American physicians who wanted specialized training had to travel to Europe. Germany was a particularly desirable place of study in the late 1800s as a result of its preeminence in scientific medicine and its acclaimed apprenticeship model of training. The first U.S. residency program, which began at Johns Hopkins Hospital in 1889, was inspired by the German model. Legendary teachers, including Sir William Osler and William Stewart Halsted, demanded exquisite attention to detail, in-depth analysis of patient problems, an attitude of critical inquiry, engage-

ment in clinical research, and devotion to patient care. Ludmerer's account notes the toll this challenging work took on the men and their families and the extremely limited opportunities for women and minorities.

In early residency programs, education occurred primarily in "charity" wards, where residents provided free care to indigent patients. Hospital conditions were often appalling, especially for racial minorities. For example, Atlanta's Grady Hospital was known at one time for having two patients per bed, placed "head to foot" in its black wards (*4*). Callous and paternalistic approaches were also commonplace in charity hospitals. The surgeon Charles W. Mayo once admitted to wondering whether it might be a good idea to arrange for a few weeks of pain as part of these clinicians'

In *Let Me Heal*, Kenneth Ludmerer deconstructs the history and current state of medical residency education in the United States.

training, speculating that it might make them more compassionate (*5*).

Ludmerer describes how residency slowly changed from an elite opportunity to a requirement for practice. Residents are now integral to the fabric and funding of teaching hospitals. But the nature of hospital care is changing. In recent years, requirements for efficiency and cost containment have vastly reduced the length of hospital stays. Inpatient work is now driven by high volumes of admissions and speedy discharges. This inevitably results in compromised attention to detail and limits residents' exposure to the full courses of illnesses. Across health care settings, an emphasis on standardization of routine processes has also led to a shift from
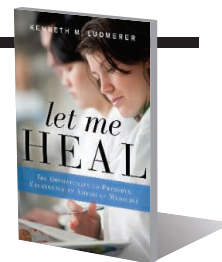
thoughtful, patient-specific care toward the memorization of guidelines, algorithms, and protocols.

With regard to the limitations to both thorough care and thoughtful learning posed by high patient volumes, Ludmerer suggests several intriguing ideas for fostering environments that enable residents and their teachers to engage in both high-quality care and high-quality education. One proposal is to provide more learning in ambulatory and community settings. Ludmerer further suggests creating inpatient teaching units with reduced patient volumes. Whether cash-strapped health care institutions will be open to this approach is by no means certain.

At present, there is minimal evidence about what truly makes a difference across most of the medical education enterprise. Although research in this arena remains substantially underfunded, one very promising development is the recent international proliferation of medical education research centers. These centers house a growing cadre of education scientists from diverse fields and have the potential to make important research contributions to the field of medical education.

A commitment to a scientifically informed reform process will require serious questioning of assumptions about what is good and what works. The research needed will draw on many disciplines relevant to health care research, including the social sciences. This book represents an important contribution to our understanding of the history and current state of American residency education and offers a strong foundation for future research.

REFERENCES

1. J. Frenk *et al.*, *Lancet* **376**, 1923 (2010).
2. J. Eden, D. Berwick, G. Wilensky, Eds., *Graduate Medical Education That Meets the Nation's Health Needs* (Institute of Medicine, Washington, DC, 2014).
3. D. Asch, D. Weinstein, *NEJM* **371**, 794 (2014).
4. M. Finland, *Maxwell Finland Papers: 1916-2003* (Francis A. Countway Library of Medicine, Harvard Medical School, Boston, MA).
5. C. W. Mayo, *Mayo: The Story of My Family and My Career* (Doubleday, New York, 1968).

The reviewer is at the Wilson Centre, Faculty of Medicine, University of Toronto, and Women's College Hospital, Toronto, ON M5S 1B2, Canada. E-mail: cynthia.whitehead@utoronto.ca

LETTERS

Edited by **Jennifer Sills**

One Medicine One Science and policy

TODAY, HUMANS, ANIMALS, and the environment are remarkably interconnected and interdependent at a global level through international commerce and movement. Thus, we have access to safe and nutritious food that fuels health, medicines and vaccines that protect us and our animals, and natural resources that support good living standards. However, conflicts arise as exponentially growing populations require more food, demand better living standards, and act to preserve the environment. How do we simultaneously produce more food, reduce disease, afford equitable living standards, and create an environment fit for humans, our animals, and wildlife? Science has played a critical role in finding solutions to many of these challenges, but difficult conflicts continue to emerge. For example, strategies that promote efficient production of food—such as concentrated farming systems, monoculture cropping, and chemical inputs of fertilizer, pesticides, and herbicides—have unintended consequences that threaten human, animal, and environmental health (1). A more integrated, holistic problem-solving approach informed by science is needed for development of public policies that address these complex problems.

A growing "one health" dialogue has focused on emerging disease surveillance, public health preparedness, and policy issues, with less attention being given to connecting these issues to the scientific foundations that underpin pathogen emergence, global health threats, food security, and environmental health. The imbalance has resulted in a compartmentalization of research and policy, sometimes diluting or compromising the efficacy of the one health movement. For example, lack of balanced scientific input imperils policies affecting antibiotic use for efficient food production and, more seriously, modification of plants for improved production of foods under adverse conditions in impoverished countries. There is value in looking back at instances of scientifically informed decision-making that have broadly benefited human, animal, and environmental health, such as the U.S. Clean Air Act of 1970 and the U.S. Clean Water Act of 1972.

Human health security today relies on finding common ground at the intersection of humans, animals, and the environment among diverse opinions and values. The concept of One Medicine One Science–based forums that bring together a diversity of scientists, policy professionals, medical experts, food producers, and other relevant stakeholders provides an important opportunity to present scientific knowledge that objectively informs public policy development (2). In the absence of scientific information, for instance, concerns about foods genetically modified for enhanced traits have resulted in European markets being closed to African farmers trying to better their standard of living. It has also led to at-risk pregnant women and children lacking access to golden rice, whose high vitamin A content could alleviate as many

Strategies to increase food production, such as pesticide use, have unintended consequences.

as 500,000 cases of irreversible blindness in pregnant women and children every year (3). Scientifically informed public policies are also needed to find solutions to foreseeable food supply limitations, prepare for existing foreign animal and human disease pathogens (e.g., African swine fever virus and Chikungunya virus), and deal with environmental implications of extreme energy production (e.g., fracking effects on water quality) or industrial pollution of agricultural land in China (4). Overall, balancing competing priorities is a major challenge as societies seek to maximize human health, animal health and welfare, and environmental integrity. Reductionism may be a natural reaction to complexity, but we are in need of evolving ways to discuss, understand, and address these complex challenges.

P. Sriramarao,[1]* Michael P. Murtaugh,[2] Kavita Berger,[3] Dominic Travis,[4] Shaun Kennedy,[4,5] Clifford J. Steer,[6] Carol Cardona[2]

[1]College of Veterinary Medicine, University of Minnesota, Saint Paul, MN 55108, USA. [2]Department of Veterinary and Biomedical Sciences, College of Veterinary Medicine, University of Minnesota, Saint Paul, MN 55108, USA. [3]Center for Science, Technology, and Security Policy, American Association for the Advancement of Science, Washington, DC 20005, USA. [4]Department of Veterinary Population Medicine, College of Veterinary Medicine, University of Minnesota, Saint Paul, MN 55108, USA. [5]The Food System Institute, LLC. [6]Departments of Medicine and Genetics, Cell Biology, and Development, University of Minnesota Medical School, Minneapolis, MN 55455, USA.

*Corresponding author. E-mail: psrao@umn.edu

REFERENCES

1. R. S. DeFries, J. A. Foley, G. P. Asner, *Front. Ecol. Environ.* **2**, 249 (2004).
2. D. A. Travis *et al.*, *Ann. N.Y. Acad. Sci.* **1334**, 26 (2014).
3. Golden Rice Project (www.goldenrice.org).
4. C. Larson, *Science* **343**, 1415 (2014).

Counting on small-scale fisheries

ON 10 JUNE 2014, the member States of the Food and Agriculture Organization of the United Nations (FAO) adopted the *Voluntary Guidelines for Securing Sustainable Small-Scale Fisheries in the Context of Food Security and Poverty Eradication* (1) ("Guidelines").

To make these Guidelines effective, it is crucial that the FAO, governments, and civil society have access to data to help understand small-scale fisheries. Currently, catches from these fisheries are not collected separately, but are lumped in with industrial catches, even though they represent about one-quarter of global catches, and the majority of catches in many developing countries. To promote the transparency needed for good governance (2, 3), the FAO ought to request from member countries a report of catch data that distinguishes between industrial and small-scale fisheries.

Many decades of debate have failed to produce one, agreed-upon definition of a "small-scale fishery," but the modest variations in definitions between countries do not preclude efforts to gather global statistics. Just as the Guidelines do not impose a single definition of small-scale fisheries, each of the FAO's member States could define their own small-scale fisheries, reflecting local realities.

These changes would help to highlight the importance of small-scale fisheries and may also help governments that still treat these fisheries as a solution to

demographic pressure and rural landlessness (4) to focus instead on their inherent value.

Daniel Pauly¹ and Anthony Charles²*

¹*Sea Around Us*, Fisheries Centre, University of British Columbia, Vancouver, BC V6T1Z4, Canada. ²School of Business and School of the Environment, Saint Mary's University, Halifax, NS B3H3C3, Canada.

*Corresponding author. E-mail: tony.charles@smu.ca

REFERENCES

1. FAO, *Voluntary Guidelines for Securing Sustainable Small-scale Fisheries in the Context of Food Security and Poverty Eradication* (Food and Agriculture Organization of the United Nations, Rome, 2014).
2. A. Charles, *Maritime Stud.* **10**, 85 (2011).
3. A. Charles, *Land Tenure J.* **1**, 9 (2013).
4. D. Pauly, *Maritime Stud.* **4**, 7 (2006).

A defense of animal welfare accreditation

IN HIS PROVOCATIVE News article, "Animal welfare accreditation called into question," (29 August, p. 988), D. Grimm discusses an article published in the *Journal of Applied Animal Welfare Sciences* (JAAWS) (*1*) analyzing the effectiveness of the accreditation system run by the Association for Assessment and Accreditation of Laboratory Animal Care (AAALAC) International. The article purports to demonstrate that institutions participating in AAALAC International's accreditation program had more U.S. Department of Agriculture (USDA)–Animal Care inspection "non-compliance items" (NCIs) than did non-accredited institutions. Grimm accurately reported my opinion that the article was not credible. My opinion is based not only on the JAAWS authors' affiliation with People for the Ethical Treatment of Animals (PETA), an organization devoted to stopping all research using animals, but also on the fact that the research was flawed.

I asked the authors to share their data set for independent analysis by AAALAC and others, but they refused. This refusal to share data fails to meet contemporary standards for the responsible conduct of research, and in itself renders the conclusions open to question. I appealed to the editors of JAAWS for assistance; however, they replied that they encourage authors to share their data, but they do not require it, in contrast to the policy of *Science* and other respected scientific publications.

To approximate the unavailable data, AAALAC International acquired a comparable data set of NCIs for one of the years analyzed in the JAAWS publication using the same search engine (*2*). Our review of these data (see Supplementary Materials) supported all the potential deficiencies in the JAAWS manuscript mentioned by me and others—including the NIH's Office of Laboratory Welfare—in Grimm's article. First, according to the Discussion section of the JAAWS paper (*1*), the authors improperly treat every NCI (or "violation," the term used by Grimm) equally, although NCIs exhibit a wide and unacceptable variability and may lack demonstrable relevance to animal care, health, or well-being outcomes. Second, the USDA designation of a licensed research animal facility does not correspond identically with the unit designation used by AAALAC in the accreditation program (*3*, *4*). This fact, and the fact that some AAALAC-accredited units are unlisted on AAALAC's public Web site, would have produced substantial data coding errors in this study. Finally, the population of non-accredited research animal facilities is vastly different in character than the population of AAALAC-accredited units. In general non-AAALAC programs are smaller in size and have narrower research missions, whereas the AAALAC-accredited programs include the largest U.S. institutions with broad research missions entailing complex and diverse animal research studies. In short, the JAAWS analysis was incapable of producing a meaningful comparison of AAALAC-accredited programs and non-accredited research animal facilities using NCIs.

The institutions and their scientists participating in AAALAC accreditation understand the stark dissimilarity between the AAALAC's performance-based, confidential, expert, peer-review accreditation program and the USDA–Animal Care regulatory inspection process. Unfortunately, Grimm's article brought wide and undeserved attention to a poorly designed and executed study. This is deeply disappointing to AAALAC International and the scientific community it serves.

Christian Newcomer

Executive Director, Association for Assessment and Accreditation of Laboratory Animal Care International, Frederick, MD 21703, USA. E-mail: cnewcomer@aaalac.org

REFERENCES

1. J. Goodman, A. Chandna, C. Borch, *J. Appl. Anim. Welfare Sci.* **18**, 82 (2015).
2. USDA Animal Care Information System Search Engine (2014); https://acissearch.aphis.usda.gov/LPASearch/faces/Warning.jspx.
3. AAALAC International, Rules of Accreditation, Section 1, Definitions, Accreditable Unit (www.aaalac.org/accreditation/rules.cfm).
4. Licensing and Registration under the Animal Welfare Act, Guidelines for Dealers, Exhibitors, Transporters and Researchers, Research Facilities (www.aphis.usda.gov/animal_welfare/downloads/aw/awlicreg.pdf).
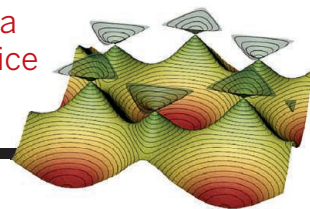
SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6219/243/suppl/DC1
Table S1

RESEARCH

Berry flux revealed in a hexagonal optical lattice
Duca et al., *p. 288*

IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

Colorized scanning electron micrograph of cardiac muscle tissue

GENETIC MEDICINE
What happens when titins are trimmed?

The most common form of heart failure, dilated cardiomyopathy, is often caused by mutations in a mammoth heart protein, called titin. Roberts *et al*. asked which titin mutations cause disease and why. They sequenced the titin gene in over 5200 people—some healthy and some with heart failure—and measured the corresponding RNA and protein levels. Many of the mutations truncated titin, causing short nonfunctional versions of the RNA or protein. These defects produced cardiomyopathy when they occurred closer to the protein's carboxyl terminus and in abundantly transcribed exons. The titin-truncating mutations that occurred in the healthy people did not have these characteristics and were usually benign. — KLK

Sci. Transl. Med. **7**, 270ra6 (2015).

NOROVIRUS
Turning viral persistence on and off

Norovirus causes >90% of the world's gastroenteritis. Norovirus can establish persistent infections, which may contribute to its spread. How does norovirus establish itself as a permanent resident of the gut and how can such persistent infections be cured (see the Perspective by Wilks and Golovkina)? Baldridge *et al.* studied mice persistently infected with norovirus and found that viral persistence required the gut microbiota: resident bacteria in the gastrointestinal tract. Antibiotics prevented persistent mouse norovirus infection in a way that depended on the secreted antiviral protein interferon λ (IFN-λ). Nice *et al.* report that IFN-λ can cure mice persistently infected with norovirus, independent of the adaptive immune system. — KLM

Science, this issue p. 266, p. 269; see also p. 233

ASIAN ARCHEOLOGY
Colonizing the roof of the world

Humans only settled permanently on the Tibetan plateau about 3600 years ago. Chen *et al.* examined archaeological crop remains unearthed in northeastern Tibet, which elucidate the timing of agricultural settlement. Although much earlier traces of humans in Tibet have been dated to 20,000 years ago, year-round presence at the highest altitudes appears to have been impossible until the advent of suitable crops, such as barley. Surprisingly, these prehistoric farming communities expanded onto the plateau at the same time as climate was cooling. — AMS

Science, this issue p. 248

TELOMERES IN CANCER
Cancer's alternative means to an end

To stay alive and proliferating, tumor cells must maintain their telomeres: the DNA sequences at the ends of chromosomes. The majority accomplish this by activating the enzyme telomerase. However, certain tumor types favor a different mechanism called alternative lengthening of telomeres (ALT), which involves DNA recombination. Flynn *et al.* delineated the molecular events that occur at the telomeres of ALT-proficient tumor cells by studying the function of a protein that is altered by mutation in these tumors. The analysis revealed a specific protein kinase
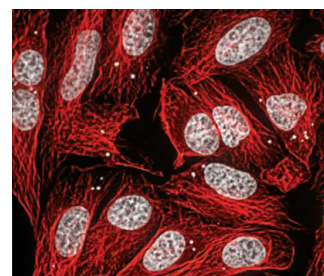
ALT-positive cells: DNA white; microtubules red.

that is essential for ALT, which could in principle be targeted to halt tumor growth. — PAK

Science, this issue p. 273

ANIMAL PHYSIOLOGY
Geese need to hug the land to fly high

Animal migrations provide numerous examples of astonishing feats. Impressive even among these is the migration of bar-headed geese across the Himalayan Mountains, which reach heights of thousands of meters. Bishop *et al.* remotely monitored birds' heart rates, movement, and body temperature during migration. The geese "hug" the landforms, taking advantage of drafting and wind patterns. This unexpected strategy conserves energy, even though it means the geese repeatedly lose, and must then regain, altitude. — SNV

Science, this issue p. 250

PALEOCEANOGRAPHY
A new tilt on predicting future ENSO variability

A new finding should improve the ability of climate models to predict the behavior of the El Niño–Southern Oscillation (ENSO) in a warmer future. Ford et al. looked at the distribution of surface and subsurface temperatures in the eastern and western equatorial Pacific 19,000 years ago and between 3000 and 6000 years ago. Temperatures fluctuated over a greater range during the older period. ENSO thus depended more on the tilt of the equatorial Pacific thermocline than on the east-to-west temperature gradient, as previously thought. — HJS

Science, this issue p. 255

WOMEN IN SCIENCE
Women's participation and attitudes to talent

Some scientific disciplines have lower percentages of women in academia than others. Leslie *et al.* hypothesized that general attitudes about the discipline would reflect the representation of women in those fields (see the Perspective by Penner). Surveys revealed that some fields are believed to require attributes such as brilliance and genius, whereas other fields are believed to require more empathy or hard work. In fields where people thought that raw talent was required, academic departments had lower percentages of women. — PJH

Science, this issue p. 262; see also p. 234

APPLIED PHYSICS
Tunnel through and emit coherently

The generation of coherent light (lasers and masers) forms the basis of a large optics industry. Liu *et al.* demonstrate a type of laser that is driven by the tunneling of single electrons in semiconductor double-quantum dots. Distinct from other existing semiconductor lasers, the emission mechanism is driven by tunneling of single charges between discrete energy levels that are electrically tunable. The ability to tune the levels by single-electron charging would allow their laser (or maser) to be turned on and off rapidly. — ISO

Science, this issue p. 285

TOPOLOGICAL MATTER
Nailing down the topology of a semimetal

Topological insulators are exotic materials that have a conducting surface state that can withstand certain types of material imperfection. Theoreticians have predicted a different kind of surface state in related three-dimensional topological Dirac semimetals, which do not have an energy gap in the band structure of the bulk. Xu *et al.* used photoemission spectroscopy to map out the band structure of the material Na_3Bi and detected the predicted surface state. Their results may lead to further insights into the physics of topological matter. — JS

Science, this issue p. 294

Lettuce plants are also a source of natural rubber

PLANT SCIENCE
Building rubber without rubber trees

Natural rubber trumps synthetic, petroleum-derived rubber in useful qualities such as elasticity and abrasion resistance. But rubber trees are quite susceptible to disease, leading scientists to search for other sources of natural rubber and to understand the specifics of its biosynthesis. Lettuce plants are a source of natural rubber, and now Qu *et al.* identify a scaffold protein called CPTL2 that keeps rubber-synthesizing enzymes from bouncing around the cell. Tethered to the cell's endoplasmic reticulum, CPTL2 anchors a protein important for rubber polymerization in place. Plants with reduced expression of CPTL2 could not synthesize rubber, revealing its essential role. — PJH

J. Biol. Chem. 10.1074/jbc.M114.616920 (2014).

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

CANCER
Visualizing active signaling in cancer

The receptor tyrosine kinase epidermal growth factor receptor (EGFR) can be aberrantly activated in diverse cancers. The ability to identify tumors with increased EGFR activity should improve the personalization of therapies. However, genetic analysis fails to identify such tumors if the increase in activity is not due to mutations in EGFR-encoding genes. Smith *et al.* developed an assay to detect the interaction between EGFR and the adaptor protein GRB2 as a marker for EGFR signaling. This assay detected active EGFR signaling in tumors with normal EGFR, which would have been undetectable by genetic analysis. Moreover, using this assay, the researchers could predict therapeutic response to EGFR inhibitors in both mice and humans. — JDB

Sci. Signal. **8**, ra4 (2015).

T CELL VACCINES
For vaccines, CD4⁺ T cells can spell trouble

The ideal vaccine elicits immune memory—either antibodies or memory T cells—to protect the host from subsequent infections. T cell–mediated immunity requires both helper CD4⁺ T cells and cytotoxic CD8⁺ T cells to kill virus-infected cells. But what happens when a vaccine only elicits CD4⁺ memory T cells? Penaloza-MacMaster *et al.* probed this question by giving mice a vaccine that generated only memory CD4⁺ T cells against lymphocytic choriomeningitis virus (LCMV). Instead of protecting mice against chronic LCMV, vaccinated mice developed massive inflammation and died. Virus-specific CD8⁺ T cells or antibodies protected mice from the pathology. These

results may have implications for vaccines against chronic viruses such as HIV. — KLM

Science, this issue p. 278

SUPERCONDUCTIVITY
Finding order in exotic superconductors

Physicists can coax some copper-oxide compounds into becoming superconducting by chemically adding extra charge carriers: holes or electrons. Concentrating on hole-doped materials, researchers have found a host of different phases in the neighborhood of or co-existing with superconductivity. One such phase is a modulation in charge density [a charge density wave (CDW)] that appears to be ubiquitous in hole-doped families. Da Silva Neto *et al.* now show that a similar phase exists in the electron-doped material $Nd_{2-x}Ce_xCuO_4$. As they cooled the material, the authors first detected the CDW at temperatures considerably higher than in the hole-doped copper-oxides. — JS

Science, this issue p. 282

QUANTUM GASES
Nailing down graphene's topology

An electron traveling along a closed path in the momentum space of the graphene crystal lattice may not end up exactly the way it started. If its path happens to include one of the special points in momentum space, it will acquire a phase shift. Physicists can detect the signatures of this process by studying the transport properties of graphene. Duca *et al.* used interferometry to directly measure this so-called Berry flux in a hexagonal optical lattice, where intersecting laser beams simulate the environment that electrons experience

in graphene (see the Perspective by Lamacraft). The high-precision technique may be useful in characterizing other topological structures. — JS

Science, this issue p. 288; see also p. 232

INTERFACIAL SOLVENTS
Structured solvents near nanoparticles

The physical properties and reactivity of nanoparticles in solution depend not only on their surface termination but also on changes in solvent ordering caused by the presence of the interface created by the nanoparticle. Zobel *et al.* used x-ray scattering to study solvent ordering for a variety of metal and metal-oxide nanoparticles in a variety of polar solvents (alcohols) and a nonpolar solvent (n-hexane). They observed layers of enhanced ordering near the nanoparticle surface relative to the bulk solvent. These trends were largely independent of surface chemistry, such as changing the surface groups from hydroxyls to carboxylates. — PDS

Science, this issue p. 292

GEOLOGY
Current impacts of the first geological map made

In the late 18th century, William Smith, an English canal surveyor, noticed that two canal branches in adjacent valleys passed through the same sequence of rock strata and that each bed of rock contained a distinctive set of fossils. These insights were central to his geological map of England, Wales, and part of Scotland, published in 1815. The map was the first of its kind, depicting stratigraphic order from oldest rocks to youngest in a way that gives an impression of three dimensions. In a Perspective, Sharpe explains how the principles used to

create the map remain in use today, for example in the identification of fossil fuel deposits and sites rich in mineral resources. — JFU

Science, this issue p. 230

MARINE CONSERVATION
Marine animals are disappearing, too

The loss of animal species in terrestrial environments has been well documented and is continuing. Loss of species in marine environments has been slower than in terrestrial systems, but appears to be increasing rapidly. McCauley *et al.* review the recent patterns of species decline and loss in marine environments. Though they note many worrying declines, they also highlight approaches that might allow us to prevent the type of massive defaunation that has occurred on land. — SNV

Science, this issue p. 247

PALEOECOLOGY
Fluctuations revealed in fossil forests

The reconstruction of past vegetation unlocks the door to understanding ecological changes associated with climatic change. But it is also difficult. Dunn *et al.* developed a method for assessing changes in vegetation openness based on epidermal cell morphology from conserved plant tissue. Applying this method to fossil assemblages from Patagonia, they show how vegetation structure changed over the Cenozoic era (49 to 11 million years ago). These changes map onto the known climate changes over this period and can also be used to track how the evolution of herbivorous mammals responded to vegetation structure. — AMS

Science, this issue p. 258

that is essential for ALT, which could in principle be targeted to halt tumor growth. — PAK

Science, this issue p. 273

ANIMAL PHYSIOLOGY
Geese need to hug the land to fly high

Animal migrations provide numerous examples of astonishing feats. Impressive even among these is the migration of bar-headed geese across the Himalayan Mountains, which reach heights of thousands of meters. Bishop *et al.* remotely monitored birds' heart rates, movement, and body temperature during migration. The geese "hug" the landforms, taking advantage of drafting and wind patterns. This unexpected strategy conserves energy, even though it means the geese repeatedly lose, and must then regain, altitude. — SNV

Science, this issue p. 250

PALEOCEANOGRAPHY
A new tilt on predicting future ENSO variability

A new finding should improve the ability of climate models to predict the behavior of the El Niño–Southern Oscillation (ENSO) in a warmer future. Ford et al. looked at the distribution of surface and subsurface temperatures in the eastern and western equatorial Pacific 19,000 years ago and between 3000 and 6000 years ago. Temperatures fluctuated over a greater range during the older period. ENSO thus depended more on the tilt of the equatorial Pacific thermocline than on the east-to-west temperature gradient, as previously thought. — HJS

Science, this issue p. 255

WOMEN IN SCIENCE
Women's participation and attitudes to talent

Some scientific disciplines have lower percentages of women in academia than others. Leslie *et al.* hypothesized that general

attitudes about the discipline would reflect the representation of women in those fields (see the Perspective by Penner). Surveys revealed that some fields are believed to require attributes such as brilliance and genius, whereas other fields are believed to require more empathy or hard work. In fields where people thought that raw talent was required, academic departments had lower percentages of women. — PJH

Science, this issue p. 262; see also p. 234

APPLIED PHYSICS
Tunnel through and emit coherently

The generation of coherent light (lasers and masers) forms the basis of a large optics industry. Liu *et al.* demonstrate a type of laser that is driven by the tunneling of single electrons in semiconductor double-quantum dots. Distinct from other existing semiconductor lasers, the emission mechanism is driven by tunneling of single charges between discrete energy levels that are electrically tunable. The ability to tune the levels by single-electron charging would allow their laser (or maser) to be turned on and off rapidly. — ISO

Science, this issue p. 285

TOPOLOGICAL MATTER
Nailing down the topology of a semimetal

Topological insulators are exotic materials that have a conducting surface state that can withstand certain types of material imperfection. Theoreticians have predicted a different kind of surface state in related three-dimensional topological Dirac semimetals, which do not have an energy gap in the band structure of the bulk. Xu *et al.* used photoemission spectroscopy to map out the band structure of the material Na_3Bi and detected the predicted surface state. Their results may lead to further insights into the physics of topological matter. — JS

Science, this issue p. 294

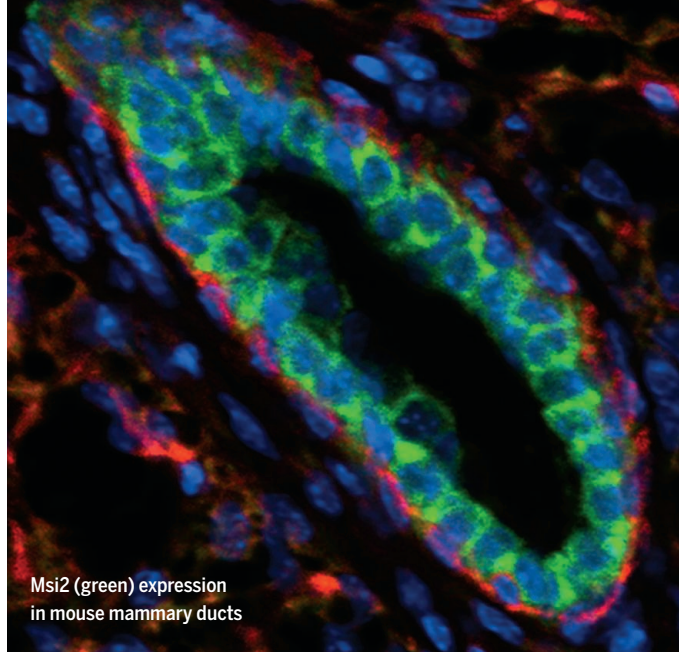*Edited by* **Kristen Mueller** *and* **Jesse Smith**

Lettuce plants are also a source of natural rubber

PLANT SCIENCE
Building rubber without rubber trees

Natural rubber trumps synthetic, petroleum-derived rubber in useful qualities such as elasticity and abrasion resistance. But rubber trees are quite susceptible to disease, leading scientists to search for other sources of natural rubber and to understand the specifics of its biosynthesis. Lettuce plants are a source of natural rubber, and now Qu *et al.* identify a scaffold protein called CPTL2 that keeps rubber-synthesizing enzymes from bouncing around the cell. Tethered to the cell's endoplasmic reticulum, CPTL2 anchors a protein important for rubber polymerization in place. Plants with reduced expression of CPTL2 could not synthesize rubber, revealing its essential role. — PJH

J. Biol. Chem. 10.1074/jbc.M114.616920 (2014).

Msi2 (green) expression in mouse mammary ducts

CANCER

Blocking tumor cell transitions

A process called the epithelial-to-mesenchymal transition allows cancer cells to invade nearby tissue and metastasize. Katz *et al*. investigated the molecular underpinnings of this process and determined that tumors with a more epithelial gene signature expressed higher amounts of the Musashi (Msi) family of RNA-binding proteins as compared to tumors with a more mesenchymal profile. High amounts of Msi proteins locked mammary cells in a state that was less mobile and incapable of differentiation. Reducing Msi expression allowed breast cancer cells to transition into a mesenchymal-like form. Msi proteins may impose an undifferentiated state on cancers, as these proteins are also abundant in neural stem cells. — LC

eLife **3**, e03915 (2014).

ENDOSOMAL SORTING

The ER gets in on an endocytic sorting act

Endosomes are organelles that carry protein cargo from the plasma membrane and the extracellular fluid into cells. Although endosomes closely associate with the endoplasmic reticulum (ER) in cells, scientists don't understand why. Rowland *et al*. now identify the ER as an unexpected player in the sorting of cargo in the endosome and when endosomes divide (fission). Studying this process in mammalian cells, the authors found that an ER tubule crosses over the endosomal membrane, forming a stable contact at positions where endosomes first constrict and then divide. Such ER wrapped-endosomal contact sites prevent cargo from diffusing inappropriately. Interfering with ER dynamics reduced the efficiency of endosome fission. — SMH

Cell **159**, 1027 (2014).

NEUROSCIENCE

Brains in synchrony help us to communicate

Humans are social beings and often understand each other quite well; however, we still only have a limited knowledge of the brain mechanisms that underlie this astonishing ability. To better understand this, Stolk *et al*. scanned the brains of pairs of people that worked together to complete a specific task by communicating only through a visual display. The participants' brain activity synchronized in an area called the right superior temporal gyrus when they completed familiar but not unfamiliar tasks. These results suggest that establishing mutual understanding relies on spatially and temporally coherent brain activity between the two people communicating. — PRS

Proc. Natl. Acad. Sci. U.S.A. **111,** 18183 (2014).

STAR CLUSTERS

An entire family with the same birthday?

Astronomers often read the history of a star cluster in the shapes that emerge when its members are plotted in color-magnitude space. Tight curves suggest a population with very similar ages, whereas more extended features imply a broader range of ages. Li *et al.* consider the main-sequence turn-off—where stars have exhausted hydrogen in their cores—for the massive cluster NGC 1651 and demonstrate the feature's breadth with plots of Hubble Space Telescope photometry. Surprisingly, their models show that the spread can be manifested only if the population shares one common age. Five other massive clusters show the same phenomenon, which the authors interpret as evidence of populations composed of rapidly rotating stars, which should revise interpretations of color-magnitude diagrams. — MMM

Nature 10.1038/nature13969 (2014).

MATERIALS SCIENCE

For stability just add some debris?

Graphene oxide membranes are made by dissolving graphene sheets—the oxidative exfoliation product of graphite—in water and then filtering the solution to form a stacked film. These membranes can exhibit long-term stability in aqueous environments, though, so why are they stable when their parent sheets are soluble? Yeh *et al.* attribute this stability to the introduction of multivalent cations during the purification process. For example, if anodized aluminum oxide filter discs are used during processing, the discs can corrode, releasing aluminum ions that crosslink the graphene oxide membranes and make them stable in water. In contrast, membranes prepared using Teflon filter discs readily disintegrated. Multivalent cations, such as Mn^{2+}, also can give them additional stability. — MSL

Nat. Chem. 10.1038/ nchem.2145 (2015).

LIPID CHEMISTRY

A more stable phase via triangulation

Cell membranes house many important proteins that are hard to study outside their native environment. Salvati Manni *et al.* now have devised a special building block that stabilizes artificial, membrane-like phases
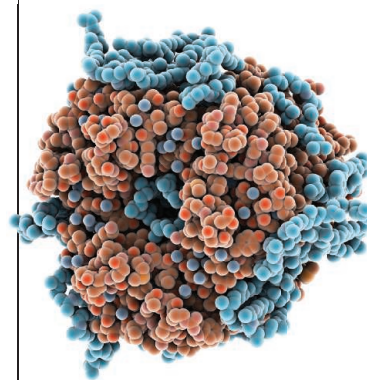
Bacteriorhodopsin, which can be crystalized ex-vivo at low temperatures

at low temperature. Their approach potentially opens the door to more detailed studies of temperature-sensitive membrane proteins. Biological membranes assemble from lipid molecules. The authors induced low-temperature stability by incorporating a rigid triangular ring, or cyclopropyl group, into a more conventional lipid structure. They validated the result by x-ray analysis of embedded bacteriorhodopsin protein at 4°C. — JSY

Angew. Chem. Int. Ed. 10.1002/ anie.201409791 (2014).

REVIEW SUMMARY

MARINE CONSERVATION

Marine defaunation: Animal loss in the global ocean

Douglas J. McCauley,* Malin L. Pinsky, Stephen R. Palumbi, James A. Estes, Francis H. Joyce, Robert R. Warner

BACKGROUND: Comparing patterns of terrestrial and marine defaunation helps to place human impacts on marine fauna in context and to navigate toward recovery. Defaunation began in earnest tens of thousands of years later in the oceans than it did on land. Although defaunation has been less severe in the oceans than on land, our effects on marine animals are increasing in pace and impact. Humans have caused few complete extinctions in the sea, but we are responsible for many ecological, commercial, and local extinctions. Despite our late start, humans have already powerfully changed virtually all major marine ecosystems.

ON OUR WEB SITE

Read the full article at http://dx.doi.org/10.1126/science.1255641

ADVANCES: Humans have profoundly decreased the abundance of both large (e.g., whales) and small (e.g., anchovies) marine fauna. Such declines can generate waves of ecological change that travel both up and down marine food webs and can alter ocean ecosystem functioning. Human harvesters have also been a major force of evolutionary change in the oceans and have reshaped the genetic structure of marine animal populations. Climate change threatens to accelerate marine defaunation over the next century. The high mobility of many marine animals offers some increased, though limited, capacity for marine species to respond to climate stress, but it also exposes many species to increased risk from other stressors. Because humans are intensely reliant on ocean ecosystems for food and other ecosystem services, we are deeply affected by all of these forecasted changes.

Three lessons emerge when comparing the marine and terrestrial defaunation experiences: (i) today's low rates of marine extinction may be the prelude to a major extinction pulse, similar to that observed on land during the industrial revolution, as the footprint of human ocean use widens; (ii) effectively slowing ocean defaunation requires both protected areas and careful management of the intervening ocean matrix; and (iii) the terrestrial experience and current trends in ocean use suggest that habitat destruction is likely to become an increasingly dominant threat to ocean wildlife over the next 150 years.

OUTLOOK: Wildlife populations in the oceans have been badly damaged by human activity. Nevertheless, marine fauna generally are in better condition than terrestrial fauna: Fewer marine animal extinctions have occurred; many geographic ranges have shrunk less; and numerous ocean ecosystems remain more wild than terrestrial ecosystems. Consequently, meaningful rehabilitation of affected marine animal populations remains within the reach of managers. Human dependency on marine wildlife and the linked fate of marine and terrestrial fauna necessitate that we act quickly to slow the advance of marine defaunation. ∎

Timeline (log scale) of marine and terrestrial defaunation. The marine defaunation experience is much less advanced, even though humans have been harvesting ocean wildlife for thousands of years. The recent industrialization of this harvest, however, initiated an era of intense marine wildlife declines. If left unmanaged, we predict that marine habitat alteration, along with climate change (colored bar: IPCC warming), will exacerbate marine defaunation.

MARINE CONSERVATION

Marine defaunation: Animal loss in the global ocean

Douglas J. McCauley,[1]* Malin L. Pinsky,[2] Stephen R. Palumbi,[3] James A. Estes,[4] Francis H. Joyce,[1] Robert R. Warner[1]

Marine defaunation, or human-caused animal loss in the oceans, emerged forcefully only hundreds of years ago, whereas terrestrial defaunation has been occurring far longer. Though humans have caused few global marine extinctions, we have profoundly affected marine wildlife, altering the functioning and provisioning of services in every ocean. Current ocean trends, coupled with terrestrial defaunation lessons, suggest that marine defaunation rates will rapidly intensify as human use of the oceans industrializes. Though protected areas are a powerful tool to harness ocean productivity, especially when designed with future climate in mind, additional management strategies will be required. Overall, habitat degradation is likely to intensify as a major driver of marine wildlife loss. Proactive intervention can avert a marine defaunation disaster of the magnitude observed on land.

Several decades of research on defaunation in terrestrial habitats have revealed a serial loss of mammals, birds, reptiles, and invertebrates that previously played important ecological roles (1). Here, we review the major advancements that have been made in understanding the historical and contemporary processes of similar defaunation in marine environments. We highlight patterns of similarity and difference between marine and terrestrial defaunation profiles to identify better ways to understand, manage, and anticipate the effects of future defaunation in our Anthropocene oceans.

Patterns of marine defaunation

Delayed defaunation in the oceans

Defaunation on land began 10,000 to 100,000 years ago as humans were expanding their range and coming into first contact with novel faunal assemblages (2–4). By contrast, the physical properties of the marine environment limited our capacity early on to access and eliminate marine animal species. This difficulty notwithstanding, humans began harvesting marine animals at least 40,000 years ago, a development that some have suggested was a defining feature in becoming "fully modern humans" (5). Even this early harvest affected local marine fauna (6). However, global rates of marine defaunation only intensified in the last century with the advent of industrial fishing and the rapid expansion of coastal populations (7). As a result, extant global marine faunal assemblages remain today more Pleistocene-like, at least with respect to species composition, than terrestrial fauna. The delayed onset of intensive global marine defaunation is most visible in a comparative chronology of faunal extinctions in which humans are likely to have directly or indirectly played a role (8) (Fig. 1).

Comparing rates of animal extinction

Despite the recent acceleration of marine defaunation, rates of outright marine extinction have been relatively low. The International Union for Conservation of Nature (IUCN) records only 15 global extinctions of marine animal species in the past 514 years (i.e., limit of IUCN temporal coverage) and none in the past five decades (8, 9). By contrast, the IUCN recognizes 514 extinctions of terrestrial animals during the same period (Fig. 1). While approximately six times more animal species have been cataloged on land than in the oceans (10), this imbalance does not explain the 36-fold difference between terrestrial and marine animal extinctions.

It is important to note that the status of only a small fraction of described marine animal species have been evaluated by the IUCN, and many assessed species were determined to be data deficient (11) (Fig. 2). This lack of information necessitates that officially reported numbers of extinct and endangered marine fauna be considered as minimum estimates (11). There remain, however, a number of data-independent explanations for the lower extinction rates of marine fauna. Marine species, for instance, tend to be more widespread, exhibit less endemism, and have higher dispersal (12, 13).

Complacency about the magnitude of contemporary marine extinctions is, however, ill-advised. If we disregard the >50,000-year head start of intense terrestrial defaunation (Fig. 1) and compare only contemporary rates of extinction on land and in the sea, a cautionary lesson emerges. Marine extinction rates today look similar to the moderate levels of terrestrial extinction observed before the industrial revolution (fig. S1). Rates of extinction on land increased dramatically after this period, and we may now be sitting at the precipice of a similar extinction transition in the oceans.

Three other kinds of extinction

The small number of species known to be permanently lost from the world's oceans inadequately
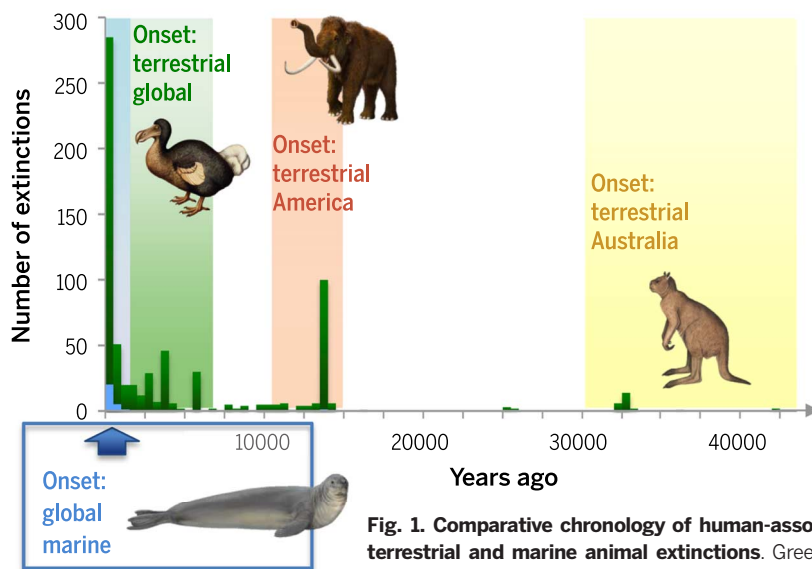

Fig. 1. Comparative chronology of human-associated terrestrial and marine animal extinctions. Green bars indicate animal extinctions that occurred on land, and blue bars indicate marine animal extinctions. Timeline measures years before 2014 CE. Only extinctions occurring less than 55,000 years ago are depicted. Defaunation has ancient origins on land but has intensified only within the last several hundred years in the oceans. See details in (8).

[1]Department of Ecology, Evolution, and Marine Biology, University of California, Santa Barbara, CA 93106, USA. [2]Department of Ecology, Evolution, and Natural Resources, Institute of Marine and Coastal Sciences, Rutgers University, New Brunswick, NJ 08901, USA. [3]Department of Biology, Stanford University, Hopkins Marine Station, Pacific Grove, CA 93950, USA. [4]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060, USA.
*Corresponding author. E-mail: douglas.mccauley@lifesci.ucsb.edu

reflects the full impacts of marine defaunation. We recognize three additional types of defaunation-induced extinction.

Local extinction

Defaunation has caused numerous geographic range constrictions in marine animal species, driving them locally extinct in many habitats. These effects have been particularly severe among large pelagic fishes, where ~90% of studied species have exhibited range contractions (8, 14) (Fig. 3). Local extinctions, however, are not unique to large pelagic predators. Close to a third of the marine fishes and invertebrates off the North American coasts that can be reliably sampled in scientific trawl surveys (often small to medium-bodied species) have also exhibited range contractions (Fig. 3). Such contractions can result from the direct elimination of vulnerable subpopulations or from region-wide declines in abundance (14). Available data suggest that the magnitude of the range contractions for this diverse set of trawl-surveyed marine species is, on average, less than the contractions observed for terrestrial animals such as mammals, birds, and butterflies. Though data deficiencies are abundant, most marine animal species also do not yet seem to exhibit some of the most extreme range contractions recorded for terrestrial animals. Asian tigers, for example, have lost ~93% of their historical range (15), whereas tiger sharks still range across the world's oceans (16).

Ecological extinction

Reductions in the abundance of marine animals have been well documented in the oceans (17). Aggregated population trend data suggest that in the last four decades, marine vertebrates (fish, seabirds, sea turtles, and marine mammals) have declined in abundance by on average 22% (18). Marine fishes have declined in aggregate by 38% (17), and certain baleen whales by 80 to 90% (19). Many of these declines have been termed ecological extinctions—although the species in question are still extant, they are no longer sufficiently abundant to perform their functional roles. Ecological extinctions are well known in terrestrial environments and have been demonstrated to be just as disruptive as species extinctions (20). On land, we know of the phenomenon of "empty forests" where ecological extinctions of forest fauna alter tree recruitment, reshape plant dispersal, and cause population explosions of small mammals (1, 20, 21). We are now observing the proliferation of "empty reefs," "empty estuaries," and "empty bays" (7, 14, 22).

Commercial extinction

Species that drop below an abundance level at which they can be economically harvested are extinct from a commercial standpoint. On land, commercial extinctions affected species ranging from chinchilla to bison (23). Cases of commercial extinction are also common in the oceans. Gray whales were commercially hunted starting in the 1840s. By 1900, their numbers were so depleted that targeted harvest of this species was no longer re-

gionally tenable (24). Likewise, the great whales in Antarctica were serially hunted to commercial extinction (25).

Not all species, however, are so "lucky" as to have human harvesters desist when they become extremely rare. Demand and prices for certain highly prized marine wildlife can continue to increase as these animals become less abundant—a phenomenon termed the anthropogenic Allee effect (26). Individual bluefin tuna can sell for >US$100,000, rare sea cucumbers >US$400/kg, and high-quality shark fins for >US$100/kg. Such species are the rhinos of the ocean—they may never be too rare to be hunted.

Differential vulnerability to defaunation

Are certain marine animals more at risk than others to defaunation? There has been considerable attention given to harvester effects on large marine animals (27). Selective declines of large-bodied animals appear to be evident in certain contexts (28, 29). As a result of such pressures, turtles, whales, sharks, and many large fishes are now ecologically extinct in many ecosystems, and the size spectra (abundance–body mass relationships) of many communities have changed considerably (7, 30, 31). Marine defaunation, however, has not caused many global extinctions of large-bodied species. Most large-bodied marine animal species still exist somewhere in the ocean. By contrast, on land, we have observed the extinction of numerous large terrestrial species and a profound restructuring of the size distribution of land-animal species assemblages. The mean body mass for the list of surviving terrestrial mammal

species, for example, is significantly smaller than the body mass of terrestrial mammal species that lived during the Pleistocene (1, 32). Such effects, however, are not evident for marine mammals (8) (fig. S2). Recent reviews have drawn attention to the fact that humans can also intensely and effectively deplete populations of smaller marine animals (29, 33). These observations have inspired a belated surge in interest in protecting small forage fishes in the oceans.

A review of modern marine extinctions and listings of species on the brink of extinction reveals further insight into aggregate patterns of differential defaunation risk in the oceans (Fig. 2). Sea turtles have the highest proportion of endangered species among commonly recognized groupings of marine fauna. No modern sea turtle species, however, have yet gone extinct. Pinnipeds and marine mustelids, followed very closely by seabirds and shorebirds, have experienced the highest proportion of species extinctions. Many of the most threatened groups of marine animals are those that directly interact with land (and land-based humans) during some portion of their life history (Fig. 2). Terrestrial contact may also explain why diadromous/brackish water fishes are more threatened than exclusively marine fishes (Fig. 2).

Although many marine animal species are clearly affected negatively by marine defaunation, there also appears to be a suite of defaunation "winners," or species that are profiting in Anthropocene oceans. Many of these winners are smaller and "weedier" (e.g., better colonizing and faster reproducing) species. Marine invertebrates, in particular, have often been cited as examples of
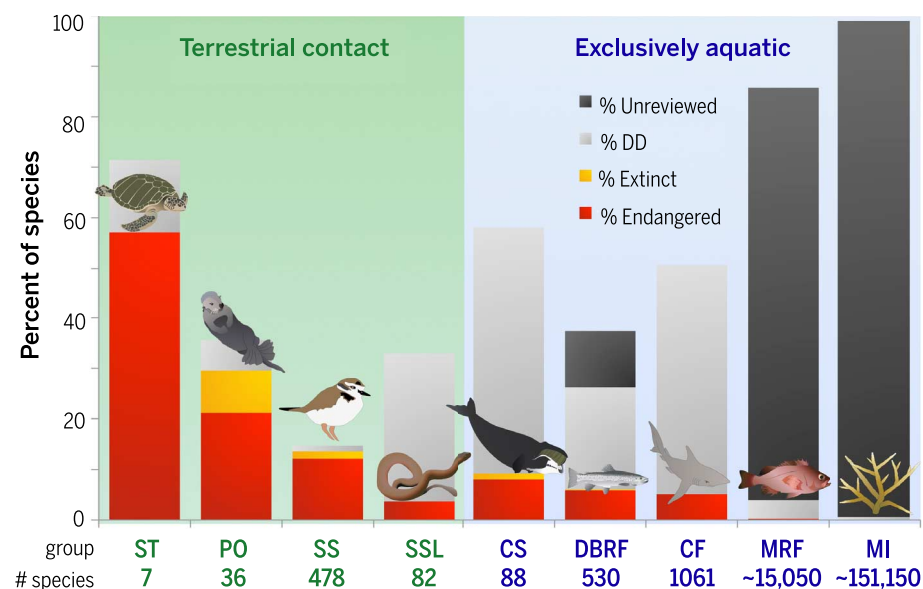

Fig. 2. Marine defaunation threat. Threat from defaunation is portrayed for different groups of marine fauna as chronicled by the IUCN Red List (113). Threat categories include "extinct" (orange), "endangered" (red; IUCN categories "critically endangered" + "endangered"), "data deficient" (light gray), and "unreviewed" (dark gray). Groups that contact land during some portion of their life history (green) are distinguished from species that do not (light blue). The total number of species estimated in each group is listed below the graph. Species groupings are coded as follows: ST, sea turtles; PO, pinnipeds and marine mustelids; SS, seabirds and shorebirds; SSL, sea snakes and marine lizard; CS, cetaceans and sirenians; DBRF, diadromous/brackish ray-finned fishes; CF, cartilaginous fishes; MRF, exclusively marine ray-finned fishes; MI, marine invertebrates. See further details in (8).

species that are succeeding in the face of intense marine defaunation: lobster proliferated as predatory groundfish declined (*34*), prawns increased and replaced the dominance of finfish in landings (*35*), and urchin populations exploded in the absence of their predators and competitors (*36*). Numerous mid-level predators also appear to benefit from the loss of top predators [e.g., small sharks and rays; (*37*)]—a phenomenon analogous to mesopredator release observed in terrestrial spheres (*38*). The status held by some of these defaunation winners in the oceans may, however, be ephemeral. Many of the marine species that have initially flourished as a result of defaunation have themselves become targets for harvest by prey-switching humans as is evidenced by the recent global expansion of marine invertebrate fisheries (*39*).

Spatial patterns of vulnerability

Patterns of marine defaunation risk track differences in the physical environment. Global assessments of human impact on marine ecosystems suggest that coastal wildlife habitats have been more influenced than deep-water or pelagic ecosystems (*40*). The vulnerability of coastal areas presumably results from ease of access to coastal zones. This relationship between access and defaunation risk manifests itself also at smaller spatial scales, with populations of marine wildlife closest to trade networks and human settlements appearing often to be more heavily defaunated (*41*, *42*). The relative insulation that animal populations in regions like the deep oceans presently experience, however, may be short-lived because depletions of shallow-water marine resources and the development of new technologies have created both the capacity and incentive to fish, mine, and drill oil in some of the deepest parts of the sea (*28*, *43*).

Coral reefs, in particular, have consistently been highlighted as marine ecosystems of special concern to defaunation. Coral reefs have been exposed to a wide range of impacts and disturbances, including sedimentation and pollution, thermal stress, disease, destructive fishing, and coastal development (*44*, *45*). Such stressors negatively influence both corals and the millions of species that live within and depend upon reefs (*46*). Risk, however, is not uniform, even across a reefscape. Shallow backreef pools, for example, routinely overheat, and consequently, corals in these parts of the reef are more resistant to ocean warming (*47*). Environmentally heterogeneous areas may in fact act as important natural factories of adaptation that will buffer against some types of marine defaunation.

Effects of marine defaunation
Extended consequences of marine defaunation

Marine defaunation has had far-reaching effects on ocean ecosystems. Depletions of a wide range of ecologically important marine fauna such as cod, sea otters, great whales, and sharks have triggered cascading effects that propagate across marine systems (*37*, *48–51*). Operating in the opposite direction from trophic cascades are changes that travel from the bottom to the top of food chains as a result of the declining abundance of lower–trophic level organisms (*52*). Depletions of fauna such as anchovies, sardines, and krill cause reductions in food for higher-trophic level animals such as seabirds and marine mammals, potentially resulting in losses in reproduction or reductions in their population size (*33*, *53*).

The extended effects of defaunation on marine ecosystems also occur beyond the bounds of these top-down or bottom-up effects. Defaunation can reduce cross-system connectivity (*54*, *55*), decrease ecosystem stability (*56*), and alter patterns of biogeochemical cycling (*57*). The ill effects of food web disarticulation can be further amplified when they occur in association with other marine disturbances. For example, mass releases of discarded plant fertilizers into marine ecosystems from which defaunation has eliminated important consumers can create "productivity explosions" by fueling overgrowth of microbes and algae that fail to be routed into food webs (*58*, *59*).

The selective force of human predation has also been sufficiently strong and protracted so as to have altered the evolutionary trajectory of numerous species of harvested marine fauna (*60*). Harvest has driven many marine animal species to become smaller and thinner, to grow more slowly, to be less fecund, and to reproduce at smaller sizes (*61*). There is also evidence that harvest can reduce the genetic diversity of many marine animal populations (*62*). The genetic effects of defaunation represent a loss of adaptive potential that may impair the resilient capacity of ocean wildlife (*63*).

Importance of marine defaunation to humans

Marine defaunation is already affecting human well-being in numerous ways by imperiling food sustainability, increasing social conflict, impairing storm protection, and reducing flows of other ecosystem services (*64*, *65*). The most conspicuous service that marine fauna make to society is the contribution of their own bodies to global diets. Marine animals, primarily fishes, make up a large proportion of global protein intake, and this contribution is especially strong for impoverished coastal nations (*66*). According to the U.N. Food and Agriculture Organization (FAO), 40 times more wild animal biomass is harvested from the oceans than from land (*67*). Declines in this source of free-range marine food represent a major source of concern (*65*).

A diverse array of nonconsumptive services are also conferred to humanity from ocean animals, ranging from carbon storage that is facilitated by whales and sea otters to regional cloud formation that appears to be stimulated by coral emissions of dimethylsulphoniopropionate (DMSP) (*57*, *68*, *69*). Another key service, given forecasts of increasingly intense weather events and sea-level rise, is coastal protection. Coral, oyster, and other living reefs can dissipate up to 97% of the wave energy reaching them, thus protecting built structures and human lives (*70*). In some cases, corals are more than just perimeter buffers; they also serve as the living platform upon which entire countries (e.g., the Maldives, Kiribati, the Marshall Islands) and entire cultures have been founded. Atoll-living human populations in these areas depend on the long-term health of these animate pedestals to literally hold their lives together.

Outlook and ways forward
Will climate change exacerbate marine defaunation?

The implications of climate change upon marine defaunation are shaped by ocean physics. Marine species live in a vast, globally connected fluid
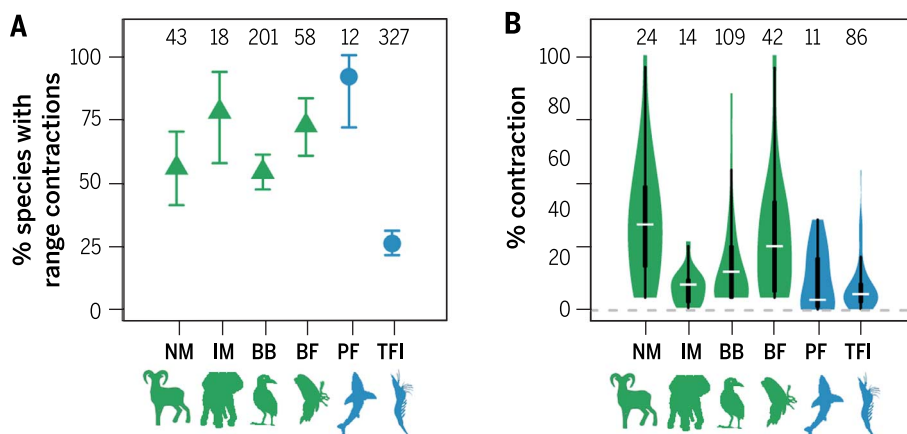

Fig. 3. Comparisons of range contractions for select marine and terrestrial fauna. Terrestrial (green) and marine cases (blue) include evaluations of geographical range change for: 43 North American mammals over the last ~200 years (NM) (*114*), 18 Indian mammals over the last 30 years (IM) (*115*), 201 British birds from ~1970 to 1997 (BB) and 58 British butterflies from ~1976 to 1997 (BF) (*116*), 12 global large pelagic fishes from the 1960s to 2000s (PF) (*14*), and 327 trawl-surveyed North American marine fish and invertebrates from the 1970s to 2000s (TFI). (**A**) Percent of species whose ranges have contracted with binomial confidence intervals and (**B**) distribution of percent contraction for those species that have contracted (violin plot). Sample sizes are shown above each data point, white horizontal lines (B) show the medians, and thick vertical black lines display the interquartile range. See details in (*8*).

medium that has immense heat-storage capacity and has exhibited a historically robust capability to buffer temperature change over daily, annual, and even decadal time scales (71). While this buffering capacity at first seems to confer an advantage to marine fauna, the thermal stability of the oceans may have left many subtidal marine animals poorly prepared, relative to terrestrial counterparts, for the temperature increases associated with global warming. The same logic supports related predictions that terrestrial fauna living in more thermally stable environments will be more vulnerable to warming than those found in areas of greater temperature variability (72).

Ocean warming presents obvious challenges to polar marine fauna trapped in thermal dead ends (73). Tropical marine species are, however, also highly sensitive to small increases in temperature. For example, coastal crabs on tropical shores live closer to their upper thermal maxima than do similar temperate species (74). Likewise, the symbiosis of corals and dinoflagellates is famously sensitive to rapid increases of only 1° to 2°C (75). Even though corals exhibit the capacity for adaptation (47), coral bleaching events are expected to be more common and consequently more stressful by the end of the century (76). The effects of rising ocean temperature extend well beyond coral reefs and are predicted to affect both the adult and juvenile stages of a diverse set of marine species (77), to reshuffle marine community composition (78), and to potentially alter the overall structure and dynamics of entire marine faunal communities (79).

The wide range of other climate change–associated alterations in seawater chemistry and physics—including ocean acidification, anoxia, ocean circulation shifts, changes in stratification, and changes in primary productivity—will fundamentally influence marine fauna. Ocean acidification, for example, makes marine animal shell building more physiologically costly, can

diminish animal sensory abilities, and can alter growth trajectories (80, 81). Climate change impacts on phytoplankton can further accentuate defaunation risk (82). At the same time that humans are reducing the abundance of marine forage fish through direct harvest, we also may be indirectly reducing the planktonic food for forage fish and related consumers in many regions.
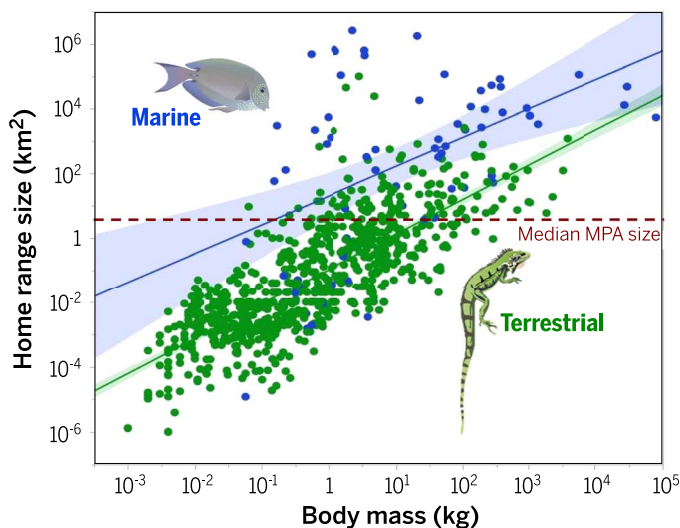
Mobility and managing defaunation

Many marine animals, on average, have significantly larger home ranges as adults [Fig. 4 and figs. S3 and S4; (8)] and disperse greater distances as juveniles than their terrestrial counterparts (13). This wide-ranging behavior of many marine species complicates the management of ocean wildlife as species often traverse multiple management jurisdictions (83–85). On the other hand, the greater mobility of many marine animal species may help them to better follow the velocity of climate change and to colonize and recolonize habitats, so long as source population refuges are kept available (71, 73, 78, 86, 87).

Marine protected areas can offer this sort of refuge for animal populations (88). The establishment of protected areas in the oceans lags far behind advancements made on land, with an upper-bound estimate of only about 3.6% of the world's oceans now protected (8) (fig. S5). One source of optimism for slowing marine defaunation, particularly for mobile species, is that the mean size of marine protected areas has increased greatly in recent years (fig. S5). However, most marine protected areas remain smaller (median 4.5 km^2) than the home range size of many marine animals (Fig. 4). Though much is lost in this type of crude comparison, this observation highlights what may be an important disconnect between the scales at which wildlife use the oceans and the scale at which we typically manage the oceans.

This spatial mismatch is just one of many reasons why protected areas cannot be the full solution for managing defaunation (83). We learned this lesson arguably too late on land. Protected areas can legitimately be viewed as some of our proudest conservation achievements on land (e.g., Yosemite, Serengeti, Chitwan National Parks), and yet with four times more terrestrial area protected than marine protected area, we have still failed to satisfactorily rein in terrestrial defaunation (1) (fig. S5). The realization that more was needed to curb terrestrial defaunation inspired a wave of effort to do conservation out of the bounds of terrestrial protected areas (e.g., conservation easements and corridor projects). The delayed implementation of these strategies has, however, often relegated terrestrial conservation to operating more as a retroactive enterprise aimed at restoring damaged habitats and triaging wildlife losses already underway. In the oceans, we are uniquely positioned to preemptively manage defaunation. We can learn from the terrestrial defaunation experience that protected areas are valuable tools, but that we must proactively introduce measures to manage our impacts on marine fauna in the vast majority of the global oceans that is unprotected.

Strategies to meet these goals include incentive-based fisheries management policies (89), spatially ambitious ecosystem-based management plans (83), and emerging efforts to preemptively zone human activities that affect marine wildlife (90, 91). There have been mixed responses among marine managers as to whether and how to embrace these tools, but more complete implementation of these strategies will help chart a sustainable future for marine wildlife (43, 90, 91). A second, complementary set of goals is to incorporate climate change into marine protected area schemes to build networks that will provide protection for ocean wildlife into the next century (92). Such built-in climate plans were unavailable, and even unthinkable, when many major terrestrial parks were laid out, but data, tools, and opportunity exist to do this thoughtfully now in the oceans.

Habitat degradation: The coming threat to marine fauna

Many early extinctions of terrestrial fauna are believed to have been heavily influenced by human hunting (2, 93), whereas habitat loss appears to be the primary driver of contemporary defaunation on land (1, 11, 86, 94). By contrast, marine defaunation today remains mainly driven by human harvest (95, 96). If the trajectory of terrestrial defaunation is any indicator, we should anticipate that habitat alteration will ascend in importance as a future driver of marine defaunation.

Signs that the pace of marine habitat modification is accelerating and may be posing a growing threat to marine fauna are already apparent (Fig. 5). Great whale species, no longer extensively hunted, are now threatened by noise disruption, oil exploration, vessel traffic, and entanglement with moored marine gear (fig. S6) (97). Habitat-modifying fishing practices (e.g., bottom trawling)

Fig. 4. Mobility of terrestrial and marine fauna. Because mobility shapes defaunation risk, we compare the size-standardized home range size of a representative selection of marine (blue) and terrestrial (green) vertebrates. Data are presented for adults over a full range of animal body sizes, plotted on a logarithmic scale. Species include seabirds, marine reptiles, marine fishes, marine mammals, terrestrial birds, terrestrial reptiles, and

terrestrial mammals (see details in (8); table S2 and fig. S3). Regression lines enclosed by shaded confidence intervals are plotted for all marine and all terrestrial species. The dotted red line demarcates the current median size of all marine protected areas (MPAs).

have affected ~50 million km² of seafloor (*40*). Trawling may represent just the beginning of our capacity to alter marine habitats. Development of coastal cities, where ~40% of the human population lives (*98*), has an insatiable demand for coastal land. Countries like the United Arab Emirates and China have elected to meet this demand by "seasteading"—constructing ambitious new artificial lands in the ocean (*99*). Technological advancement in seafloor mining, dredging, oil and gas extraction, tidal/wave energy generation, and marine transport is fueling rapid expansion of these marine industries (*43*, *100*). Even farming is increasing in the sea. Projections now suggest that in less than 20 years, aquaculture will provide more fish for human consumption than wild capture fisheries (*101*). Fish farming, like crop farming, can consume or drastically alter natural habitats when carried out carelessly (*102*). Many of these emerging marine development activities are reminiscent of the types of rapid environmental change observed on land during the industrial revolution that were associated with pronounced increases in rates of terrestrial defaunation. Marine habitats may eventually join the ranks of terrestrial frontier areas, such as the American West, the Brazilian Amazon, and Alaska, which were once believed to be impervious to development, pollution, and degradation.

Land to sea defaunation connections

The ecologies of marine and terrestrial systems are dynamically linked. Impacts on terrestrial fauna can perturb the ecology of marine fauna (*54*) and vice versa (*103*). Furthermore, the health of marine animal populations is interactively connected to the health of terrestrial wildlife populations—and to the health of society. People in West Africa, for example, exploit wild terrestrial fauna more heavily in years when marine fauna are in short supply (*104*). It is not yet clear how these linkages between marine and terrestrial defaunation will play out at the global level. Will decreasing yields from marine fisheries, for example, require that more terrestrial wildlands be brought into human service as fields and pastures to meet shortfalls of ocean-derived foods? Marine ecosystem managers would do well to better incorporate considerations of land-to-sea defaunation connections in decision making.

Not all bad news

It is easy to focus on the negative course that defaunation has taken in the oceans. Humans have, however, demonstrated a powerful capacity to reverse some of the most severe impacts that we have had on ocean fauna, and many marine wildlife populations demonstrate immense potential for resilience (*47*, *105–107*). The sea otter, the ecological czar of many coastal ecosystems, was thought to be extinct in the early 1900s but was rediscovered in 1938, protected, and has resumed its key ecological role in large parts of the coastal North Pacific and Bering Sea (*108*). The reef ecosystems of Enewetak and Bikini Atolls present another potent example. The United States detonated 66 nuclear explosions above and below the water of these coral reefs in the 1940s and 1950s. Less than 50 years later, the coral and reef fish fauna on these reefs recovered to the point where they were being described as remarkably healthy (*109*).

There is great reason to worry, however, that we are beginning to erode some of the systemic resilience of marine animal communities (*110*). Atomic attacks on local marine fauna are one thing, but an unimpeded transition toward an era of global chemical warfare on marine ecosystems (e.g., ocean acidification, anoxia) may retard or arrest the intrinsic capacity of marine fauna to bounce back from defaunation (*75*, *111*).

Conclusions

On many levels, defaunation in the oceans has, to date, been less severe than defaunation on land. Developing this contrast is useful because our more advanced terrestrial defaunation experience can serve as a harbinger for the possible future of marine defaunation (*3*). Humans have had profoundly deleterious impacts on marine animal populations, but there is still time and there exist mechanisms to avert the kinds of defaunation disasters observed on land. Few marine extinctions have occurred; many subtidal marine habitats are today less developed, less polluted, and more wild than their terrestrial counterparts; global body size distributions of extant marine animal species have been mostly unchanged in the oceans; and many marine fauna have not yet experienced range contractions as severe as those observed on land.

We are not necessarily doomed to helplessly recapitulate the defaunation processes observed on land in the oceans: intensifying marine hunting until it becomes untenable and then embarking on an era of large-scale marine habitat modification. However, if these actions move forward in tandem, we may finally trigger a wave of marine extinctions of the same intensity as that observed on land. Efforts to slow climate change, rebuild affected animal populations, and intelligently engage the coming wave of new marine development activities will all help to change the
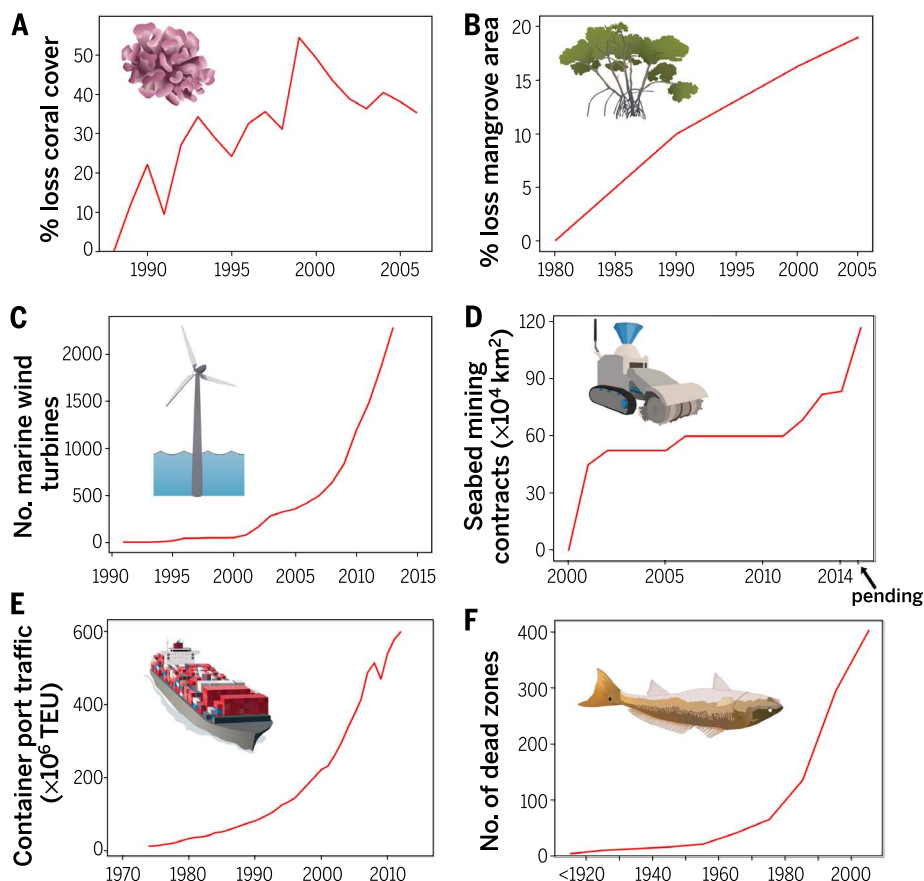

Fig. 5. Habitat change in the global oceans. Trends in six indicators of marine habitat modification suggest that habitat change may become an increasingly important threat to marine wildlife: (**A**) change in global percent cover of coral reef outside of marine protected areas [percent change at each time point measured relative to percent coral cover in 1988 (*44*)]; (**B**) global change in mangrove area (percent change each year measured relative to mangrove area in 1980) (*117*); (**C**) change in the cumulative number of marine wind turbines installed worldwide (*118*); (**D**) change in the cumulative area of seabed under contract for mineral extraction in international waters (*119*); (**E**) trends in the volume of global container port traffic (*120*); and (**F**) change in the cumulative number of oxygen depleted marine "dead zones." See details and data sources in (*8*).

present course of marine defaunation. We must play catch-up in the realm of marine protected area establishment, tailoring them to be operational in our changing oceans. We must also carefully construct marine spatial management plans for the vast regions in between these areas to help ensure that marine mining, energy development, and intensive aquaculture take important marine wildlife habitats into consideration, not vice versa. All of this is a tall order, but the oceans remain relatively full of the raw faunal ingredients and still contain a sufficient degree of resilient capacity so that the goal of reversing the current crisis of marine defaunation remains within reach. The next several decades will be those in which we choose the fate of the future of marine wildlife.

REFERENCES AND NOTES

1. R. Dirzo et al., Defaunation in the Anthropocene. Science **345**, 401–406 (2014). doi: 10.1126/science.1251817; pmid: 25061202

2. A. D. Barnosky, P. L. Koch, R. S. Feranec, S. L. Wing, A. B. Shabel, Assessing the causes of late Pleistocene extinctions on the continents. Science **306**, 70–75 (2004). doi: 10.1126/science.1101476; pmid: 15459379

3. P. L. Koch, A. D. Barnosky, Late Quaternary extinctions: State of the debate. Annu. Rev. Ecol. Evol. Syst. **37**, 215–250 (2006). doi: 10.1146/annurev.ecolsys.34.011802.132415

4. A. D. Barnosky et al., Prelude to the Anthropocene: Two new North American land mammal ages (NALMAs). Anthropol. Rev. (2014). doi: 10.1177/2053019614547433.

5. S. O'Connor, R. Ono, C. Clarkson, Pelagic fishing at 42,000 years before the present and the maritime skills of modern humans. Science **334**, 1117–1121 (2011). doi: 10.1126/science.1207703; pmid: 22116883

6. H. K. Lotze et al., Depletion, degradation, and recovery potential of estuaries and coastal seas. Science **312**, 1806–1809 (2006). doi: 10.1126/science.1128035; pmid: 16794081

7. J. B. C. Jackson et al., Historical overfishing and the recent collapse of coastal ecosystems. Science **293**, 629–637 (2001). doi: 10.1126/science.1059199; pmid: 11474098

8. Materials and methods are available as supplementary materials on Science Online.

9. IUCN, The IUCN Red List of Threatened Species, Version 2014.2 (2014); available at www.iucnredlist.org/.

10. C. Mora, D. P. Tittensor, S. Adl, A. G. B. Simpson, B. Worm, How many species are there on Earth and in the ocean? PLOS Biol. **9**, e1001127 (2011). doi: 10.1371/journal.pbio.1001127; pmid: 21886479

11. S. L. Pimm et al., The biodiversity of species and their rates of extinction, distribution, and protection. Science **344**, 1246752 (2014). doi: 10.1126/science.1246752; pmid: 24876501

12. E. Alison Kay, S. R. Palumbi, Endemism and evolution in Hawaiian marine invertebrates. Trends Ecol. Evol. **2**, 183–186 (1987). doi: 10.1016/0169-5347(87)90017-6; pmid: 21227847

13. B. P. Kinlan, S. D. Gaines, Propagule dispersal in marine and terrestrial environments: A community perspective. Ecology **84**, 2007–2020 (2003). doi: 10.1890/01-0622

14. B. Worm, D. P. Tittensor, Range contraction in large pelagic predators. Proc. Natl. Acad. Sci. U.S.A. **108**, 11942–11947 (2011). doi: 10.1073/pnas.1102353108; pmid: 21693644

15. E. Dinerstein et al., The Fate of wild tigers. Bioscience **57**, 508–514 (2007). doi: 10.1641/B570608

16. S. L. Fowler et al., Eds., Sharks, Rays and Chimaeras: The Status of the Chondrichthyan Fishes. Status Survey (IUCN, Cambridge, UK, 2005).

17. J. A. Hutchings, C. Minto, D. Ricard, J. K. Baum, O. P. Jensen, Trends in the abundance of marine fishes. Can. J. Fish. Aquat. Sci. **67**, 1205–1210 (2010). doi: 10.1139/F10-081

18. WWF, Living Planet Report (WWF International, Gland, Switzerland, 2012); http://www.panda.org/lpr.

19. A. M. Springer et al., Sequential megafaunal collapse in the North Pacific Ocean: An ongoing legacy of industrial whaling? Proc. Natl. Acad. Sci. U.S.A. **100**, 12223–12228 (2003). doi: 10.1073/pnas.1635156100; pmid: 14526101

20. K. H. Redford, The empty forest. Bioscience **42**, 412–422 (1992). doi: 10.2307/1311860

21. H. S. Young et al., Declines in large wildlife increase landscape-level prevalence of rodent-borne disease in Africa. Proc. Natl. Acad. Sci. U.S.A. **111**, 7036–7041 (2014). doi: 10.1073/pnas.1404958111; pmid: 24778215

22. D. J. McCauley et al., Positive and negative effects of a threatened parrotfish on reef ecosystems. Conserv. Biol. **28**, 1312–1321 (2014). doi: 10.1111/cobi.12314; pmid: 25065396

23. J. E. Jiménez, The extirpation and current status of wild chinchillas Chinchilla lanigera and C. brevicaudata. Biol. Conserv. **77**, 1–6 (1996). doi: 10.1016/0006-3207(95)00116-6

24. R. R. Reeves, T. D. Smith, Commercial whaling, especially for gray whales, Eschrichtius robustus, and humpback whales, Megaptera novaeangliae, at California and Baja California shore stations in the 19th century (1854–1899). Mar. Fish. Rev. **72**, 1–25 (2010).

25. R. Hilborn et al., State of the world's fisheries. Annu. Rev. Environ. Resour. **28**, 359–399 (2003). doi: 10.1146/annurev.energy.28.050302.105509

26. F. Courchamp et al., Rarity value and species extinction: The anthropogenic Allee effect. PLOS Biol. **4**, e415 (2006). doi: 10.1371/journal.pbio.0040415; pmid: 17132047

27. D. Pauly, V. Christensen, J. Dalsgaard, R. Froese, F. Torres Jr., Fishing down marine food webs. Science **279**, 860–863 (1998). doi: 10.1126/science.279.5352.860; pmid: 9452385

28. S. A. Sethi, T. A. Branch, R. Watson, Global fishery development patterns are driven by profit but not trophic level. Proc. Natl. Acad. Sci. U.S.A. **107**, 12163–12167 (2010). doi: 10.1073/pnas.1003236107; pmid: 20566867

29. M. L. Pinsky, O. P. Jensen, D. Ricard, S. R. Palumbi, Unexpected patterns of fisheries collapse in the world's oceans. Proc. Natl. Acad. Sci. U.S.A. **108**, 8317–8322 (2011). doi: 10.1073/pnas.1015313108; pmid: 21536889

30. J. B. C. Jackson, Ecological extinction and evolution in the brave new ocean. Proc. Natl. Acad. Sci. U.S.A. **105** (suppl. 1), 11458–11465 (2008). doi: 10.1073/pnas.0802812105; pmid: 18695220

31. S. Jennings, J. L. Blanchard, Fish abundance with no fishing: Predictions based on macroecological theory. J. Anim. Ecol. **73**, 632–642 (2004). doi: 10.1111/j.0021-8790.2004.00839.x

32. S. K. Lyons, F. A. Smith, J. H. Brown, Of mice, mastodons and men: Human-mediated extinctions on four continents. Evol. Ecol. Res. **6**, 339–358 (2004).

33. A. D. Smith et al., Impacts of fishing low-trophic level species on marine ecosystems. Science **333**, 1147–1150 (2011). doi: 10.1126/science.1209395; pmid: 21778363

34. R. S. Steneck, J. Vavrinec, A. V. Leland, Accelerating trophic-level dysfunction in kelp forest ecosystems of the western North Atlantic. Ecosystems **7**, 323–332 (2004). doi: 10.1007/s10021-004-0240-6

35. B. Worm, R. A. Myers, Meta-analysis of cod–shrimp interactions reveals top-down control in oceanic food webs. Ecology **84**, 162–173 (2003). doi: 10.1890/0012-9658(2003)084[0162:MAOCSI]2.0.CO;2

36. D. O. Duggins, C. A. Simenstad, J. A. Estes, Magnification of secondary production by kelp detritus in coastal marine ecosystems. Science **245**, 170–173 (1989). doi: 10.1126/science.245.4914.170; pmid: 17787876

37. R. A. Myers, J. K. Baum, T. D. Shepherd, S. P. Powers, C. H. Peterson, Cascading effects of the loss of apex predatory sharks from a coastal ocean. Science **315**, 1846–1850 (2007). doi: 10.1126/science.1138657; pmid: 17395829

38. L. R. Prugh et al., The Rise of the Mesopredator. Bioscience **59**, 779–791 (2009). doi: 10.1525/bio.2009.59.9.9

39. S. C. Anderson, J. Mills Flemming, R. Watson, H. K. Lotze, Rapid global expansion of invertebrate fisheries: Trends, drivers, and ecosystem effects. PLOS ONE **6**, e14735 (2011). doi: 10.1371/journal.pone.0014735; pmid: 21408090

40. B. S. Halpern et al., A global map of human impact on marine ecosystems. Science **319**, 948–952 (2008). doi: 10.1126/science.1149345; pmid: 18276889

41. D. J. McCauley et al., Conservation at the edges of the world. Biol. Conserv. **165**, 139–145 (2013). doi: 10.1016/j.biocon.2013.05.026

42. J. E. Cinner, N. A. J. Graham, C. Huchery, M. A. Macneil, Global effects of local human population density and distance to markets on the condition of coral reef fisheries. Conserv. Biol. **27**, 453–458 (2013). doi: 10.1111/j.1523-1739.2012.01933.x; pmid: 23025334

43. K. J. Mengerink et al., A call for deep-ocean stewardship. Science **344**, 696–698 (2014). doi: 10.1126/science.1251458; pmid: 24833377

44. E. R. Selig, K. S. Casey, J. F. Bruno, Temperature-driven coral decline: The role of marine protected areas. Glob. Change Biol. **18**, 1561–1570 (2012). doi: 10.1111/j.1365-2486.2012.02658.x

45. J. F. Bruno, E. R. Selig, Regional decline of coral cover in the Indo-Pacific: Timing, extent, and subregional comparisons. PLOS ONE **2**, e711 (2007). doi: 10.1371/journal.pone.0000711; pmid: 17684557

46. C. M. Roberts et al., Marine biodiversity hotspots and conservation priorities for tropical reefs. Science **295**, 1280–1284 (2002). doi: 10.1126/science.1067728; pmid: 11847338

47. S. R. Palumbi, D. J. Barshis, N. Traylor-Knowles, R. A. Bay, Mechanisms of reef coral resistance to future climate change. Science **344**, 895–898 (2014). doi: 10.1126/science.1251336; pmid: 24762535

48. J. A. Estes et al., Trophic downgrading of planet Earth. Science **333**, 301–306 (2011). doi: 10.1126/science.1205106; pmid: 21764740

49. D. J. McCauley et al., Acute effects of removing large fish from a near-pristine coral reef. Mar. Biol. **157**, 2739–2750 (2010). doi: 10.1007/s00227-010-1533-2; pmid: 24391253

50. J. A. Estes, Whales, Whaling, and Ocean Ecosystems (Univ. of California Press, Berkeley, CA, 2006).

51. J. Terborgh, J. A. Estes, Trophic Cascades: Predators, Prey, and the Changing Dynamics of Nature (Island Press, Washington, DC, 2013).

52. D. J. McCauley, F. Keesing, T. P. Young, B. F. Allan, R. M. Pringle, Indirect effects of large herbivores on snakes in an African savanna. Ecology **87**, 2657–2663 (2006). doi: 10.1890/0012-9658(2006)87[2657:IEOLHO]2.0.CO;2; pmid: 17089673

53. P. M. Cury et al., Global seabird response to forage fish depletion—one-third for the birds. Science **334**, 1703–1706 (2011). doi: 10.1126/science.1212928; pmid: 22194577

54. D. J. McCauley et al., From wing to wing: The persistence of long ecological interaction chains in less-disturbed ecosystems. Sci. Rep. **2**, 409 (2012). doi: 10.1038/srep00409

55. D. J. McCauley et al., Assessing the effects of large mobile predators on ecosystem connectivity. Ecol. Appl. **22**, 1711–1717 (2012). doi: 10.1890/11-1653.1; pmid: 23092009

56. G. L. Britten et al., Predator decline leads to decreased stability in a coastal fish community. Ecol. Lett. **17**, 1518–1525 (2014). doi: 10.1111/ele.12354

57. J. Roman et al., Whales as marine ecosystem engineers. Front. Ecol. Environ **12**, 377–385 (2014). doi: 10.1890/130220

58. J. M. Pandolfi et al., Ecology. Are U.S. coral reefs on the slippery slope to slime? Science **307**, 1725–1726 (2005). doi: 10.1126/science.1104258; pmid: 15774744

59. S. R. Palumbi, A. R. Palumbi, The Extreme Life of the Sea (Princeton Univ. Press, Princeton, NJ, 2014).

60. S. R. Palumbi, Humans as the world's greatest evolutionary force. Science **293**, 1786–1790 (2001). doi: 10.1126/science.293.5536.1786; pmid: 11546863

61. C. Jørgensen et al., Ecology: Managing evolving fish stocks. Science **318**, 1247–1248 (2007). doi: 10.1126/science.1148089; pmid: 18033868

62. M. L. Pinsky, S. R. Palumbi, Meta-analysis reveals lower genetic diversity in overfished populations. Mol. Ecol. **23**, 29–39 (2014). doi: 10.1111/mec.12509; pmid: 24372754

63. M. R. Walsh, S. B. Munch, S. Chiba, D. O. Conover, Maladaptive changes in multiple traits caused by fishing: Impediments to population recovery. Ecol. Lett. **9**, 142–148 (2006). doi: 10.1111/j.1461-0248.2005.00858.x; pmid: 16958879

64. B. Worm et al., Impacts of biodiversity loss on ocean ecosystem services. Science **314**, 787–790 (2006). doi: 10.1126/science.1132294; pmid: 17082450

65. J. S. Brashares et al., Wildlife decline and social conflict. Science **345**, 376–378 (2014). doi: 10.1126/science.1256734; pmid: 25061187

66. Fisheries and Aquaculture Department, FAO, "The State of World Fisheries and Aquaculture 2012" (FAO, Rome, 2012).

67. FAO, FAOSTAT (2012); available at http://faostat3.fao.org/faostat-gateway/go/to/home/E.

68. C. C. Wilmers, J. A. Estes, M. Edwards, K. L. Laidre, B. Konar, Do trophic cascades affect the storage and flux of atmospheric carbon? An analysis of sea otters and kelp forests. Front. Ecol. Environ **10**, 409–415 (2012). doi: 10.1890/110176

69. J.-B. Raina et al., DMSP biosynthesis by an animal and its role in coral thermal stress response. Nature **502**, 677–680 (2013). doi: 10.1038/nature12677; pmid: 24153189

70. F. Ferrario et al., The effectiveness of coral reefs for coastal hazard risk reduction and adaptation. Nat. Commun. 5, 3794 (2014). doi: 10.1038/ncomms4794

71. M. T. Burrows et al., The pace of shifting climate in marine and terrestrial ecosystems. Science 334, 652–655 (2011). doi: 10.1126/science.1210288; pmid: 22053045

72. J. J. Tewksbury, R. B. Huey, C. A. Deutsch, Ecology. Putting the heat on tropical animals. Science 320, 1296–1297 (2008). doi: 10.1126/science.1159328; pmid: 18535231

73. W. W. Cheung et al., Projecting global marine biodiversity impacts under climate change scenarios. Fish Fish. 10, 235–251 (2009). doi: 10.1111/j.1467-2979.2008.00315.x

74. J. H. Stillman, Acclimation capacity underlies susceptibility to climate change. Science 301, 65–65 (2003). doi: 10.1126/science.1083073; pmid: 12843385

75. O. Hoegh-Guldberg et al., Coral reefs under rapid climate change and ocean acidification. Science 318, 1737–1742 (2007). doi: 10.1126/science.1152509; pmid: 18079392

76. C. A. Logan, J. P. Dunne, C. M. Eakin, S. D. Donner, Incorporating adaptive responses into future projections of coral bleaching. Glob. Change Biol. 20, 125–139 (2014). doi: 10.1111/gcb.12390

77. M. I. O'Connor et al., Temperature control of larval dispersal and the implications for marine ecology, evolution, and conservation. Proc. Natl. Acad. Sci. U.S.A. 104, 1266–1271 (2007). doi: 10.1073/pnas.0603422104; pmid: 17213327

78. M. L. Pinsky, B. Worm, M. J. Fogarty, J. L. Sarmiento, S. A. Levin, Marine taxa track local climate velocities. Science 341, 1239–1242 (2013). doi: 10.1126/science.1239352; pmid: 24031017

79. W. W. L. Cheung et al., Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. Nat. Clim. Change 3, 254–258 (2013). doi: 10.1038/nclimate1691

80. T. A. Branch, B. M. DeJoseph, L. J. Ray, C. A. Wagner, Impacts of ocean acidification on marine seafood. Trends Ecol. Evol. 28, 178–186 (2013). doi: 10.1016/j.tree.2012.10.001; pmid: 23122878

81. G. E. Hofmann et al., The effect of ocean acidification on calcifying organisms in marine ecosystems: An organism-to-ecosystem perspective. Annu. Rev. Ecol. Evol. Syst. 41, 127–147 (2010). doi: 10.1146/annurev.ecolsys.110308.120227

82. D. G. Boyce, M. Dowd, M. R. Lewis, B. Worm, Estimating global chlorophyll changes over the past century. Prog. Oceanogr. 122, 163–173 (2014). doi: 10.1016/j.pocean.2014.01.004

83. T. Agardy, G. N. di Sciara, P. Christie, Mind the gap: Addressing the shortcomings of marine protected areas through large scale marine spatial planning. Mar. Policy 35, 226–232 (2011). doi: 10.1016/j.marpol.2010.10.006

84. L. B. Crowder et al., Sustainability. Resolving mismatches in U.S. ocean governance. Science 313, 617–618 (2006). doi: 10.1126/science.1129706; pmid: 16888124

85. D. J. McCauley et al., Reliance of mobile species on sensitive habitats: A case study of manta rays (Manta alfredi) and lagoons. Mar. Biol. 161, 1987–1998 (2014). doi: 10.1007/s00227-014-2478-7

86. M. H. Carr et al., Comparing marine and terrestrial ecosystems: Implications for the design of coastal marine reserves. Ecol. Appl. 13 (suppl.), 90–107 (2003). doi: 10.1890/1051-0761(2003)013[0090:CMATEI]2.0.CO;2

87. M. T. Burrows et al., Geographical limits to species-range shifts are suggested by climate velocity. Nature 507, 492–495 (2014). doi: 10.1038/nature12976; pmid: 24509712

88. G. J. Edgar et al., Global conservation outcomes depend on marine protected areas with five key features. Nature 506, 216–220 (2014). doi: 10.1038/nature13022; pmid: 24499817

89. C. Costello, S. D. Gaines, J. Lynham, Can catch shares prevent fisheries collapse? Science 321, 1678–1681 (2008). doi: 10.1126/science.1159478; pmid: 18801999

90. C. White, B. S. Halpern, C. V. Kappel, Ecosystem service tradeoff analysis reveals the value of marine spatial planning for multiple ocean uses. Proc. Natl. Acad. Sci. U.S.A. 109, 4696–4701 (2012). doi: 10.1073/pnas.1114215109; pmid: 22392996

91. W. Qiu, P. J. S. Jones, The emerging policy landscape for marine spatial planning in Europe. Mar. Policy 39, 182–190 (2013). doi: 10.1016/j.marpol.2012.10.010

92. E. McLeod, R. Salm, A. Green, J. Almany, Designing marine protected area networks to address the impacts of climate change. Front. Ecol. Environ 7, 362–370 (2009). doi: 10.1890/070211

93. S. T. Turvey, Holocene Extinctions (Oxford Univ. Press, New York, 2009).

94. J. Schipper et al., The status of the world's land and marine mammals: Diversity, threat, and knowledge. Science 322, 225–230 (2008). doi: 10.1126/science.1165115; pmid: 18845749

95. N. K. Dulvy, J. K. Pinnegar, J. D. Reynolds, in Holocene Extinctions, S. T. Turvey, Ed. (Oxford University Press, New York, 2009), pp. 129–150.

96. C. V. Kappel, Losing pieces of the puzzle: Threats to marine, estuarine, and diadromous species. Front. Ecol. Environ 3, 275–282 (2005). doi: 10.1890/1540-9295(2005)003[0275:LPOTPT]2.0.CO;2

97. S. D. Kraus et al., Ecology. North Atlantic right whales in crisis. Science 309, 561–562 (2005). doi: 10.1126/science.1111200; pmid: 16040692

98. M. L. Martínez et al., The coasts of our world: Ecological, economic and social importance. Ecol. Econ. 63, 254–272 (2007). doi: 10.1016/j.ecolecon.2006.10.022

99. P. F. Sale et al., The growing need for sustainable ecological management of marine communities of the Persian Gulf. Ambio 40, 4–17 (2011). doi: 10.1007/s13280-010-0092-6; pmid: 21404819

100. A. B. Gill, Offshore renewable energy: Ecological implications of generating electricity in the coastal zone. J. Appl. Ecol. 42, 605–615 (2005). doi: 10.1111/j.1365-2664.2005.01060.x

101. The World Bank, "Fish to 2030: Prospects for fisheries and aquaculture" (83177, The World Bank, 2013), pp. 1–102.

102. D. H. Klinger, R. Naylor, Searching for Solutions in Aquaculture: Charting a Sustainable Course. Annu. Rev. Environ. Resour. 37, 247–276 (2012). doi: 10.1146/annurev-environ-021111-161531

103. M. D. Rose, G. A. Polis, The distribution and abundance of coyotes: The effects of allochthonous food subsidies from the sea. Ecology 79, 998–1007 (1998). doi: 10.1890/0012-9658(1998)079[0998:TDAAOC]2.0.CO;2

104. J. S. Brashares et al., Bushmeat hunting, wildlife declines, and fish supply in West Africa. Science 306, 1180–1183 (2004). doi: 10.1126/science.1102425; pmid: 15539602

105. S. R. Palumbi, C. Sotka, The Death and Life of Monterey Bay: A Story of Revival (Island Press, Washington, DC, 2010).

106. H. K. Lotze, M. Coll, A. M. Magera, C. Ward-Paige, L. Airoldi, Recovery of marine animal populations and ecosystems. Trends Ecol. Evol. 26, 595–605 (2011). doi: 10.1016/j.tree.2011.07.008

107. B. Worm et al., Rebuilding global fisheries. Science 325, 578–585 (2009). doi: 10.1126/science.1173146; pmid: 19644114

108. K. W. Kenyon, The sea otter in the Eastern Pacific Ocean. North Am. Fauna 68, 1–352 (1969). doi: 10.3996/nafa.68.0001

109. J. S. Davis, Scales of Eden: Conservation and pristine devastation on Bikini Atoll. Environ. Plan. Soc. Space 25, 213–235 (2007). doi: 10.1068/d1405

110. B. deYoung et al., Regime shifts in marine ecosystems: Detection, prediction and management. Trends Ecol. Evol. 23, 402–409 (2008). doi: 10.1016/j.tree.2008.03.008; pmid: 18501990

111. T. P. Hughes et al., Phase shifts, herbivory, and the resilience of coral reefs to climate change. Curr. Biol. 17, 360–365 (2007). doi: 10.1016/j.cub.2006.12.049; pmid: 17291763

112. T. Saxby, IAN Image and Video Library, IAN, UMCES; available at ian.umces.edu/imagelibrary/.

113. IUCN, The IUCN Red List of Threatened Species, Version 2013.2 (2013); available at www.iucnredlist.org/.

114. A. S. Laliberte, W. J. Ripple, Range contractions of North American carnivores and ungulates. Bioscience 54, 123–138 (2004). doi: 10.1641/0006-3568(2004)054[0123:RCONAC]2.0.CO;2

115. R. Pillay, A. J. T. Johnsingh, R. Raghunath, M. D. Madhusudan, Patterns of spatiotemporal change in large mammal distribution and abundance in the southern Western Ghats, India. Biol. Conserv. 144, 1567–1576 (2011). doi: 10.1016/j.biocon.2011.01.026

116. J. A. Thomas et al., Comparative losses of British butterflies, birds, and plants and the global extinction crisis. Science 303, 1879–1881 (2004). doi: 10.1126/science.1095046; pmid: 15031508

117. FAO, "The world's mangroves 1980-2005. FAO Forestry Paper 153" (FAO, Rome, 2007).

118. The European Wind Energy Association, Offshore Statistics; available at www.ewea.org/statistics/offshore/.

119. International Seabed Authority, Contractors. Int. Seabed Auth.; available at www.isa.org.jm/deep-seabed-minerals-contractors/overview.

120. The World Bank, Container port traffic (TEU: 20 foot equivalent units); available at http://data.worldbank.org/indicator/IS.SHP.GOOD.TU.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6219/1255641/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 and S2
References (121–298)

10.1126/science.1255641

REPORTS

ASIAN ARCHAEOLOGY

Agriculture facilitated permanent human occupation of the Tibetan Plateau after 3600 B.P.

F. H. Chen,[1]* G. H. Dong,[1]* D. J. Zhang,[1] X. Y. Liu,[7] X. Jia,[1] C. B. An,[1] M. M. Ma,[1] Y. W. Xie,[1] L. Barton,[3] X. Y. Ren,[4] Z. J. Zhao,[5] X. H. Wu,[6] M. K. Jones[2]

Our understanding of when and how humans adapted to living on the Tibetan Plateau at altitudes above 2000 to 3000 meters has been constrained by a paucity of archaeological data. Here we report data sets from the northeastern Tibetan Plateau indicating that the first villages were established only by 5200 calendar years before the present (cal yr B.P.). Using these data, we tested the hypothesis that a novel agropastoral economy facilitated year-round living at higher altitudes since 3600 cal yr B.P. This successful subsistence strategy facilitated the adaptation of farmers-herders to the challenges of global temperature decline during the late Holocene.

The Tibetan Plateau retains traces of an intermittent human presence from at least 20,000 years ago. Much of this comes in the form of surface finds of worked stone (1), but among them are a series of finds with a secure scientific date. Handprints and footprints from this date have been found in the southern plateau (Quesang site) at 4200 m above sea level (masl) (2). Archaeological traces are found from 14,600 calendar years before the present (cal yr B.P.) at Jiangxigou 1 (~3200 masl) and 13,100 cal yr B.P. at Heimahe 1 (~3200 masl) (3), 9200 cal yr B.P. at Xidatan 2 (~4300 masl) (4), and 7500 cal yr B.P. at Yeniugou (3800 masl) in the northeastern part of the plateau (5). The evidence in each case comprises animal bones, stone artefacts, and small-scale hearths. Humans clearly reached those altitudes, plausibly in pursuit of game. Although various potential models for human activity in this early episode have been considered (1), evidence securely dated to this period at these altitudes reflects the use of stone tools, the lighting of fires, and the processing of hunted carcasses. It may be equated with hunting camps, in most cases used for a single episode. Evidence for sustained agricultural and artisanal activity is lacking.

From the 6th millennium B.P. onward, the northeastern Tibetan Plateau (hereafter NETP)

became the principal region of human settlement on the Tibetan Plateau, accounting for 72.4% of its known prehistoric sites (Fig. 1) (6–9). These are primarily the sites of farming settlements, associated with the reaches of the Yellow River and its tributaries. The NETP constitutes

an altitudinal entry point into the higher plateau from the adjacent Loess Plateau (Fig. 1), with which it shares a series of Neolithic and Bronze Age material cultures (10). These include the late Yangshao (5500 to 5000 cal yr B.P.), Majiayao (5300 to 4000 cal yr B.P.), Qijia (4100 to 3600 cal yr B.P.), Xindian (3400 to 2700 cal yr B.P.), Kayue (3600 to 2600 cal yr B.P.), and Nuomuhong (3400 to 2800 cal yr B.P.) cultures (text S1). In order to ascertain during what period, and at what altitude, sustained food production first enabled an enduring human presence, we collected artefacts, animal bones, and plant remains from a selection of sites within these cultures. Fifty-three NETP sites (text S2) were thus selected to provide an optimal chronological and geographical range. To establish a secure chronology, 63 charred grains were collected for accelerator mass spectrometry (AMS) radiocarbon dating (11) (table S1).

We identified charred cereal grains from all 53 sites (foxtail millet, broomcorn millet, barley, and wheat) and animal bones and teeth from 10 sites (sheep, cattle, and pig) (table S1). Among the 53 sites, an earlier group of 25 sites dates to 5200 to 3600 cal yr B.P. (Fig. 2C and table S1) and reaches a maximum elevation of 2527 masl. A later group of 29 sites dating to 3600 to 2300 cal yr B.P. approaches an elevation of 3400 masl, among which 12 sites lie between 2500 to 1700 masl, 9 sites between 3000 to 2500 masl, and 8 sites between 3400 to 3000 masl.
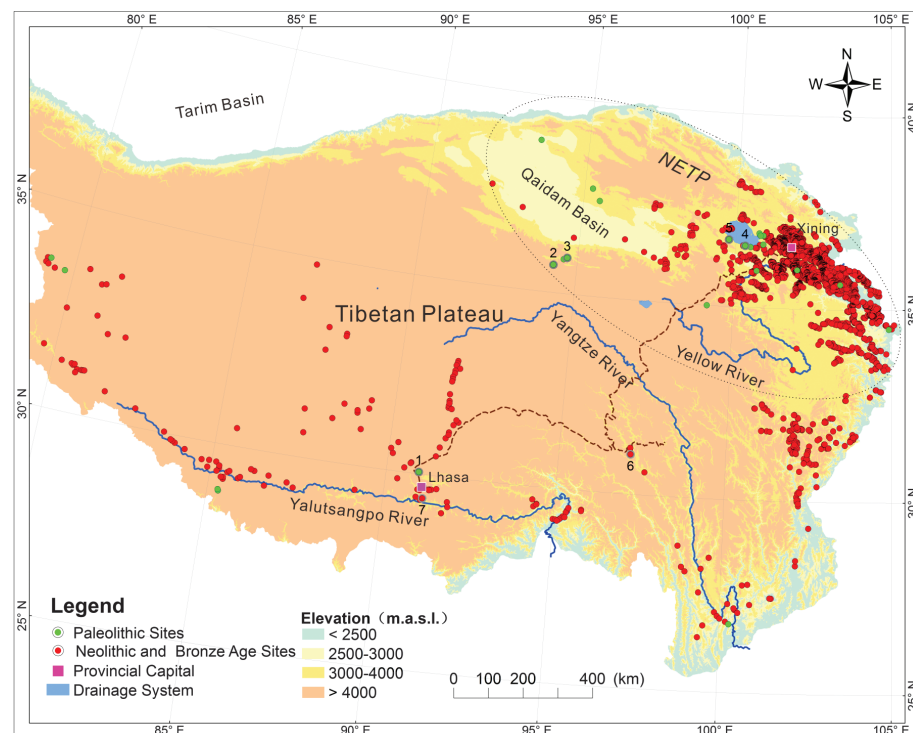
Fig. 1. Distribution of prehistoric sites on the Tibetan Plateau (2–10). The brown dashed line indicates the Tang-Tibetan routeway in use over the past 2000 years. Solid green circles with numbers indicate dated Paleolithic sites in the Tibetan Plateau mentioned in the text: 1, Quesang; 2, Xidatan 2; 3, Yeniugou; 4, Jiangxigou 1; 5, Heimahe 1. Solid red circles with numbers indicate Neolithic and Bronze Age sites mentioned in the text: 6, Karuo; 7, Changguogou.

[1]Key Laboratory of Western China's Environmental Systems (Ministry of Education), Lanzhou University, Lanzhou 730000, China. [2]McDonald Institute of Archaeological Research, University of Cambridge, Cambridge CB2 3ER, UK. [3]Center for Comparative Archaeology, Department of Anthropology, University of Pittsburgh, Pittsburgh, PA 15260, USA. [4]Qinghai Provincial Institute of Cultural Relics and Archaeology, Xining 810007, China. [5]Institute of Archaeology, Chinese Academy of Social Sciences, Beijing 100710, China. [6]School of Archaeology and Museology, Peking University, Beijing 100871, China. [7]Department of Anthropology, Washington University in St. Louis, St. Louis, MO 63130-48, USA.
*Corresponding author. E-mail: fhchen@lzu.edu.cn (F.H.C.); ghdong@lzu.edu.cn (G.H.D.)

The earlier group reflects the widespread settlement within the NETP of farming communities along the Yellow River and its tributaries at elevations below ~2500 masl. Foxtail millet and broomcorn millet account for 98.1% of the recovered charred cereal grains (table S1), indicating that millets constituted the primary crops during this period, a conclusion also supported by stable carbon isotope studies (text S3 and fig. S5). Taken together with the material culture evidence, we can view this occupation of the higher reaches of the Yellow River as an extension of the expansion of millet agriculture across the middle and lower reaches in preceding millennia (12). The apparent limits of farming settlement around 2500 masl in NETP before 3600 cal yr B.P. may in turn be related to the frost sensitivity of millet crops (13, 14), as has been shown for the southwestern Tibetan Plateau (15).

The later groups of sites do not share the same altitudinal constraint. Among the selected samples are sites reaching 3400 masl, and in the wider landscape are contemporary sites with a similar material culture reaching 4700 masl (6). These higher-altitude sites moreover display a shift in the balance of crops among the charred grains. Although the same suite of crops may be found at both lower and higher altitudes, the lower-altitude assemblages are dominated by millet but the higher-altitude assemblages by barley, with an occasional record of wheat and broomcorn millet. Sites located above 3000 masl are also marked by the presence of sheep bones.

The presence of crops and livestock in itself indicates a more sustained human presence than what is needed to hunt game at high altitudes. Although more frost-hardy than millets, barley has a longer growing season, typically requiring 6 months between sowing and harvest (16). Other evidence [for example, of house and tomb construction (10)] further endorses the notion of a sustained and probably year-round human presence.

On the basis of the above evidence, the prehistoric human occupation of the NETP can be subdivided into three phases. During the first phase (pre–5200 cal yr B.P.), hunter-gatherers made occasional forays to altitudes reaching above 4300 masl, presumably tracking game. During the second phase (5200 to 3600 cal yr B.P.), a longstanding tradition of millet farming that had become widely established along the middle and lower reaches of the Yellow River extended upstream into the NETP. Millet farming had spread across the Loess Plateau after 5900 cal yr B.P. (17) and subsequently spread across these lower reaches of the NETP from 5200 cal yr B.P. Toward the end of the second phase (4000 and 3600 cal yr B.P.), two significant additions are observed in the crop repertoire (text S4 and fig. S6). The North Chinese crops of broomcorn and foxtail millet were joined or displaced on some sites by the principal cereals of the Fertile Crescent, barley and wheat. There has been much interest in the chronology and consequences of the meeting of east and west staple crops in prehistory (18–20). Here, its notable consequence was to facilitate the sustained settlement of the Tibetan Plateau's higher altitudes. The importation of wheat and barley enabled human communities to adapt to the harsher conditions of higher altitudes in the Tibetan Plateau, a possibility raised in previous studies (15, 21).

The key addition was barley. During phase three, from around 3600 cal yr B.P., sites can be divided into those that lie above or below 2500 masl. In the lower-altitude group, the longstanding crops, broomcorn and foxtail millet, are joined by barley as a third component in an otherwise traditional dietary repertoire. In the higher-altitude group, however, the frost-sensitive millet is absent, and the cold-tolerant barley has moved to a primary position (Fig. 2D). Alongside the presence of wheat (also relatively cold-tolerant) and sheep, the diet at these high altitudes has clearly been transformed, but in a manner that enabled sustained settlement at unprecedented altitudes.

Fewer sites have been investigated in detail in areas beyond our study area on the plateau; the evidence they yield is consistent with the pattern that has been noted here and in (22). Ecological niche modeling on the southeastern Tibetan Plateau has shown that the warmer condition of more southerly latitudes raised the altitude that millets were able to reach at Karuo between 4700 and 4300 cal BP to 3100 masl (15, 22, 23). During our phase three, agriculture reached even higher elevations in the south-central Tibetan Plateau. For example, Changguogou (Fig. 1) was occupied after 3500 cal yr B.P. at an elevation of 3600 masl, growing a range of Fertile Crescent crops, including naked barley, wheat, oat, rye, and pea alongside foxtail millet (24).

Turning to the climatic context, during the early and middle Holocene the summer monsoon was strong in north China and the climate was generally warm in the Northern Hemisphere (Fig. 2, A and B) (25, 26). These relatively favorable conditions provided a context both for the earlier forays of hunter-gatherers into higher altitudes and the broader expansion of millet agriculture in northern China. The Northern Hemisphere temperature curve displays a significant temperature drop throughout our second phase, reaching a minimum at the start of our third phase (26). In other words, the human expansion into the higher, colder altitudes took place as the continental temperatures had themselves become colder.

In our third phase, the evidence displays two aspects of the human response to this cooling of climatic change. The established farming landscapes of the lower altitudes retained their
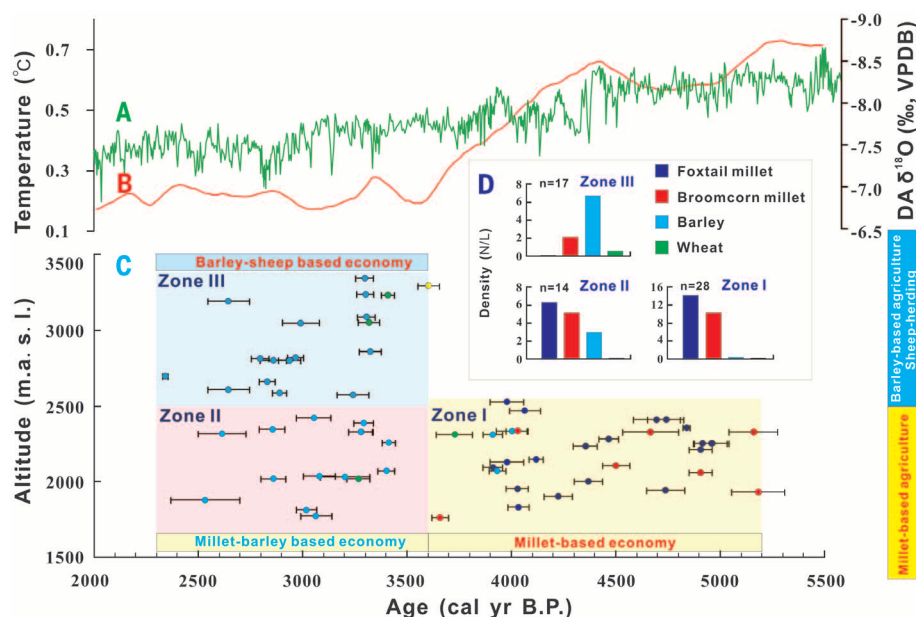

Fig. 2. Climatic records, radiocarbon dates, and charred cereal grain records from 53 investigated sites on the NETP. (A) Asian summer monsoon changes indicated by Dongge Cave speleothem oxygen isotopes (25). (B) Northern Hemisphere (30° to 90°N) temperature record compared to 1961–1990 instrumental mean temperature (26). (C) Calibrated AMS radiocarbon dates of charred grains (solid symbols with 2σ error bar) from 53 investigated sites of different archaeological cultures on the NETP and their altitudes (table S1). Zone I includes 25 sites dated between 5200 and 3600 cal yr B.P., and zones II and III include 12 sites and 17 sites dated between 3600 and 2300 cal yr B.P., below and above 2500 m.a.s.l., respectively. Circle colors indicate crops as in (D), with the addition of capers indicated in yellow. (D) Density variation of crop remains from flotation samples from zones I, II and III. N = number of charred grains, n = number of flotation samples.

essential crop repertoire, buffered against temperature change with the significant addition of cold-hardy barley. That same combination of crops additionally enabled the establishment of farms at altitudes hitherto uncultivated, taking farming in some places to elevations above 4000 masl.

Several features of this high-altitude farming prompt further questions about adaptive response. As indicated at the outset, these may include genetic resistance in humans to altitude sickness (27); genetic response in crop plants that is observable in the genetics of barley, in relation to such attributes as grain vernalization, flowering time response, and ultraviolet radiation tolerance (28); and the identity, genetic and ethnic, of the human communities themselves (1, 29). Such genetic outcomes are all consequent upon the ecological trajectories of cross-continental crop movement. Elsewhere in Europe, Asia, and Africa, that movement has been seen to have a wide variety of outcomes. In the NETP, the data presented here document its facilitation of cultivating the "roof of the world."

REFERENCES AND NOTES

1. M. Aldenderfer, *High Alt. Med. Biol.* **12**, 141–147 (2011).
2. D. D. Zhang, S. H. Li, *Geophys. Res. Lett.* **29**, 16-1–16-3 (2002).
3. P. J. Brantingham, X. Gao, *World Archaeol.* **38**, 387–414 (2006).
4. P. J. Brantingham et al., *Geoarchaeology* **28**, 413–431 (2013).
5. H. S. Tang, C. L. Zhou, Y. Q. Li, Z. Liang, *Chin. Sci. Bull.* **58**, 247–253 (2013).
6. Bureau of National Cultural Relics, *Atlas of Chinese Cultural Relics-Fascicule of Qinghai Province* (China Cartographic Press, Beijing, 1996).
7. Bureau of National Cultural Relics, *Atlas of Chinese Cultural Relics-Fascicule of the Sichuan Province* (Cultural Relics Press, Beijing, 2009).
8. Bureau of National Cultural Relics, *Atlas of Chinese Cultural Relics-Fascicule of the Tibet Autonomous Region* (Cultural Relics Press, Beijing, 2010).
9. G. H. Dong et al., *J. Archaeol. Sci.* **40**, 2538–2546 (2013).
10. D. J. Xie, *Prehistoric Archaeology of Gansu Province and Qinghai Province* (Cultural Relics Press, Beijing, 2002).
11. Materials and methods are available as supplementary materials on *Science* Online.
12. Z. J. Zhao, *Curr. Anthropol.* **52**, S295–S306 (2011).
13. X. Y. Wang, *Zhongguo Shuji (Common Millet in China)* (China Agricultural Press, Beijing, 1994).
14. Y. Chai, *Mizi (Common Millet)* (China Agricultural Press, Beijing, 1999).
15. J. D'Alpoim Guedes, E. E. Butler, *Quat. Int.* **349**, 29–41 (2014).
16. E. Páldi et al., *Biol. Plant.* **44**, 145–147 (2001).
17. L. Barton et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5523–5528 (2009).
18. R. Flad, L. Shuicheng, W. Xiaohong, Z. Zhijun, *Holocene* **20**, 955–965 (2010).
19. J. R. Dodson et al., *Quat. Sci. Rev.* **72**, 108–111 (2013).
20. M. Jones et al., *World Archaeol.* **43**, 665–675 (2011).
21. Z. J. Zhao, *Kaogu yu Wenwu (Archaeolgy and Relics)* **2**, 85–91 (2004).
22. J. d'Alpoim Guedes et al., *Archaeol. Anthropol. Sci.* **6**, 255–269 (2014).
23. M. Brink, *Setaria italica* (L.), P. Beauv, PROTA4U; in PROTA (Plant Resources of Tropical Africa), M. Brink, G. Belay, Eds. (Wageningen University, Wageningen, Netherlands, 2006); www.prota4u.org/search.asp.
24. D. X. Fu, *Kaogu (Archael.)* **47**, 66–74 (2001).
25. Y. Wang et al., *Science* **308**, 854–857 (2005).
26. S. A. Marcott, J. D. Shakun, P. U. Clark, A. C. Mix, *Science* **339**, 1198–1201 (2013).
27. C. M. Beall et al., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11459–11464 (2010).
28. J. Cockram et al., *Theor. Appl. Genet.* **115**, 993–1001 (2007).
29. M. Zhao et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 21230–21235 (2009).

___

ANIMAL PHYSIOLOGY

The roller coaster flight strategy of bar-headed geese conserves energy during Himalayan migrations

C. M. Bishop,[1]*‡ R. J. Spivey,[1]* L. A. Hawkes,[1]†‡ N. Batbayar,[2] B. Chua,[3] P. B. Frappell,[4] W. K. Milsom,[3] T. Natsagdorj,[5] S. H. Newman,[6] G. R. Scott,[7] J. Y. Takekawa,[8] M. Wikelski,[9,10] P. J. Butler[11]

The physiological and biomechanical requirements of flight at high altitude have been the subject of much interest. Here, we uncover a steep relation between heart rate and wingbeat frequency (raised to the exponent 3.5) and estimated metabolic power and wingbeat frequency (exponent 7) of migratory bar-headed geese. Flight costs increase more rapidly than anticipated as air density declines, which overturns prevailing expectations that this species should maintain high-altitude flight when traversing the Himalayas. Instead, a "roller coaster" strategy, of tracking the underlying terrain and discarding large altitude gains only to recoup them later in the flight with occasional benefits from orographic lift, is shown to be energetically advantageous for flights over the Himalayas.

Migrating birds must overcome many challenging environmental obstacles, such as arid deserts (1, 2) and featureless oceans (3–5), but few are capable of negotiating the formidably high mountains separating the Indian subcontinent from central Asia. Famously, one species that manages this feat is the bar-headed goose (*Anser indicus*), which bi-annually traverses the high passes of the Tibetan massif and snow-capped Himalayan mountains (6–8). Over the years, there has been much debate as to how high these birds might fly and what physiological mechanisms could be involved at the highest altitudes (8–12), but, although one goose has been directly tracked as high as 7290 m for a brief period (12), no measurements of their physiological or biomechanical flight performance have been made in the wild.

To investigate the flight dynamics and energetics of migratory bar-headed geese, we used custom-designed implantable instruments (13) to measure abdominal temperature and pressure (every 30 s), tri-axial acceleration (100 Hz in 18-s bursts every 2 min), and electrocardiography (180 Hz in the same 18-s period) from seven birds, collecting data totaling 391 hours of migratory flight (Fig. 1). The data loggers weighed 32 g and were housed in biocompatible tubing (dimensions 7 × 2 cm) capped by titanium electrodes.

Abdominal body temperature during flight (40.2°C ± 1.2 SD) tended to increase in tandem with flight activity, especially during times of

[1]School of Biological Sciences, Bangor University, Bangor, Gwynedd, UK. [2]Wildlife Science and Conservation Center of Mongolia, Ulaanbaatar, Mongolia. [3]Department of Zoology, University of British Columbia, Vancouver, British Columbia, Canada. [4]Office of the Dean of Graduate Research, University of Tasmania, Tasmania, Australia. [5]Mongolian Academy of Sciences, Ulaanbaatar, Mongolia. [6]Emergency Prevention System (EMPRES) Wildlife and Ecology Unit, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy. [7]Department of Biology, McMaster University, Ontario, Ontario, Canada. [8]San Francisco Bay Estuary Field Station, Western Ecological Research Center, U.S. Geological Survey, Vallejo, CA 94592 USA. [9]Max Planck Institüt für Ornithologie, Radolfzell, Germany. [10]Department of Biology, University of Konstanz, Konstanz, Germany. [11]School of Biosciences, University of Birmingham, Birmingham, UK.
*These authors contributed equally to this work. †Present address: Centre for Ecology and Conservation, University of Exeter, Cornwall Campus, UK. ‡Corresponding author. E-mail: c.bishop@bangor.ac.uk (C.M.B.); l.hawkes@exeter.ac.uk (L.A.H.)

intense effort (Fig. 1) but was generally insensitive to changes in altitude (fig. S1). The frequency distribution of all pressure-determined altitude measurements recorded during the migratory flights is shown in Fig. 2A. The median altitude while traversing the Tibetan plateau was 4707 m (maximum 6443 m, 90% of observations <5600 m). Thus, pressure-derived altitudes do not provide evidence for a general paradigm of extreme high-altitude (>8000 m) migratory flight in this species (*12*).

In order to estimate rate of oxygen consumption ($\dot{V}_{O_2}$, ml min^{-1}) during flight from measures of heart rate (f_h, beats min^{-1}) (*14–17*), we apply an allometric proportionality derived for 12 species of birds during flight (*14*) to data obtained from bar-headed geese flying in a wind tunnel (*17*) (fig. S2), and obtain the calibration relationship:

$$\dot{V}_{O_2} = 0.07 \pm 0.002 \, M_b^{0.24 \pm 0.01} \, M_h f_h^2 \qquad (1)$$

For wild migratory geese, we substitute values for body mass (M_b) of 2.8 kg and heart mass (M_h) of 1% of body mass (*18*). We then converted estimates of $\dot{V}_{O_2}$ to estimates of metabolic flight power (P_m, W kg^{-1}) by assuming 1 ml $O_2 \cong 20.9$ J. Additionally, we estimate bio-mechanical body power (P_b, W kg^{-1}) during flight, using measures of dynamic body acceleration (*19–22*). Here, we show that a single P_b component is dominant when empirically correlating several theoretical terms (*22*) for P_b against our estimates of P_m, which determines that time-averaged body power during the flapping flight of geese could be predicted by

$$P_b = \frac{\ddot{Z}_{rms}^2}{2p^2 f_w} \qquad (2)$$

where $\ddot{Z}_{rms}^2$ is root-mean-square dorsoventral acceleration (z axis) and f_w is wingbeat frequency.
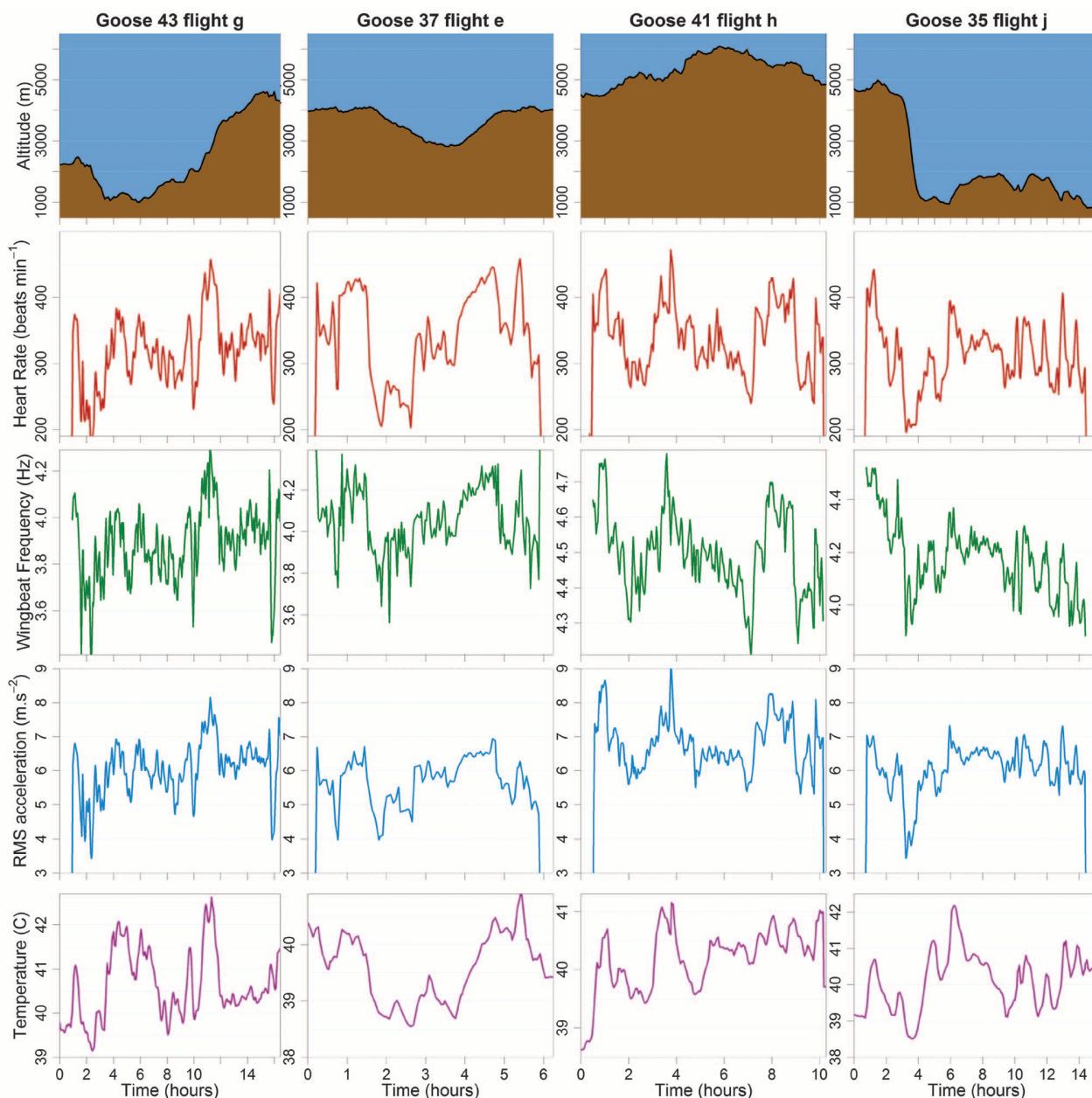
Fig. 1. Examples of autumn migratory flights. Bar-headed goose (*Anser indicus*) P43 travelled South from Mongolia and ascended onto the Tibetan Plateau (column 1); goose P37 (column 2) and goose P41 (column 3) were traversing the Tibetan Plateau; goose P35 (column 4) crossed the Himalayas and descended into India. Pressure altitude (row 1), f_h (row 2), f_w (row 3), $\ddot{Z}_{rms}^2$ (row 4), abdominal body temperature (row 5).

This simple term maximized correlations between the independently derived biomechanical P_b and metabolic P_m (mean $r^2 = 0.91 \pm 0.05$ SD) (Fig. 2B).

During flight, heart rate and wingbeat frequency were significantly correlated (mean $r^2 > 0.86 \pm 0.11$ SD) (Fig. 2, C and D, and fig. S3A), as well as heart rate and $\ddot{Z}^2_{rms}$ (mean $r^2 = 0.91 \pm 0.05$ SD) (Fig. 2C and fig. S3B) and wingbeat frequency and $\ddot{Z}^2_{rms}$ (mean $r^2 = 0.89 \pm 0.09$ SD) (fig. S3C). Median wingbeat frequency increased with pressure-derived altitude as air density declined (median $f_w = 3.94$ Hz at altitude < 2300 m; $f_w = 4.35$ Hz at altitude >4800 m) (Fig. 2E). Similarly, median heart rate during flight increased with altitude and was generally higher on the Tibetan plateau ($f_h = 364$ beats min^{-1} at altitude >4800 m) (Fig. 2F) than at lower altitudes ($f_h = 300$ beats min^{-1} at altitude <2300 m). Although the partial pressure of oxygen decreases with increasing altitude, up to around

5000 m, any potential desaturation of oxygen-bound hemoglobin in the blood of bar-headed geese should still be relatively small, at around 10% (18, 23). Indeed, captive bar-headed geese are able to run for 15 min at similar maximum speeds, whether exposed to atmospheres of 21, 10.5, or 7% oxygen, the last-mentioned condition resulting in a desaturation of between 20 and 23% (18).

Our data show that median heart rate during flight scales with air density (ρ) as $f_h \propto \rho^{-0.64}$ (Fig. 2G) and, therefore, that estimated P_m should scale approximately as $P_m \propto \rho^{-0.91}$ (if one assumes that $P_m \propto f_h^2$ but allowing for a 10% additional increase of f_h for a given value of $\dot{V}_{O_2}$ at 5500 m due to a hemoglobin desaturation of 10%). Thus, the relative metabolic flight power of the geese at 5000 m compared with that at sea level is estimated to be around 1.7-fold. This is higher than the anticipated sensitivity of flight power to air density of $P_m \propto \rho^{-0.54}$

predicted by aerodynamic theory (24). Similarly, flight theory predicts that wingbeat frequency should be $\propto \rho^{-0.38}$, whereas the present results for bar-headed geese show median $f_w \propto \rho^{-0.23}$ (Fig. 2H). This is at the lower end of the predicted range but in keeping with the observations of large Ciconiiformes (herons, spoonbill, ibis) migrating high above the Negev Desert in Israel (25).

Bar-headed geese exhibit an extreme sensitivity of heart rate and, therefore, metabolic flight power to small changes in wingbeat frequency, when a precise method is used for extracting values of f_w (26). For example, a 5% increase in f_w from 4.0 to 4.2 Hz equates to a 19% increase in f_h and, therefore, a 41% increase in estimated P_m. Across all migratory flights, f_h correlated in the range of $f_h \propto f_w^{1.95 \text{ to } 6.65}$ and estimated P_m as $P_m \propto f_w^{3.9 \text{ to } 13.3}$, the latter exponent exceeding 3 in every case (median exponent 6.96) (Fig. 2I). For steady horizontal flight, the inertial

Fig. 2. Descriptive flight statistics. Frequency histograms of (**A**) altitude reported during migratory flights of bar-headed geese (*Anser indicus*) and (**B**) correlation of estimated P_m versus estimated P_b. (**C**) Correlation of f_h versus f_w plotted against correlation of $\ddot{Z}^2_{rms}$ versus f_w. (**D**) Examples of f_h against f_w for four individual flights. Frequency distribution of (**E**) f_w and (**F**) f_h within three altitude zones. Scatter plots of (**G**) f_h and (**H**) f_w plotted against altitude. (**I**) Frequency distribution of power exponents for f_w against estimated P_m.
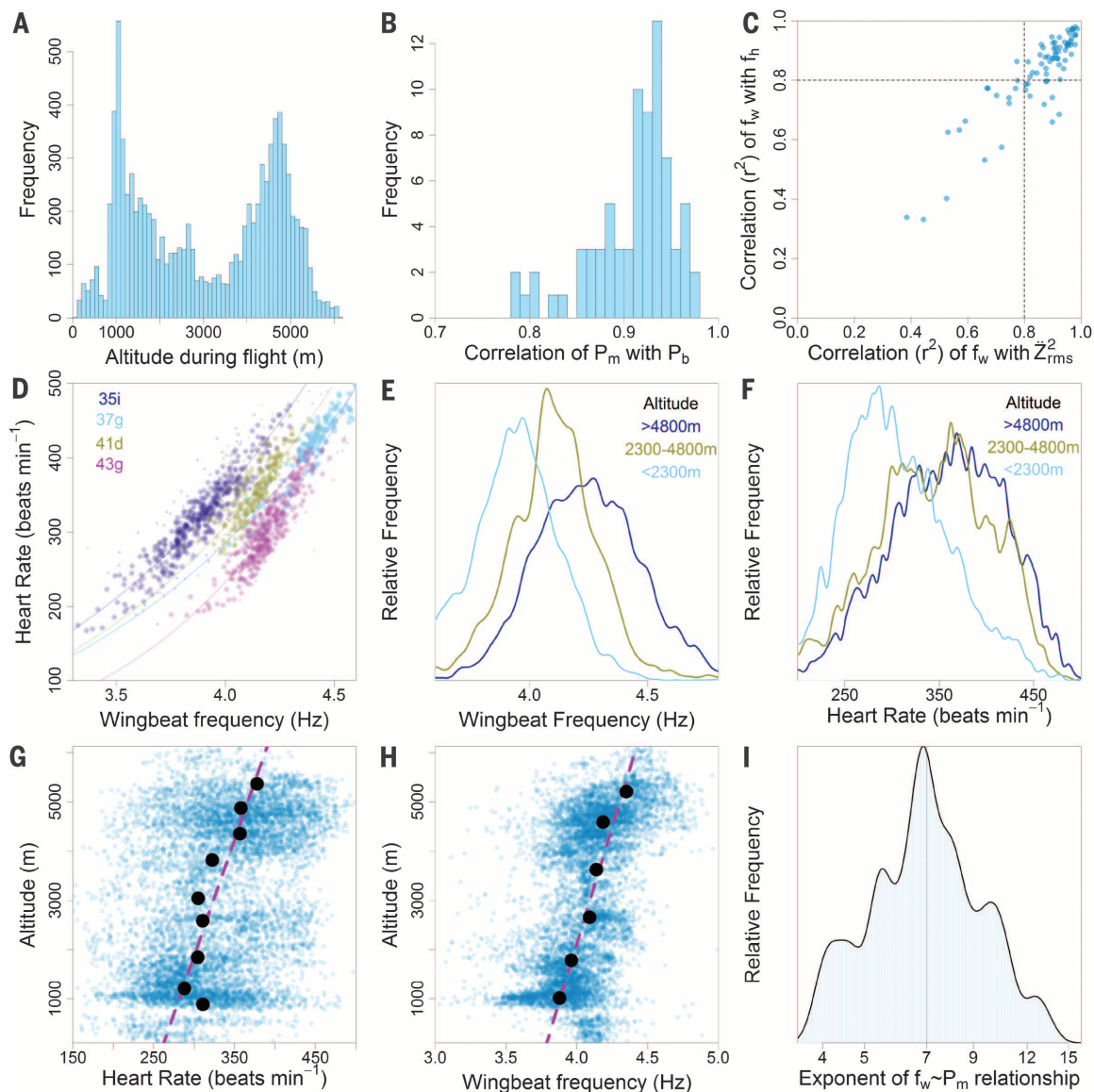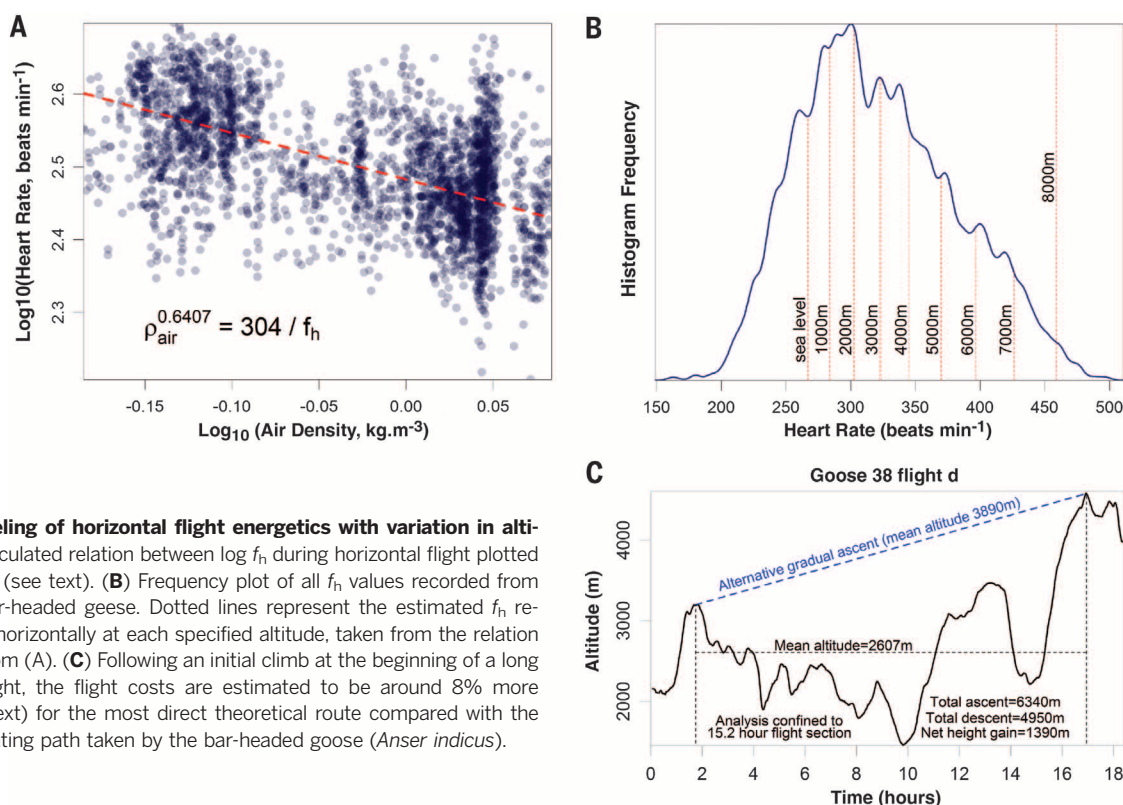
Fig. 3. Modeling of horizontal flight energetics with variation in altitude. (**A**) Calculated relation between log f_h during horizontal flight plotted against log ρ (see text). (**B**) Frequency plot of all f_h values recorded from the same bar-headed geese. Dotted lines represent the estimated f_h required to fly horizontally at each specified altitude, taken from the relation calculated from (A). (**C**) Following an initial climb at the beginning of a long migratory flight, the flight costs are estimated to be around 8% more costly (see text) for the most direct theoretical route compared with the actual undulating path taken by the bar-headed goose (*Anser indicus*).
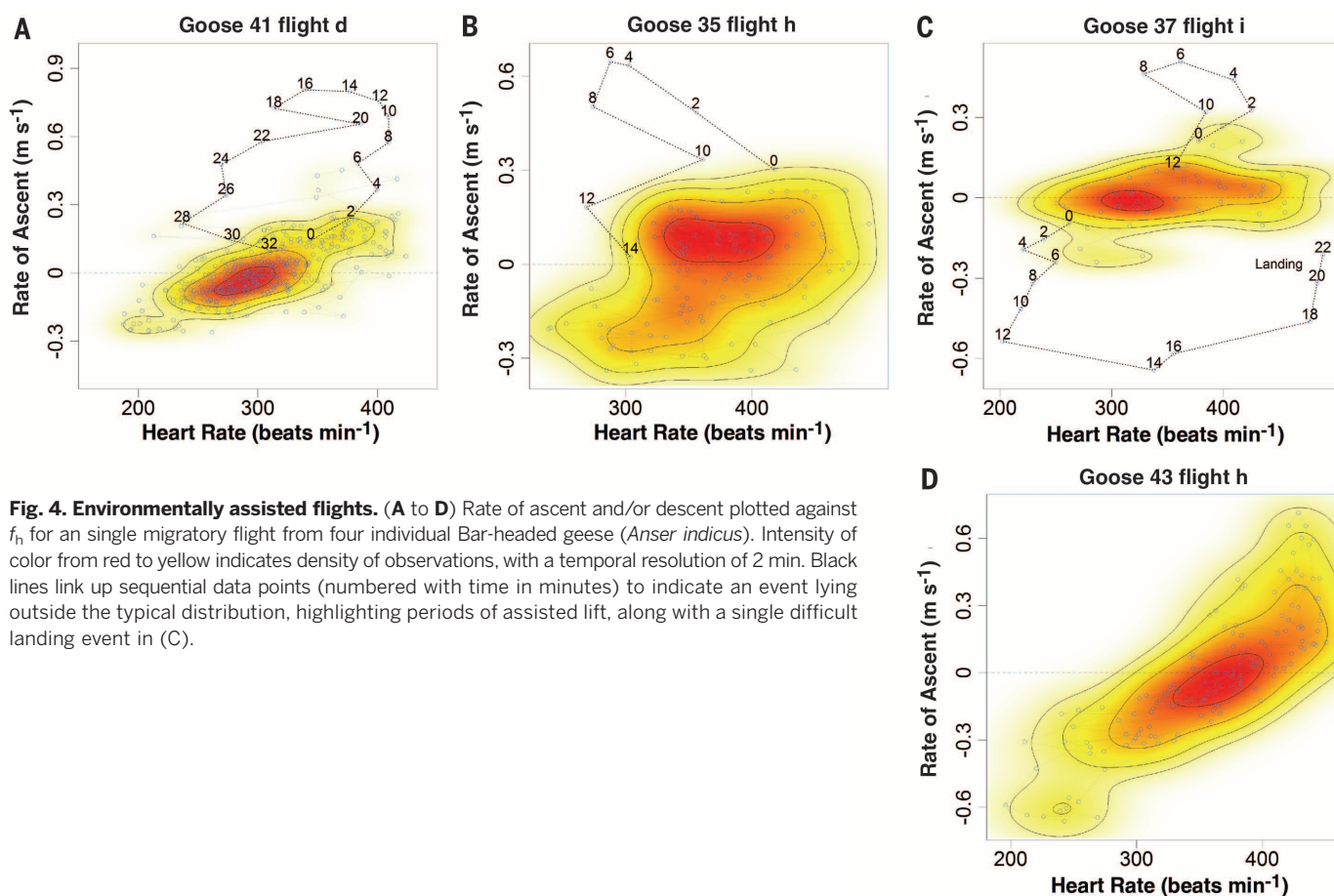
Fig. 4. Environmentally assisted flights. (**A** to **D**) Rate of ascent and/or descent plotted against f_h for an single migratory flight from four individual Bar-headed geese (*Anser indicus*). Intensity of color from red to yellow indicates density of observations, with a temporal resolution of 2 min. Black lines link up sequential data points (numbered with time in minutes) to indicate an event lying outside the typical distribution, highlighting periods of assisted lift, along with a single difficult landing event in (C).

costs of flapping the wings should be proportional to the product of wingbeat frequency cubed and the wing amplitude squared. If the body of the bird undergoes sinusoidal amplitude displacements on the vertical axis (B) then $\ddot{Z}^2_{rms} = 2\sqrt{2}\pi^2 B f_w^2$ (22) and so Eq. 2 can be rewritten

$$P_b = 4\pi^2 B^2 f_w^3 \quad (3).$$

Because B should be positively correlated with wingbeat amplitude, the implication of our experimental data, showing that $P_m \propto f_w^{6.96}$, is that the angular travel of the wing increases with higher f_w. Thus, the exquisite sensitivity of P_m to f_w in geese stems from wingbeat amplitude that is positively correlated with changes in wingbeat frequency.

In the present study, there was no evidence of gliding behavior in bar-headed geese, even when descending rapidly from the Himalayas into India (fig. S4). During the steepest descent phases, f_w remained above 3.6 Hz for 98% of observations, whereas f_h decreased to between 150 and 200 beats min^{-1}. Indeed, f_h was surprisingly low in general throughout the entire migration (overall mean f_h = 328 ± 64 beats min^{-1}) (Fig. 2F), with geese only spending 2.3% of their flight time at altitudes above 4800 m with a f_h greater than 455 beats min^{-1} (and 0.37% of their flight time when below 2300 m altitude). A simple extrapolation of the relations between heart rate and air density (Fig. 3A), with data filtered so that only rates of ascent or descent lying between ±0.1 m s^{-1} are included (an approximation of horizontal flight), demonstrates that a minimum heart rate of around 460 beats min^{-1} might just suffice at around 8000 m in still air conditions (Fig. 3B). However, even this assessment might seem unduly optimistic, given that it ignores the energetics and time required to make the climb itself and the steepness of the relation for hemoglobin desaturation once the partial pressures of oxygen fall below a critical value (18, 23). Thus, unaided horizontal flights over 8000 m are likely to be approaching the limit for sustained aerobic capacity in this species.

Previous low temporal-resolution global positioning system altitude data (12) indicated that bar-headed geese tend to fly closest to the ground when traversing the Tibetan massif, with a median height of only 62 m. This is consistent with the high-resolution pressure altitude results of the present study, which imply that geese opt repeatedly to shed hard-won altitude only subsequently to regain height later in the same flight. An example of this tactic can be seen in a 15.2-hour section of a 17-hour flight (Fig. 3C) in which, after an initial climb to 3200 m, the goose followed an undulating profile involving a total ascent of 6340 m with a total descent of 4950 m for a net altitude gain of only 1390 m. Revealingly, calculations show that steadily ascending in a straight line would have increased the journey cost by around 8%. As even horizontal flapping flight is relatively ex-

pensive, the increase in energy consumption due to occasional climbs is not as important as the effect of reducing the general costs of flying by seeking higher-density air at lower altitudes.

Rates of ascent and descent during four migratory flights are plotted against f_h (Fig. 4) and against f_w (fig. S5), with maximum ascent rates of up to at least 0.8 m s^{-1}, lasting for several minutes. However, such extreme ascent rates were generally not associated with increases in f_h and f_w. A particularly clear example of such an episode that occurred during a 13-hour migratory flight is shown in Fig. 4A. The central cluster of Fig. 4A exhibits a sloping relation between f_h and rate of ascent (typical of a number of flights), but there was a dramatic departure from this pattern lasting ~30 min involving unusually high rates of ascent despite "normal" values of heart rate. Although the degree of central clustering varied between flights, presumably according to the prevailing wind conditions and underlying terrain, similar unusually high ascent rates occurred on other flights (Fig. 4, B to D). These unique results are interpreted as evidence of sustained assistance from updrafts due to orographic lift (27, 28), presumably indicative of geese flying along the windward side of a ridge. Thus, it is logical to conclude that weaker vertical updrafts could also provide more gentle assistance during other phases of the migratory flights, perhaps comparable in magnitude to the assistance geese might at times receive from V-formation flight (29, 30).

When traversing mountainous areas, a terrain-tracking strategy or flying in the cool of the night (12) can reduce the cost of flight in bar-headed geese through exposure to higher air density. Ground-hugging flight may also confer additional advantages including maximizing the potential of any available updrafts of air, reduced exposure to crosswinds and headwinds, greater safety through improved ground visibility, and increased landing opportunities. The atmospheric challenges encountered at the very highest altitudes, coupled with the need for near-maximal physical performance in such conditions, likely explains why bar-headed geese rarely fly close to their altitude ceiling, typically remaining below 6000 m. Given that aerodynamic mass-specific flight costs are thought to increase with body mass and that bar-headed geese are heavier than 98% of avian species, it is particularly impressive that these birds are able to migrate across the world's highest land massif while remaining comfortably within their physiological capabilities.

REFERENCES AND NOTES

1. H. Schmaljohann, F. Liechti, B. Bruderer, *Proc. Biol. Sci.* **274**, 735–739 (2007).
2. R. H. G. Klaassen, T. Alerstam, P. Carlsson, J. W. Fox, A. Lindström, *Biol. Lett.* **7**, 833–835 (2011).
3. P. J. Butler, A. J. Woakes, C. M. Bishop, *J. Avian Biol.* **29**, 536–545 (1998).
4. S. A. Shaffer et al., *Proc. Natl. Acad. Sci. U.S.A.* **103**, 12799–12802 (2006).
5. R. E. Gill Jr. et al., *Proc. Biol. Sci.* **276**, 447–457 (2009).
6. L. A. Hawkes et al., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9516–9519 (2011).
7. A. Blum, *Annapurna: A Woman's Place* (Sierra Club, San Francisco, 1998).
8. L. W. Swan, *Nat. Hist.* **70**, 68–75 (1970).
9. F. M. Faraci, *Annu. Rev. Physiol.* **53**, 59–70 (1991).
10. G. R. Scott, W. K. Milsom, *Respir. Physiol. Neurobiol.* **154**, 284–301 (2006).
11. B. Pinshow, M. H. Bernstein, Z. Arad, *Am. J. Physiol.* **249**, R758–R764 (1985).
12. L. A. Hawkes et al., *Proc. Biol. Sci.* **280**, 20122114 (2013).
13. R. J. Spivey, C. M. Bishop, *Rev. Sci. Instrum.* **85**, 014301 (2014).
14. C. M. Bishop, R. J. Spivey, *J. Theor. Biol.* **323**, 11–19 (2013).
15. P. J. Butler, J. A. Green, I. L. Boyd, J. R. Speakman, *Funct. Ecol.* **18**, 168–183 (2004).
16. J. A. Green, *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **158**, 287–304 (2011).
17. S. Ward, C. M. Bishop, A. J. Woakes, P. J. Butler, *J. Exp. Biol.* **205**, 3347–3356 (2002).
18. L. A. Hawkes et al., *PLOS ONE* **9**, e94015 (2014).
19. L. G. Halsey, S. J. Portugal, J. A. Smith, C. P. Murn, R. P. Wilson, *J. Field Ornithol.* **80**, 171–177 (2009).
20. K. H. Elliott, M. Le Vaillant, A. Kato, J. R. Speakman, Y. Ropert-Coudert, *Biol. Lett.* **9**, 20120919 (2013).
21. O. Duriez et al., *PLOS ONE* **9**, e84887 (2014).
22. R. J. Spivey, C. M. Bishop, *J. R. Soc. Interface* **10**, 20130404 (2013).
23. J. U. Meir, W. K. Milsom, *J. Exp. Biol.* **216**, 2172–2175 (2013).
24. C. Pennycuick, *J. Exp. Biol.* **199**, 1613–1618 (1996).
25. F. Liechti, E. Schaller, *Naturwissenschaften* **86**, 549–551 (1999).
26. R. J. Spivey, S. Stansfield, C. M. Bishop, *Prog. Oceanogr.* **125**, 62–73 (2014).
27. P. J. Butler, *Comp. Biochem. Physiol. A Mol. Integr. Physiol.* **156**, 325–329 (2010).
28. G. Bohrer et al., *Ecol. Lett.* **15**, 96–103 (2012).
29. H. Weimerskirch, J. Martin, Y. Clerquin, P. Alexandre, S. Jiraskova, *Nature* **413**, 697–698 (2001).
30. S. J. Portugal et al., *Nature* **505**, 399–402 (2014).

PALEOCEANOGRAPHY

Reduced El Niño–Southern Oscillation during the Last Glacial Maximum

Heather L. Ford,[1,2]* A. Christina Ravelo,[1] Pratigya J. Polissar[2]

El Niño–Southern Oscillation (ENSO) is a major source of global interannual variability, but its response to climate change is uncertain. Paleoclimate records from the Last Glacial Maximum (LGM) provide insight into ENSO behavior when global boundary conditions (ice sheet extent, atmospheric partial pressure of CO_2) were different from those today. In this work, we reconstruct LGM temperature variability at equatorial Pacific sites using measurements of individual planktonic foraminifera shells. A deep equatorial thermocline altered the dynamics in the eastern equatorial cold tongue, resulting in reduced ENSO variability during the LGM compared to the Late Holocene. These results suggest that ENSO was not tied directly to the east-west temperature gradient, as previously suggested. Rather, the thermocline of the eastern equatorial Pacific played a decisive role in the ENSO response to LGM climate.

The equatorial Pacific mean climate state (average oceanic and atmospheric properties across the basin) is characterized by a strong east-west sea surface temperature (SST) gradient that is tightly coupled to the thermocline and the winds that drive warm water to the west and cause cold water to upwell in the east (1). Because wind strength and the zonal SST gradient are mutually dependent, perturbations in this ocean-atmosphere link initiate and propagate El Niño–Southern Oscillation (ENSO) events (2). Theoretical and modeling studies suggest that the mean state should strongly affect ENSO (3–5) by altering the balance of several positive and negative ocean-atmosphere feedbacks that determine ENSO behavior (2, 3, 6, 7). However, climate models disagree on how these feedbacks interact when the mean state changes (8). Here we examine climate variability during the Last Glacial Maximum (LGM) (~20,000 years ago) as an opportunity to investigate ENSO behavior during an altered mean state when ice sheets covered North America and partial pressure of CO_2 (P_{CO_2}) levels were ~100 parts per million lower than during preindustrial times (9).

We use deep-sea sediment samples from Ocean Drilling Program (ODP) site 806 in the western equatorial Pacific (WEP) warm pool and ODP site 849 in the eastern equatorial Pacific (EEP) cold tongue to examine tropical variability during discrete time intervals (Fig. 1). Site 806 is located in the heart of the warm pool on the equator, where SST variability is small and primarily at interannual-to-decadal frequencies. In contrast, site 849 is located in the EEP cold tongue extension, where variability is large and dominated equally by seasonal and canonical ENSO frequencies (2). Here we use the distribution of Mg/Ca-based temperatures measured on individual shells of surface- (Globigerinoides sacculifer) and subsurface-dwelling (Globorotalia tumida) forami-

nifera in a sediment sample to quantify tropical variability during the late Holocene (<6000 years ago) and LGM. This distribution has been used to accurately reconstruct the mean and seasonal variability at several locations, and our sample size (40 to 70 individuals) is sufficient to capture climate variability at our study sites (10). Three prior individual foraminifera studies of the LGM suggest either increased [site V21-30 (11); site CD38-17P (12)] or decreased [site MD02-2529 (13)] ENSO variability compared with the late Holocene. We synthesized these apparently di-

vergent results and our newly generated data by considering geographic location, choice of foraminifera species, and changes in thermocline depth (see supplementary materials).

ENSO variability is asymmetric (the El Niño warm phase is more extreme than the La Niña cold phase) (14), so temperature variations in the equatorial Pacific are not normally distributed (7, 15), and statistical tests that assume normality (e.g., standard deviation) can lead to erroneous conclusions with respect to changes in variance. Therefore, we use quantile-quantile (Q-Q) plots—a simple, yet powerful way to visualize distribution data—to compare the temperature range and distribution recorded by two populations of individual foraminifera shells to interpret possible climate forcing mechanisms. Sensitivity studies using modern hydrographic data show how changes in ENSO and seasonality modify temperature distributions that can be diagnosed as changed slopes on Q-Q plots (Fig. 2) (also see supplementary materials). For our locations, the sensitivity studies indicate that seasonality weakly affects the temperature distributions, whereas ENSO has a large influence (Fig. 2); consequently, changes in temperature distribution between the late Holocene and LGM can be attributed to changes in ENSO.

In the WEP, our data show that LGM surface and subsurface temperatures were cooler by ~2.3° to 2.4°C compared with the late Holocene (Fig. 3). Cooler surface waters are consistent with adjustment to reduced P_{CO_2} forcing (16). Cooler subsurface temperatures could be interpreted
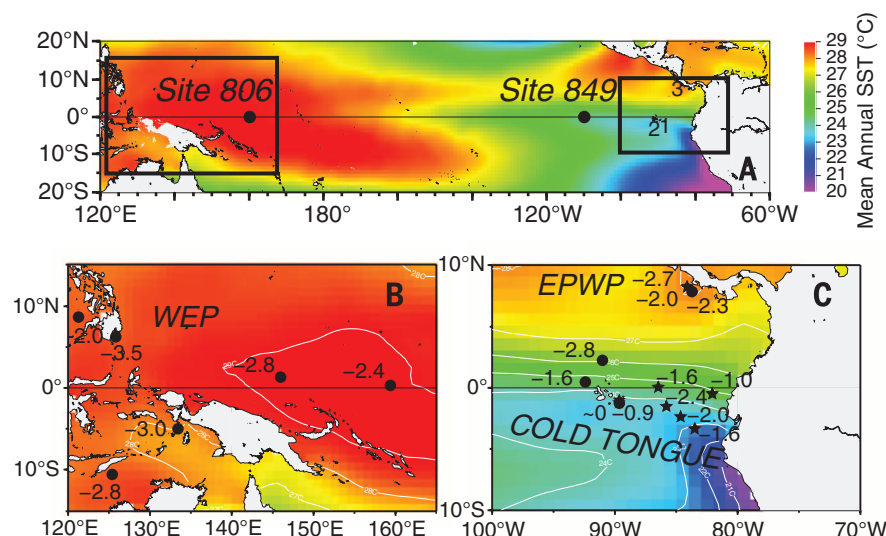

Fig. 1. Reconstruction of the zonal Holocene-LGM temperature gradient. (**A**) Ocean Drilling Program sites 806 and 849 indicated on a mean annual SST map, using the Met Office Hadley Centre's HadISST 1.1 data set (32). Numbers indicate individual foraminifera study sites: V21-30 (1), CD38-17P (2), and MD02-2529 (3). Inset maps of (**B**) WEP and (**C**) EEP show anomaly of published LGM [18 to 20 thousand years ago (ka)] minus Holocene (4 to 6 ka) temperatures reconstructing the zonal SST gradient. Locations using Mg/Ca proxy are denoted by circles, whereas locations using the alkenone proxy are denoted with stars. Generally, there was a reduced zonal temperature across the Pacific, though there is some spatial heterogeneity in the EEP. The WEP and EPWP have similar cooling magnitudes (~2.7° and ~2.3°C, respectively), which suggests that LGM cooling was largely dominated by radiative cooling. In contrast, the EEP cold tongue region cooled less (~1.6°C) during the LGM, indicating that radiative cooling was partially compensated by dynamic components (a deep thermocline and reduced upwelling).

[1]Ocean Sciences Department, University of California, Santa Cruz, CA 95064, USA. [2]Biology and Paleo Environment, Lamont-Doherty Earth Observatory, Palisades, NY 10964, USA.
*Corresponding author. E-mail: hford@ldeo.columbia.edu

as a colder or shallower thermocline (17). However, because modeling (18), faunal (19), and geochemical (20) studies suggest that the thermocline was actually deeper during the LGM in comparison with today, our cool subsurface temperatures indicate a colder thermocline. Equatorial thermocline waters originate in mid-latitude regions where they subduct, move equatorward, and upwell along the equator (21). Our *G. tumida* subsurface temperatures suggest that these mid-latitude thermocline source water regions were cooler during the LGM, also probably as a response to P_{CO_2} forcing. Surface and subsurface cooling in the WEP occurred without changes in variability, consistent with P_{CO_2} radiative forcing as the most likely agent of WEP temperature change (16).

In the EEP, our data from site 849 show that average surface temperatures were only ~1.2° to 1.3°C cooler during the LGM compared with the Holocene and that SST variability was reduced such that the cooling was greater during the warm season and smaller during the cold season (Fig. 4). Given the sensitivity of our site to ENSO, this reduction in surface ocean variability during the LGM reflects greatly reduced ENSO amplitude (Figs. 2C and 4, B and C). In contrast, site V21-30 located ~2200 km to the east (Fig. 1) shows high surface ocean variability during the LGM, as recorded by the distribution of $\delta^{18}O$ values of individual surface dwelling foraminifera (*Globigerinoides ruber*). An increase in variance has been interpreted as enhanced ENSO (11); however, a recent statistical study (22) suggests that the geographical location of site V21-30 is unlikely to capture changes in ENSO,

and our Q-Q reanalysis of these data (fig. S13) suggests that the high variance reported at V21-30 (11) reflects enhanced seasonality during the LGM in comparison to the late Holocene. Together, these data indicate that ENSO was reduced and seasonality was enhanced during the LGM in comparison with today, which is corroborated by a recent modeling study (23). Some climate models indicate a strong inverse relationship between the amplitude of the seasonal cycle and ENSO (3); that is, when the seasonal cycle is strong, ENSO is weak. Our interpretation suggests that this inverse relationship may be a robust behavior over long time scales.

The EEP *G. tumida* subsurface temperatures exhibit no change in mean temperature but do show an increase in variability (Fig. 4). Although an increase in thermocline variability is consistent with enhanced ENSO (Fig. 2E), this signal is conflated with a deepening of the EEP cold tongue thermocline during the LGM (24, 25). Deepening of the thermocline probably led *G. tumida* to inhabit a shallower portion of the thermocline with a steeper temperature gradient (24, 25), resulting in increased reconstructed temperature variability during the LGM that obscures any true change in thermocline variability. Additionally, the lack of cooling in the LGM subsurface *G. tumida* measurements is consistent with this interpretation, as calcification in upper, warmer thermocline waters would offset LGM cooling. A long-term change in the depth habitat of *G. tumida* is unlikely because in the modern ocean it calcifies at a relatively constant depth range, regardless of vertical temperature gra-

dients (26). Although subsurface variability is a strong indicator of ENSO in today's ocean (6, 7), comparing the magnitude of subsurface variability from foraminifera proxies may not be the best indicator of ENSO strength when there are substantial changes in thermocline depth.

At site CD38-17P (east of site 849 and south of V21-30) (Fig. 1), *Neogloboquadrina dutertrei* show no change in mean temperature but greater subsurface temperature variability during the LGM in comparison with the Late Holocene (12). The lack of a change in mean temperature agrees with our *G. tumida* observation, and we similarly suggest that *N. dutertrei* must have also occupied a shallower portion of the thermocline during the LGM in comparison to today. This change in habitat explains the apparent increase in variability, and consequently, *N. dutertrei* from site CD38-17P do not support increased ENSO variability during the LGM (figs. S14 and S15).

Fig. 3. Q-Q plots of WEP temperature variability reconstructed from planktonic foraminifera. (**A**) Surface temperature variability. (**B**) Subsurface temperature variability. The inset within each Q-Q plot shows the dpdfs of single-shell analyses plotted for Holocene (~3.0 ka, orange) and LGM (~19.0 ka, blue) intervals. Q-Q plots show that SSTs are uniformly shifted ~2.4°C colder and subsurface temperatures are shifted ~2.3°C colder during the LGM in comparison with the Holocene.
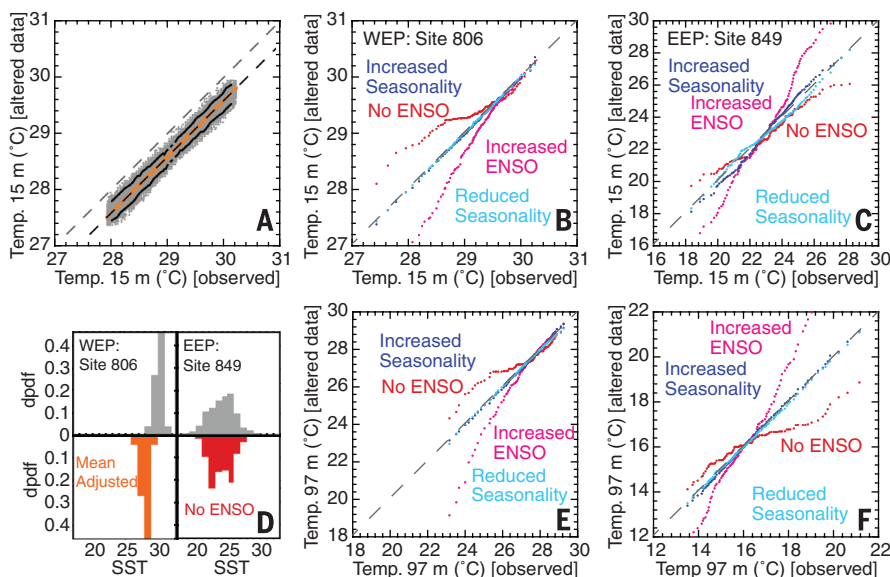
Fig. 2. Schematic framework for the interpretation of Q-Q plots of WEP (site 806) and EEP (site 849) surface and subsurface temperature variability reconstructed from planktonic foraminifera. (**A**) Example of a mean shift in temperature without a change in variability, as seen in a Q-Q plot. Cumulative distribution function (orange line), 1:1 line (gray dashed line), mean adjusted 1:1 line (black dashed line), Monte Carlo simulation (gray dots), and 90% confidence limits (black lines) are shown (see supplementary materials for further details). For WEP site 806 (**B** and **E**) and EEP site 849 (**C** and **F**), SODA v2.4 monthly temperature data at 15 and 97 m are altered with no ENSO (red), increased ENSO (pink), increased seasonality (blue), and reduced seasonality (light blue). (**D**) Example discrete probability density function (dpdf) for unaltered data (grey), altered mean [plotted in (A), orange], and altered "No ENSO" scenario [plotted in (C), red].
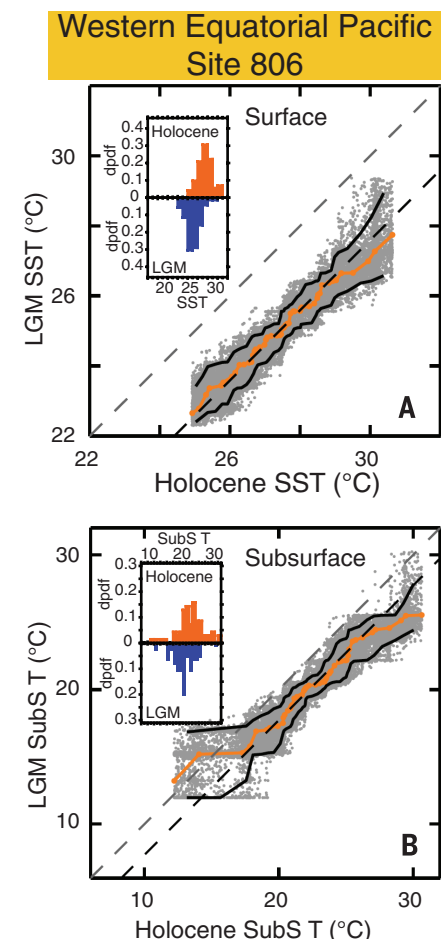
Within the eastern Pacific warm pool (EPWP) (Fig. 1C), *N. dutertrei* at site MD02-2529 show reduced shallow subsurface temperature variability and ~2°C cooler average temperatures during the LGM in comparison with the late Holocene (13). *N. dutertrei* are known to alter their calcification depth, preferring a shallower habitat in places with a well-defined shallow thermocline, such as the EPWP, and a deeper habitat in places such as the EEP (27). The cool LGM subsurface temperatures recorded by *N. dutertrei* are matched by a similar magnitude cooling of the mean surface temperature in the EPWP (discussed below), suggesting that *N. dutertrei* maintained their calcification position near the top of the thermocline. Therefore, a deepening of the thermocline cannot explain reduced shallow subsurface temperature variability at site MD02-2529 and necessitates a reduction of ENSO amplitude, supporting our findings (fig. S16).

Past studies have attempted to characterize the tropical Pacific during the LGM by measuring the WEP-EEP zonal gradient with mixed results (16, 28). Our compilation of available SST data in the EEP and WEP reveals spatial heterogeneity in SST cooling during the LGM relative to to-

day (table S3). We find that the EPWP and WEP have a similar temperature change (~2.3° and 2.7°C, respectively), which suggests that cooling was a function of lower P_{CO_2} and associated radiative forcing in these regions (Fig. 1). In contrast, the EEP cold tongue was only ~1.6°C cooler in the LGM. Our single-shell analyses demonstrate that any radiative forcing signal was offset by dynamic changes within the EEP cold tongue, including a deep thermocline (24, 25), related changes in upwelling, and reduced ENSO. Overall, this suggests that the western warm pool–cold tongue tropical Pacific zonal SST gradient was reduced.

Our findings do not support previous studies suggesting an inverse relationship between the zonal SST gradient and ENSO variability during the LGM (11, 12). Instead, we find weaker ENSO variability when the zonal SST gradient is reduced, consistent with recent modeling results (29). We suggest that an altered thermocline mean state during the LGM reduced ENSO variability by shifting the balance of feedbacks important to ENSO dynamics. Theoretical studies suggest that there is a thermocline mode and an SST mode that give rise to ENSO (3, 5). These modes differ in the propagation of subsurface and sur-

face temperature anomalies, and the balance between these modes changes with mean thermocline depth (3, 5). A deep mean thermocline favors the thermocline mode and reduced ENSO amplitude (3). ENSO amplitude is reduced because a deep thermocline inhibits upwelling of anomalous subsurface temperatures, weakening the thermocline feedback and ocean-atmosphere coupling (4, 6, 29). During the LGM, the zonal SST gradient was reduced, the basin-wide winds were enhanced, and the thermocline was deep (18, 19, 24, 25). With this major change in mean state, the balance of dynamic feedbacks probably changed; one possibility is that this LGM thermocline mean state favored thermocline-mode ENSO behavior and a weakened thermocline feedback that reduced ENSO amplitude.

Although zonal SST gradient reconstructions are often used to infer the tropical mean state and the strength of ocean-atmospheric feedbacks (11, 12, 16), subsurface temperatures and the equatorial thermocline mean state may more accurately reflect Walker circulation (18) and affect ENSO dynamics (30). Several positive and negative feedbacks determine ENSO behavior (4), and individual feedback strength changes under climate states with different mean thermocline depth and mean wind stress (3, 4, 29). For example, a deeper thermocline may reduce the strength of the thermocline feedback, thereby reducing ENSO amplitude, but simultaneous changes in the strength of other feedbacks may counteract this effect. Climate models have strong biases in simulating thermocline conditions and tropical upwelling (5, 9), and these biases influence the strength and balance of positive and negative feedbacks that determine ENSO behavior. Although model simulations of anthropogenic climate change suggest that the equatorial thermocline will shoal, the effect on the individual feedbacks determining ENSO variability is uncertain (31). The LGM mean state is radically different from future climate projections and thus cannot be directly used to predict future ENSO behavior. However, our results linking observations of the mean state and ENSO variability can be used to test theoretical and numerical models that are pivotal to understanding ENSO behavior in the face of mean state climate changes in the future.
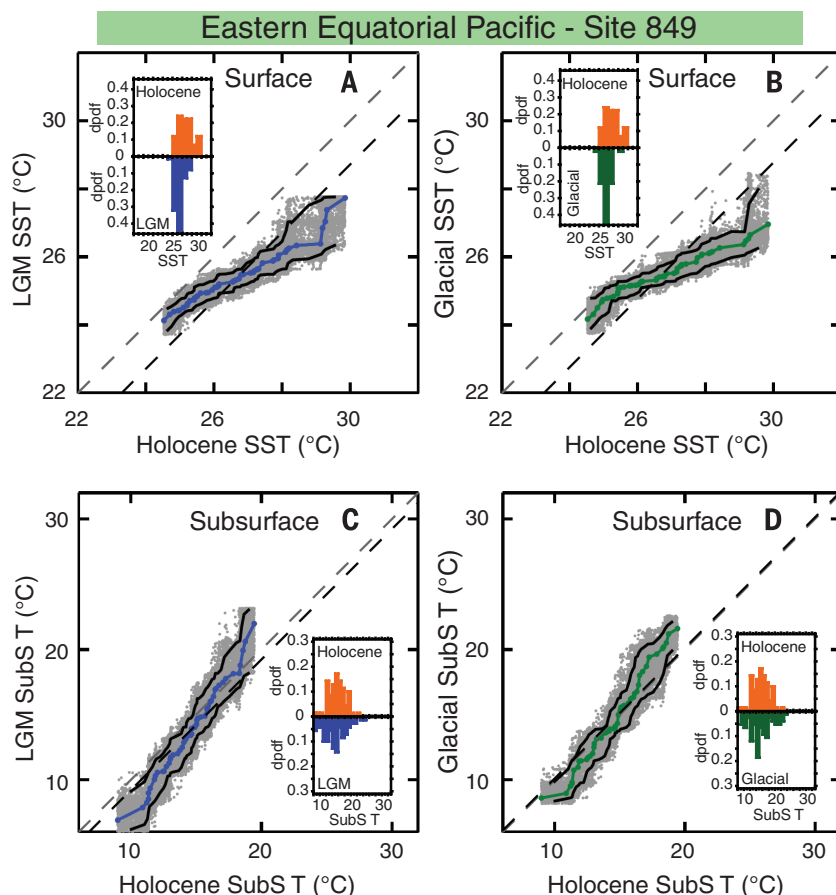

Fig. 4. Q-Q plots of EEP temperature variability reconstructed from planktonic foraminifera from the LGM and from one glacial sample. (**A** and **B**) Surface temperature variability. (**C** and **D**) Subsurface temperature variability. The inset within each Q-Q plot shows the dpdfs of single-shell analyses plotted for Holocene (~2.8 and ~5.6 ka, orange), LGM (~20.3 ka, blue), and glacial (~21.5 ka, green) intervals. Q-Q plots of SSTs show a reduction in ENSO, as demonstrated in Fig. 2. Q-Q plots of subsurface temperatures show greater variability, probably related to a change in the thermocline mean state.

REFERENCES AND NOTES

1. W. Wang, M. J. McPhaden, *J. Phys. Oceanogr.* **29**, 1812–1831 (1999).
2. W. Wang, M. J. McPhaden, *J. Phys. Oceanogr.* **30**, 2989–3008 (2000).
3. A. V. Fedorov, S. Philander, *J. Clim.* **14**, 3086–3101 (2001).
4. F.-F. Jin, S. T. Kim, L. Bejarano, *Geophys. Res. Lett.* **33**, L23708 (2006).
5. E. Guilyardi *et al.*, *Bull. Am. Meteorol. Soc.* **90**, 325–340 (2009).
6. S.-I. An, F.-F. Jin, *J. Clim.* **14**, 3421–3432 (2001).
7. J. Choi, S.-I. An, S.-W. Yeh, *Clim. Dyn.* **38**, 2631–2644 (2012).
8. G. A. Schmidt *et al.*, *Clim. Past* **10**, 221–250 (2014).
9. B. L. Otto-Bliesner *et al.*, *Clim. Dyn.* **32**, 799–815 (2009).
10. J. C. Wit, G.-J. Reichart, S. J. A. Jung, D. Kroon, *Paleoceanography* **115**, PA4220 (2010).
11. A. Koutavas, S. Joanides, *Paleoceanography* **27**, PA4208 (2012).
12. A. Y. Sadekov *et al.*, *Nat. Commun.* **4**, 2692 (2013).
13. G. Leduc, L. Vidal, O. Cartapanis, E. Bard, *Paleoceanography* **24**, PA3202 (2009).
14. S.-I. An, F.-F. Jin, *J. Clim.* **17**, 2399–2412 (2004).
15. J. Boucharel *et al.*, *Clim. Dyn.* **37**, 2045–2065 (2011).

16. D. W. Lea, *J. Clim.* **17**, 2170–2179 (2004).
17. T. Sagawa, Y. Yokoyama, M. Ikehara, M. Kuwae, *Geophys. Res. Lett.* **38**, L00F02 (2011).
18. P. N. DiNezio *et al.*, *Paleoceanography* **26**, PA3217 (2011).
19. D. H. Andreasen, A. C. Ravelo, *Paleoceanography* **12**, 395–413 (1997).
20. J. Xu, W. Kuhnt, A. Holbourn, M. Regenberg, N. Andersen, *Paleoceanography* **25**, PA4230 (2010).
21. P. Fiedler, L. Talley, *Prog. Oceanogr.* **69**, 143–180 (2006).
22. K. Thirumalai, J. W. Partin, C. S. Jackson, T. M. Quinn, *Paleoceanography* **28**, 401–412 (2013).
23. E. C. Brady, B. L. Otto-Bliesner, J. E. Kay, N. Rosenbloom, *J. Clim.* **26**, 1901–1925 (2013).
24. A. Patrick, R. Thunell, *Paleoceanography* **12**, 649–657 (1997).
25. H. Spero, K. Mielke, E. Kalve, D. Lea, D. Pak, *Paleoceanography* **18**, 1022 (2003).
26. D. Rincón-Martínez, S. Steph, F. Lamy, A. Mix, R. Tiedemann, *Mar. Micropaleontol.* **79**, 24–40 (2011).
27. H. M. Benway, A. C. Mix, B. A. Haley, G. P. Klinkhammer, *Paleoceanography* **21**, PA3008 (2006).
28. A. Koutavas, J. Lynch-Stieglitz, T. M. Marchitto Jr., J. P. Sachs, *Science* **297**, 226–230 (2002).
29. G. E. Manucharyan, A. V. Fedorov, *J. Clim.* **27**, 5836–5850 (2014).
30. Z.-Z. Hu *et al.*, *J. Clim.* **26**, 2601–2613 (2013).
31. M. Collins *et al.*, *Nat. Geosci.* **3**, 391–397 (2010).
32. N. A. Rayner *et al.*, *J. Geophys. Res.* **108**, 4407 (2003).

PALEOECOLOGY

Linked canopy, climate, and faunal change in the Cenozoic of Patagonia

Regan E. Dunn,[1]* Caroline A. E. Strömberg,[1] Richard H. Madden,[2] Matthew J. Kohn,[3] Alfredo A. Carlini[4]

Vegetation structure is a key determinant of ecosystems and ecosystem function, but paleoecological techniques to quantify it are lacking. We present a method for reconstructing leaf area index (LAI) based on light-dependent morphology of leaf epidermal cells and phytoliths derived from them. Using this proxy, we reconstruct LAI for the Cenozoic (49 million to 11 million years ago) of middle-latitude Patagonia. Our record shows that dense forests opened up by the late Eocene; open forests and shrubland habitats then fluctuated, with a brief middle-Miocene regreening period. Furthermore, endemic herbivorous mammals show accelerated tooth crown height evolution during open, yet relatively grass-free, shrubland habitat intervals. Our Patagonian LAI record provides a high-resolution, sensitive tool with which to dissect terrestrial ecosystem response to changing Southern Ocean conditions during the Cenozoic.

Vegetation structure—the degree of canopy openness—is a fundamental aspect of ecosystems, influencing productivity, hydrological and carbon cycling, erosion, and composition of faunal communities (*1, 2*). However, methods to quantify ancient vegetation structure have eluded paleoecologists. Here, we present a method with which to reconstruct vegetation openness, specifically leaf area index [LAI = foliage area (m²)/ground area (m²)], using the morphology of leaf epidermal cells preserved as phytoliths (plant biosilica) (Fig. 1). LAI quantifies vegetation structure in ecological and climate modeling studies (*1, 3*). In modern ecosystems, LAI relates primarily to soil moisture (*4*), by which vegetation becomes more closed with increasing soil water availability; ultimately, soil moisture is determined by temperature, precipitation, and atmospheric partial pressure of CO_2 (Pco_2) (*4, 5*). Disturbance in the form of fire and herbivory can offset this relationship, resulting in open habitats in areas with relatively high rainfall (*6*).

Using this paleobotanical archive, we reconstructed a LAI record for the middle Cenozoic [49 million to 11 million years ago (Ma)] of Patagonia to test predictions about vegetation

[1]Department of Biology and Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA. [2]Department of Organismal Biology and Anatomy, University of Chicago, Chicago, IL 60637, USA. [3]Department of Geosciences, Boise State University, Boise, ID 83725, USA. [4]Paleontología de Vertebrados, Universidad Nacional de La Plata, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), La Plata, Argentina.
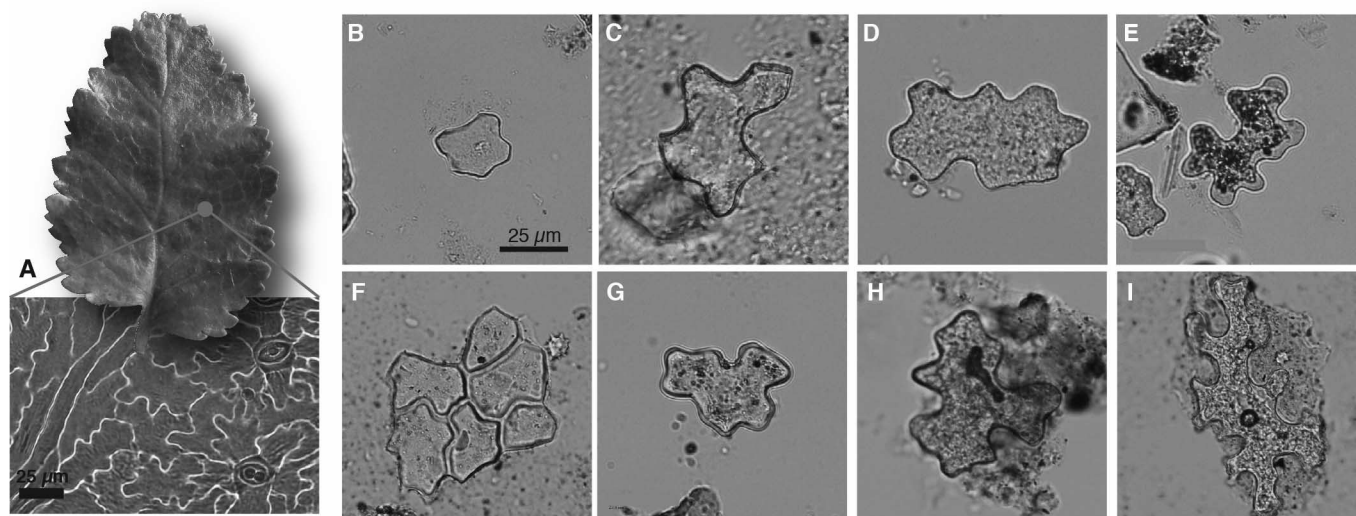*Corresponding author. E-mail: dunnr@u.washington.edu

Fig. 1. Leaf epidermis and examples of epidermal phytoliths. (**A**) *Nothofagus* leaf and epidermis. (**B** to **E**) Fossil phytoliths from Patagonia. (**F** to **I**) Modern soil phytoliths from Costa Rica.

response to Cenozoic climate fluctuations and how changes in vegetation structure relate to the evolution of high-crowned (hypsodont) and ever-growing (hypselodont) tooth morphologies in South American herbivores (7).

Climatic cooling beginning in the middle Eocene (49 Ma) and major warming events in the late Oligocene (~26 Ma) and middle Miocene (17 to 14 Ma) (8) have been linked to tectonics, ocean circulation (9), atmospheric P_{CO_2} (10), and ice volume after the onset of extensive Antarctic

glaciation at the Eocene–Oligocene Transition (EOT; 33.9 Ma) (8). A narrow landmass, Patagonia is sensitive to Southern Ocean climate and provides an ideal test case for the influence of global climate on vegetation structure.

It has long been assumed that hypsodonty in endemic South American herbivores beginning in the middle Eocene (~40 Ma) evolved in response to the spread of Earth's first grasslands (11), but recent work found that grasses were rare (12). When grasses are rare, "traditional" phyto-

lith analysis cannot resolve habitat openness (13, 14), so it has remained unclear whether hypsodonty evolved in forests or in open but grass-free habitats, possibly in conjunction with tooth abrasion during ingestion of exogenous grit (12).

Our proxy for reconstructing ancient LAI [reconstructed LAI (rLAI)] is based on the well-known influence of sunlight on the size and shape of pavement cells in leaf epidermis (Fig. 1A). In shade, these cells are larger and have more undulated outlines than those of cells exposed to full sun (15, 16). Silica mineralization produces a precision cast of pavement cells in living plants that can be preserved as fossils (Fig. 1B). Because sunlight filtering through a canopy is a function of LAI, we hypothesized that leaf epidermal cells and their phytoliths are on average larger and more undulated in closed forests than in open habitats and that the relationship is linear across a canopy density gradient. Because these phytolith types are taxonomically nondiagnostic we cannot control for phylogenetic variation in cell morphology. Instead, we tested our hypotheses using modern assemblages of phytoliths extracted from soils collected across an LAI gradient from 0 (completely open) to 5 (dense forest) in Costa Rica (Fig. 1C).

Cenozoic-aged floras from Patagonia contain a nonanalog mix of mesic and xeric taxa of tropical and sub-Antarctic lineages (such as Arecaceae, Anacardiaceae, Fabaceae, Zingiberaceae, Proteaceae, *Nothofagus*, Podocarpaceae, and Araucariaceae). We chose to sample phytoliths from modern tropical soils in Costa Rica because it has wet and dry forests, savanna, and shrubby areas containing many of the reported fossil genera (41%) and families (85%) (table S1). We assume that relative change in epidermal cell morphology is based on canopy density and is independent of taxonomy and latitudinal variation in light regime. Using light microscopy, nongrass epidermal phytoliths in extracted samples were photographed and measured for the calibration data set. Phytolith undulation was standardized by using an undulation index (UI = circumference of cell/circumference of a circle with cell area) (16), and mean site values for phytolith UI (PUI) and phytolith area (PA) were compared with field measurements of LAI from hemispherical photographs (Fig. 2A). Measurements of fossil phytoliths followed the same protocol.

In the modern training data set of 45 sites (table S2), LAI correlates with PUI [coefficient of determination (R^2) = 0.59, $P < 0.0001$] (Fig. 2B) and PA ($R^2 = 0.44$, $P < 0.0001$) (Fig. 2C). A linear multiple regression analysis including both variables improves the correlation (Fig. 2D and table S3):

$$rLAI = 0.0012 \times PA(\mu m^2) + 10.4118 \times PUI - 13.1621 \qquad (1)$$

($R^2 = 0.63$, $P < 0.0001$, SE = 0.695, $F_{2,42} = 39$). Using Eq. 1, we reconstructed rLAI for 46 fossil phytolith assemblages from Patagonian paleosols spanning 49 to 11 Ma (Fig. 3A, fig. S3, and
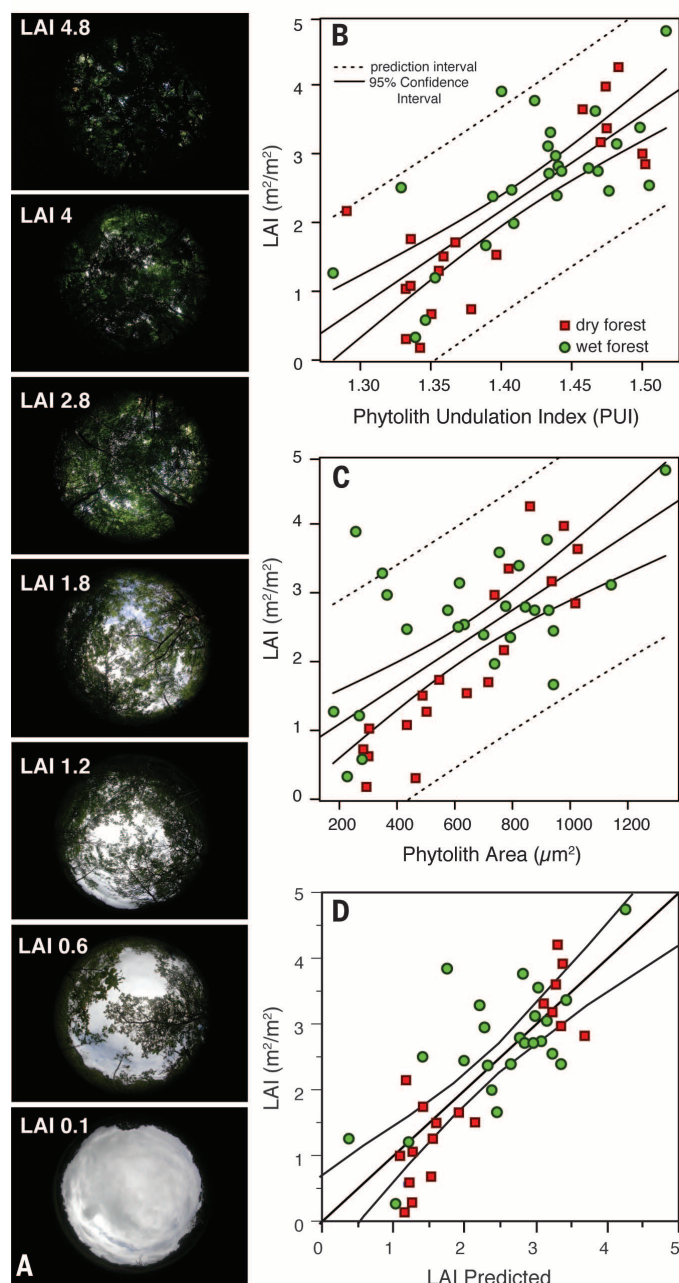

Fig. 2. Modern soil phytolith morphology and LAI. (**A**) Hemispherical photographs from Costa Rica illustrating LAI values. (**B**) Linear regressions for mean PUI and LAI (rLAI = 13.92 × PUI – 17.31); and (**C**) mean Phytolith Area (PA) and LAI (rLAI = 0.0028 × PA + 0.531). (**D**) Plot of simulated versus measured values of LAI. Simulated values are derived from Eq. 1.

table S4). Data from different times should be comparable because all samples are from the same region with the same moisture resources (for example, similar vegetation occupy all sites today). From the middle Eocene to early Oligocene (49.0 to 32.3 Ma), rLAI values decline from

Fig. 3. Middle Cenozoic rLAI and comparisons to climate and biotic records. (**A**) rLAI reconstruction with 95% confidence intervals. Shown are raw values (open circles) and binned values by age or geologic unit (red diamonds) (table S4). (**B**) Foraminiferal $\delta^{18}O$ records for sea surface (red) (21) and deep sea temperature (blue) (8). (**C**) MAT (blue), MAP (green), and dry-season precipitation (orange) estimates from megafloras (27, 33), with corrected ages (table S5). Length of the solid bars indicate age uncertainty. Yellow boxes are MAP estimates from paleosols (34). (**D**) Proportion hypsodont+hypselodont taxa of notoungulates, with 95% confidence intervals (table S6).

~4 to 0.6, indicating an opening of the landscape, from dense vegetation (such as broadleaf forest; LAI = 4 to 3) to progressively more open vegetation (such as dry forest, woodland, and scrub; LAI = 2 to 1); and last, to very open habitats (such as shrubland or desert; LAI < 1) (Table 1 and Fig. 3). High rLAI values during the middle Eocene correspond in age with the highly diverse megathermal floras 380 km farther north (17), and middle Eocene–early Oligocene decreases in rLAI correspond with increased abundances of sub-Antarctic taxa such as *Nothofagus* (18, 19).

Between 38.5 and 38.0 Ma and 35.0 and 32.2 Ma, habitats were maximally open (rLAI <1). Diagnostic phytoliths from these assemblages indicate abundant palms, woody eudicotyledons, and sparse grasses (<1 to 15%) (12). We interpret these habitats as a nonanalog palm shrubland with a discontinuous canopy. During the late Oligocene, rLAI increases, indicating dry forest, woodland, or scrub until the earliest Miocene (~21.1 and 18.8 Ma), when rLAI drops again. Early Miocene rLAI fluctuations (<1 to 2.4) suggest shifts between palm shrubland and open forest without a continuous grassy understory (12). The middle Miocene (~15.7 to 14.6 Ma) saw a brief spike in rLAI values (>3.5) at 14.6 Ma, after which (~14.2 Ma) rLAI values trend downward again. Pollen records from late Oligocene–middle Miocene marine strata of Patagonia indicate humid forest conditions dominated by *Nothofagus* and podocarps, with low abundances of arid-adapted (inferred as open habitat) taxa before 10 Ma (20). These pollen data do not contradict our findings because they reflect regional vegetation and are biased toward prolific pollen producers, whereas our phytolith samples represent in situ vegetation of the central Patagonian lowlands.

Broad changes in rLAI track the Southern Ocean $\delta^{18}O$ temperature record (8, 21); rLAI decreases during mid–late Eocene cooling and increases during late Oligocene warming. The middle Miocene (~15.7 to 14.6 Ma) regreening of Patagonia coincides with increased atmospheric P_{CO_2} (22) and reduced Antarctic ice sheet volume (23). After ~14.2 Ma, declining rLAI values are consistent with middle Miocene cooling, ice-growth, and modeled changes in meridional heat and vapor transport (24).

Vegetation-climate decoupling occurred during pulses of maximum openness at 38.5 to 38.0 Ma and at ~35 Ma, as marine temperatures gradually declined. This second pulse predates the EOT by >1 million years; in contrast, during abrupt EOT cooling, rLAI remained relatively unchanged. Quasi-constant rLAI during the EOT is consistent with phytolith abundance data, suggesting compositionally stable vegetation (12) and isotopically inferred invariant temperatures (25). However, aeolian sedimentation rates dramatically increased at 33.6 Ma (26), which is consistent with Oi-1 glaciation.

Intervals of open vegetation likely reflect reduced precipitation, although disturbance such as fire, volcanism, or herbivory may have contributed. Megafloral records from elsewhere in
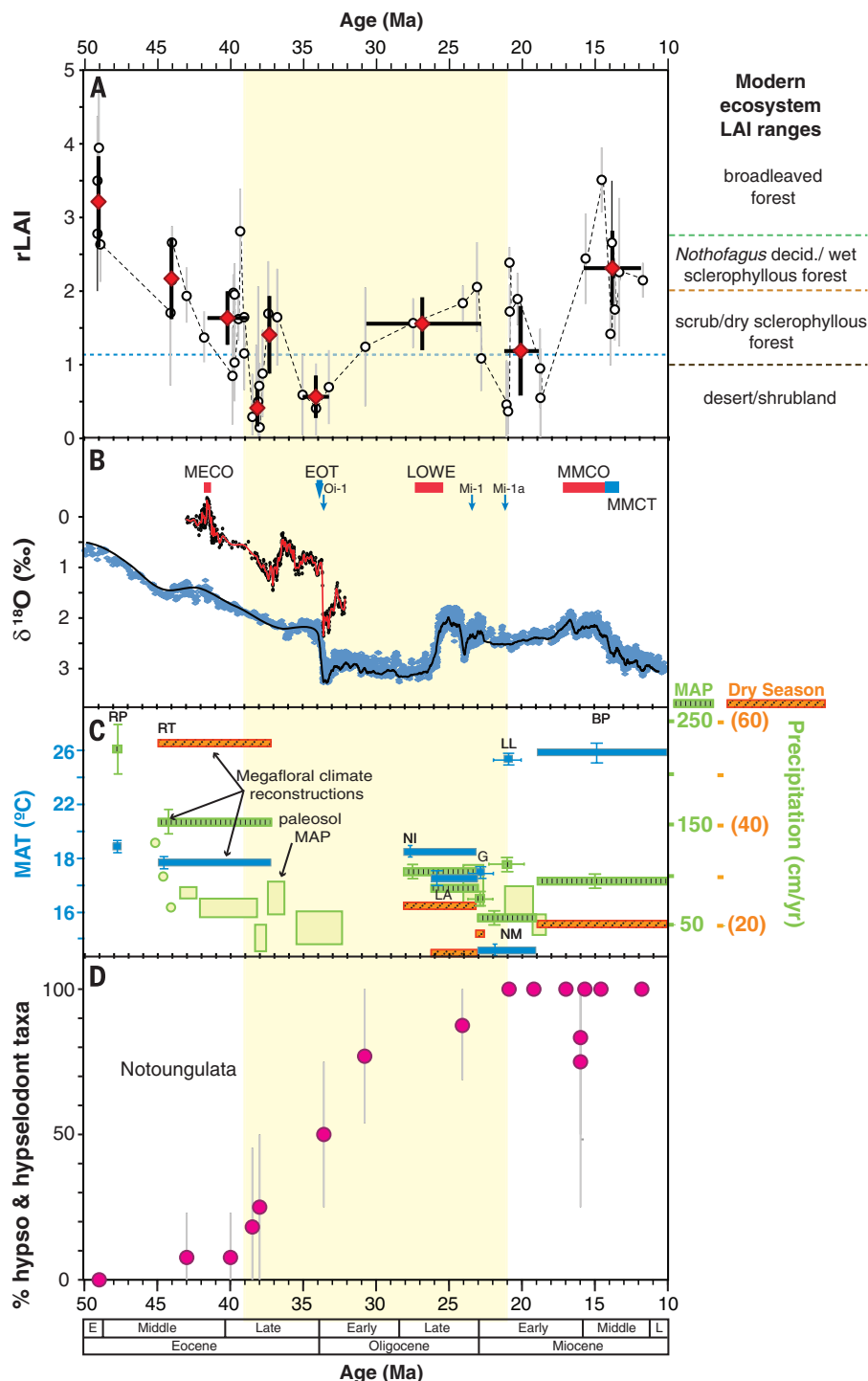
Table 1. Modern habitat LAI ranges from literature. Dashes indicate no reported data.

Biome type	Mean LAI	Minimum	Maximum	Reference
Wetlands	6.30	2.5	8.4	(1)
Tropical evergreen broad	4.80	1.5	8	(1)
Temperate evergreen broadleaf	4.33	0.8	11.6	(1, 35)
Tropical deciduous broadleaf	3.90	0.6	8.9	(1)
Temperate deciduous broadleaf	3.56	0.6	5.08	(1, 35)
Nothofagus evergreen	3.50	2.5	4.5	(36)
Nothofagus deciduous	2.55	2.3	2.8	(36)
Sclerophyllous forest Australia	2.40	—	—	(37)
Grasslands	1.75	0.08	5	(1, 35)
Monte	1.70	0.5	2.9	(38)
Mediterranean scrub	1.50	1	4	(39)
Chaco (dry forest)	1.5	1	3	(40)
Savanna	1.49	0.78	1.72	(35)
Sclerophyllous forest Australia	0.95	—	—	(37)
Shrublands	0.77	0.1	4.5	(1, 35, 38)
Desert	0.55	0.9	0.2	(1)

Patagonia document stable mean annual temperatures (MATs) of ~18°C, but decreasing mean annual precipitation (MAP) from the middle Eocene onward (Fig. 3C) (27). By at least the late Oligocene, decreased MAP values reflect reduced dry-season rainfall (27). Locally, episodes of low rLAI correspond to the lowest MAP estimates from paleosols and shifts to aeolian sedimentation (28). Additionally, phytoliths of water-demanding gingers become very rare (0.4%) by 38.1 Ma and disappear after ~38 Ma (12). Our climate interpretation is seemingly at odds with phytolith evidence for abundant palms, which in modern South America is linked to warm, humid climates (29). However, Patagonian fossils indicate that a largely dry-adapted palm clade (Attaleinae) had diversified in South America by the Paleocene (fig. S5). We hypothesize that water-use efficiency in these palms was further enhanced under elevated Eocene atmospheric P_{CO_2} (30).

Increasing openness (rLAI < 1) ~40 Ma coincided with initiation of tooth crown height increases in several clades of notoungulates (Fig. 3D). Proportions of hypsodont+hypselodont taxa continued to rise from 38 to 20 Ma, as rLAI remained low (between 0 and 2; average ≤ 1.5). The hypsodonty trend may have reversed during more forested middle Miocene conditions, but errors are large, and constant hypsodonty proportions cannot be ruled out. In modern South American environments, the proportion of hypsodont+hypselodont species dramatically increases under a LAI value of ~1.2 (fig. S6 and table S7). These areas experience low precipitation, frequent dust storms, and erosion of tephric materials (31).

Evidently, feeding in drier, more open Eocene–early Miocene ecosystems provided evolutionary pressure to drive hypsodonty and hypselodonty in Patagonia. The temporal coincidence of wind-blown ash, low rLAI, and increased rates of hypsodonty+hypselodonty further suggests that ash played a key role in this process (12). In low

LAI habitats today (such as shrublands), sparse vegetation includes both bare ground (dust source areas) and shrubs (traps for dust) (32). Thus, ingestion of dust adhering to plants growing on highly erodible surfaces (tephra-rich soils) plausibly drove this pattern of tooth evolution in South America.

Taken together, these patterns indicate that long-term climate changes that predated the EOT drove ecosystem changes in Patagonia. Specifically, we propose that Southern Ocean instability during the protracted opening of Drake Passage beginning in the middle Eocene (9) and associated cooling sea surface temperatures resulted in reduced rainfall on land and triggered successive opening-up of landscapes during the middle-late Eocene. Our method for estimating rLAI allows for quantification of vegetation structure through time, and because it relies on microfossils, extremely high-resolution records of habitat change are possible. Additionally, because leaf epidermis is a highly conserved tissue, the method should be applicable across a broad range of temporal scales to test many outstanding hypotheses in paleoecology.

REFERENCES AND NOTES

1. G. P. Asner, J. M. O. Scurlock, J. A. Hicke, *Glob. Ecol. Biogeogr.* **12**, 191–205 (2003).
2. R. F. Kay, R. H. Madden, C. Van Schaik, D. Higdon, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 13023–13027 (1997).
3. R. Betts, P. Cox, S. Lee, F. Woodward, *Nature* **387**, 796–799 (1997).
4. O. Dermody, J. F. Weltzin, E. C. Engel, P. Allen, R. J. Norby, *Plant Soil* **301**, 255–266 (2007).
5. A. Iio, K. Hikosaka, N. P. R. Anten, Y. Nakagawa, A. Ito, *Glob. Ecol. Biogeogr.* **23**, 274–285 (2013).
6. A. C. Staver, S. Archibald, S. A. Levin, *Science* **334**, 230–232 (2011).
7. Materials and methods are available as supplementary materials on *Science* Online.
8. J. Zachos, M. Pagani, L. Sloan, E. Thomas, K. Billups, *Science* **292**, 686–693 (2001).
9. R. Livermore, C.-D. Hillenbrand, M. Meredith, G. Eagles, *Geochem. Geophys. Geosyst.* **8**, Q01005 (2007).
10. M. Pagani, J. C. Zachos, K. H. Freeman, B. Tipple, S. Bohaty, *Science* **309**, 600–603 (2005).
11. B. F. Jacobs, J. D. Kingston, L. L. Jacobs, *Ann. Mo. Bot. Gard.* **86**, 590–643 (1999).
12. C. A. E. Strömberg, R. E. Dunn, R. H. Madden, M. J. Kohn, A. A. Carlini, *Nat. Commun.* **4**, 1478 (2013).
13. C. A. E. Strömberg, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **207**, 239–275 (2004).
14. L. Bremond, A. Alexandre, C. Hély, J. Guiot, *Global Planet. Change* **45**, 277–293 (2005).
15. R. W. Watson, *New Phytol.* **41**, 223–229 (1942).
16. W. M. Kürschner, *Rev. Palaeobot. Palynol.* **96**, 1–30 (1997).
17. P. Wilf *et al.*, *Am. Nat.* **165**, 634–650 (2005).
18. V. Barreda, L. Palazzesi, *Bot. Rev.* **73**, 31–50 (2007).
19. E. J. Romero, *Ann. Mo. Bot. Gard.* **73**, 449–461 (1986).
20. L. Palazzesi, V. Barreda, *Nat. Commun.* **3**, 1294 (2012).
21. S. M. Bohaty, J. C. Zachos, *Geology* **31**, 1017–1020 (2003).
22. Y. G. Zhang, M. Pagani, Z. Liu, S. M. Bohaty, R. Deconto, *Philos. Trans. A Math. Phys. Eng. Sci.* **371**, 20130096 (2013).
23. S. Warny *et al.*, *Geology* **37**, 955–958 (2009).
24. A. E. Shevenell, J. P. Kennett, D. W. Lea, *Science* **305**, 1766–1770 (2004).
25. M. J. Kohn *et al.*, *Geology* **32**, 621 (2004).
26. R. E. Dunn *et al.*, *Geol. Soc. Am. Bull.* **125**, 539–555 (2013).
27. L. Hinojosa, *Rev. Geológica Chile* **32**, 95–115 (2005).
28. E. S. Bellosi, in *The Paleontology of Gran Barranca: Evolution and Environmental Change Through the Middle Cenozoic of Patagonia*, R. H. Madden, A. A. Carlini, M. G. Vucetich, R. F. Kay, Eds. (Cambridge Univ Press, Cambridge, UK, 2010), pp. 278–292.
29. S. W. Punyasena, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **265**, 226–237 (2008).
30. M. H. Ibrahim, H. Z. E. Jaafar, M. H. Harun, M. R. Yusop, *Acta Physiol. Plant.* **32**, 305–313 (2009).
31. R. H. Madden, *Hypsodonty in Mammals: Evolution, Geomorphology and the Role of Earth Surface Processes* (Cambridge Univ. Press, Cambridge, UK, 2014).
32. D. D. Breshears, J. J. Whicker, M. P. Johansen, J. E. Pinder, *Earth Surf. Process. Landf.* **28**, 1189–1209 (2003).
33. L. F. Hinojosa, C. Villagrán, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **217**, 1–23 (2005).
34. E. S. Bellosi, M. G. Gonzalez, in *The Paleontology of Gran Barranca: Evolution and Environmental Change through the Middle Cenozoic of Patagonia*, R. H. Madden, A. A. Carlini, M. G. Vucetich, R. F. Kay, Eds. (Cambridge Univ. Press, 2010), pp. 293–305.
35. S. Ganguly *et al.*, *Remote Sens. Environ.* **112**, 4318–4332 (2008).
36. E. Gutiérrez, V. Vallejo, *Oecologica Aquat.* **10**, 351–366 (1991).
37. W. Woodgate, *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **XXXIX**, 457–462 (2012).
38. G. P. Asner, C. E. Borghi, R. A. Ojeda, *Ecol. Appl.* **13**, 629–648 (2003).
39. M. Barradas, F. Novo (*Folia Geobot. Phytotaxon.* **22**, 415–433 (1987).
40. V. Barraza *et al.*, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **7**, 421–430 (2013).

WOMEN IN SCIENCE

Expectations of brilliance underlie gender distributions across academic disciplines

Sarah-Jane Leslie,[1]*† Andrei Cimpian,[2]*† Meredith Meyer,[3] Edward Freeland[4]

The gender imbalance in STEM subjects dominates current debates about women's underrepresentation in academia. However, women are well represented at the Ph.D. level in some sciences and poorly represented in some humanities (e.g., in 2011, 54% of U.S. Ph.D.'s in molecular biology were women versus only 31% in philosophy). We hypothesize that, across the academic spectrum, women are underrepresented in fields whose practitioners believe that raw, innate talent is the main requirement for success, because women are stereotyped as not possessing such talent. This hypothesis extends to African Americans' underrepresentation as well, as this group is subject to similar stereotypes. Results from a nationwide survey of academics support our hypothesis (termed the field-specific ability beliefs hypothesis) over three competing hypotheses.

Laboratory, observational, and historical evidence reveals pervasive cultural associations linking men but not women with raw intellectual talent (1–4). Given these ambient stereotypes, women may be underrepresented in academic disciplines that are thought to require such inherent aptitude. We term this the field-specific ability beliefs hypothesis (fig. S1).

Current discourse about women in academia focuses mainly on women's underrepresentation in (natural) science, technology, engineering, and mathematics (STEM) (5). However, STEM disciplines vary in their female representation (fig. S2) (5, 6). Recently, women have earned approximately half of all Ph.D.'s in molecular biology and neuroscience in the United States, but fewer than 20% of all Ph.D.'s in physics and computer science (7). The social sciences and humanities (SocSci/Hum) exhibit similar variability. Women are currently earning more than 70% of all Ph.D.'s in art history and psychology, but fewer than 35% of all Ph.D.'s in economics and philosophy (7). Thus, broadening the scope of inquiry beyond STEM fields might reveal new explanations and solutions for gender gaps (8). We offer evidence that the field-specific ability beliefs hypothesis can account for the distribution of gender gaps across the entire academic spectrum.

Individuals' beliefs about what is required for success in an activity vary in their emphasis on fixed, innate talent (9). Similarly, practitioners of different disciplines may vary in the extent to which they believe that success in their discipline requires such talent. Because women are often negatively stereotyped on this dimension (1–4), they may find the academic fields that emphasize such talent to be inhospitable. There are several mechanisms by which these field-specific ability beliefs might influence women's participation. The practitioners of disciplines that emphasize raw aptitude may doubt that women possess this sort of aptitude and may therefore exhibit biases against them (10). The emphasis on raw aptitude may activate the negative stereotypes in women's own minds, making them vulnerable to stereotype threat (11). If women internalize the stereotypes, they may also decide that these fields are not for them (12). As a result of these processes, women may be less represented in "brilliance-required" fields.

We used a large-scale, nationwide study of academics from 30 disciplines to evaluate the field-specific ability beliefs hypothesis, along with three competing hypotheses. The first competitor concerns possible gender differences in willingness or ability to work long hours (13): The more demanding a discipline in terms of work hours, the fewer the women. The second competing hypothesis concerns possible gender differences at the high end of the aptitude distribution [(14, 15); but see (16, 17) for criticism]. Such differences might cause greater gender gaps in fields that, by virtue of their selectivity, sample from the extreme right of the aptitude distribution: The more selective a discipline, the fewer the women. The third competing hypothesis concerns possible differences among

fields in the extent to which they require systemizing (the ability to think systematically and abstractly) or empathizing (the ability to understand thoughts and emotions in an insightful way): The more a discipline prioritizes systemizing over empathizing, the fewer the women (14, 18, 19). Our findings suggest that the field-specific ability beliefs hypothesis, unlike these three competitors, is able to predict women's representation across all of academia, as well as the representation of other similarly stigmatized groups (e.g., African Americans).

We surveyed faculty, postdoctoral fellows, and graduate students (N = 1820) from 30 disciplines (12 STEM, 18 SocSci/Hum) (table S1) at geographically diverse high-profile public and private research universities across the United States. Participants were asked questions concerning their own discipline (table S2); responses in each discipline were averaged (tables S3 and S4), and analyses were conducted over disciplines (not individuals). As our dependent measure, we used the percentage of female Ph.D. recipients in each discipline (7).

To assess field-specific ability beliefs, we asked participants to rate their agreement with four statements concerning what is required for success in their field (e.g., "Being a top scholar of [discipline] requires a special aptitude that just can't be taught") (table S2). Respondents rated both the extent to which they personally agreed with these statements, and the extent to which they believed other people in their field would agree with the statements. Because answers to these eight questions displayed very similar patterns (α = 0.90), they were averaged to produce a field-specific ability belief score for each discipline (with higher scores indicating more emphasis on raw ability). As predicted, the more a field valued giftedness, the fewer the female Ph.D.'s. Field-specific ability belief scores were significantly correlated with female representation across all 30 fields [correlation coefficient $r(28) = -0.60, P < 0.001$], in STEM alone [$r(10) = -0.64, P = 0.025$], and in SocSci/Hum alone [$r(16) = -0.62, P = 0.006$] (Fig. 1). In a hierarchical regression with a STEM indicator variable entered in the first step and field-specific ability belief scores entered in the second (Table 1, models 1 and 2), adding the ability belief variable significantly increased the variance accounted for, $\Delta R^2 = 0.29, P < 0.001$.

To assess work demands, we asked participants to report the number of hours they worked per week, on-campus and off-campus (table S2). There was no correlation between the total number of hours worked (on- plus off-campus) and female representation, $r(28) = -0.03, P = 0.895$. Women tended to be underrepresented in fields whose practitioners worked more on-campus hours, but this correlation was not significant either, $r(28) = -0.32, P = 0.088$. No significant correlations with on-campus hours were found either within STEM, $r(10) = 0.46, P = 0.131$ (note the positive coefficient here), or within SocSci/Hum, $r(16) = -0.07, P = 0.772$. Adding on-campus hours to the hierarchical regression predicting

[1]Department of Philosophy, Princeton University, Princeton, NJ 08544, USA. [2]Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA. [3]Department of Psychology, Otterbein University, Westerville, OH 43081, USA. [4]Survey Research Center, Princeton University, Princeton, NJ 08544, USA.
*These authors contributed equally to the work. †Corresponding author. E-mail: sjleslie@princeton.edu (S.-J.L.) or acimpian@illinois.edu (A.C.)
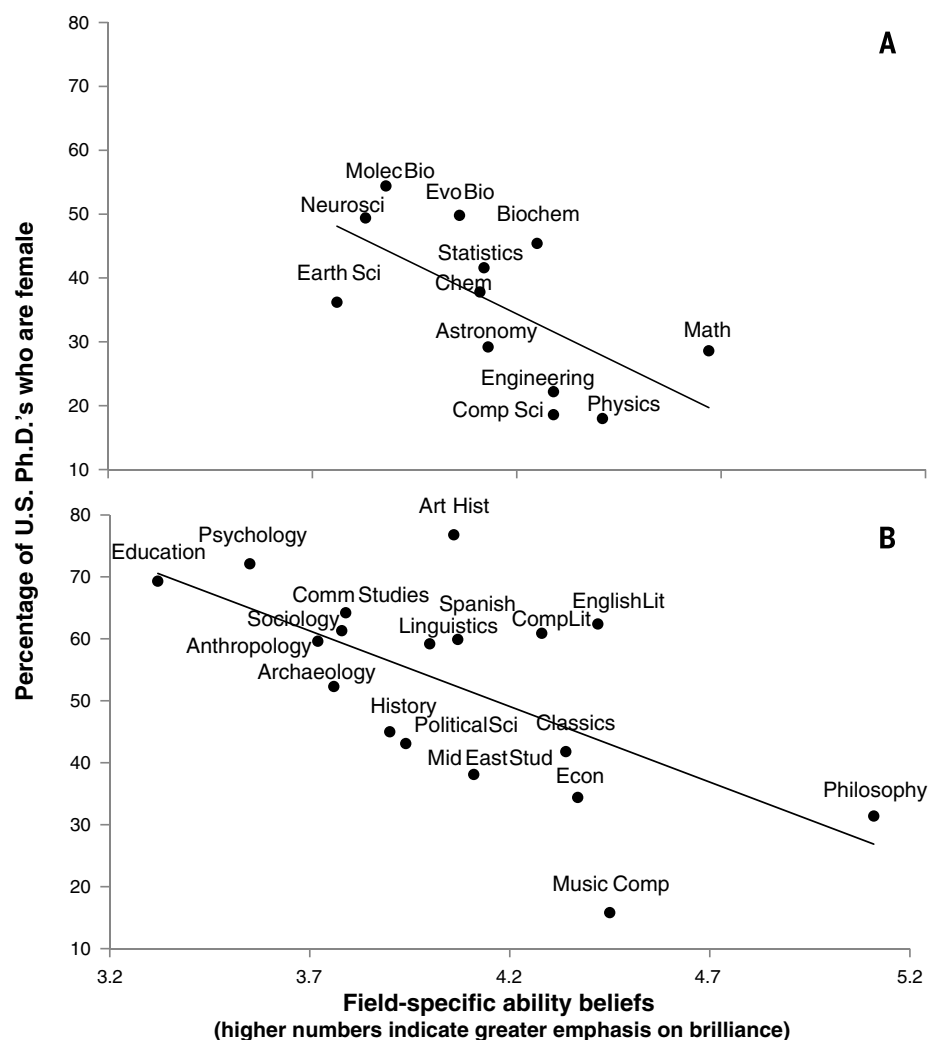
Fig. 1. Field-specific ability beliefs and the percentage of female 2011 U.S. Ph.D.'s in (**A**) **STEM and** (**B**) **Social Science and Humanities.**

women's representation did not significantly increase the variance accounted for, $\Delta R^2 < 0.01$, $P = 0.687$ (Table 1, model 3) [Similar results were obtained with total hours worked, as detailed in the supplementary materials (SM).] Thus, differences between fields in hours worked did not explain variance in the distribution of gender gaps beyond that explained by field-specific ability beliefs and the STEM indicator variable.

To assess selectivity, we asked faculty participants to estimate the percentage of graduate applicants admitted each year to their department. We then reverse-coded this measure so that higher values indicate more selectivity. Fields that were more selective tended to have higher, rather than lower, female representation, but this correlation did not reach significance, $r(28) = 0.34$, $P = 0.065$. Further, this selectivity measure did not predict female representation in STEM alone or in SocSci/Hum alone (both $Ps > 0.478$), and adding it to the hierarchical regression did not result in a statistically significant increase in the variance accounted for, $\Delta R^2 = 0.04$, $P = 0.134$ (Table 1, model 4). (An analysis considering only selectivity measures from top-10% departments produced the same pattern of results; see the SM.) To account for potential differences in the strength of the applicant pools across disciplines, we compared the 2011–2012 Graduate Record Examination (GRE) General Test scores of Ph.D. applicants. These data were available for only 19 of the disciplines in our study (7 STEM and 12 SocSci/Hum) (*20*). A composite measure of GRE scores was not significantly correlated with female representation, $r(17) = -0.24$, $P = 0.333$, and so provided no evidence that fields with more women have weaker applicant pools. Further, the relation between field-specific ability beliefs and female representation remained significant when adjusting for GRE scores, $r(16) = -0.57$, $P = 0.013$.

Table 1. Hierarchical regression models predicting female representation. $N = 30$ disciplines. Significant statistics are bold. R^2 comparisons are always with the preceding model (to the left).

Predictor	Model 1			Model 2			Model 3			Model 4			Model 5		
	β	t	P	β	t	P	β	t	P	β	t	P	β	t	P
STEM indicator	**−0.50****	**−3.03**	**0.005**	**−0.42****	**−3.20**	**0.003**	−0.35	−1.49	0.148	−0.30	−1.34	0.193	−0.28	−1.07	0.297
Field-specific ability beliefs				**−0.55*****	**−4.13**	**<0.001**	**−0.56*****	**−3.98**	**<0.001**	**−0.58*****	**−4.17**	**<0.001**	**−0.56****	**−3.46**	**0.002**
On-campus hours worked							−0.09	−0.41	0.687	−0.01	−0.03	0.975	0.02	0.07	0.945
Selectivity										0.24	1.55	0.134	0.24	1.54	0.137
Systemizing versus empathizing													−0.06	−0.23	0.817
R^2	**0.25**			**0.54**			0.54			0.58			0.58		
F for change in R^2	**9.19****			**17.08*****			0.17			2.40			0.06		
P for change in R^2	**0.005**			**<0.001**			0.687			0.134			0.817		

**$P < 0.01$.   ***$P < 0.001$.

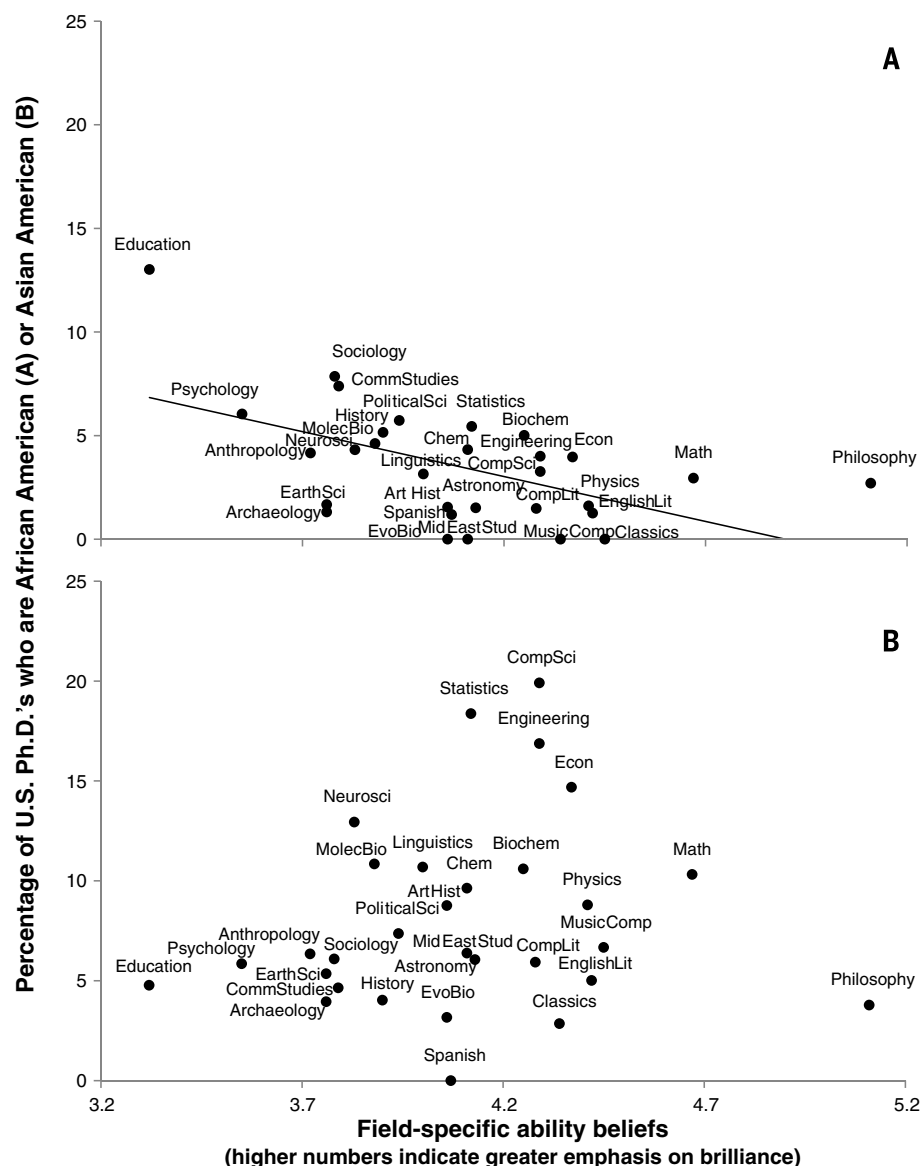**Fig. 2. Field-specific ability beliefs and the percentage of 2011 U.S. Ph.D.'s who are** (**A**) **African American and** (**B**) **Asian American.**

To assess systemizing versus empathizing, we asked participants to evaluate the extent to which scholarly work in their discipline requires these two abilities (two questions each, table S2) (αs = 0.63 and 0.90, respectively). A composite systemizing-minus-empathizing score was significantly correlated with female representation across all disciplines, $r(28) = -0.53$, $P = 0.003$. However, this score did not significantly predict female representation in STEM alone, $r(10) = -0.27$, $P = 0.402$, or in SocSci/Hum alone, $r(16) = -0.25$, $P = 0.310$. Adding systemizing versus empathizing composite scores to the hierarchical regression did not increase the variance accounted for, $\Delta R^2 < 0.01$, $P = 0.817$ (Table 1, model 5). Indeed, field-specific ability beliefs were the sole significant predictor of female representation in this final model, $\beta = -0.56$, $P = 0.002$.

To further compare our hypothesis to its competitors, we performed another hierarchical regression in which the STEM indicator, on-campus hours, selectivity, and systemizing versus empathizing were all added together as a first step. The model was significant, $R^2 = 0.38$, $P = 0.016$, although no predictor in it was. When field-specific ability belief scores were added, the variance accounted for increased to 58%, $\Delta R^2 = 0.21$, $P = 0.002$. This finding reflects once again the predictive power of the field-specific ability beliefs hypothesis. (However, we do not claim that field-specific ability beliefs are the sole determiner of gender gaps or that these three are the only alternative hypotheses; other factors undoubtedly play a role. Further, this study was not designed to eliminate the three competing hypotheses but

rather to use them as a benchmark for our hypothesis.)

To check for robustness, we duplicated our analyses using weights created by comparing the demographic characteristics of our respondents against the demographics of the population of academics we initially contacted. Adding these post-stratification weights to our analyses helped us to check for the influence of bias resulting from differential nonresponse (*21, 22*) (for details, see the SM). In a weighted version of the final model in the hierarchical regression, field-specific ability beliefs were again the only significant predictor of female representation, $\beta = -0.40$, $P = 0.029$ (see table S7 for weighted versions of all models).

There are many potential mechanisms by which field-specific ability beliefs may influence women's representation. To assess some possibilities, we asked participants to evaluate the statement, "Even though it's not politically correct to say it, men are often more suited than women to do high-level work in [discipline]." Participants rated their own agreement and the extent to which they thought that other people in their field would agree. These two scores were averaged into a single measure ($\alpha = 0.80$). Disciplines that emphasized raw talent were more likely to endorse the idea that women are less suited for high-level scholarly work, $r(28) = 0.38$, $P = 0.036$. In turn, higher endorsement of this idea was associated with lower female representation, $r(28) = -0.67$, $P < 0.001$. We also asked participants to rate whether they thought that their discipline was welcoming to women (table S2). Disciplines that valued giftedness over dedication rated themselves as less welcoming to women, $r(28) = -0.58$, $P = 0.001$, and fields that viewed themselves as less welcoming had fewer female Ph.D.'s, $r(28) = 0.74$, $P < 0.001$. Together, ratings of whether women were suitable for and welcome in a discipline mediated 70.2% of the relation between field-specific ability beliefs and the percentage of female Ph.D.'s (bootstrapped $P < 0.001$) (*23*). Thus, field-specific ability beliefs may lower women's representation at least in part by fostering the belief that women are less well-suited than men to be leading scholars and by making the atmosphere in these fields less welcoming to women.

Like women, African Americans are stereotyped as lacking innate intellectual talent (*24*). Thus, field-specific ability belief scores should predict the representation of African Americans across academia. Indeed, African Americans were less well represented in disciplines that believed giftedness was essential for success, $r(28) = -0.54$, $P = 0.002$ (Fig. 2). However, field-specific ability belief scores should not predict the representation of Asian Americans, who are not stereotyped in the same way (*25*). Indeed, Asian American representation was not correlated with field-specific ability beliefs, $r(28) = 0.16$, $P = 0.386$. A hierarchical regression with the STEM indicator, total hours worked (the same results are found with on-campus hours only), selectivity, and systemizing versus empathizing, all entered together in

Table 2. Hierarchical regression models predicting African American representation. $N = 30$ disciplines. Significant statistics are bold. R^2 comparisons are always with the preceding model (to the left).

Predictor	Model 1			Model 2		
	β	t	P	β	t	P
STEM indicator	0.05	0.15	0.878	−0.06	−0.21	0.833
Total hours worked	−0.13	−0.58	0.571	−0.26	−1.30	0.207
Selectivity	−0.21	−0.88	0.387	−0.19	−0.93	0.359
Systemizing versus empathizing	−0.23	−0.71	0.487	0.10	0.35	0.733
Field-specific ability beliefs				**−0.61****	**−3.28**	**0.003**
R^2		0.06			**0.35**	
F for change in R^2		0.37			**10.76****	
P for change in R^2		0.827			**0.003**	

**$P < 0.01$.

the first step predicted very little variance in African Americans' representation, $R^2 = 0.06$, $P = 0.827$ (Table 2, model 1). Adding field-specific ability beliefs in the second step significantly increased the variance accounted for, $\Delta R^2 = 0.29$, $P = 0.003$ (Table 2, model 2). As with women, field-specific ability beliefs were the only significant predictor of African American representation (with total hours worked: β = −0.61, $P = 0.003$; with on-campus hours: β = −0.64, $P = 0.005$).

Finally, we considered alternative explanations for our results. If women believe more strongly than men in the value of hard work, disciplines with fewer women may have higher field-specific ability belief scores precisely for that reason—because they have fewer women. Relative to men, women in our survey did report lower field-specific ability belief scores and thus more belief in the importance of dedication ($M_{women} = 3.86$ versus $M_{men} = 4.24$), $t(1812) = 7.31$, $P < 0.001$. However, contrary to this alternative hypothesis, the field-specific ability belief scores derived separately from each discipline's female and male respondents were independently predictive of the percentage of female Ph.D.'s across the 30 fields: $r(28) = −0.40$, $P = 0.028$, for women's scores, and $r(28) = −0.50$, $P = 0.005$, for men's. We also constructed a gender-balanced field-specific ability belief score for each discipline by computing the average scores for men and women in that discipline and then averaging the two gender-specific scores. This measure weights women's and men's scores equally, regardless of their actual representation in a field, and was again the only variable that significantly predicted women's representation in the regression model that included all competitors plus the STEM indicator, β = −0.49, $P = 0.009$. Thus, the relation between field-specific ability beliefs and women's representation is not a simple matter of men and women valuing effort differently.

Is it possible that people simply infer what is required for success in a field on the basis of their estimates of that field's diversity, so that they assume a field requires less brilliance if women are well represented in it? Contrary to this alternative hypothesis, a regression analysis performed on the individual-level data revealed that academics' perceptions of the diversity of their field (see table S2) were in fact not a significant predictor of their field-specific ability beliefs, β = −0.05, $P = 0.277$. This analysis included indicator variables for each discipline so as to minimize the influence of discipline-specific unobserved variables (*26*). Thus, the relation between field-specific ability beliefs and women's representation is not a simple matter of using women's representation in a field to infer what is required for success.

Is natural brilliance truly more important to success in some fields than others? The data presented here are silent on this question. However, even if a field's beliefs about the importance of brilliance were to some extent true, they may still discourage participation among members of groups that are currently stereotyped as not having this sort of brilliance. As a result, fields that wished to increase their diversity may nonetheless need to adjust their achievement messages.

Are women and African Americans less likely to have the natural brilliance that some fields believe is required for top-level success? Although some have argued that this is so, our assessment of the literature is that the case has not been made that either group is less likely to possess innate intellectual talent (as opposed to facing stereotype threat, discrimination, and other such obstacles) (*10, 16, 17, 24, 27*).

The extent to which practitioners of a discipline believe that success depends on sheer brilliance is a strong predictor of women's and African Americans' representation in that discipline. Our data suggest that academics who wish to diversify their fields might want to downplay talk of innate intellectual giftedness and instead highlight the importance of sustained effort for top-level success in their field. We expect that such easily implementable changes would enhance the diversity of many academic fields.

REFERENCES AND NOTES

1. M. Bennett, *J. Soc. Psychol.* **136**, 411–412 (1996).
2. B. Kirkcaldy, P. Noack, A. Furnham, G. Siefen, *Eur. Psychol.* **12**, 173–180 (2007).
3. A. Lecklider, *Inventing the Egghead* (Univ. of Pennsylvania Press, Philadelphia, 2013).
4. J. Tiedemann, *Educ. Stud. Math.* **41**, 191–207 (2000).
5. S. J. Ceci, W. M. Williams, *Why Aren't More Women in Science?* (APA Books, Washington, DC, 2007).
6. S. J. Ceci, W. M. Williams, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 3157–3162 (2011).
7. National Science Foundation, Survey of Earned Doctorates, (2011); www.nsf.gov/statistics/srvydoctorates/.
8. S. J. Ceci, D. K. Ginther, S. Kahn, W. M. Williams, *Psychol. Sci. Public Interest* **15**, 75(2014).
9. C. S. Dweck, *Mindset: The New Psychology of Success* (Random House, New York, 2006).
10. V. Valian, *Why So Slow? The Advancement of Women* (MIT Press, Cambridge, MA, 1998).
11. I. Dar-Nimrod, S. J. Heine, *Science* **314**, 435 (2006).
12. A. Wigfield, J. S. Eccles, *Contemp. Educ. Psychol.* **25**, 68–81 (2000).
13. K. Ferriman, D. Lubinski, C. P. Benbow, *J. Pers. Soc. Psychol.* **97**, 517–532 (2009).
14. D. C. Geary, *Male, Female: The Evolution of Human Sex Differences* (APA Books, Washington, DC, ed. 2, 2010).
15. L. V. Hedges, A. Nowell, *Science* **269**, 41–45 (1995).
16. J. S. Hyde, *Am. Psychol.* **60**, 581–592 (2005).
17. A. M. Penner, *Am. J. Sociol.* **114** (suppl.), S138–S170 (2008).
18. J. Billington, S. Baron-Cohen, S. Wheelwright, *Learn. Individ. Differ.* **17**, 260–268 (2007).
19. R. Lippa, *J. Pers. Soc. Psychol.* **74**, 996–1009 (1998).
20. Educational Testing Service, *GRE: Guide to the Use of Scores* (ETS, Princeton, NJ, 2012).
21. J. Bethlehem, *Int. Stat. Rev.* **78**, 161–188 (2010).
22. R. J. A. Little, *J. Am. Stat. Assoc.* **88**, 1001–1012 (1993).
23. A. F. Hayes, *Introduction to Mediation, Moderation, and Conditional Process Analysis* (Guilford Press, New York, 2013).
24. C. M. Steele, J. Aronson, *J. Pers. Soc. Psychol.* **69**, 797–811 (1995).
25. M. Shih, T. L. Pittinsky, N. Ambady, *Psychol. Sci.* **10**, 80–83 (1999).
26. J. D. Angrist, J. S. Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton Univ. Press, Princeton, NJ, 2008).
27. S. J. Gould, *The Mismeasure of Man* (Norton, New York, 1996).

NOROVIRUS

Commensal microbes and interferon-λ determine persistence of enteric murine norovirus infection

Megan T. Baldridge,[1] Timothy J. Nice,[1] Broc T. McCune,[1] Christine C. Yokoyama,[1] Amal Kambal,[1] Michael Wheadon,[1] Michael S. Diamond,[1,2] Yulia Ivanova,[1] Maxim Artyomov,[1] Herbert W. Virgin[1]*

The capacity of human norovirus (NoV), which causes >90% of global epidemic nonbacterial gastroenteritis, to infect a subset of people persistently may contribute to its spread. How such enteric viruses establish persistent infections is not well understood. We found that antibiotics prevented persistent murine norovirus (MNoV) infection, an effect that was reversed by replenishment of the bacterial microbiota. Antibiotics did not prevent tissue infection or affect systemic viral replication but acted specifically in the intestine. The receptor for the antiviral cytokine interferon-λ, *Ifnlr1*, as well as the transcription factors *Stat1* and *Irf3*, were required for antibiotics to prevent viral persistence. Thus, the bacterial microbiome fosters enteric viral persistence in a manner counteracted by specific components of the innate immune system.

The microbiota of the intestine, including bacteria, fungi, viruses, and the meiofauna, contributes to both enteric disease and homeostatic immune function (*1–7*). Given the ability of human norovirus (NoV) to establish persistent infections in people (*8–10*), it is important to identify mechanisms of enteric NoV persistence. This task is made possible by the identification and molecular cloning of murine NoV (MNoV) strains that are capable of only acute enteric infection (strain CW3) or both acute and persistent infection (strain CR6) (*11–14*).

Because bacteria can interact with viruses in the intestine to alter intestinal physiology and cause pathology (*2, 6, 15, 16*), we tested the hypothesis that bacteria influence persistent enteric infection in vivo by treating C57BL/6J mice (control mice herein) for 2 weeks with broad-spectrum oral antibiotics (vancomycin, neomycin, ampicillin, and metronidazole; Abx or anti-

otics herein), then inoculating them orally with 10^6 plaque-forming units (pfu) of MNoV CR6. Antibiotics prevented persistent enteric infection in the majority of animals as measured by fecal viral shedding (Fig. 1, A and B) and by levels of virus in intestinal tissues 3 and 14 days after inoculation (Fig. 1, C and D) but had no direct inhibitory effect on virus replication in cultured cells (fig. S1, A and B). The source of differential MNoV levels at 4 and 24 hours (Fig. 1A) is the delayed intestinal transit time of the viral inoculum in Abx-treated mice (fig. S2). We confirmed the durability and efficacy of antibiotic effects by the lack of viral shedding and diminished tissue infection 35 days after inoculation (fig. S3, A and B) and reduced antiviral antibody response, which correlated with reduced virus levels (fig. S3C).

Viral persistence could not be consistently prevented with a single antibiotic targeting different bacterial species (fig. S4). In comparison, vancomycin plus ampicillin had an effect similar to that of the antibiotic cocktail (fig. S4). Successful inhibition of viral persistence in general correlated with a substantial reduction in detectable 16S ribosomal DNA copies (Fig. 2A). In some cases, antibiotic treatment did not decrease 16S copies, but this did not correlate with a failure to prevent MNoV persistence (fig. S5).

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA. [2]Departments of Medicine and Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA.
*Corresponding author. E-mail: virgin@wustl.edu

Fig. 1. Pretreatment with an antibiotic cocktail prevents establishment of persistent intestinal infection by murine norovirus strain CR6. Mice were treated with Abx (vancomycin, neomycin, ampicillin, and metronidazole) for 2 weeks before oral infection with 10^6 pfu of CR6. (**A** and **B**) Time course of MNoV genome copies shed into fecal pellets with time points at 4 hours and 1, 3, 7, and 14 days, with individual data points at day 14 in (B). Results were analyzed by two-way analysis of variance (ANOVA) with Sidak's multiple-comparisons test. ANOVA $P < 0.0001$; $N = 7$ to 33 mice per cohort per time point combined from three independent experiments. (**C** and **D**) MNoV genome copies detected in ileum, colon, or MLN at day 3 (C) or day 14 (D) after CR6 infection. Analyzed by Mann-Whitney test; $N = 10$ to 16 mice per cohort combined from three or four independent experiments. **$P < 0.01$, ***$P < 0.001$.
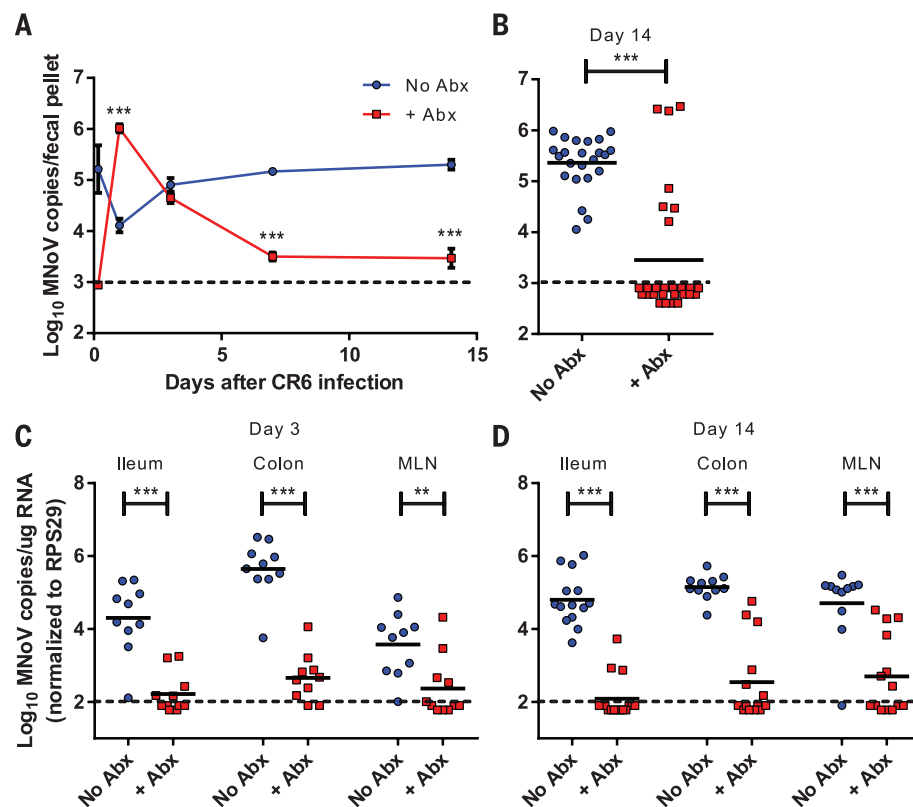
To confirm the importance of bacteria, we replaced antibiotics with drinking water for 3 days prior to CR6 infection to allow for drug clearance. Mice were still resistant to CR6 infection in this setting (Fig. 2, B and C). We then performed fecal transplantation to replace the intestinal microbiota. Fecal transplantation from control mice was sufficient to rescue the ability of CR6 to infect mice persistently, as measured by both fecal shedding and establishment of infection in intestinal tissues (Fig. 2, B and C). However, fecal transplants from antibiotic-treated mice did not rescue CR6 infection (Fig. 2, B and C). These findings strongly implicate the intestinal microbiota in the establishment of persistent enteric NoV infection.

It was possible that the effects of antibiotics were specific to the intestine, the only tissue in which persistent CR6 infection is supported in control mice (*12*). To measure the effect of antibiotics on systemic viral infection, we used three approaches. First, we inoculated CR6 by intraperitoneal injection to allow systemic spread (fig. S6, A and B), a method that did not consistently
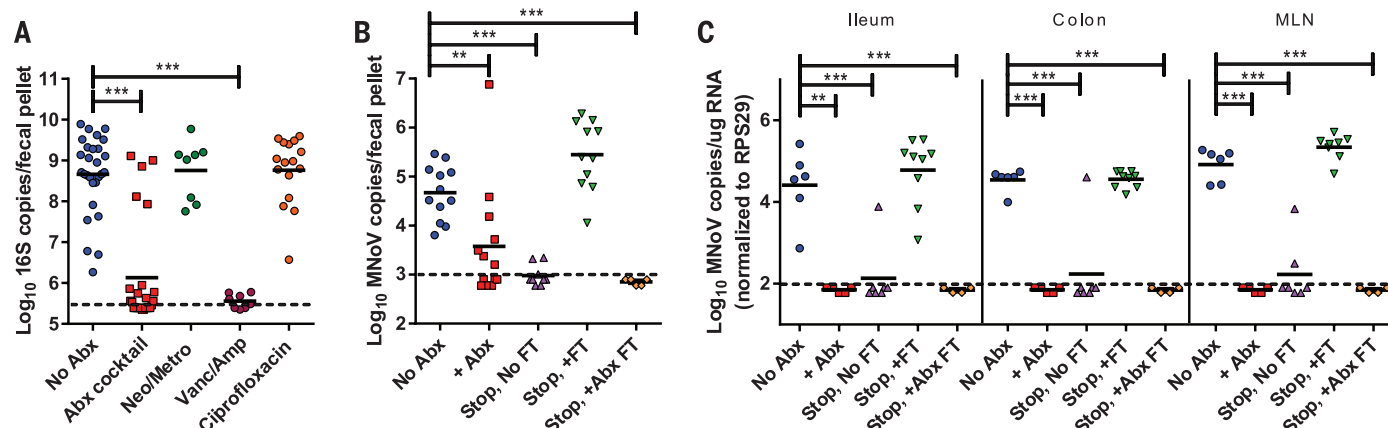
Fig. 2. Bacterial depletion prevents CR6 infection, and reconstitution of the intestinal microbiota rescues infection. (**A**) 16*S* rDNA copies per fecal pellet as detected by quantitative real-time polymerase chain reaction of the V4 hypervariable region of the 16*S* rRNA gene. Fecal pellets were collected after 2 weeks of the indicated antibiotic treatment. Neo, neomycin; Metro, metronidazole; Vanc, vancomycin; Amp, ampicillin. ANOVA *P* < 0.0001; *N* = 8 to 22 mice per cohort combined from two to five independent experiments. (**B** and **C**) Mice received no Abx, continuous Abx treatment (+Abx), or discontinuance of Abx with no fecal transplantation (Stop, No FT), transplantation from control untreated mice (Stop, +FT), or transplantation from control Abx-treated mice (Stop, +Abx FT). MNoV genome copies were detected at day 14 in fecal pellets (B) or indicated tissues (C). ANOVA *P* < 0.0001; *N* = 4 to 9 mice per cohort combined from three independent experiments. All data were analyzed by one-way ANOVA followed by Tukey's multiple-comparisons test. **P* < 0.01, ***P* < 0.001.
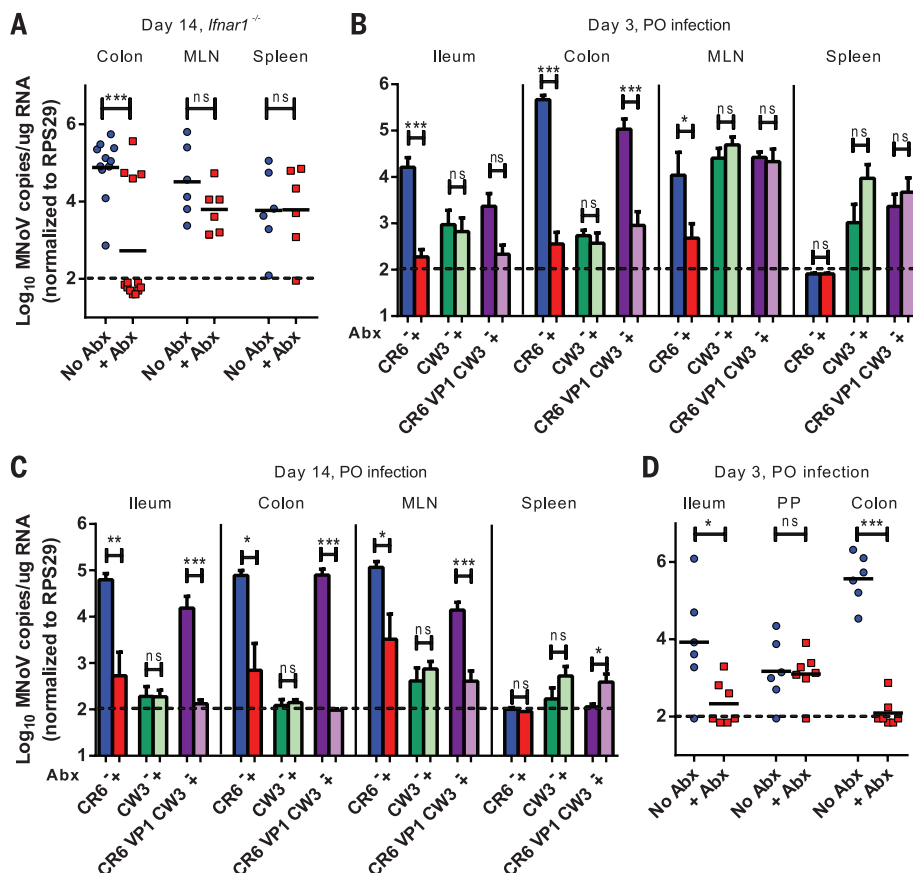
Fig. 3. The intestinal microbiota is dispensable for extraintestinal infection of MNoV and initial viral trafficking. (**A**) MNoV genome copies in indicated tissues at day 14 after oral CR6 infection of *Ifnar1*^−/− mice. Analyzed by Mann-Whitney test; *N* = 6 to 14 mice per cohort combined from three independent experiments. (**B** and **C**) MNoV genome copies detected in tissues at day 3 (B) or day 14 (C) after oral CR6, CW3, or CR6^VP1-CW3 infection. Results from each tissue type were analyzed by one-way ANOVA with Tukey's multiple-comparisons test. ANOVA *P* < 0.0001 except day 14 spleen, where ANOVA *P* < 0.01; *N* = 6 to 11 mice per cohort combined from three independent experiments. (**D**) MNoV genome copies detected in the indicated tissues (PP, Peyer's patches; Ileum, ileal tissue with no PP) at day 3 after oral CR6 infection. Analyzed by Mann-Whitney test; *N* = 6 or 7 mice per cohort combined from three independent experiments. **P* < 0.05, ***P* < 0.001; ns, not significant.

Fig. 4. The IFN-λ pathway regulates antibiotic-dependent clearance of CR6. (**A**) MNoV was detected in fecal pellets of control B6 mice at day 7 after infection of 10^5, 10^6, or 10^7 pfu of CR6. Percentages of infected mice are shown. Numbers of mice are shown in table S1. Analyzed by contingency table analysis with Fisher's exact test. (**B**) Mice deficient for the indicated genes or control mice (B6) were pretreated with Abx, then infected with 10^6 pfu of CR6. Infectivity was assessed according to detection of CR6 in fecal pellets at day 7 after infection; percentages of mice that became infected are shown. Numbers of mice are shown in table S1. Analyzed by contingency table analysis with Fisher's exact test. Each genotype was assessed in at least three independent experiments. (**C**) MNoV genome copies detected in ileum at day 14 after CR6 infection. Analyzed by one-way ANOVA with Tukey's multiple-comparisons test. ANOVA $P < 0.0001$; $N = 6$ to 15 mice per cohort combined from two to seven independent experiments. (**D**) MNoV was detected in fecal pellets at day 7 after infection with 10^5 or 10^6 pfu of CR6. Percentages of infected mice are shown. Numbers of mice are shown in table S1. Analyzed by contingency table analysis with Fisher's exact test. Each genotype was assessed in at least four independent experiments. Comparisons of +Abx, 10^5 pfu of CR6 infections to B6, +Abx, 10^5 pfu of CR6 infections shown in (A) by Fisher's exact test are all $P < 0.01$. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.
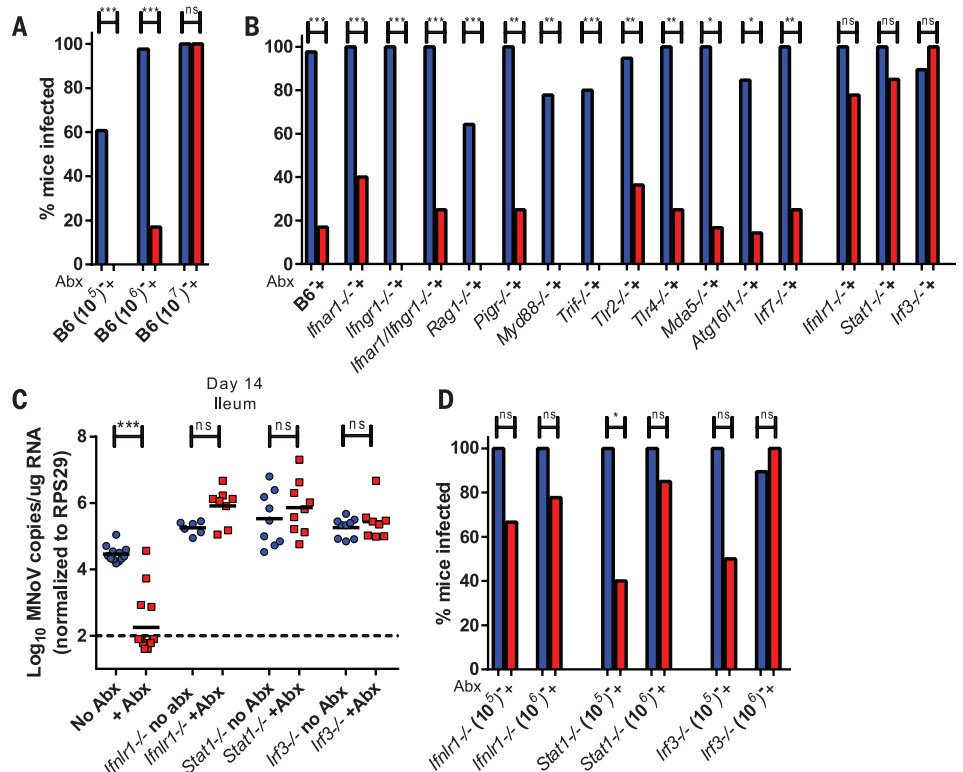
lead to persistent infection (fig. S6B). Antibiotic treatment prior to intraperitoneal infection with CR6 did not alter viral levels at day 3 (fig. S6A). Second, we infected $Ifnar1^{-/-}$ mice, which lack the gene that encodes the α chain of the interferon (IFN) α and β receptor and so are unable to respond to type I IFNs, and in which CR6 spread to and replicated persistently in the mesenteric lymph node (MLN) and spleen (17) (Fig. 3A). Levels of MNoV RNA in the MLN, the draining lymph node for the intestine, likely reflect both systemic and intestinal infection. Antibiotics prevented intestinal replication in these mice but had no effect on replication in the MLN and spleen (Fig. 3A). Third, we used a genetic approach to determine whether antibiotics acted specifically in intestinal tissues. MNoV strain CW3 spreads to systemic tissues after oral inoculation, a property conferred by the CW3 capsid protein VP1 (12, 18). We therefore tested the effects of antibiotics on infection with a chimeric virus with the CR6 backbone and VP1 of CW3 (CR6$^{VP1-CW3}$), which effectively replicates in both intestinal and systemic tissues (12, 18). Antibiotics had no effect on acute CW3 or CR6$^{VP1-CW3}$ replication in the spleen and MLN after oral inoculation. This indicated that antibiotics did not prevent viral entry into host tissues. However, antibiotics prevented replication of both CR6 and CR6$^{VP1-CW3}$ in the intestine at 3 and 14 days (Fig. 3, B and C). Therefore, the effects of antibiotic treatment are specific to the intestine.

To test further whether antibiotics inhibit viral trafficking from the lumen into tissues,

we examined viral levels in Peyer's patches (PP), which have been identified as the site of MNoV entry via M cells (19). Virus was detectable in PP at equivalent levels in antibiotic-treated and control mice 1 and 3 days after infection (Fig. 3D and fig. S7), although by day 3, antibiotic treatment substantially lowered viral levels in the ileum and colon. We conclude that virus successfully trafficked from the intestinal lumen to the mucosa via PP even in the presence of antibiotics. Antibiotic-mediated prevention of persistent infection must therefore be via an effect of the bacterial microbiome on viral clearance from intestinal tissues.

Control mice infected with either 10^5 or 10^6 pfu of MNoV demonstrated antibiotic-mediated inhibition of infection, whereas a higher dose (10^7 pfu) overcame the effects of antibiotics (Fig. 4A and fig. S8A). This suggests that host pathways that regulate viral replication might play a role in antibiotic-mediated prevention of persistent enteric norovirus infection. To define the immunologic mechanism(s) responsible for antibiotic-mediated prevention of enteric persistence, we tested the involvement of candidate host pathways by treating mice carrying null mutations in specific immune genes with antibiotics followed by oral challenge with 10^6 pfu of CR6 (table S1). To address the well-established role of IFN signaling in antiviral host defense, we examined $Ifnar1^{-/-}$, $Ifngr1^{-/-}$ (encodes IFN-γ receptor 1), and $Irf7^{-/-}$ (encodes the transcription factor IFN regulatory factor 7, which is re-

quired for the IFN-α amplification loop) mice. To determine whether adaptive immunity or secretory immunoglobulin A (IgA), both previously implicated in NoV control (20, 21), was involved, we examined $Rag1^{-/-}$ and $Pigr^{-/-}$ mice lacking all T and B cell responses or the capacity to secrete IgA, respectively. Because lipopolysaccharide (LPS) signaling can increase virus infectivity (22–24), we assessed $Tlr4^{-/-}$ mice lacking the innate immune receptor for LPS as well as mice deficient in the downstream signaling molecules Myd88 and Trif (encoded by $Ticam1$) required for Toll-like receptor signaling. We also examined mice lacking the genes that encode the bacterial sensor $Tlr2$ and the double-stranded RNA recognition molecule $Mda5$, which is involved in MNoV triggering of type I IFN responses (25). Finally, because autophagy has been reported to contribute to epithelial cell defenses against enteric bacterial infection (26, 27), we examined $Atg16l1$ HM mice, which are hypomorphic for expression of the autophagy protein Atg16L1 and which exhibit exacerbated intestinal pathology after CR6 infection (2). Remarkably, none of these genes were required for antibiotic-mediated prevention of persistent enteric MNoV infection, and indeed $Rag1^{-/-}$ mice exhibited mild resistance to infection at baseline (Fig. 4B and fig. S8B)—a finding consistent with a role for B cells in MNoV infection (28).

Because IFN-λ signaling was reported to control acute intestinal rotavirus infection (29) and has a key role in controlling persistent MNoV

infection (17), we examined the requirement for this pathway (30) in antibiotic suppression of persistent infection. $Ifnlr1^{-/-}$ mice (lacking IFN-λ receptor 1), $Stat1^{-/-}$ mice (lacking a key transcription factor that conveys IFN-αβ, IFN-γ, and IFN-λ signals), and $Irf3^{-/-}$ mice (lacking the transcription factor IFN regulatory factor 3 that induces expression of IFN-αβ and IFN-λ) (31, 32) were vulnerable to persistent CR6 infection even in the presence of antibiotics (Fig. 4, B to D; fig. S8, A and B; and fig. S9, A to G). Whereas antibiotic treatment prevented persistent infection of control mice inoculated with 10^5 pfu of MNoV (Fig. 4A and fig. S8A), some mice lacking components of the IFN-λ induction or signaling pathway became persistently infected even at this low dose (Fig. 4D and fig. S8A, $P < 0.01$). This vulnerability to MNoV infection did not correspond with differential effects of antibiotics on bacterial depletion in these mutant mice, as shown by $16S$ rDNA sequencing studies and by experiments in which fecal transplantation from antibiotic-treated mutant mice did not support establishment of persistent infection (fig. S10, A to D). Resistance to MNoV infection in antibiotic-treated control mice was not explained by up-regulation of $Ifnlr1$ or a selected set of IFN-stimulated genes including $Isg15$, which has direct antiviral activity against MNoV (33) (fig. S11, A to I).

Our data suggest that the bacterial component of the enteric microbiome plays an essential role in controlling the capacity of a virus to establish persistent infection and that antibiotics can substantially alter the pathogenesis of enteric viral infection. Prior reports have shown that bacterial LPS can alter viral infectivity (22–24), but the effects of antibiotics reported here did not require TLR4 or other signaling molecules involved in host responses to LPS. Although the absence of B cells may explain the partial resistance to MNoV infection seen in $Rag1^{-/-}$ mice (28), adaptive immunity was not required for antibiotic treatment to significantly inhibit persistent MNoV infection. Our results indicate that innate immunity and in particular the IFN-λ pathway, is required for the effects of antibiotic treatment on persistent infection. These observations suggest that the bacterial microbiota limits the efficacy of IFN-λ–dependent innate immunity or alters some yet-undefined innate immune pathway that renders viruses susceptible to the effects of IFN-λ. These data indicate that the clinical use of antibiotics in humans may alter the enteric virome (6, 34) and that the effects of antibiotics in the treatment of infectious diseases may not be entirely attributable to their antibacterial properties. Given the contribution of the virome to host physiology (4, 6), these data suggest the importance of considering the effects of trans-kingdom interactions for understanding the pathogenesis of infectious diseases.

REFERENCES AND NOTES

1. Y. Belkaid, T. W. Hand, *Cell* **157**, 121–141 (2014).
2. K. Cadwell *et al.*, *Cell* **141**, 1135–1145 (2010).
3. A. Reyes *et al.*, *Nature* **466**, 334–338 (2010).
4. H. W. Virgin, E. J. Wherry, R. Ahmed, *Cell* **138**, 30–50 (2009).
5. I. D. Iliev *et al.*, *Science* **336**, 1314–1317 (2012).
6. H. W. Virgin, *Cell* **157**, 142–150 (2014).
7. A. L. Kau, P. P. Ahern, N. W. Griffin, A. L. Goodman, J. I. Gordon, *Nature* **474**, 327–336 (2011).
8. S. M. Karst, C. E. Wobus, I. G. Goodfellow, K. Y. Green, H. W. Virgin, *Cell Host Microbe* **15**, 668–680 (2014).
9. R. L. Atmar *et al.*, *Emerg. Infect. Dis.* **14**, 1553–1557 (2008).
10. F. H. Sukhrie, J. J. Siebenga, M. F. Beersma, M. Koopmans, *J. Clin. Microbiol.* **48**, 4303–4305 (2010).
11. L. B. Thackray *et al.*, *J. Virol.* **81**, 10460–10473 (2007).
12. T. J. Nice, D. W. Strong, B. T. McCune, C. S. Pohl, H. W. Virgin, *J. Virol.* **87**, 327–334 (2013).
13. C. E. Wobus *et al.*, *PLOS Biol.* **2**, e432 (2004).
14. S. M. Karst, C. E. Wobus, M. Lay, J. Davidson, H. W. Virgin 4th, *Science* **299**, 1575–1578 (2003).
15. Y. G. Kim *et al.*, *Cell Host Microbe* **9**, 496–507 (2011).
16. M. Basic *et al.*, *Inflamm. Bowel Dis.* **20**, 431–443 (2014).
17. T. J. Nice *et al.*, *Science* **347**, 269–273 (2015).
18. D. W. Strong, L. B. Thackray, T. J. Smith, H. W. Virgin, *J. Virol.* **86**, 2950–2958 (2012).
19. M. B. Gonzalez-Hernandez *et al.*, *J. Virol.* **88**, 6934–6943 (2014).
20. K. A. Chachu *et al.*, *J. Virol.* **82**, 6610–6617 (2008).
21. L. Lindesmith *et al.*, *Nat. Med.* **9**, 548–553 (2003).
22. M. Kane *et al.*, *Science* **334**, 245–249 (2011).
23. S. K. Kuss *et al.*, *Science* **334**, 249–252 (2011).
24. C. M. Robinson, P. R. Jesudhasan, J. K. Pfeiffer, *Cell Host Microbe* **15**, 36–46 (2014).
25. S. A. McCartney *et al.*, *PLOS Pathog.* **4**, e1000108 (2008).
26. J. L. Benjamin, R. Sumpter Jr., B. Levine, L. V. Hooper, *Cell Host Microbe* **13**, 723–734 (2013).
27. K. L. Conway *et al.*, *Gastroenterology* **145**, 1347–1357 (2013).
28. M. K. Jones *et al.*, *Science* **346**, 755–759 (2014).
29. J. Pott *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7944–7949 (2011).
30. D. E. Levy, I. J. Marié, J. E. Durbin, *Curr. Opin. Virol.* **1**, 476–486 (2011).
31. P. I. Osterlund, T. E. Pietilä, V. Veckman, S. V. Kotenko, I. Julkunen, *J. Immunol.* **179**, 3434–3442 (2007).
32. K. Onoguchi *et al.*, *J. Biol. Chem.* **282**, 7576–7581 (2007).
33. M. R. Rodriguez, K. Monte, L. B. Thackray, D. J. Lenschow, *J. Virol.* **88**, 9277–9286 (2014).
34. J. M. Norman, S. A. Handley, H. W. Virgin, *Gastroenterology* **146**, 1459–1469 (2014).

NOROVIRUS

Interferon-λ cures persistent murine norovirus infection in the absence of adaptive immunity

Timothy J. Nice,[1] Megan T. Baldridge,[1] Broc T. McCune,[1] Jason M. Norman,[1] Helen M. Lazear,[2] Maxim Artyomov,[1] Michael S. Diamond,[1,2,3] Herbert W. Virgin[1]*

Norovirus gastroenteritis is a major public health burden worldwide. Although fecal shedding is important for transmission of enteric viruses, little is known about the immune factors that restrict persistent enteric infection. We report here that although the cytokines interferon-α (IFN-α) and IFN-β prevented the systemic spread of murine norovirus (MNoV), only IFN-λ controlled persistent enteric infection. Infection-dependent induction of IFN-λ was governed by the MNoV capsid protein and correlated with diminished enteric persistence. Treatment of established infection with IFN-λ cured mice in a manner requiring nonhematopoietic cell expression of the IFN-λ receptor, Ifnlr1, and independent of adaptive immunity. These results suggest the therapeutic potential of IFN-λ for curing virus infections in the gastrointestinal tract.

Human noroviruses (HNoVs) are a leading cause of gastroenteritis worldwide (1, 2). Asymptomatic fecal shedding of HNoVs may be important epidemiologically, as it provides a reservoir between outbreaks (1, 3–9). Some strains of murine norovirus (MNoV) also establish persistent enteric infection, provid-

ing a model for analyzing mechanisms of enteric NoV persistence and immunity in a natural host (1, 10, 11). Interferons (IFNs) are critical for control of both murine and human NoV replication (12–18). Interferon-α (IFN-α) and IFN-β (also called type I IFNs and hereafter IFN-αβ), IFN-γ (also called type II IFN), and IFN-λ (also called type

III IFN or interleukin-28/9) signal through the distinct heterodimeric receptors Ifnar1/Ifnar2, Ifngr1/Ifngr2, and Ifnlr1/Il10rb to regulate gene expression through phosphorylation of Stat proteins (19, 20). Although the roles of IFNs in control of persistent enteric infection have not been elucidated, it is of interest that IFN-λ, but not IFN-αβ, is important for control of acute rotavirus infection in the intestine of mice (21).

To define the role of IFNs in MNoV enteric persistence, we measured levels of the persistent MNoV strain CR6 in different tissues and in feces after oral inoculation of control mice and mice deficient in *Ifnar1*, *Ifngr1*, *Ifnlr1*, or *Stat1* (Fig. 1) (also see supplementary materials and methods). As expected, *Ifnar1* and *Stat1* were important for limiting replication in the spleen and mesenteric lymph node (MLN) (12, 13, 16, 17), whereas *Stat1* rather than *Ifnar1* controlled levels of replication in the colon (Fig. 1A), suggesting that IFN-αβ responses did not explain Stat1-dependent control of replication in the intestine. Consistent with this, comparison of the requirement for each IFN receptor in control of fecal shedding revealed that only *Stat1* and *Ifnlr1* limited levels of fecal

[1]Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO 63110, USA. [2]Department of Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA. [3]Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63110, USA.
*Corresponding author. E-mail: virgin@wustl.edu

shedding of MNoV (Fig. 1B). Furthermore, we observed increased fecal shedding compared with controls in *Ifnlr1*⁻/⁻ but not *Ifnar1*⁻/⁻ mice over 35 days of infection (Fig. 1C).

To define the basis for the intestine-specific role of IFN-λ in control of enteric persistence, we inoculated mice with the persistent MNoV strain CR6 or the nonpersistent MNoV strain CW3 and compared viral replication and induction of IFN-λ and IFN-β in Peyer's patches, MLNs, and the colon. As expected, CW3 replicated preferentially in MLN, CR6 and CW3 replicated equivalently in Peyer's patches, and CR6 replicated preferentially in colon (fig. S1) (11, 22). CW3 induced both IFN-β and IFN-λ (Fig. 2, A and B) in MLN and Peyer's patches. In contrast, CR6 did not induce detectable IFN-β or IFN-λ mRNA in any organ, despite the high level of replication in the intestine (Fig. 2, A and B). Both strains induced equivalent amounts of IFN-β from bone marrow–derived dendritic cells (BMDCs) in vitro (fig. S2). The capacity of strain CW3 to infect systemic organs maps to the protruding domain of the viral capsid protein (11, 22), whereas a single coding change (Asp⁹⁴→Glu⁹⁴, hereafter D94E) in the NS1-2 protein confers the capacity for enteric persistence upon CW3 (11, 23). In chimeric viruses, the presence of the entire CW3 capsid gene or the protruding domain of the CW3 capsid gene correlated with IFN-β and IFN-λ induction (fig. S3, A to F). Furthermore, in CW3-derived viruses carrying the NS1-2 D94E mu-

tation that confers persistence (CW3^D94E), the presence of the CR6 capsid lessened IFN-β and IFN-λ induction in MLNs despite similar levels of viral replication (Fig. 2C). This phenotype allowed us to use a chimeric virus to test the hypothesis that IFN-λ responses are required for prevention of persistence. The CW3^D94E strain is capable of efficiently establishing enteric persistence only at low doses (fig. S4). When control mice were inoculated with a high dose of CW3^D94E, many mice failed to establish persistence (fig. S4 and Fig. 2D). This failure of CW3^D94E to persist was rescued either by the CR6 capsid protein, which is associated with diminished IFN-β and IFN-λ responses (fig. S4), or by infection of *Ifnlr1*⁻/⁻ mice (Fig. 2D). Persistence of parental CW3, lacking intestinal tropism conferred by the D94E mutation, was not rescued in *Ifnlr1*⁻/⁻ mice (fig. S5). These data indicate that induction of IFN-λ interferes with the establishment of enteric MNoV persistence.

These findings suggest a primary role for IFN-λ in control of enteric MNoV persistence. Consistent with this, intraperitoneal treatment of mice with IFN-λ 1 day after oral inoculation with CR6 prevented persistent enteric MNoV infection (Fig. 3A). When cured mice were rechallenged with CR6 2 weeks later, persistent infection was established (Fig. 3B), indicating that IFN-λ acted by stimulating innate immunity rather than promoting adaptive immunity. These data also indicate that therapeutic levels of IFN-λ wane in animals within 2 weeks of administration.

Fig. 1. Systemic and intestinal persistence of MNoV are controlled by IFN-αβ and IFN-λ, respectively. Mice were orally inoculated with 10⁶ plaque-forming units (PFUs) of MNoV strain CR6, and genome copies in the indicated tissue (**A**) or feces (**B** and **C**) were determined by quantitative reverse transcription polymerase chain reaction (qRT-PCR). Genome copies were compared between control, *Stat1*⁻/⁻, and *Ifnar1*⁻/⁻ mice at day 21 in tissues (A); between control, *Stat1*⁻/⁻, *Ifngr1*⁻/⁻, *Ifnar1*⁻/⁻, *Ifnlr1*⁻/⁻, and *Ifnar1*⁻/⁻ × *Ifngr1*⁻/⁻ mice at day 14 in feces (B); and between control, *Ifnar1*⁻/⁻, and *Ifnlr1*⁻/⁻ over time in feces (C). Data shown are pooled from at least two independent experiments, with each point representing an individual animal in (A) and (B). Points in (C) represent at least four animals pooled from two to four independent experiments. Dashed lines represent limit of detection. Statistical significance was determined by one-way (A and B) or two-way (C) analysis of variance (ANOVA). n.s., not significant (P > 0.05); *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. Error bars in (C) denote SD.
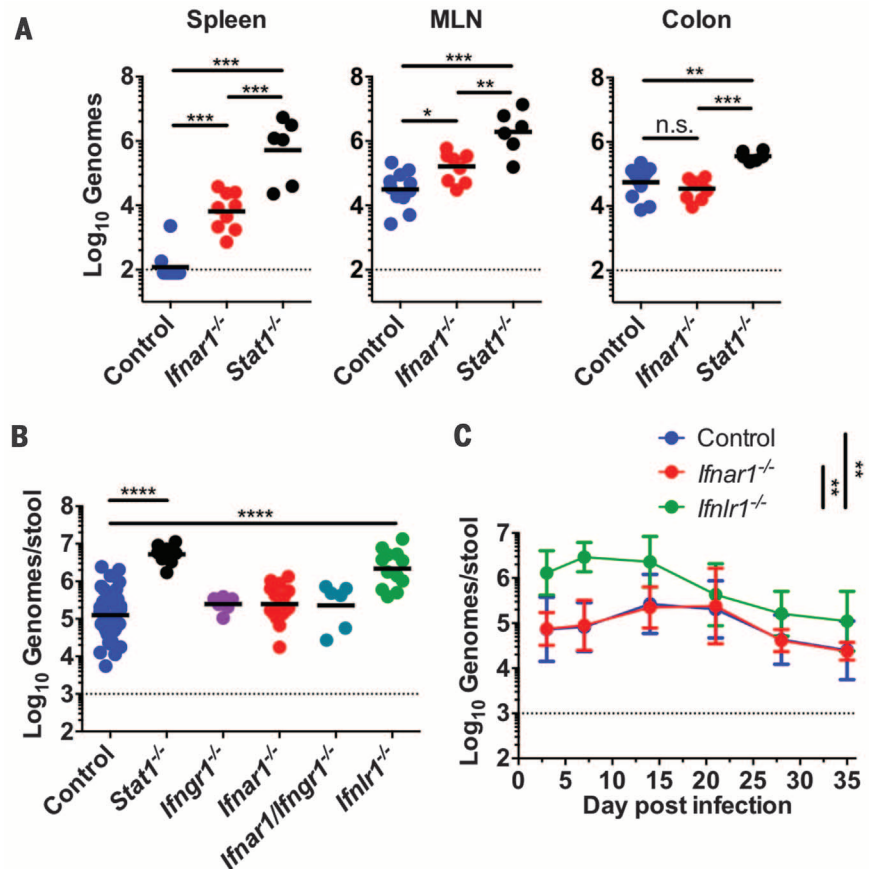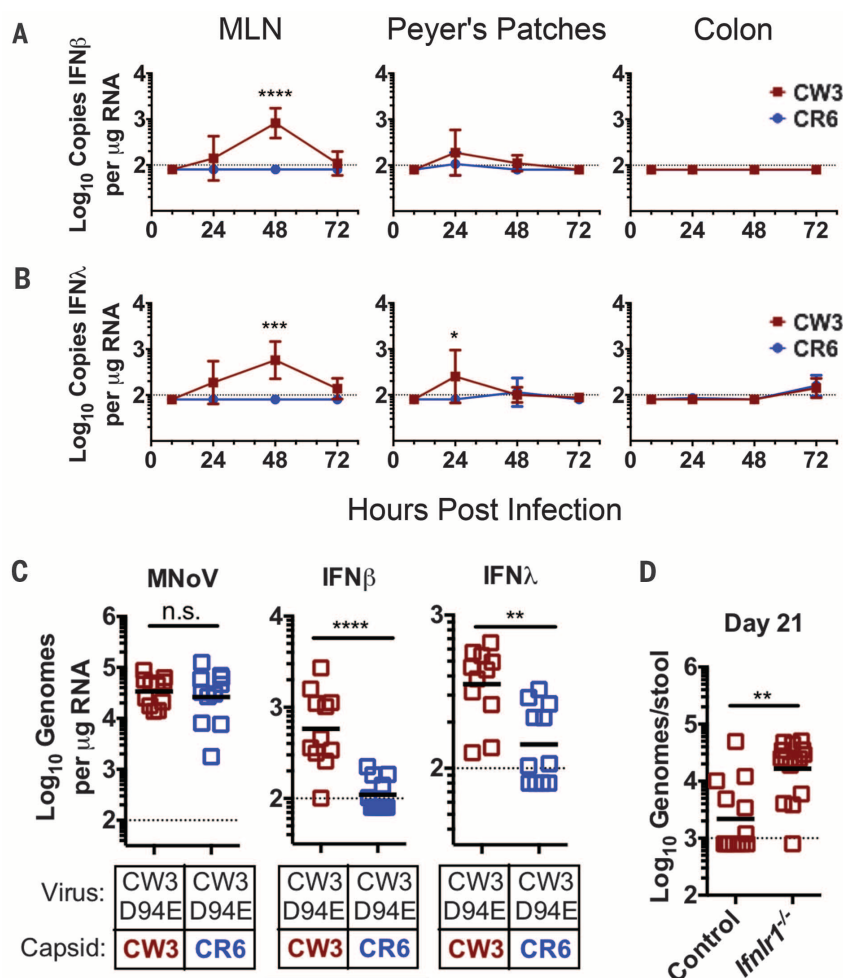
Fig. 2. Induction of IFN-λ prevents enteric NoV persistence. (**A** and **B**) Control mice were orally inoculated with 10^6 PFUs of either MNoV CW3 or CR6, and total RNA was isolated from MLN, Peyer's patches, or colon at the indicated times. Relative copy numbers for IFN-β transcripts (A) and IFN-λ transcripts (B) were quantified by qRT-PCR. (**C**) MNoV genomes, IFN-β transcripts, and IFN-λ transcripts in MLNs 48 hours after inoculation were compared between mice infected with CW3^{D94E} containing either the CW3 or the CR6 capsid gene. (**D**) MNoV genomes in feces from control or $Ifnlr1^{-/-}$ mice were measured at day 21 after inoculation with 10^6 PFUs of CW3^{D94E}. Data in (A) to (C) are pooled from three independent experiments for a total of three to four mice per time point. Data points in (D) are individual mice pooled from three independent experiments. Dashed lines represent limit of detection. Statistical significance was determined by two-way ANOVA (A and B) or Mann-Whitney test (C and D). n.s. indicates $P > 0.05$; *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$, ****$P \leq 0.0001$. Error bars in (A) and (B) denote SD.
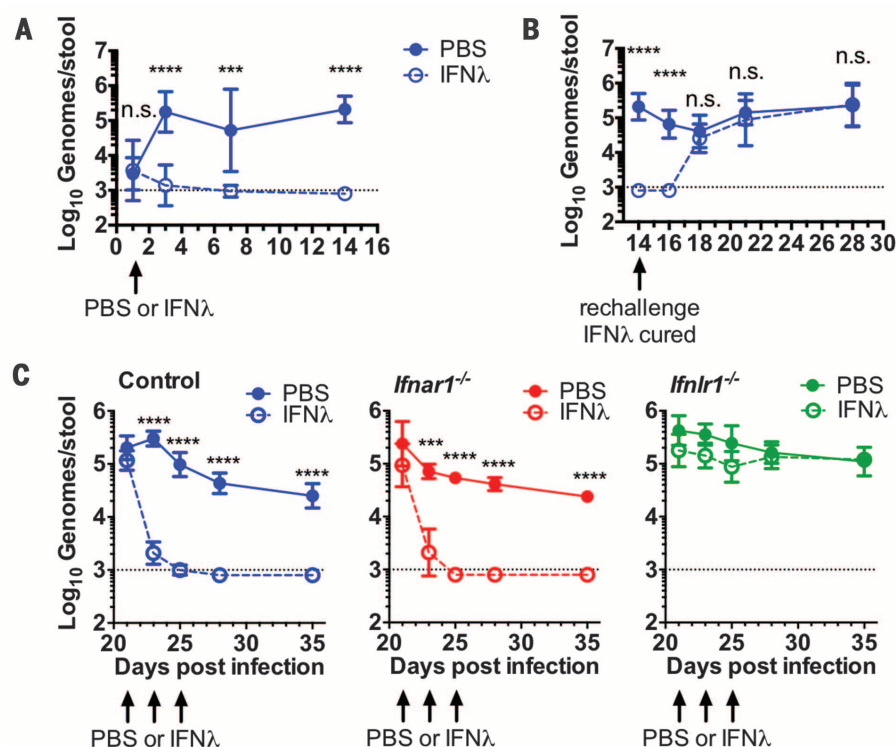
Fig. 3. IFN-λ treatment prevents and cures persistent enteric MNoV infection. (**A** to **C**) Feces were collected at the indicated day after oral inoculation, and MNoV genomes were quantified by qRT-PCR. (A) Mice were injected with 25 μg of IFN-λ or phosphate-buffered saline (PBS) intraperitoneally 1 day after oral inoculation with 10^6 PFUs of CR6. (B) The IFN-λ–treated mice from (A) were rechallenged with 10^6 PFUs of CR6 at day 14 after initial infection. (C) Persistent infection with CR6 was established in control, $Ifnar1^{-/-}$, or $Ifnlr1^{-/-}$ mice, followed by intraperitoneal injection of 25 μg of IFN-λ on days 21, 23, and 25. Data shown are pooled from two (A) or three (B and C) independent experiments for a total of four to eight mice per time point. Dashed lines represent limit of detection. Statistical significance was determined by two-way ANOVA. n.s. indicates $P > 0.05$; ***$P \leq 0.001$, ****$P \leq 0.0001$. Error bars in (A) and (B) denote SD and in (C) denote SEM.
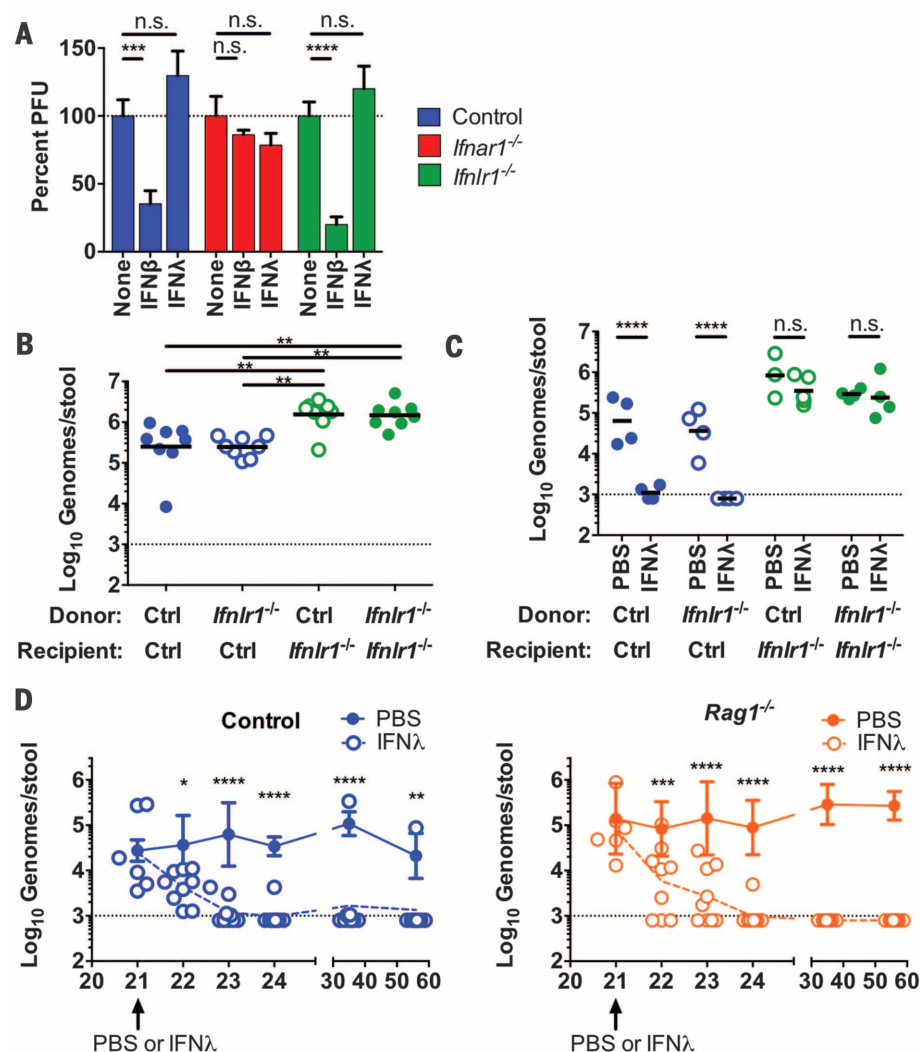
Fig. 4. IFN-λ durably clears enteric MNoV persistence through effects on radiation-insensitive cells in the absence of adaptive immunity. (**A**) BMDCs were treated with media, 100 IU/ml IFN-β, or 100 ng/ml IFN-λ for 24 hours, then inoculated with CR6 at a multiplicity of infection of five, and viral titers were determined by plaque assay 12 hours later. Data are pooled from three independent experiments performed in triplicate and normalized to untreated. (**B** and **C**) Bone marrow chimeras were generated using control and $Ifnlr1^{-/-}$ donor and recipient mice as indicated and 8 to 10 weeks later were orally inoculated with 10^6 PFUs of CR6. (B) MNoV shedding in feces was quantified on day 14. (C) A single 25-μg dose of IFN-λ or PBS was administered intraperitoneally 21 days after inoculation, and MNoV shedding was quantified on day 23. Data in (B) and (C) are pooled from two independent experiments and shown as individual mice. (**D**) Control or $Rag1^{-/-}$ mice orally inoculated with CR6 were treated 21 days later with a single 25-μg dose of IFN-λ. Shedding was monitored for 35 days postinjection. Dashed lines in (B) to (D) represent limit of detection. Statistical significance was determined by one-way (B) or two-way (A, C, D) ANOVA. n.s. indicates $P > 0.05$; $*P \le 0.05$, $**P \le 0.01$, $***P \le 0.001$, $****P \le 0.0001$. Error bars in (A) denote SEM and in (D) denote SD.

To determine if therapeutic IFN-λ was effective against established persistent infection, we treated mice with IFN-λ at 21, 23, and 25 days after oral inoculation with CR6, a time at which stable enteric persistence is established (*11*). Shedding of virus into the feces was reduced ~100-fold within 2 days and was undetectable within 1 week (Fig. 3C). Analysis of viral genomes in tissues 2 weeks after IFN-λ treatment revealed reduced or undetectable levels of virus in the MLN and undetectable virus levels in the colon (fig. S6). As little as a single 1-μg dose of IFN-λ was sufficient

to clear persistent MNoV (fig. S7A). Ifnlr1 was required for IFN-λ activity, whereas Ifnar1 was dispensable (Fig. 3C and fig. S6), indicating that IFN-λ does not require IFN-αβ signaling to eliminate enteric persistence.

Whereas MNoV is known to replicate in immune cells, attempts to cultivate MNoV or HNoV in epithelial cells have not been successful (*1, 24*). Treatment of cultured BMDCs with IFN-β, but not IFN-λ, inhibited MNoV replication (Fig. 4A). Taken together with previously published data (*21, 25–27*), this result raised the possibility that

IFN-λ acts on nonhematopoietic cells to control enteric persistence. Experiments in reciprocal bone marrow chimeras revealed that the control of fecal MNoV levels (Fig. 4B) and the capacity for IFN-λ to cure enteric persistence (Fig. 4C) mapped to the nonhematopoietic rather than radiation-sensitive hematopoietic cells. These data were consistent with an Ifnar1-independent effect of IFN-λ on enteric persistence through stimulation of immunity via signaling in radiation-insensitive cells.

Despite the capacity of mice to mount protective adaptive immune responses to MNoV infection (*28, 29*), their susceptibility to reinfection after treatment with IFN-λ (Fig. 3B) suggested that IFN-λ might act preferentially through the innate immune system. However, it is generally thought that in mammals, viral clearance requires adaptive B and T cell immunity. To test whether IFN-λ can clear an established infection in the absence of an adaptive immune response, we inoculated $Rag1^{-/-}$ mice with CR6 and 21 days later treated them with a single dose of IFN-λ. Enteric persistence of CR6 was cured by IFN-λ treatment of both control and $Rag1^{-/-}$ mice (Fig. 4D). All $Rag1^{-/-}$ mice and the majority of control mice remained MNoV-free for 35 days after treatment, well after therapeutic levels of IFN-λ waned (Fig. 3B). The absence of infectious MNoV was confirmed by fecal transplantation from IFN-λ–cleared $Rag1^{-/-}$ mice to naïve $Rag1^{-/-}$ or $Stat1^{-/-}$ mice (fig. S7, B and C). Together, these data demonstrate that IFN-λ treatment both prevents and cures established enteric persistence of MNoV in the absence of an adaptive immune response.

Our study establishes the importance of IFN-λ in innate immunity to persistent enteric viral infection. Establishment of persistent enteric infection by certain strains of MNoV was related to failure to induce IFN-λ responses. Whereas MNoV replicates in hematopoietic cells, IFN-λ was found to act on nonhematopoietic cells, suggesting that the mechanism of action is indirect. The efficacy of IFN-λ in the prevention and cure of enteric persistence did not require adaptive immunity. These findings provide a different view of the immune system because it is generally believed that the development of an adaptive immune response is required for clearance of viral infection by antigen-specific targeting. One implication of our findings is that immune therapies may be able to control persistent virus infection regardless of their effects on adaptive immune responses. We propose that this is an example of sterilizing innate immunity wherein viral clearance can be determined by immune responses apart from adaptive immunity. A similar observation of viral clearance by an independent mechanism of a second enteric virus, rotavirus, supports the existence of sterilizing innate immunity (*30*). Analogously, metazoan organisms lacking adaptive immune systems are probably capable of clearing some pathogens and preventing persistent infection. We speculate that evolutionarily conserved innate immune mechanisms with sterilizing potential exist, perhaps with particular importance in protection against persistent infection at mucosal surfaces such as the intestine.

REFERENCES AND NOTES

1. S. M. Karst, C. E. Wobus, I. G. Goodfellow, K. Y. Green, H. W. Virgin, *Cell Host Microbe* **15**, 668–680 (2014).
2. R. I. Glass, U. D. Parashar, M. K. Estes, *N. Engl. J. Med.* **361**, 1776–1785 (2009).
3. G. Phillips, C. C. Tam, L. C. Rodrigues, B. Lopman, *Epidemiol. Infect.* **138**, 1454–1458 (2010).
4. M. Saito *et al.*, *Clin. Infect. Dis.* **58**, 483–491 (2014).
5. J. Ayukekbong *et al.*, *J. Med. Virol.* **83**, 2135–2142 (2011).
6. D. S. Cheon *et al.*, *Foodborne Pathog. Dis.* **7**, 1427–1430 (2010).
7. F. H. Sukhrie, J. J. Siebenga, M. F. Beersma, M. Koopmans, *J. Clin. Microbiol.* **48**, 4303–4305 (2010).
8. F. H. Sukhrie *et al.*, *Clin. Infect. Dis.* **54**, 931–937 (2012).
9. B. Lopman, K. Simmons, M. Gambhir, J. Vinjé, U. Parashar, *Am. J. Epidemiol.* **179**, 507–512 (2014).
10. V. T. Tomov *et al.*, *J. Virol.* **87**, 7015–7031 (2013).
11. T. J. Nice, D. W. Strong, B. T. McCune, C. S. Pohl, H. W. Virgin, *J. Virol.* **87**, 327–334 (2013).
12. S. M. Karst, C. E. Wobus, M. Lay, J. Davidson, H. W. Virgin IV, *Science* **299**, 1575–1578 (2003).
13. C. E. Wobus *et al.*, *PLOS Biol.* **2**, e432 (2004).
14. K. O. Chang, D. W. George, *J. Virol.* **81**, 12111–12118 (2007).
15. K. O. Chang, S. V. Sosnovtsev, G. Belliot, A. D. King, K. Y. Green, *Virology* **353**, 463–473 (2006).
16. L. B. Thackray *et al.*, *J. Virol.* **86**, 13515–13523 (2012).
17. S. A. McCartney *et al.*, *PLOS Pathog.* **4**, e1000108 (2008).
18. S. Hwang *et al.*, *Cell Host Microbe* **11**, 397–409 (2012).
19. T. Taniguchi, A. Takaoka, *Curr. Opin. Immunol.* **14**, 111–116 (2002).
20. A. J. Sadler, B. R. Williams, *Nat. Rev. Immunol.* **8**, 559–568 (2008).
21. J. Pott *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 7944–7949 (2011).
22. D. W. Strong, L. B. Thackray, T. J. Smith, H. W. Virgin, *J. Virol.* **86**, 2950–2958 (2012).
23. B. N. Borin *et al.*, *Proteins* **82**, 1200–1209 (2014).
24. M. K. Jones *et al.*, *Science* **346**, 755–759 (2014).
25. N. Ank *et al.*, *J. Immunol.* **180**, 2474–2485 (2008).
26. M. Mordstein *et al.*, *J. Virol.* **84**, 5670–5677 (2010).
27. C. Sommereyns, S. Paul, P. Staeheli, T. Michiels, *PLOS Pathog.* **4**, e1000017 (2008).
28. G. Liu, S. M. Kahan, Y. Jia, S. M. Karst, *J. Virol.* **83**, 6963–6968 (2009).
29. K. A. Chachu, A. D. LoBue, D. W. Strong, R. S. Baric, H. W. Virgin, *PLOS Pathog.* **4**, e1000236 (2008).
30. B. Zhang *et al.*, *Science* **346**, 861–865 (2014).

TELOMERES IN CANCER

Alternative lengthening of telomeres renders cancer cells hypersensitive to ATR inhibitors

Rachel Litman Flynn,[1,2]* **Kelli E. Cox,**[2]† **Maya Jeitany,**[3]† **Hiroaki Wakimoto,**[4] **Alysia R. Bryll,**[2] **Neil J. Ganem,**[2] **Francesca Bersani,**[1,5] **Jose R. Pineda,**[3] **Mario L. Suvà,**[1,6] **Cyril H. Benes,**[1] **Daniel A. Haber,**[1,5] **Francois D. Boussin,**[3] **Lee Zou**[1,6]*

Cancer cells rely on telomerase or the alternative lengthening of telomeres (ALT) pathway to overcome replicative mortality. ALT is mediated by recombination and is prevalent in a subset of human cancers, yet whether it can be exploited therapeutically remains unknown. Loss of the chromatin-remodeling protein ATRX associates with ALT in cancers. Here, we show that ATRX loss compromises cell-cycle regulation of the telomeric noncoding RNA TERRA and leads to persistent association of replication protein A (RPA) with telomeres after DNA replication, creating a recombinogenic nucleoprotein structure. Inhibition of the protein kinase ATR, a critical regulator of recombination recruited by RPA, disrupts ALT and triggers chromosome fragmentation and apoptosis in ALT cells. The cell death induced by ATR inhibitors is highly selective for cancer cells that rely on ALT, suggesting that such inhibitors may be useful for treatment of ALT-positive cancers.

Cancer cells overcome replicative senescence by activating telomerase or the alternative lengthening of telomeres (ALT) pathway (*1–3*). ALT is used in ~5% of all human cancers and is prevalent in specific cancer types, including osteosarcoma and glioblastoma (*4*). Currently, there are no therapies specifically targeting ALT. ALT relies on recombination to elongate telomeres (*3*), but how the recombinogenic state of ALT telomeres is established remains elusive. In contrast to cancer cells defective for homologous recombination (HR) and susceptible to poly(ADP-ribose) polymerase (PARP) inhibition (*5, 6*), ALT-positive cells are HR-proficient (*7*). Thus, the reliance of ALT on recombination raises the question as to whether recombination can be exploited in ALT-positive cancers as a means for targeted therapy.

Single-stranded DNA (ssDNA) coated by replication protein A (RPA) is a key intermediate in both DNA replication and HR (*8*). RPA transiently associates with telomeres during DNA replication, but is released from telomeres after S phase (*9, 10*). The release of RPA may be an important mechanism to suppress HR at telomeres. The association of RPA with telomeres in S phase is facilitated by TERRA, the telomere repeat-containing RNA, which is also present at telomeres during this period (*9, 11–13*). To investigate how ALT is established, we determined whether the association of TERRA with telomeres is altered in ALT cells. TERRA colocalized with the telomere-binding protein TRF2 in telomerase-positive HeLa cervical cancer cells (fig. S1) (*9*). However, in both HeLa and telomerase-positive SJSA1 osteosarcoma cells (fig. S24B), the number of TERRA foci declined from S phase to G_2 (Fig. 1, A and B, and fig. S2) (*9, 12*). Although in ALT-positive U2OS osteosarcoma cells TERRA also colocalized with the telomere marker TRF2 (fig. S3, A and B), neither the levels of TERRA, nor the colocalization of TERRA and TRF2, declined from S to G_2 (figs. S2; S3, B and C; and S4, A and B). Furthermore, in ALT-positive U2OS and HUO9 osteosarcoma cells (Fig. 3D and fig. S25, A and B), the number of TERRA foci increased significantly in S phase and remained high into G_2 (Fig. 1, A and B, and fig. S2). Thus, in contrast to telomerase-positive cells, ALT cells are defective in the cell-cycle regulation of TERRA.

We next explored why TERRA persistently associates with telomeres in ALT cells. Recent studies have revealed a correlation of ALT with mutations in the *ATRX* gene and loss of the chromatin-remodeling protein ATRX in cancer (*14–17*). ATRX was detected in HeLa but not U2OS cells (figs. S5A and S25C) (*14*), prompting us to investigate whether the dysregulation of TERRA in ALT cells is a result of ATRX loss. Indeed, knockdown of ATRX in HeLa cells resulted in persistent TERRA foci and elevated TERRA levels in G_2/M (Fig. 1, C and D, and figs. S5 and S6). Furthermore, the levels of TERRA derived from individual telomeres (15q and Xp/Yp) declined from S phase to mitosis in control HeLa

[1]Massachusetts General Hospital Cancer Center, Harvard Medical School, Charlestown, MA 02129, USA. [2]Departments of Pharmacology & Experimental Therapeutics, and Medicine, Cancer Center, Boston University School of Medicine, Boston, MA 02118, USA. [3]Laboratoire de Radiopathologie, UMR 967, Institut de Radiobiologie Cellulaire et Moleculaire, CEA Fontenay-aux-Roses, France. [4]Deparment of Surgery and Brain Tumor Center, Massachusetts General Hospital, Boston, MA 02115, USA. [5]Howard Hughes Medical Institute, Massachusetts General Hospital, Charlestown, MA 02129, USA. [6]Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02115, USA.
*Corresponding author. E-mail: zou.lee@mgh.harvard.edu (L.Z.); rlflynn@bu.edu (R.L.F.) †These authors contributed equally to this work.
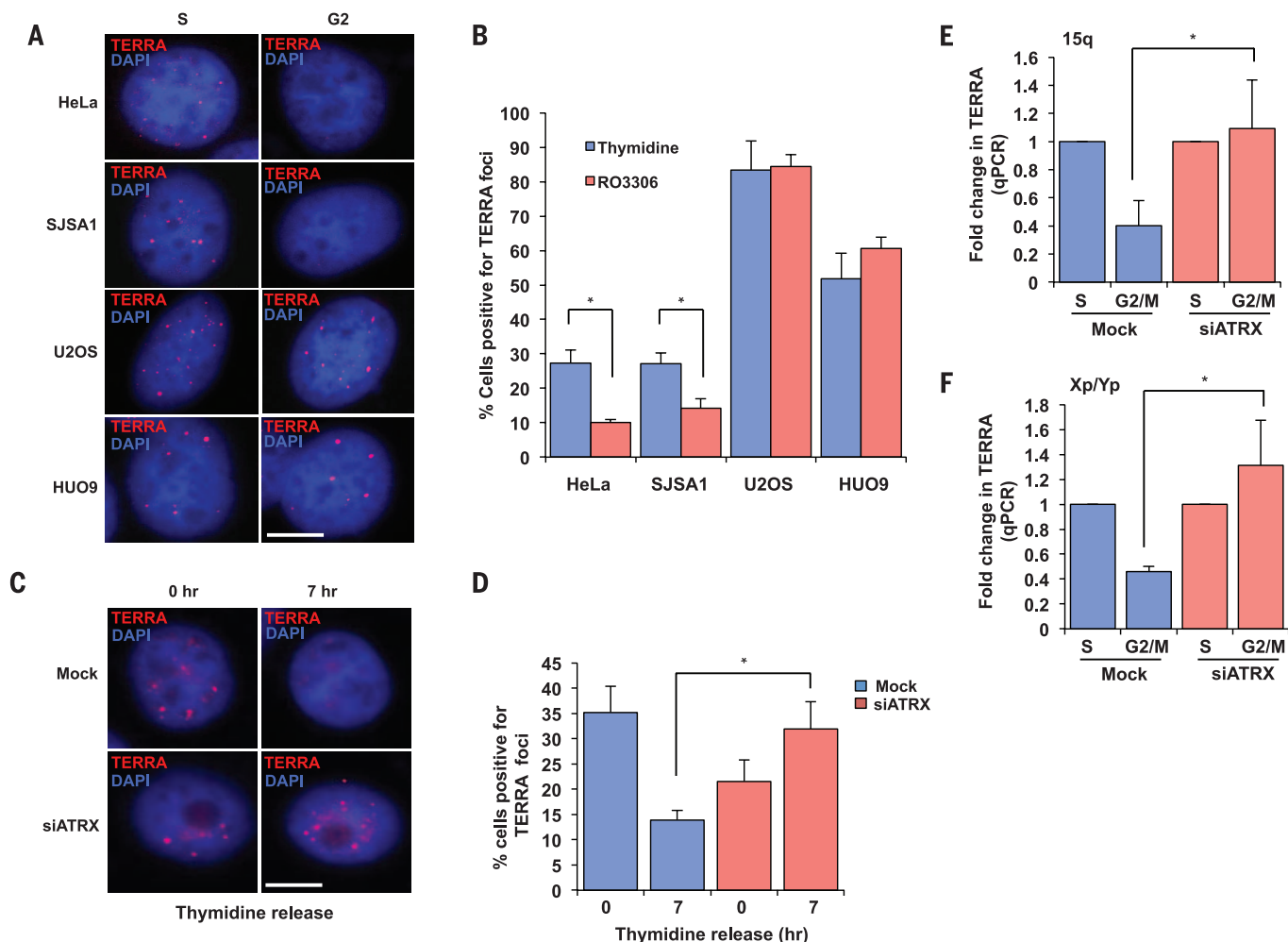
Fig. 1. Loss of ATRX compromises the cell-cycle regulation of TERRA. (**A**) RNA fluorescence in situ hybridization (FISH) analyses of TERRA in HeLa, SJSA1, U2OS, and HUO9 cells during the cell cycle. TERRA foci colocalized with TRF2 at telomeres (figs. S1 and S3, A and B). To enrich cells in S phase, cells were treated with thymidine alone. To enrich cells in G_2, cells were first arrested in S phase with thymidine and then released into medium containing the CDK1 inhibitor RO3306 (fig. S2). Scale bar: 10 μm. (**B**) The percentage of cells positive for TERRA foci (>5 foci) was graphed as the mean ± SD (n = 2). (**C**) HeLa cells were mock treated or treated with ATRX siRNA, and RNA FISH analysis

of TERRA was performed after thymidine release. The knockdown of ATRX was confirmed by Western blot (fig. S5A). Cells were enriched in late S and G_2 phases 7 hours after thymidine release (fig. S5B). Scale bar: 10 μm. (**D**) The percentage of cells positive for TERRA foci was graphed as the mean ± SD (n = 3). (**E** and **F**) HeLa cells were mock treated or treated with ATRX siRNA, and were enriched in S or M phase with thymidine and nocodazole, respectively (fig. S5B). TERRA was analyzed by reverse transcription quantitative polymerase chain reaction (RT-qPCR) using the subtelomeric primers of chromosome 15q or Xp/Yp. The results are graphed as the mean fold change ± SD (15q n = 3, Xp/Yp n = 4). *P < 0.05.

cells but not in ATRX knockdown cells (Fig. 1, E and F). These results suggest that TERRA is repressed by ATRX in G_2/M.

Considering that RPA is released from telomeres in G_2/M when TERRA is repressed by ATRX (9), we examined whether ATRX is required for the release of RPA. In HeLa cells, numerous small replication-associated RPA foci (type-A RPA foci) were detected in S phase (fig. S7). As cells progressed from S to G_2, type-A RPA foci became largely undetectable (Fig. 2A). However, upon ATRX knockdown, bright damage-associated RPA foci (type-B RPA foci) were detected at telomeres in a fraction of G_2 cells (Fig. 2A and figs. S7 and S8). Knockdown of ATRX with two independent small interfering RNAs (siRNAs) led to a significant increase of type-B RPA foci in G_2 cells (Fig. 2B). To examine the release of RPA from telomeric ssDNA biochemically, we fol-

lowed this process in cell extracts using an in vitro assay that we previously established (9). A biotinylated ssDNA oligomer of telomeric repeats (ssTEL) was coated with recombinant RPA and incubated in extracts from S-phase or mitotic HeLa cells. Consistent with the release of RPA from telomeres in G_2/M, RPA was released from ssTEL more efficiently in mitotic extracts than in S-phase extracts (Fig. 2C) (9). Knockdown of ATRX reduced the release of RPA from ssTEL in mitotic extracts (Fig. 2C), demonstrating that ATRX contributes to the RPA release in G_2/M. To determine whether the loss of ATRX in ALT cells affects RPA release, we analyzed IMR90 myofibroblast-derived SW39^{TEL} (telomerase-positive) and SW26^{ALT} (ALT-positive) cells (fig. S9) (7). ATRX was detected in SW39^{TEL} but not SW26^{ALT} (Fig. 2D). Moreover, the loss of ATRX in SW26^{ALT} was associated with a four-fold in-

crease in TERRA compared with SW39^{TEL} (Fig. 2E-F). Notably, RPA was released from ssTEL more efficiently in SW39^{TEL} cell extracts than in SW26^{ALT} cell extracts (Fig. 2G), showing that ALT cells lacking ATRX indeed have a reduced ability to release RPA from telomeric ssDNA.

Given that RPA-ssDNA is a key HR intermediate, we asked if ATRX loss induces HR. Knockdown of ATRX in HeLa cells did not inactivate telomerase, nor did it induce telomere lengthening (fig. S10, A and B). These results are consistent with a previous study (14) and suggest that loss of ATRX is insufficient to establish ALT. Nevertheless, ATRX knockdown in HeLa cells promoted some features of ALT, such as the persistent association of TERRA and RPA with telomeres. A recent study showed that loss of the histone chaperone ASF1 led to the acquisition of several ALT phenotypes, including accumulation
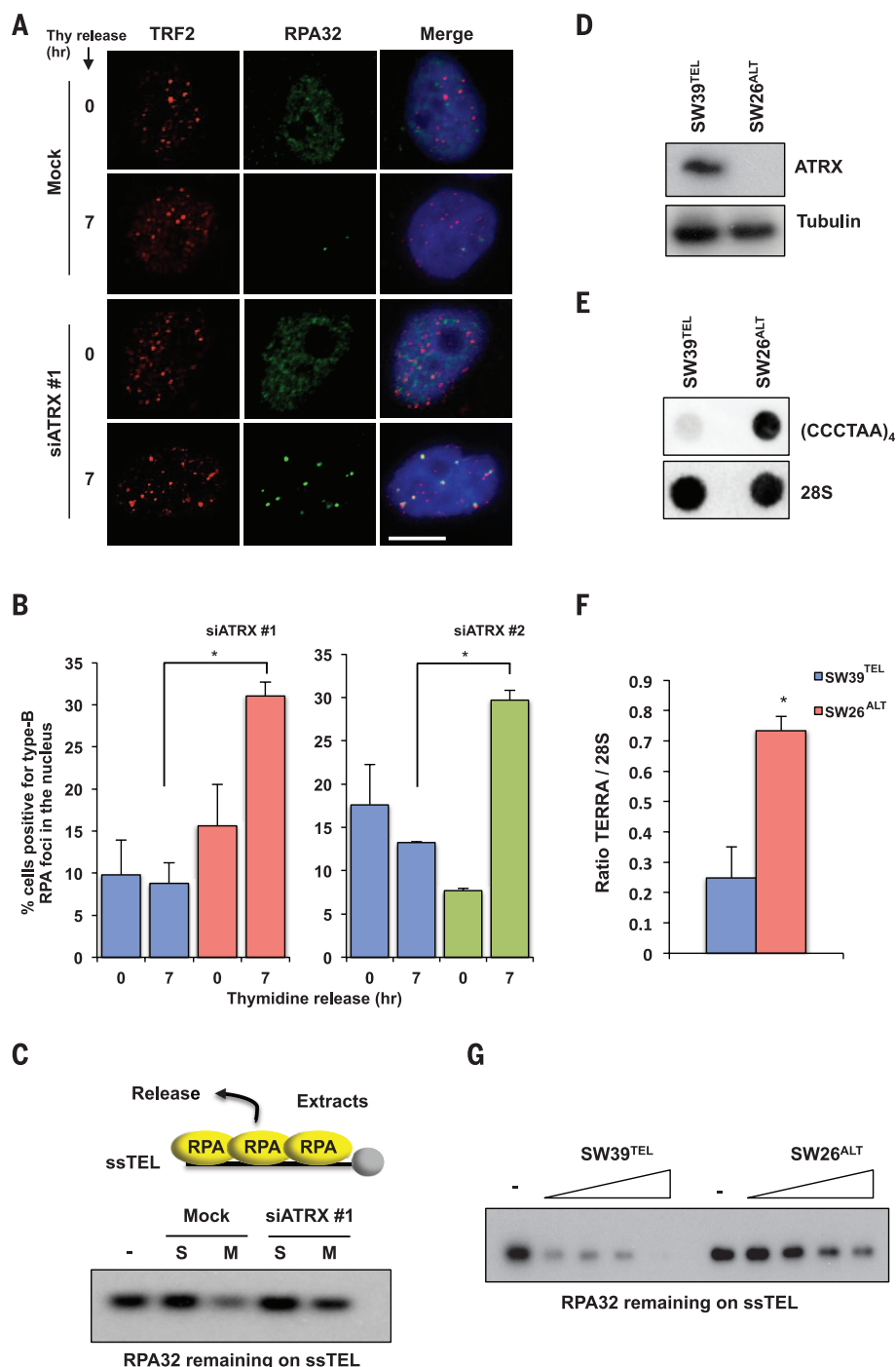
Fig. 2. Loss of ATRX compromises RPA release from telomeres. (**A**) HeLa cells were mock treated or treated with ATRX siRNA, and RPA and TRF2 foci were analyzed in S and G_2 phases as in Fig. 1C. Scale bar: 10 μm. (**B**) The percentage of cells positive for RPA foci was graphed as the mean ± SD ($n = 2$) *P = 0.008 (left) and *P = 0.002 (right). (**C**) HeLa cells were either mock treated or treated with ATRX siRNA, and whole-cell extracts (WCE) were generated from cells in S or M phase. Biotinylated ssTEL was coated with RPA and incubated with the WCE. After the incubation, ssTEL was retrieved, and the remaining RPA32 on ssTEL was analyzed by Western blot. (**D**) SW39[TEL] and SW26[ALT] cells were analyzed for ATRX protein expression by Western blot. (**E**) SW39[TEL] and SW26[ALT] cells were analyzed for TERRA transcript by dot blot with digoxigenin-labeled anti-TERRA or 28S RNA probes. (**F**) Quantification of dot blots for TERRA transcript in SW39[TEL] and SW26[ALT] cells. TERRA signal was normalized to 28S signal, and ratios were graphed as the mean ± SD ($n = 2$). *P = 0.001. (**G**) RPA-ssTEL was incubated in WCE from SW39[TEL] or SW26[ALT] cells. The RPA32 remaining on ssTEL was analyzed by Western blot.

of RPA at telomeres (*18*). We postulate that ALT is established via a multistep process in which loss of ATRX poises telomeres for ALT, but additional genetic or epigenetic changes are needed to fully activate the ALT pathway (fig. S29).

RPA-ssDNA is not only an HR intermediate, but also the nucleoprotein structure that recruits ATR, a protein kinase that is a key regulator of HR (*19, 20*). The defective RPA release from telomeres in ATRX knockdown cells and ALT cells suggests that ATR may be recruited to telomeres during the establishment of ALT. Consistent with our hypothesis, ATR colocalizes with promyelocytic leukemia protein (PML) in U2OS cells but not in HeLa cells (*21*), suggesting its presence in ALT-associated PML bodies (APBs) (*22*). Furthermore, ATRIP, the regulatory partner of ATR, associates with telomeres in ALT-positive WI38-VA13 cells but not in HeLa cells (*23*). These findings prompted us to explore whether ATR is functionally required for ALT. The ATR inhibitor VE-821 (*24*) and ATR siRNA significantly reduced APBs in U2OS and SW26[ALT] cells (Fig. 3, A and B, and figs. S11, A and B, and S12A). VE-821 also disrupted APBs in U2OS cells synchronized in G2 (fig. S12B) (*25*), ruling out cell-cycle changes as the cause of APB dispersal. In marked contrast, the ATM inhibitor KU-55933 and ATM siRNA did not affect APBs in U2OS cells (Fig. 3, A and B, and fig. S12, B and C), highlighting the role for ATR, but not ATM, in the maintenance of APBs in ALT cells.

To determine whether VE-821 affects the recombinogenic state of ALT telomeres, we analyzed telomere sister-chromatid exchange (T-SCE) and extrachromosomal telomeric C-rich DNA (C-circles) in ALT cells. VE-821 not only decreased T-SCE in U2OS cells (Fig. 3C and fig. S13A), but also reduced C-circle levels in U2OS and HUO9 cells (Fig. 3, D and E), showing that ALT is indeed inhibited. Furthermore, VE-821 increased the frequency of telomere loss in U2OS cells (Fig. 3F, S13B), suggesting that the stability of ALT telomeres is compromised. Consistent with the idea that TERRA acts upstream of ATR to promote RPA retention at ALT telomeres, VE-821 did not affect TERRA levels and telomere association in U2OS cells (fig. S14, A and B).

The effects of VE-821 on ALT telomeres prompted us to investigate whether VE-821 selectively kills ALT cells. SW26[ALT] was indeed more sensitive to VE-821 than SW39[TEL] (fig. S15). Notably, SW26[ALT] and SW39[TEL] were similarly sensitive to a panel of DNA-damaging agents (fig. S16), demonstrating that the effects of VE-821 are specific to ATR inhibition but not a result of general genotoxicity. Moreover, VE-821 induced accumulation of the phosphorylated form of histone H2AX (γH2AX) more efficiently in SW26[ALT] than in SW39[TEL] cells (fig. S17), suggesting that it inflicts more DNA damage in ALT cells. At a concentration that kills U2OS cells, VE-821 only modestly reduced the proliferation of untransformed RPE-1 retinal pigment epithelial cells (fig. S18). Using H2B-mRFP and live-cell imaging, we followed the chromosome segregation in U2OS, HeLa, and RPE-1 cells after VE-821 treatment. Furthermore,
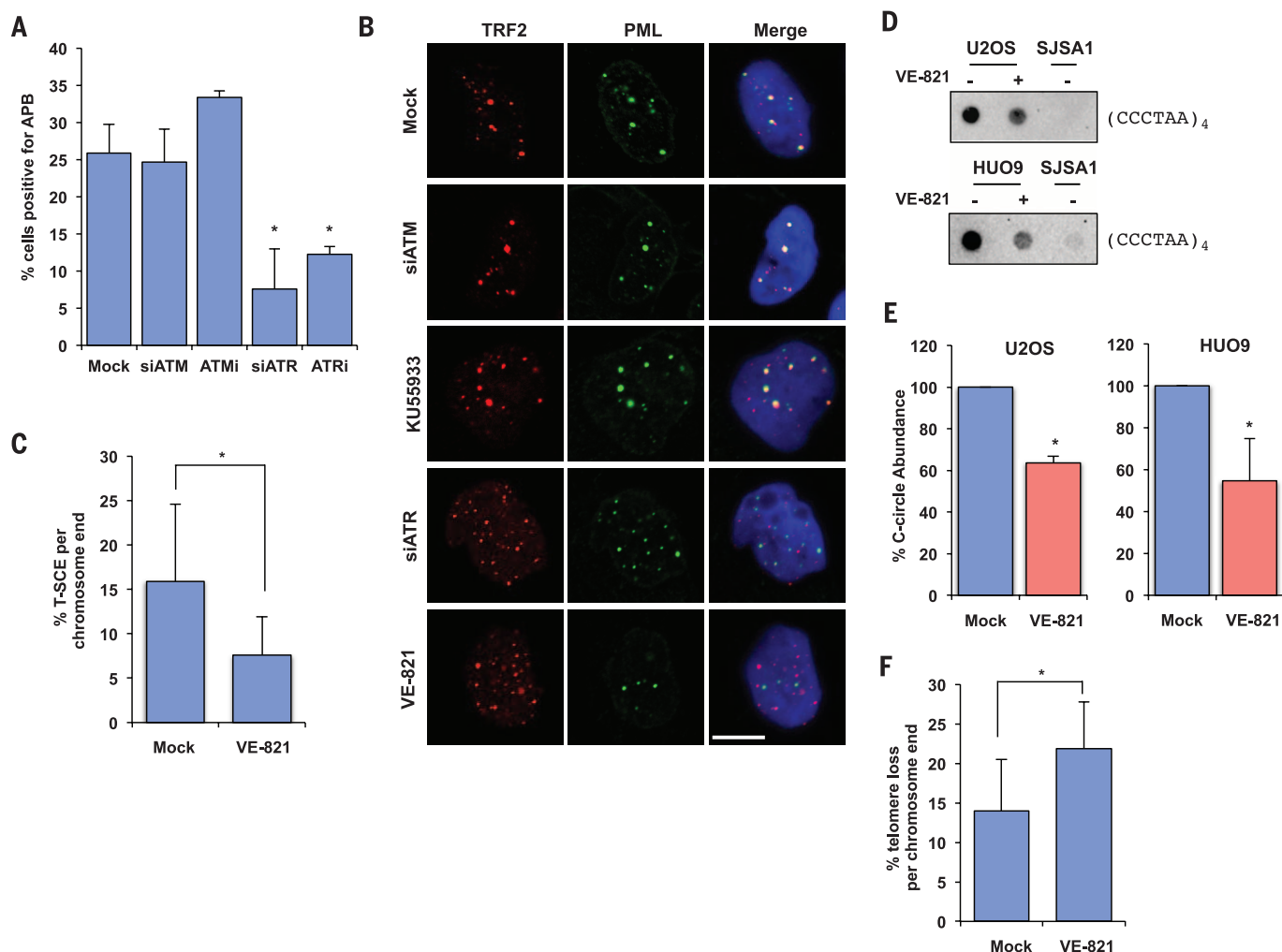
Fig. 3. ATR inhibitor disrupts ALT activity. (**A**) U2OS cells were mock treated, treated with 5 μM VE-821 or 5 μM KU-55933, or treated with siRNA for ATR or ATM, and then immunostained for TRF2 and PML. The percentage of cells positive for APBs was graphed as the mean ± SD; experiment was performed in triplicate (n = 3). P < 0.02. (**B**) Representative images from cells quantified in (A). Scale bar: 10 μm. (**C**) U2OS cells were mock treated or treated with 2.5 μM VE-821 for 4 days and analyzed for T-SCE events with G-rich (green) and C-rich (red) PNA probes. The fraction of chromosome ends with T-SCE was quantified and graphed as the mean ± SD (Mock n = 1032, VE-821 n =1556). *P < 0.01. (**D** and **E**) HUO9 and U2OS cells were mock treated or treated with 5 μM VE-821 for 24 and 48 hours, respectively. C-circle amplification products were detected by dot blot in (D). The levels of C-circles were graphed in (E) as the mean ± SD (n = 2). Telomerase-positive SJSA1 cells were used as a negative control. P < 0.02. (**F**) The fraction of chromosome ends with telomere loss was quantified and graphed as the mean ± SD (Mock n = 1032, VE-821 n = 1556). *P < 0.01.

we used 53BP1-GFP to visualize DNA double-stranded breaks (DSBs) in U2OS cells. VE-821 induced pronounced errors in anaphase chromosome segregation in U2OS cells but not in HeLa or RPE-1 cells (Fig. 4, A and B, fig. S19, and movie S1). In the subsequent interphase, U2OS cells displayed an increase in micronuclei compared to HeLa or RPE-1 cells (fig. S20A and movie S1). Moreover, U2OS cells exhibited numerous 53BP1 foci (Fig. 4, A and C, and movie S1). A fraction of the 53BP1 foci in U2OS cells colocalized with telomeres (Fig. 4C and fig. S20B), although only a minority of telomeres associated with 53BP1. The colocalization of 53BP1 with telomeres but not centromeres was significantly induced by VE-821 (fig. S21, A and B), suggesting that ALT telomeres are particularly fragile upon ATR inhibition. Knockdown of ATRX in HeLa cells and BJ fibroblasts did not increase the induction of γH2AX by VE-821 or VE-821 sensitivity (fig. S22, A to C), suggesting that although ATRX loss may prime cells for ALT, it is not directly responsible for the vulnerability of ALT cells to ATR inhibition.

Given the prevalence of ALT in osteosarcoma (26), we tested the effects of VE-821 on a panel of osteosarcoma cell lines. These cell lines clearly clustered into two groups (Fig. 4D). The mean 50% inhibition concentration (IC_{50}) of VE-821 for one group (U2OS, SAOS2, CAL72, NOS1, and HUO9) was ~0.8 μM, whereas the mean IC_{50} for the other group (MG63 and SJSA1) was ~9 μM (Fig. 4D and S23A). Among these lines, U2OS and SAOS2 are known ALT lines without detectable ATRX protein (fig. S24A) (14). CAL72, NOS1, and HUO9 lacked detectable telomerase activity and

ATRX protein, and displayed APBs (fig. S24, A to C, and fig. S25, A to C), suggesting that they are also ALT-positive. In contrast, MG63 and SJSA1 were positive for telomerase activity and ATRX protein, and negative for APBs (fig. S24, A to C). In this panel of cell lines, VE-821 induced substantially higher levels of apoptosis in the ALT lines than in the telomerase-positive lines (Fig. 4E). The hypersensitivity of ALT cells to ATR inhibition was confirmed with a second ATR inhibitor, AZ20 (fig. S23B). In contrast to ATR inhibitors, neither the ATM inhibitor KU-55933 nor the DNA replication inhibitor gemcitabine showed significant selectivity toward ALT cells (Fig. 4D and fig. S23, C and D). Notably, several ATRX-expressing ALT lines were also hypersensitive to VE-821 (figs. S25, A to D, and S26, A and B) (14), again suggesting that the state of
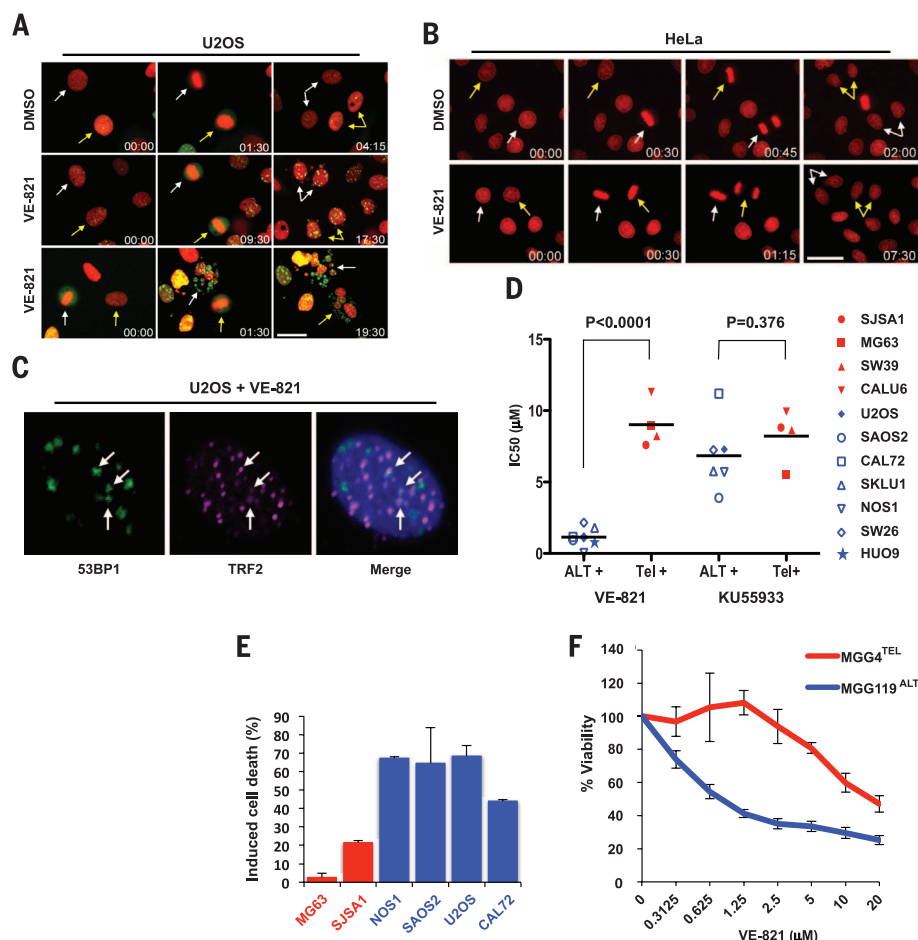
Fig. 4. Selective killing of ALT cells by ATR inhibitor. (**A** and **B**) Stills from time-lapse live-cell imaging experiments of (A) U2OS cells stably expressing H2B-mRFP and 53BP1-GFP or (B) HeLa cells expressing H2B-mRFP after treatment with either 5 μM VE-821 or vehicle control (dimethyl sulfoxide, DMSO). Colored arrows mark individual cells as they progress through mitosis. Time scale: hours:minutes. Scale bar: 30 μm. At least 150 cells were scored for each condition over two independent experiments. (**C**) U2OS cells were treated with VE-821 as in (A) and analyzed by immunofluorescence with 53BP1 and TRF2 antibodies. Scale bar: 10 μm. (**D**) A panel of cell lines were mock treated, treated with VE-821, or treated with KU-55933 for 4 to 6 days. Cell viability was determined with CellTiter Glo. Dots represent IC_{50} concentrations calculated from experiments performed in triplicate ($n = 3$). (**E**) The osteosarcoma cell lines were treated with 3 μM VE-821 for 6 days, and cell death was quantified by annexin V staining. Induced cell death is shown as the mean ± SD ($n = 2$). (**E**) MGG4[TEL] and MGG119[ALT] cells were treated with increasing concentrations of VE-821 for 6 days. Cell viability was determined with CellTiter Glo. Error bars represent SDs; experiment was performed in duplicate ($n = 2$).

ALT telomeres but not ATRX loss per se renders cells hypersensitive to ATR inhibitors.

ALT is prevalent not only in osteosarcoma but also in pediatric glioblastoma (27). MGG119, a newly developed glioma stem cell (GSC) line (28), lacked detectable telomerase activity and ATRX protein, but expressed high levels of TERRA and displayed APBs (fig. S27, A to D), suggesting that it is ALT-positive. In contrast, the GSC line MGG4 was positive for telomerase activity and ATRX protein, but expressed low levels of TERRA and lacked APBs (fig. S27, A to D) (29). Although MGG119[ALT] and MGG4[TEL] were similarly sensitive to a panel of DNA-damaging agents (fig. S28, A to C), MGG119[ALT] was significantly more sensitive to VE-821 than MGG4[TEL] (Fig. 4F), suggesting that VE-821 is uniquely effective in killing ALT GSCs.

The HR defects of specific cancers have offered an opportunity for targeted therapy using PARP inhibitors (5, 6). However, in contrast to HR-defective cancers, ALT-positive cancers actively use recombination to sustain immortality. We show that ATR inhibitors disrupt ALT (fig. S29) and selectively kill ALT cells in vitro, suggesting a rational strategy for the treatment of ALT-positive cancers. Several ATR inhibitors are entering clinical trials for cancer treatment (24, 30–33). Our findings suggest that cancers reliant on recombination, which include but are not limited to ALT-positive cancers, are hypersensitive to ATR inhibitors, offering an unexplored direction for future preclinical and clinical studies.

REFERENCES AND NOTES

1. J. W. Shay, W. E. Wright, *Semin. Cancer Biol.* **21**, 349–353 (2011).
2. S. E. Artandi, R. A. DePinho, *Carcinogenesis* **31**, 9–18 (2010).
3. A. J. Cesare, R. R. Reddel, *Nat. Rev. Genet.* **11**, 319–330 (2010).
4. C. M. Heaphy et al., *Am. J. Pathol.* **179**, 1608–1615 (2011).
5. H. E. Bryant et al., *Nature* **434**, 913–917 (2005).
6. H. Farmer et al., *Nature* **434**, 917–921 (2005).
7. O. E. Bechter, Y. Zou, J. W. Shay, W. E. Wright, *EMBO Rep.* **4**, 1138–1143 (2003).
8. M. S. Wold, *Annu. Rev. Biochem.* **66**, 61–92 (1997).
9. R. L. Flynn et al., *Nature* **471**, 532–536 (2011).
10. V. Schramke et al., *Nat. Genet.* **36**, 46–54 (2004).
11. C. M. Azzalin, P. Reichenbach, L. Khoriauli, E. Giulotto, J. Lingner, *Science* **318**, 798–801 (2007).
12. A. Porro, S. Feuerhahn, P. Reichenbach, J. Lingner, *Mol. Cell. Biol.* **30**, 4808–4817 (2010).
13. S. Schoeftner, M. A. Blasco, *Nat. Cell Biol.* **10**, 228–236 (2008).
14. C. A. Lovejoy et al., *PLOS Genet.* **8**, e1002772 (2012).
15. K. Bower et al., *PLOS ONE* **7**, e50062 (2012).
16. C. M. Heaphy et al., *Science* **333**, 425 (2011).
17. J. Schwartzentruber et al., *Nature* **482**, 226–231 (2012).
18. R. J. O'Sullivan et al., *Nat. Struct. Mol. Biol.* **21**, 167–174 (2014).
19. H. Wang, H. Wang, S. N. Powell, G. Iliakis, Y. Wang, *Cancer Res.* **64**, 7139–7143 (2004).
20. L. Zou, S. J. Elledge, *Science* **300**, 1542–1548 (2003).
21. S. M. Barr, C. G. Leung, E. E. Chang, K. A. Cimprich, *Curr. Biol.* **13**, 1047–1051 (2003).
22. T. R. Yeager et al., *Cancer Res.* **59**, 4175–4179 (1999).
23. J. Déjardin, R. E. Kingston, *Cell* **136**, 175–186 (2009).
24. P. M. Reaper et al., *Nat. Chem. Biol.* **7**, 428–430 (2011).
25. P. R. Potts, H. Yu, *Nat. Struct. Mol. Biol.* **14**, 581–590 (2007).
26. C. Scheel et al., *Oncogene* **20**, 3835–3844 (2001).
27. V. Hakin-Smith et al., *Lancet* **361**, 836–838 (2003).
28. H. Wakimoto et al., *Clin. Cancer Res.* **20**, 2898–2909 (2014).
29. H. Wakimoto et al., *Cancer Res.* **69**, 3472–3481 (2009).
30. E. Fokas et al., *Cell Death Dis.* **3**, e441 (2012).
31. J. D. Charrier et al., *J. Med. Chem.* **54**, 2320–2330 (2011).
32. K. M. Foote et al., *J. Med. Chem.* **56**, 2125–2138 (2013).
33. L. I. Toledo et al., *Nat. Struct. Mol. Biol.* **18**, 721–727 (2011).

T CELL VACCINES

Vaccine-elicited CD4 T cells induce immunopathology after chronic LCMV infection

Pablo Penaloza-MacMaster,[1] Daniel L. Barber,[2] E. John Wherry,[3]
Nicholas M. Provine,[1] Jeffrey E. Teigler,[1]* Lily Parenteau,[1] Stephen Blackmore,[1]
Erica N. Borducchi,[1] Rafael A. Larocca,[1] Kathleen B. Yates,[4] Hao Shen,[3]
W. Nicholas Haining,[4] Rami Sommerstein,[5] Daniel D. Pinschewer,[5,6]
Rafi Ahmed,[7] Dan H. Barouch[1,8]†

CD4 T cells promote innate and adaptive immune responses, but how vaccine-elicited
CD4 T cells contribute to immune protection remains unclear. We evaluated whether
induction of virus-specific CD4 T cells by vaccination would protect mice against
infection with chronic lymphocytic choriomeningitis virus (LCMV). Immunization with
vaccines that selectively induced CD4 T cell responses resulted in catastrophic
inflammation and mortality after challenge with a persistent strain of LCMV.
Immunopathology required antigen-specific CD4 T cells and was associated with a
cytokine storm, generalized inflammation, and multi-organ system failure. Virus-specific
CD8 T cells or antibodies abrogated the pathology. These data demonstrate that
vaccine-elicited CD4 T cells in the absence of effective antiviral immune responses
can trigger lethal immunopathology.

CD4 T cells play an essential role in facilitating innate and adaptive immune responses. Absence of CD4 T cells at the time of priming results in impaired memory CD8 T cell responses (1–4) and severe CD8 T cell dysfunction with uncontrolled viral replication after persistent viral infections (5–8). Moreover, adoptive transfer of virus-specific CD4 T cells during chronic lymphocytic choriomeningitis virus (LCMV) infection has been shown to rescue cytotoxic and humoral responses, resulting in enhanced viral control (9). As a result, developing strategies that preferentially elicit CD4 T cell responses by candidate vaccines has been a research priority, and several CD4 T cell–based vaccines against smallpox and HIV are being tested (10–13). However, little is known about the role of vaccine-elicited CD4 T cells after viral challenge.

We explored whether a vaccine that elicited CD4 T cell responses would afford protective immunity against LCMV infection in mice. We first vaccinated C57BL/6 mice with a *Listeria monocytogenes* vector expressing the LCMV glycoprotein-specific I-A^b–restricted CD4 T cell epitope GP61-80 (LM-GP61) (14). Vaccination elicited durable GP61-specific CD4 T cell responses (fig. S1A) that peaked at day 8 and persisted for more than 60 days after immunization (fig. S1B).

Vaccinated mice were then challenged with LCMV Clone-13 (Cl-13), which causes a systemic infection that lasts for 60 to 90 days (15). As expected, control mice (LM-wt) exhibited modest weight loss after challenge, followed by recovery (16) (Fig. 1A). In contrast, LM-GP61–vaccinated mice exhibited immunopathology characterized by >20% weight loss ($P < 0.0001$) (Fig. 1A) and 90% mortality by day 20 after challenge ($P = 0.0005$) (Fig. 1B), which was associated with a cytokine storm (Fig. 1C). Gross pathology of vaccinated mice after challenge showed widespread inflammation and hemorrhage (Fig. 1D), and histopathology revealed involution of lymphoid tissues, impaired development of B cell follicles, and severe tissue destruction (Fig. 1E), consistent with multi-organ system failure.

We next determined the generalizability of these observations. Immunization of C57BL/6 mice with dendritic cells coated with various I-A^b–restricted CD4 T cell epitopes (GP6, GP126, and NP309 with or without GP61) (fig. S2A) also resulted in mortality upon LCMV Cl-13 challenge (fig. S2B). Moreover, immunization of BALB/c mice with dendritic cells pulsed with the I-A^d–restricted NP116 epitope (fig. S2C) similarly led to mortality upon challenge (fig. S2D). These data dem-

onstrate that the CD4 T cell immunopathology observed with the LM-GP61 vaccine was not specific to the vaccine platform, target epitope, or host genetic background.

We next analyzed adaptive immune responses after challenge. Mice vaccinated with LM-GP61 and challenged with LCMV Cl-13 exhibited elevated GP66-specific CD4 T cell responses in tissues and blood at day 8 (factor of 25 increase relative to controls, $P < 0.0001$) (Fig. 2, A and B). By day 15, these vaccinated mice showed a factor of 21 reduction in immunoglobulin G (IgG) responses ($P = 0.02$) (Fig. 2C), a factor of 153 reduction in the number of germinal center B cells ($P = 0.001$) (Fig. 2, D and E), and a factor of 76 reduction in the number of antibody-secreting cells ($P = 0.002$) relative to controls (Fig. 2F). This decrease in humoral responses in mice that received the CD4 T cell vaccine paralleled the observations from our histological analyses, which showed an absence of germinal centers in the lymph nodes and spleen (Fig. 1F). Moreover, there was a factor of 5.2 reduction in the number of GP276-specific CD8 T cells in the spleen ($P = 0.05$) (Fig. 2G) (gating scheme shown in fig. S3, A to C), which may have been due to a greater CD8 T cell deletion in the context of higher viral loads. Vaccinated mice also exhibited a factor of 6.3 increase in viremia at day 8 ($P = 0.02$) (Fig. 2H). Tissue viral loads were also increased ($P ≤ 0.05$) (fig. S4A), and the pattern of infected cells was similar between vaccinated and control mice at day 8 (fig. S4, B and C).

Despite the massive expansion of GP66-specific CD4 T cells, the lethal immunopathology was associated with a factor of 2.7 reduction in the total number of CD4 T cells ($P = 0.05$) (fig. S5, A and B), suggesting impaired maintenance of CD4 T cells. Mice that received the LM-GP61 vaccine also showed a factor of 3.6 reduction in the frequencies of regulatory T cells (T$_{regs}$) ($P = 0.03$; fig. S5, A and C) and a factor of 15.4 increase in the effector/T$_{reg}$ ratio ($P = 0.003$; fig. S5D), and these mice were moderately lymphopenic (fig. S5E) relative to controls after LCMV Cl-13 challenge. It is unlikely that partial T$_{reg}$ collapse alone caused the observed mortality, because complete T$_{reg}$ ablation typically induces immunopathology after 2 to 3 weeks (17, 18), and the mortality reported here was fulminant, consistent with a cytokine storm rather than autoimmunity.

We next explored the mechanism of the observed lethal immunopathology. First, LM-GP61–vaccinated mice that were challenged with a mutant LCMV Cl-13 virus strain that specifically lacked the GP61-80 epitope (rCl-13/WE-GP ΔGP61) demonstrated no mortality, and depletion of CD4 T cells before LCMV Cl-13 challenge abrogated the immunopathology (Fig. 3A). These data suggest that virus-specific CD4 T cells are required for the observed immunopathology. Second, we assessed whether the immunopathology could be recapitulated simply by increasing the precursor frequency of virus-specific CD4 T cells. We challenged mice with LCMV Cl-13 1 day after adoptive transfer with 10^3 to 10^5 naïve SMARTA cells (TCR-transgenic

[1]Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, MA 02215, USA. [2]Laboratory of Parasitic Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892, USA. [3]Department of Microbiology and Institute for Immunology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA. [4]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA 02115, USA. [5]Department of Pathology and Immunology, WHO Collaborating Centre for Vaccine Immunology, University of Geneva, 1211 Geneva, Switzerland. [6]Department of Biomedicine–Haus Petersplatz, Division of Experimental Virology, University of Basel, 4009 Basel, Switzerland. [7]Emory Vaccine Center and Department of Microbiology and Immunology, Emory University School of Medicine, Atlanta, GA 30322, USA. [8]Ragon Institute of MGH, MIT, and Harvard, Boston, MA 02114, USA.
*Present address: U.S. Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA.
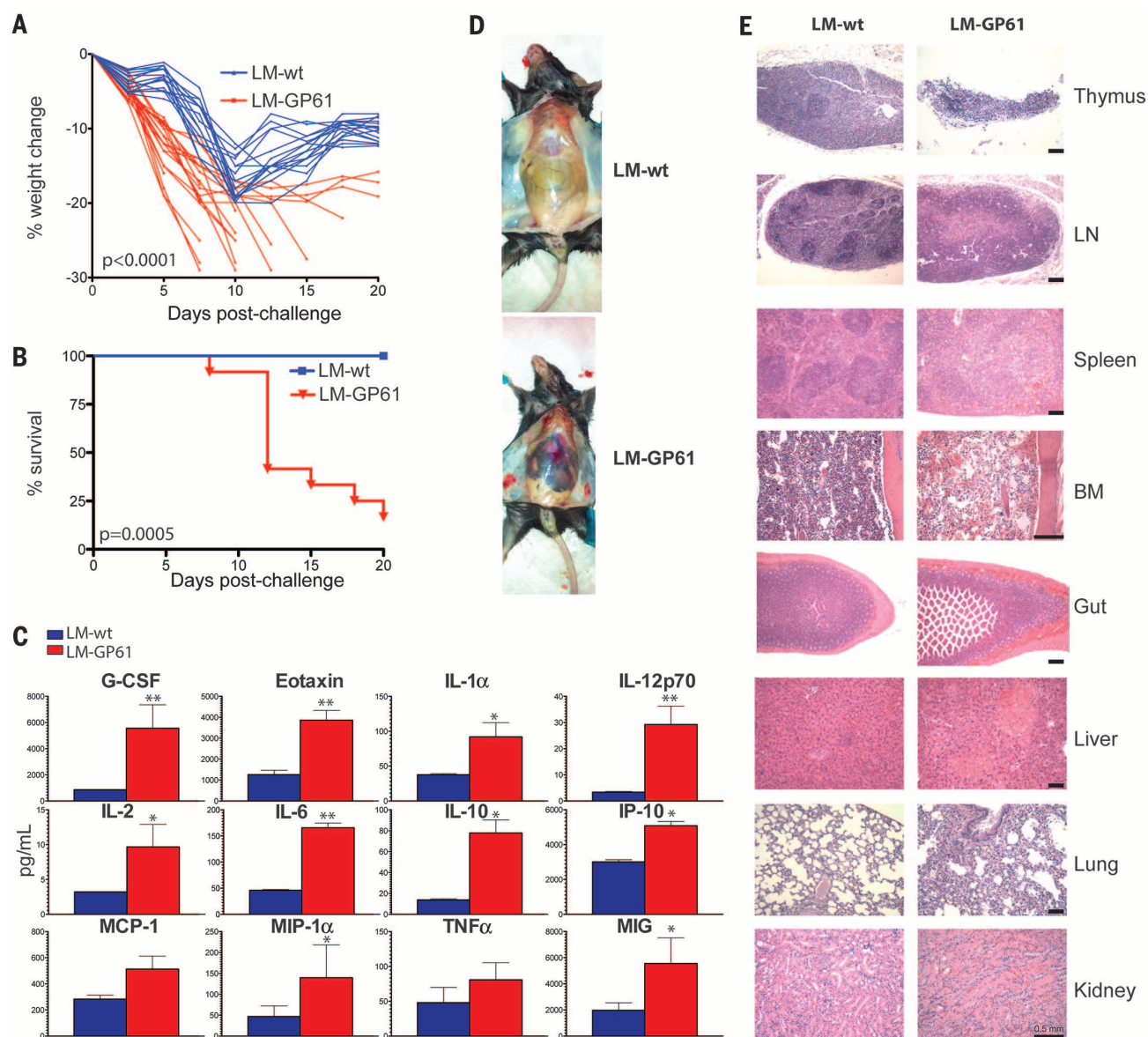†Corresponding author. E-mail: dbarouch@bidmc.harvard.edu

Fig. 1. CD4 T cell vaccines induce lethal immunopathology and systemic inflammation after LCMV Cl-13 challenge. LM-wt or LM-GP61 immune C57BL/6 mice were challenged with 2 × 10⁶ plaque-forming units of LCMV Cl-13. (**A**) Weight loss. *P* value on day 7 is indicated. (**B**) Percent survival. Statistical analysis for survival plot was performed using the Mantel-Cox test. (**C**) Average cytokine levels in serum by luminex assays at day 8. (**D**) Gross pathology of inflammation and hemorrhage from representative mice at day 8. (**E**) Hematoxylin and eosin (H&E) staining of lymphoid and nonlymphoid tissues at day 8 (BM, bone marrow; LN, lymph node). Scale bars, 0.5 mm. In (A) and (B), data from four experiments are combined; *N* = 3 to 5 mice per group per experiment. In (C) to (E), representative data from one of three experiments are shown; *N* = 4 mice per group per experiment. *P = 0.05, **P = 0.02 (Mann-Whitney test). Error bars indicate SEM.

CD4 T cells specific for the LCMV GP66-77 epitope) (fig. S6A). Transfer of 10⁵ SMARTA cells resulted in significant mortality (fig. S6B) and impaired antiviral immunity (fig. S6, C and D), similar to what we observed after vaccination with LM-GP61 or peptide-pulsed dendritic cells. These data suggest that the lethal immunopathology could be recapitulated by increasing the precursor frequency of CD4 T cells.

To explore the mechanism further, we assessed whether suppressing viral replication with LCMV-specific CD8 T cells or antibodies would abrogate this pathology. All mice that were co-immunized with LM-GP61 and various vaccines that expressed the CD8 epitope GP33 or the full-length LCMV glycoprotein survived the LCMV Cl-13 challenge (Fig. 3B). The abrogation of the immunopathology was specifically due to vaccine-elicited CD8 T cells, because depletion of CD8 T cells in co-immunized mice before LCMV Cl-13 challenge recapitulated the observed mortality (Fig. 3C). In addition, adoptive transfer of 10⁶ P14 CD8 T cells (TCR transgenic CD8 T cells specific for LCMV GP33) or purified CD8 T cells from mice

that cleared LCMV Armstrong (a strain that is acutely cleared and induces functional responses) prevented the immunopathology (Fig. 3C). To test the role of antibodies in preventing the observed immunopathology, we challenged vaccinated mice with a recombinant LCMV Cl-13 strain expressing LCMV WE-GP (LCMV Cl-13/WE-GP), which can be neutralized by administering the monoclonal antibody KL25 (19). Similarly, administration of this neutralizing antibody, but not an isotype-matched control antibody, abrogated the lethal pathology in LM-GP61–vaccinated mice after
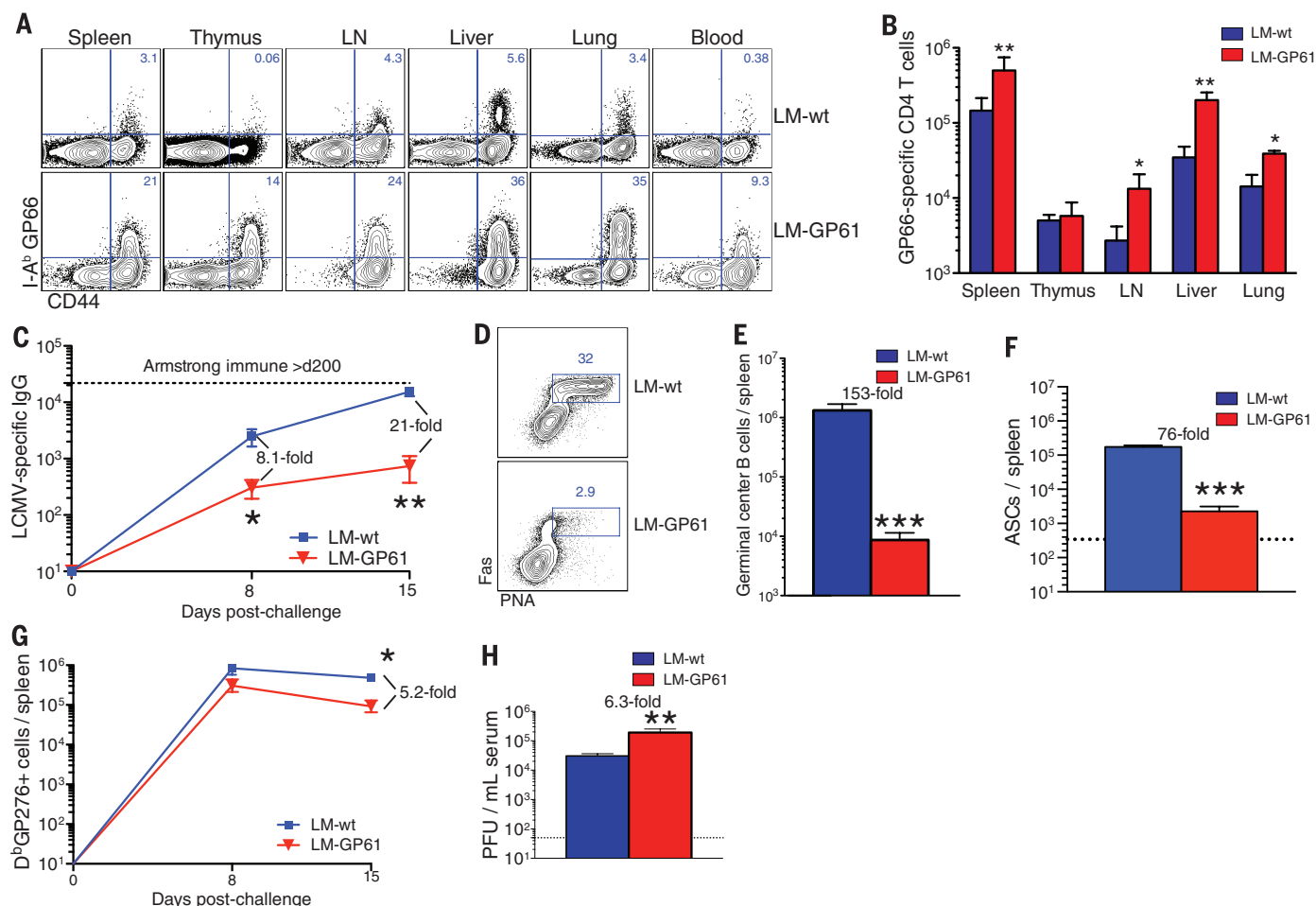
Fig. 2. Uncontrolled anamnestic LCMV-specific CD4 T cells and impaired adaptive immunity after LCMV Cl-13 challenge. (**A**) Representative fluorescence-activated cell sorting (FACS) plot showing I-A^b–restricted GP66-specific CD4 T cells in lymphoid and nonlymphoid tissues at day 8. (**B**) Numbers of I-A^b–restricted GP66-specific CD4 T cells in lymphoid and nonlymphoid tissues at day 8. (**C**) Longitudinal analysis of LCMV-specific IgG responses in sera. (**D**) Representative FACS plot showing germinal center B cell responses in spleen at day 15. (**E**) Number of germinal center B cells in spleen at day 15. (**F**) Number of antibody-secreting cells in spleen at day 15. (**G**) Longitudinal analysis of LCMV-specific (D^bGP276^+) CD8 T cell responses in spleen. (**H**) Viremia on day 8 after infection. Experiment was performed similarly to Fig. 1. In (B), (C), and (E) to (H), data from five experiments are combined; n = 3 or 4 mice per group per experiment. *P = 0.05, **P = 0.02, ***$P \leq 0.002$ (Mann-Whitney test). Error bars indicate SEM.

challenge with neutralization-sensitive LCMV Cl-13/WE-GP (Fig. 3D).

Mice that were co-immunized with vaccines that induce both CD4 and CD8 T cell responses demonstrated more robust CD8 T cell recall responses (Fig. 3E), which was associated with a factor of 40.3 reduction in CD4 T cell responses (P = 0.04) (Fig. 3, F and G) and complete virological control (P = 0.007) by day 8 after challenge (Fig. 3G). These data support the proposed model of antigen-driven hyperstimulation of vaccine-elicited CD4 T cells. Sufficient antiviral CD8 T cells or antibodies limit viral replication and thereby reduce the antigen-dependent activation of memory CD4 T cells, thus abrogating the observed immunopathology. The absence of mortality in LM-wt–immunized mice was likely due to the low numbers of virus-specific CD4 T cell responses relative to LM-GP61–immunized mice.

Many CD4 T cell epitopes incorporate smaller CD8 T cell epitopes (20–24). For example, within the LCMV GP61-80 CD4 epitope lies an em-

bedded H-2K^b–restricted GP70-77 CD8 T cell epitope (25) (fig. S7). However, this CD8 T cell response was too subdominant to control viral replication; H-2K^b–deficient mice (which cannot generate GP70-specific CD8 T cells) and wild-type mice vaccinated with LM-GP61 similarly succumbed after LCMV Cl-13 challenge (fig. S8).

Finally, we assessed differences in the transcriptional profiles of CD4 T cells on day 8 after chronic viral challenge by gene expression profiling. Transcriptional analysis of purified GP66-specific CD4 T cells (Fig. 4A) identified numerous differentially expressed genes (Fig. 4, B and C, and tables S1 to S3). After LCMV Cl-13 challenge, LCMV-specific CD4 T cells from both experimental groups remained FoxP3^− (fig. S9A), and some exhibited T follicular helper (T_FH) differentiation (fig. S9B), as expected (26). GP66-specific CD4 T cells from LM-GP61–vaccinated mice showed lower levels of Eomes (a transcription factor associated with expression of inhibitory receptors and exhaustion) (27) and expressed higher levels of CCR5 (the HIV

co-receptor) relative to controls (Fig. 4, B and C, and fig. S9, C and D).

GP66-specific CD4 T cells from controls exhibited the expected CD4 T cell exhaustion signature characterized by high Eomes expression (Fig. 4, B to D). In contrast, GP66-specific CD4 T cells from LM-GP61–vaccinated mice showed marked enrichment of genes and expression of cytokines associated with highly functional, effector T helper 1 (T_H1) responses (Fig. 4, C to E, and fig. S10). Moreover, by gene set enrichment analysis, immunopathologic CD4 T cells displayed an activated T_H1 CD4 T cell signature, which suggests that these cells failed to undergo normal physiologic exhaustion (Fig. 4, E to G). Furthermore, the gene expression profiles in LM-GP61–vaccinated mice after challenge showed enrichment for cellular processes involved in T cell activation and lymphocyte activation (fig. S11 and tables S1 to S3). Moreover, the lethal immunopathology required ongoing viral replication, as LM-GP61–vaccinated mice challenged with LCMV Armstrong

Fig. 3. Prevention of CD4 T cell–mediated pathology by antiviral CD8 T cells and antibodies. (**A**) Percent survival in LM-GP61–vaccinated mice after LCMV CI-13 challenge with or without CD4 T cell depletion or after rCI-13/WE-GP ΔGP61 challenge. (**B**) Percent survival after co-immunization with LM-GP61 and *Listeria*-, poxvirus-, or adenovirus-based vaccines that elicited CD8 T cell responses before challenge. (**C**) Percent survival in co-immunized mice after CD8 T cell depletion or in LM-GP61–immunized mice after adoptive transfer of P14 transgenic CD8 T cells or LCMV-specific CD8 T cells from LCMV Armstrong immune mice

before chronic viral challenge. (**D**) Percent survival after infection with rCI-13/WE-GP expressing the glycoprotein of the WE strain, with or without infusion with the WE-GP–specific monoclonal antibody KL25. (**E**) Representative FACS plots showing the kinetics of GP33-specific CD8 T cell responses before and after chronic viral challenge. (**F**) Representative FACS plots showing the kinetics of GP61-specific CD4 T cell responses before and after chronic viral challenge. (**G**) Summary of CD8 T cell responses, CD4 T cell responses, and viral loads after chronic viral infection. Statistical analyses for survival plots were performed using the Mantel-Cox test. Experiment was performed similarly to Fig. 1. In (A) to (D) and (G), data from two experiments are combined; *n* = 5 or 6 mice per group per experiment. Error bars indicate SEM.
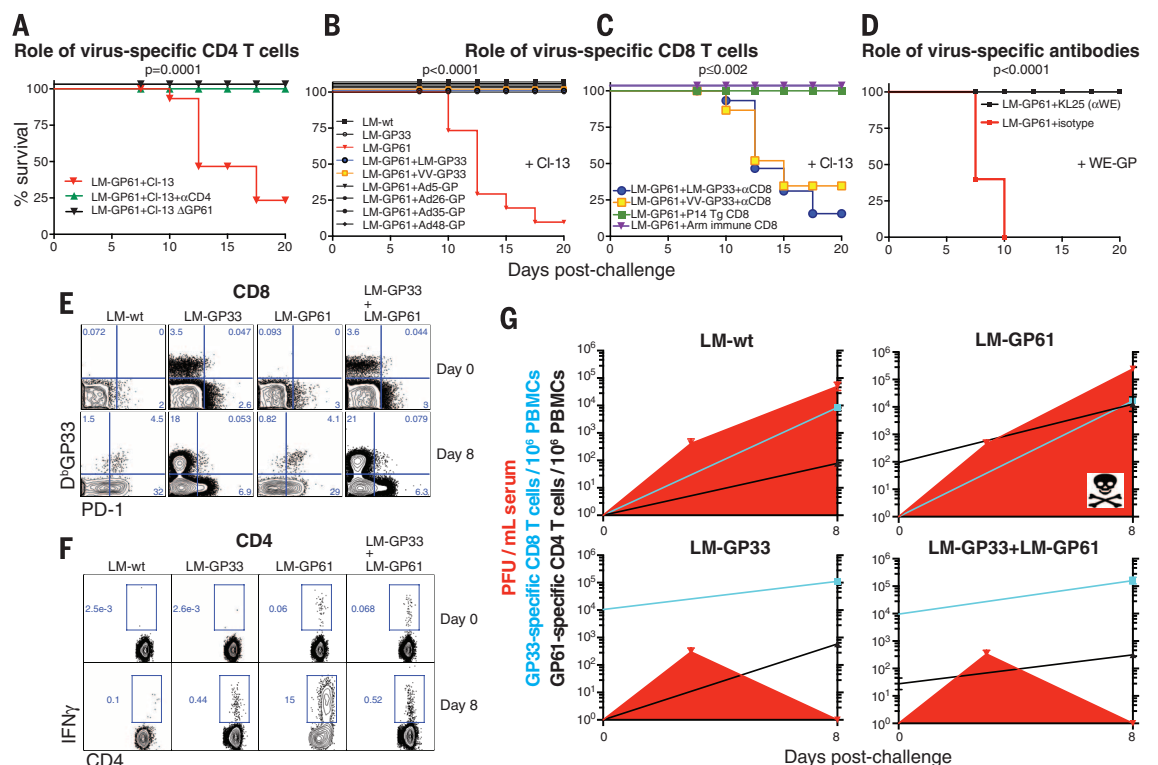
Fig. 4. Vaccine-elicited CD4 T cells are highly functional and bypass typical functional exhaustion after LCMV CI-13 challenge. Microarray analysis was performed on sorted GP66-specific CD4 T cells at day 8. (**A**) Cell purity after FACS sorting of I-A^b–restricted GP66-specific CD44^+ CD4 T cells. (**B**) Heat map of the most differentially expressed genes. (**C**) Heat map comparing the expression of transcription factors involved in CD4 T cell differentiation. (**D**) Relative change in the expression of transcription factors involved in

CD4 T cell differentiation. Dotted line represents the cutoff for significance (factor of ≥1.5 change). (**E**) Enrichment score for different CD4 T cell subsets by gene set enrichment analysis (GSEA). Asterisks represent significant values [*P* value and false discovery rate (FDR) *q* value <0.01]. (**F**) GSEA demonstrating enrichment for a CD4 T cell exhaustion signature in chronically infected mice that had received the control vaccine. (**G**) GSEA demonstrating enrichment for a functional T_H1 CD4 T cell response signature in chronically infected mice that had received the CD4 T cell vaccine. Experiment was performed similarly to Fig. 1. Data are from LM-wt (*n* = 3) and LM-GP61–vaccinated (*n* = 4) mice at day 8 after challenge with LCMV CI-13.
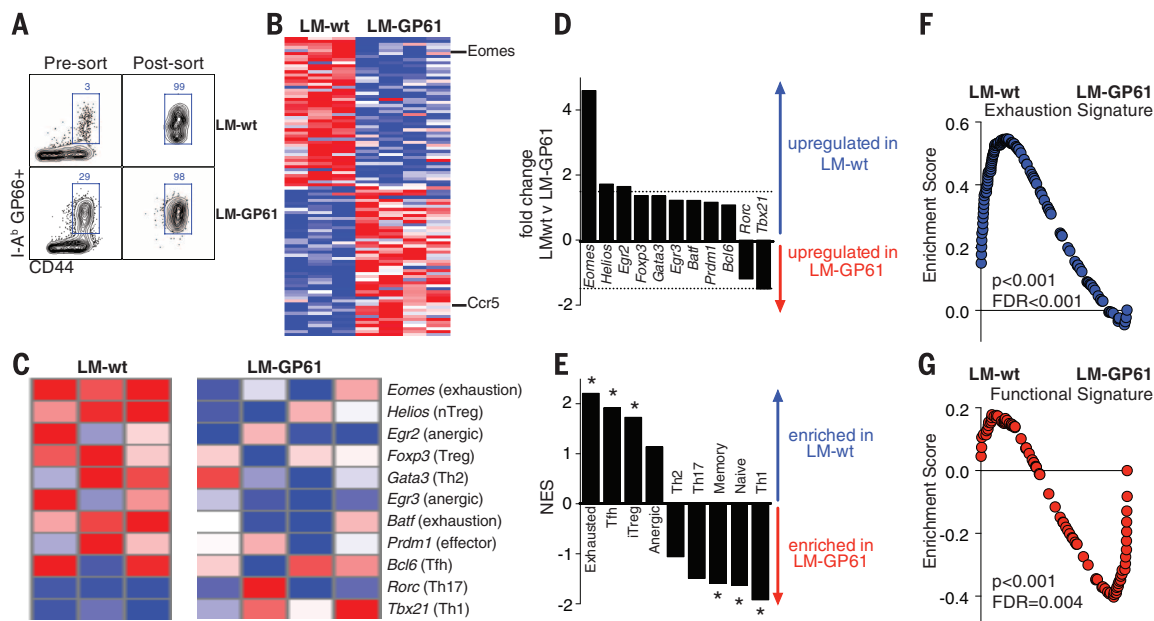
showed no mortality, with a modestly enhanced memory CD8 T cell differentiation and viral control (fig. S12). Taken together, these data are consistent with a model in which uncontrolled viral replication resulted in overstimulation of vaccine-elicited T_H1 CD4 cells, leading to generalized inflammation and multi-organ system failure (fig. S13).

Our data demonstrate that a vaccine that elicits primarily CD4 T cells can result in lethal immunopathology after challenge with a persistently replicating virus by a mechanism that involves hyperstimulation of vaccine-elicited CD4 T cells by uncontrolled viral replication. Both antiviral CD8 T cells and antibodies that limit viral replication abrogate this pathology. These data show that vaccine-elicited CD4 T cells can trigger immunopathology and mortality in certain settings.

Although the extent to which this phenomenon may occur in humans has not yet been determined, this mechanism is potentially generalizable to other vaccines that primarily induce CD4 T cells in the absence of other effective antiviral immune responses. A previous study reported that a vaccine that encoded an SIV CD4 T cell epitope led to higher viral loads and accelerated AIDS progression relative to controls after SIV challenge in rhesus monkeys (28), although SIV-specific CD4 T cell responses were not directly measured in that study. Moreover, because activated CD4 T cells also can serve directly as targets for HIV, vaccine-elicited CD4 T cells could theoretically have multifactorial negative effects (29). These findings warrant a thorough reevaluation of CD4 T cell responses, especially in the context of chronic infection.

REFERENCES AND NOTES

1. J. C. Sun, M. J. Bevan, *Science* **300**, 339–342 (2003).
2. D. J. Shedlock, H. Shen, *Science* **300**, 337–339 (2003).
3. E. M. Janssen *et al.*, *Nature* **421**, 852–856 (2003).
4. E. M. Janssen *et al.*, *Nature* **434**, 88–93 (2005).
5. M. Matloubian, R. J. Concepcion, R. Ahmed, *J. Virol.* **68**, 8056–8063 (1994).
6. M. Battegay *et al.*, *J. Virol.* **68**, 4700–4704 (1994).
7. R. D. Cardin, J. W. Brooks, S. R. Sarawar, P. C. Doherty, *J. Exp. Med.* **184**, 863–871 (1996).
8. A. Grakoui *et al.*, *Science* **302**, 659–662 (2003).
9. R. D. Aubert *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 21182–21187 (2011).
10. M. Lichterfeld *et al.*, *J. Acquir. Immune Defic. Syndr.* **59**, 1–9 (2012).
11. R. R. Amara, P. Nigam, S. Sharma, J. Liu, V. Bostik, *J. Virol.* **78**, 3811–3816 (2004).
12. F. Rodriguez, S. Harkins, J. M. Redwine, J. M. de Pereda, J. L. Whitton, *J. Virol.* **75**, 10421–10430 (2001).
13. H. Streeck, M. P. D'Souza, D. R. Littman, S. Crotty, *Nat. Med.* **19**, 143–149 (2013).
14. M. A. Williams, E. V. Ravkov, M. J. Bevan, *Immunity* **28**, 533–545 (2008).
15. E. J. Wherry, J. N. Blattman, K. Murali-Krishna, R. van der Most, R. Ahmed, *J. Virol.* **77**, 4911–4927 (2003).
16. S. N. Waggoner, M. Cornberg, L. K. Selin, R. M. Welsh, *Nature* **481**, 394–398 (2012).
17. P. Penaloza-MacMaster *et al.*, *J. Exp. Med.* **211**, 1905–1918 (2014).
18. J. M. Kim, J. P. Rasmussen, A. Y. Rudensky, *Nat. Immunol.* **8**, 191–197 (2007).
19. J. M. Bruns, J. Cihak, G. Müller, F. Lehmann-Grube, *Virology* **130**, 247–251 (1983).
20. T. Dao *et al.*, *PLOS ONE* **4**, e6730 (2009).
21. C. Dow *et al.*, *J. Virol.* **82**, 11734–11741 (2008).
22. R. J. May *et al.*, *Clin. Cancer Res.* **13**, 4547–4555 (2007).
23. D. Ou, L. A. Mitchell, D. Décarie, S. Gillam, A. J. Tingle, *Virology* **235**, 286–292 (1997).
24. S. I. Abrams, S. F. Stanziale, S. D. Lunin, S. Zaremba, J. Schlom, *Eur. J. Immunol.* **26**, 435–443 (1996).
25. D. Homann *et al.*, *Virology* **363**, 113–123 (2007).
26. L. M. Fahey *et al.*, *J. Exp. Med.* **208**, 987–999 (2011).
27. A. Crawford *et al.*, *Immunity* **40**, 289–302 (2014).
28. S. I. Staprans *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 13026–13031 (2004).
29. A. S. Fauci, M. A. Marovich, C. W. Dieffenbach, E. Hunter, S. P. Buchbinder, *Science* **344**, 49–51 (2014).

SUPERCONDUCTIVITY

Charge ordering in the electron-doped superconductor Nd2−xCexCuO4

Eduardo H. da Silva Neto,[1,2,3,4]*† Riccardo Comin,[1,2]* Feizhou He,[5] Ronny Sutarto,[5] Yeping Jiang,[6] Richard L. Greene,[6] George A. Sawatzky,[1,2] Andrea Damascelli[1,2]†

In cuprate high-temperature superconductors, an antiferromagnetic Mott insulating state can be destabilized toward unconventional superconductivity by either hole or electron doping. In hole-doped (p-type) cuprates, a charge ordering (CO) instability competes with superconductivity inside the pseudogap state. We report resonant x-ray scattering measurements that demonstrate the presence of charge ordering in the n-type cuprate Nd2−xCexCuO4 near optimal doping. We find that the CO in Nd2−xCexCuO4 occurs with similar periodicity, and along the same direction, as in p-type cuprates. However, in contrast to the latter, the CO onset in Nd2−xCexCuO4 is higher than the pseudogap temperature, and is in the temperature range where antiferromagnetic fluctuations are first detected. Our discovery opens a parallel path to the study of CO and its relationship to antiferromagnetism and superconductivity.

Copper oxide superconductors are susceptible to a number of instabilities, but the relevance of these phases to the superconducting pairing mechanism is unclear. Charge ordering (CO) has emerged as a universal feature of hole-doped (p-type) cuprates, but it has so far not been detected in n-type cuprates (1) (Fig. 1A). Early evidence for a CO in the cuprates came from the detection in La-based cuprates of a periodic organization of spins and charge known as stripes (2–5), where the charge is periodic every four lattice constants along the Cu-O bond direction. More recently, following evidence for Fermi surface reconstruction from quantum oscillations (6, 7), nuclear magnetic resonance (8) and x-ray scattering measurements (9, 10) have directly shown the presence of a similar CO competing with superconductivity in Y-based cuprates. The opportunity to directly probe CO in reciprocal space has further propelled several resonant x-ray scattering measurements of the Y-based family (11–13) as well as the detection of CO in Bi cuprates (14–16)—substantiating earlier surface evidence by scanning tunneling microscopy (17–20)—and also in the single-layer Hg compound (21).

Studies of Bi-based cuprates, for which a considerable amount of angle-resolved photoemission spectroscopy (ARPES) data are available, show that the CO wave vector connects the ends of the Fermi arcs (14, 15)—an observation that links the existence of CO to the pseudogap in hole-doped systems. Additionally, doping-dependent measurements on bilayer systems (9, 13, 22, 23) find charge ordering to be most pronounced in a region of hole doping near $x = 1/8$, where stripes are predominant in La-based cuprates (2, 3). These results raise the questions of whether the

[1]Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia V6T 1Z1, Canada. [2]Quantum Matter Institute, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada. [3]Max Planck Institute for Solid State Research, D-70569 Stuttgart, Germany. [4]Quantum Materials Program, Canadian Institute for Advanced Research, Toronto, Ontario M5G 1Z8, Canada. [5]Canadian Light Source, Saskatoon, Saskatchewan S7N 2V3, Canada. [6]Center for Nanophysics and Advanced Materials and Department of Physics, University of Maryland, College Park, MD 20742, USA.
*These authors contributed equally to this work. †Corresponding author. E-mail: ehda@physics.ubc.ca (E.H.d.S.N.); damascelli@physics.ubc.ca (A.D.)
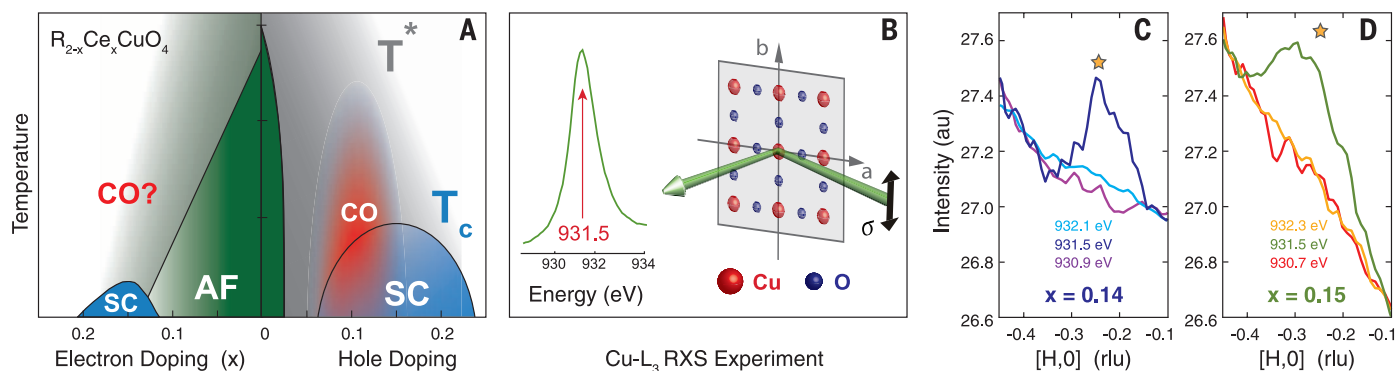
Fig. 1. Charge ordering in electron-doped cuprates. (**A**) Temperature-doping phase diagram for the cuprates, including the AF parent state (green), the superconductivity (SC, blue), and distinct n-type (faded green) and p-type (gray) pseudogap phases. The CO phase observed in p-type cuprates is marked in red. (**B**) The Cu-L_3 absorption edge at 931.5 eV ($2p \rightarrow 3d$ transition) and a schematic of the scattering geometry. (**C** and **D**) On- and off-resonance θ scans at 22 K, showing the RXS diffraction signal as a function of in-plane momentum transfer (H) along the Cu-O bond direction [see (B)] for $x = 0.14$ and $x = 0.15$, respectively. To provide a better comparison, the off-resonance scans were rescaled to match the tails of the on-resonance θ scans. The yellow stars mark the H values of highest intensity for the two samples (obtained from Fig. 3).

particular phenomenology of the hole-doped cuprates such as the pseudogap-induced Fermi arcs, or the propensity toward stripe formation, are necessary ingredients for CO formation, or whether CO is a generic electronic property of the CuO_2 layer that is ubiquitous to all cuprates including n-type materials.

Here, we report resonant x-ray scattering (RXS) measurements on the electron-doped cuprate superconductor $Nd_{2-x}Ce_xCuO_4$ (24). Our studies were performed on samples with doping levels ($x = 0.14 \pm 0.01$ and $x = 0.15 \pm 0.01$) for which quantum oscillations indicate a small Fermi surface (25, 26). We use the standard scattering geometry (Fig. 1B) (24), similar to previous studies (9, 14, 15). The tetragonal b axis of the sample is positioned perpendicular to the scattering plane, allowing the in-plane components of momentum transfer to be accessed by rotating the sample around the b axis (θ scan). For RXS measurements, the energy of the incoming photons is fixed to the maximum of the Cu-L_3 absorption edge, which is at $E \approx 931.5$ eV (Fig. 1B).

Our main finding is summarized in Fig. 1, C and D. An RXS peak is observed at an in-plane momentum transfer of $H \approx -0.24$ rlu (reciprocal lattice units) along the Cu-O bond direction; this is notably similar in periodicity and direction to the x-ray scattering peaks found in the hole-doped materials (3–5, 9–16, 21, 23). The use of photons tuned to the Cu-L_3 edge is expected to greatly enhance the sensitivity in our measurement to charge modulations involving the valence electrons in the CuO_2 planes (3). As the photon energy is tuned away from resonance, the distinct peak near $H = -0.24$ disappears, thus confirming its electronic origin (Fig. 1, C and D) (24). This shows the presence of charge ordering in an electron-doped cuprate.

Further insights into charge ordering formation are obtained by temperature-dependent measurements. The distinct CO peak observed at low temperatures (Fig. 2, A and B) weakens as the temperature is raised, but disappears only above
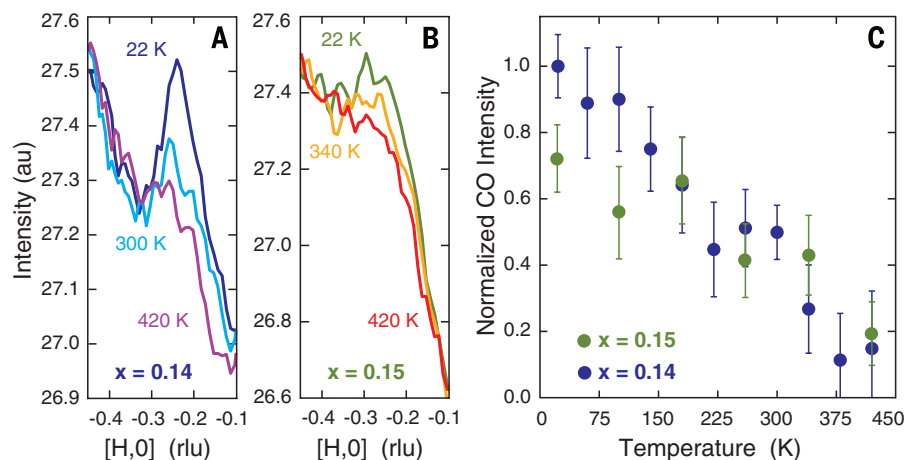
Fig. 2. Temperature dependence of the CO. (**A** and **B**) On-resonance θ scans for $x = 0.14$ and $x = 0.15$ samples, respectively, at select temperatures, showing that the onset of the charge ordering occurs above 300 K. (**C**) Temperature dependence of the RXS intensity for the two samples in (A) and (B) obtained from the maxima of the background-subtracted peaks. The intensity in (C) is normalized to the maximum value between the two samples; the error bars represent the standard errors from Lorentzian fits to the background-subtracted peaks (24).

300 K. Although a temperature evolution is clearly seen in the raw data (Fig. 2, A and B, and fig. S3), the small size of the peak relative to the high-temperature background precludes a precise determination of an onset temperature. Nonetheless, within the detection limits of the experiment, the CO seems to gradually develop with lowering of temperature starting around 340 K (Fig. 2C). Note that this temperature is much higher than the pseudogap onset in $Nd_{2-x}Ce_xCuO_4$ [~80 to 170 K in the $x = 0.14$ to 0.15 doping range (1, 27, 28)], in clear contrast to observations in hole-doped cuprates, where the p-type pseudogap either precedes or matches the emergence of CO (9–15, 21–23). This dichotomy is not completely unexpected given that the pseudogaps observed in p- and n-type cuprates are dissimilar in many ways (1). In particular, the n-type pseudogap has been asso-

ciated with the buildup of antiferromagnetic (AF) correlations that first appear below 320 K (for $x = 0.145$ samples), as determined by inelastic neutron scattering measurements (27–29). Interestingly, the temperature evolution of the CO resembles the soft onset of AF correlations (28)—an observation that suggests a connection between CO and AF fluctuations in electron-doped cuprates.

We now use the available knowledge of the Fermi surface of $Nd_{2-x}Ce_xCuO_4$ to further investigate the connection between AF and CO formation. We find that the CO peak, although broad, is centered around an in-plane momentum transfer $Q_{CO} = 0.23 \pm 0.04$ and $Q_{CO} = 0.24 \pm 0.04$ for $x = 0.14$ and $x = 0.15$, respectively (Fig. 3, A and B). Comparison of Q_{CO} to the Fermi surface topology measured by ARPES (Fig. 3C, left panel) shows that its value is consistent with scattering between

the parallel segments near (π, 0). Thanks to the relative robustness of the AF phase in n-type cuprates, the Fermi surface has often been interpreted to undergo (π, π) folding along the AF zone boundary—a scenario that is consistent with both ARPES (30) and quantum oscillation results (25). In this context, Q_{CO} would connect opposite sides of electron pockets centered at (π, 0) (Fig. 3C, right panel). Alternatively, Q_{CO} might instead connect the intersections between the AF zone boundary and the underlying Fermi surface, the so-called hot spots where the effect of AF scattering and the pseudogap are maximal (30). However, the conventional expectation that the onset of CO above room temperature should gap the Fermi surface seems to contradict both scenarios, because the pseudogap opens only at the hot spots below 180 K, whereas no gapping is observed near (π, 0) above the superconducting transition (1); this suggests that Fermi surface nesting might not be the origin of the CO. Unfortunately, however, this kind of comparison between temperature scales might be rendered inconclusive by the possibility that the CO never becomes sufficiently long-ranged, or large enough in amplitude, to induce a detectable reconstruction of the Fermi surface (at least in the absence of an applied magnetic field). Indeed, the widths of the CO peaks shown in Fig. 3, A and B, indicate a short correlation length (15 to 27 Å) (24), again similar to what has been observed in Bi-based cuprates (14–16). Perhaps further measurements, spanning larger doping ranges, will be able to test exactly which momentum states are involved in the CO, although the broadness of the CO peak in reciprocal space might ultimately limit the precision to which the location of Q_{CO} on the Fermi surface can be determined.

The fact that CO never develops into a long-ranged electronic ground state might also hinder the ability of transport or thermodynamic probes to detect it. However, we find that the presence of CO might be relevant to the interpretation of experiments that probe the inelastic excitations of $Nd_{2-x}Ce_xCuO_4$. We start by observing that the value of Q_{CO} is consistent with the phonon anomaly near $H \approx 0.2$ observed by inelastic x-ray scattering in $Nd_{2-x}Ce_xCuO_4$ (31). More recently, Hinton et al. (32) reported time-resolved reflectivity studies that show the presence of a fluctuating order competing with superconductivity, although they could not determine which electronic degrees of freedom (i.e., charge or spin) were responsible for such order. Additionally, resonant inelastic x-ray scattering measurements (33, 34) have recently shown the presence of an inelastic mode—above a minimum energy transfer of 300 ± 30 meV [comparable to the pseudogap (30)]—which is distinct from the well-characterized AF fluctuations reminiscent of the Mott-insulating parent state (28, 33–35). Whereas Ishii et al. (34) ascribed this new mode to particle-hole charge excitations, Lee et al. (33) proposed that the mode might be the consequence of an unspecified broken symmetry—a scenario supported by their observation that this mode disappears above 270 K for x = 0.166. Our discovery of charge ordering in $Nd_{2-x}Ce_xCuO_4$ might provide the missing piece of information to interpret the aforementioned studies by identifying the actual broken symmetry.

Finally, on a fundamental level, some degree of electron-hole asymmetry should be expected in the cuprate phase diagram. In fact, whereas doped hole states below the charge transfer gap have a strong O-2p character, n-type doping creates low-energy electronic states of predominantly Cu-3d character in the upper Hubbard band (36–38). This dichotomy, together with recent RXS reports of a bond-centered CO in p-type ma-

terials (39), suggests that the n-type CO observed here may instead be centered on the Cu sites—an idea that requires further investigation. However, despite this underlying electron-hole asymmetry, the CO uncovered in $Nd_{2-x}Ce_xCuO_4$ by our study shows several similarities to its p-type equivalent, such as its direction, periodicity, and short correlation length (14, 15). In addition, our observation of a connection between the onset of CO and AF fluctuations suggests that the latter might generally lead to an accompanying intertwined charge order in unconventional superconductors, regardless of which orbitals are involved in the CO (40, 41). If such is the case, detailed studies will be necessary to understand the role of antiferromagnetism in charge order formation, perhaps even beyond the cuprates. Nonetheless, our discovery of charge ordering in n-type cuprates expands the universality of this phenomenon to the electron-doped side of the phase diagram, and provides a new avenue to understand its microscopic origin by exploiting the differences between p- and n-type cuprates.

REFERENCES AND NOTES

1. N. P. Armitage, P. Fournier, R. L. Greene, Rev. Mod. Phys. 82, 2421–2487 (2010).
2. J. M. Tranquada, B. J. Sternlieb, J. D. Axe, Y. Nakamura, S. Uchida, Nature 375, 561–563 (1995).
3. P. Abbamonte et al., Nat. Phys. 1, 155–158 (2005).
4. J. Fink et al., Phys. Rev. B 79, 100502 (2009).
5. J. Fink et al., Phys. Rev. B 83, 092503 (2011).
6. N. Doiron-Leyraud et al., Nature 447, 565–568 (2007).
7. S. E. Sebastian et al., Nat. Commun. 2, 471 (2011).
8. T. Wu et al., Nature 477, 191–194 (2011).
9. G. Ghiringhelli et al., Science 337, 821–825 (2012).
10. J. Chang et al., Nat. Phys. 8, 871–876 (2012).
11. A. J. Achkar et al., Phys. Rev. Lett. 109, 167001 (2012).
12. S. Blanco-Canosa et al., Phys. Rev. Lett. 110, 187001 (2013).
13. S. Blanco-Canosa et al., Phys. Rev. B 90, 054513 (2014).
14. R. Comin et al., Science 343, 390–392 (2014).
15. E. H. da Silva Neto et al., Science 343, 393–396 (2014).
16. M. Hashimoto et al., Phys. Rev. B 89, 220511 (2014).
17. J. E. Hoffman et al., Science 295, 466–469 (2002).
18. M. Vershinin et al., Science 303, 1995–1998 (2004).
19. C. Howald, H. Eisaki, N. Kaneko, A. Kapitulnik, Proc. Natl. Acad. Sci. U.S.A. 100, 9705–9709 (2003).
20. W. D. Wise et al., Nat. Phys. 4, 696–699 (2008).
21. W. Tabis et al., http://arxiv.org/abs/1404.7658 (2014).
22. C. V. Parker et al., Nature 468, 677–680 (2010).
23. M. Hücker et al., Phys. Rev. B 90, 054514 (2014).
24. See supplementary materials on Science Online.
25. T. Helm et al., Phys. Rev. Lett. 103, 157002 (2009).
26. T. Helm et al., Phys. Rev. Lett. 105, 247002 (2010).
27. Y. Onose, Y. Taguchi, K. Ishizaka, Y. Tokura, Phys. Rev. B 69, 024504 (2004).
28. E. M. Motoyama et al., Nature 445, 186–189 (2007).
29. B. Kyung, V. Hankevych, A.-M. Daré, A.-M. S. Tremblay, Phys. Rev. Lett. 93, 147004 (2004).
30. N. P. Armitage et al., Phys. Rev. Lett. 87, 147003 (2001).
31. M. d'Astuto et al., Phys. Rev. Lett. 88, 167002 (2002).
32. J. P. Hinton et al., Phys. Rev. Lett. 110, 217002 (2013).
33. W. S. Lee et al., Nat. Phys. 10, 883–889 (2014).
34. K. Ishii et al., Nat. Commun. 5, 3714 (2014).
35. K. Yamada et al., Phys. Rev. Lett. 90, 137004 (2003).
36. J. Zaanen, G. A. Sawatzky, J. W. Allen, Phys. Rev. Lett. 55, 418–421 (1985).
37. C. T. Chen et al., Phys. Rev. Lett. 66, 104–107 (1991).
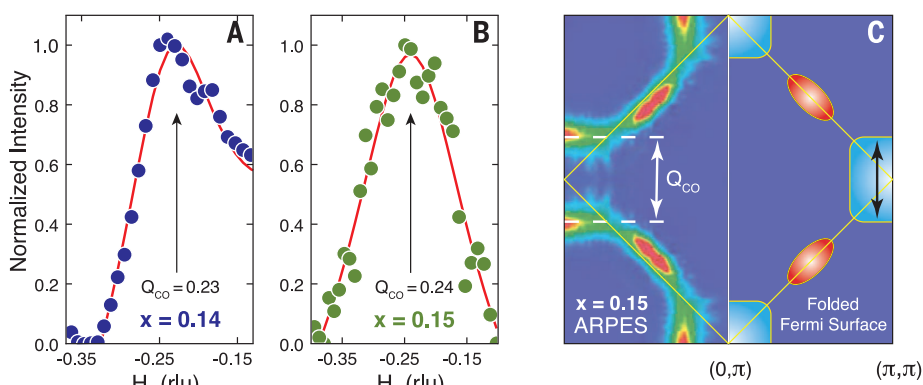38. E. Pellegrin et al., Phys. Rev. B 47, 3354–3367 (1993).

Fig. 3. Electronic origin of the CO. (**A** and **B**) CO peak extracted by subtraction of the highest-temperature θ scan from an average of the lowest-temperature θ scans (22 to 180 K) (24). A fit of the data to a Lorentzian plus linear background function (red line) is used to indicate the H value of highest intensity, which is −0.23 ± 0.04 rlu for the x = 0.14 sample (A) and −0.24 ± 0.04 rlu for the x = 0.15 sample (B). The extracted peaks in (A) and (B) are normalized to their respective maxima. (**C**) Fermi surface of $Nd_{2-x}Ce_xCuO_4$ (x = 0.15) measured by ARPES (30) (left) and a schematic of the expected Fermi surface reconstruction (right) due to AF folding (yellow diamond). The folded Fermi surface is composed of hole (red) and electron (cyan) pockets. The arrows (white and black) and dashed lines represent Q_{CO} = 0.24 rlu, and connect either the parallel segments of the Fermi surface near (π, 0) or the intersection with the AF zone boundary.

39. R. Comin *et al.*, http://arxiv.org/abs/1402.5415 (2014).
40. E. Fradkin, S. A. Kivelson, *Nat. Phys.* **8**, 864–866 (2012).
41. J. C. S. Davis, D.-H. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17623–17630 (2013).

APPLIED PHYSICS

Semiconductor double quantum dot micromaser

Y.-Y. Liu,[1] J. Stehlik,[1] C. Eichler,[1] M. J. Gullans,[2] J. M. Taylor,[2,3] J. R. Petta[1,4]*

The coherent generation of light, from masers to lasers, relies upon the specific structure of the individual emitters that lead to gain. Devices operating as lasers in the few-emitter limit provide opportunities for understanding quantum coherent phenomena, from terahertz sources to quantum communication. Here we demonstrate a maser that is driven by single-electron tunneling events. Semiconductor double quantum dots (DQDs) serve as a gain medium and are placed inside a high-quality factor microwave cavity. We verify maser action by comparing the statistics of the emitted microwave field above and below the maser threshold.

A conventional laser uses an ensemble of atoms that are pumped into the excited state to achieve population inversion (*1*, *2*). Enabled by advances in semiconductor device technology, semiconductor lasers quickly evolved from p-i-n junctions (*3*, *4*), to quantum well structures (*5*) and quantum cascade lasers (QCLs) (*6*). In QCLs, an electrical bias is applied across exquisitely engineered multiple quantum well structures, resulting in cascaded intraband transitions between confined two-dimensional electronic states that lead to photon emission (*7*). However, QCL emission frequencies are set by heterostructure growth profiles and cannot be easily tuned in situ. At the same time, in atomic physics, researchers demonstrated a single-atom maser, where atoms prepared in the excited state transit through a microwave cavity for a precisely controlled period of time, such that the atom "swaps" its excitation to the microwave cavity, generating a large photon field (*8*). These early experiments were extended to a single atom trapped in a high-finesse optical cavity (*9*), as well as condensed-matter systems, where artificial atoms were strongly coupled to cavities (*10–14*).

Here we demonstrate a maser that is driven by single-electron tunneling events. The gain medium consists of semiconductor double quantum dots (DQDs) that support zero-dimensional electronic states (*15*). Electronic tunneling through the DQDs generates photons that are coupled to a cavity mode (*16*). In contrast to optically pumped systems, population inversion is generated in the DQD system through the application of a bias voltage that results in sequential single-electron tunneling.

The maser consists of two semiconductor DQDs (referred to as the left DQD and right DQD, Fig. 1), which are electric-dipole coupled to a microwave cavity. The cavity is formed from a half-wavelength ($\lambda/2$) superconducting Nb transmission line resonator with a center frequency f_c = 7880.55 MHz and a loaded quality factor $Q_c \approx 3000$ (*17*, *18*). Two lithographically defined InAs nanowire DQDs serve as the maser gain medium (*16*, *19*). Each DQD is fabricated by placing a single InAs nanowire over five Ti/Au bottom gate electrodes (Fig. 1C) (*20*, *21*). The bottom gates create a tunable DQD confinement potential in the nanowire (*21*). Electrostatically defined DQDs, often regarded as artificial molecules (*15*), are a unique gain medium. They are fully reconfigurable, with electronic transitions that can be tuned from gigahertz to terahertz frequencies.

A source-drain bias voltage V_{SD} = 2 mV is applied across the DQDs to drive a current. The energy levels of each DQD can be separately tuned and are described by the left (right) DQD detuning ε_L (ε_R). Current will flow in a nanowire DQD through a series of downhill (in energy) single-electron tunneling events (see level diagrams in Fig. 1B). In contrast with quantum well structures, current results from single-electron tunneling events between electrically tunable zero-dimensional states in the DQD (*15*, *22*). Electron tunneling results in microwave gain, which is accessed by measuring the transmission through the cavity (*16*).

To measure the gain, the cavity is driven with a coherent field at frequency $f_{in} = f_c$ with a power P_{in}. Measurements of the output power P_{out} yield the power gain $G = CP_{out}/P_{in}$, where C is a normalization constant set such that $G = 1$ when both DQDs are in Coulomb blockade (no current flow) (*16*, *23*). With $V_{SD} = 0$, charge dynamics within the DQD result in an effective microwave admittance that damps the electromagnetic field inside the cavity, yielding $G < 1$ (*18*, *24*, *25*). Application of a source-drain bias that drives sequential tunneling through the DQD can lead to gain in the cavity transmission, $G > 1$ (*16*). In Fig. 1D, we plot G as a function of ε_L for V_{SD} = 2 mV and $f_{in} = f_c$. For downhill electron tunneling ($\varepsilon_L > 0$), we measure a maximum gain $G \approx 7$ (*23*). In contrast, for $\varepsilon_L < 0$, the left DQD can absorb a photon from the cavity, leading to loss $G \approx 0.2$ (*18*, *25*). These data are acquired with the right DQD configured in Coulomb blockade such that the current is zero (*15*). For simplicity, we refer to a DQD as "on" when its detuning is set to achieve maximum gain and "off" when the DQD is configured in Coulomb blockade with $G = 1$.

We investigate the cavity response by measuring G as a function of f_{in} with $P_{in} = -120$ dBm (Fig. 2). The black curve is the "cold cavity transmission" obtained with both DQDs configured in the off state, where the maximum $G = 1$. Here the gain curve is a Lorentzian with a width set by the cavity decay rate $\kappa/2\pi = 2.6$ MHz. When ε_L is set to the gain peak shown in Fig. 1D, we observe a maximum $G \approx 16$ at $f_{in} = 7880.30$ MHz. Similarly, with the right DQD on and the left DQD off, we observe a maximum $G \approx 6$ at $f_{in} = 7880.41$ MHz. In both configurations, the observed gain rate is too small to reach the maser threshold. In contrast, the red curve in Fig. 2 shows the gain curve with both DQDs in the on state. Here the cavity response is sharply peaked at $f_{in} = 7880.25$ MHz, yielding a maximum gain $G \approx 1000$, which is much larger than the product of the individual gains.

We next examine the characteristics of the device in free-running mode (with no cavity drive tone). Figure 3 shows the power spectral density $S(f)$ of microwave radiation emitted from the cavity in the on/on configuration. The spectrum is

[1]Department of Physics, Princeton University, Princeton, NJ 08544, USA. [2]Joint Quantum Institute, University of Maryland–National Institute of Standards and Technology, College Park, MD 20742, USA. [3]Joint Center for Quantum Information and Computer Science, University of Maryland and NIST, College Park, MD 20742, USA. [4]Department of Physics, University of California, Santa Barbara, CA 93106, USA.
*Corresponding author. E-mail: petta@princeton.edu

39. R. Comin *et al.*, http://arxiv.org/abs/1402.5415 (2014).
40. E. Fradkin, S. A. Kivelson, *Nat. Phys.* **8**, 864–866 (2012).
41. J. C. S. Davis, D.-H. Lee, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17623–17630 (2013).

APPLIED PHYSICS

Semiconductor double quantum dot micromaser

Y.-Y. Liu,[1] J. Stehlik,[1] C. Eichler,[1] M. J. Gullans,[2] J. M. Taylor,[2,3] J. R. Petta[1,4]*

The coherent generation of light, from masers to lasers, relies upon the specific structure of the individual emitters that lead to gain. Devices operating as lasers in the few-emitter limit provide opportunities for understanding quantum coherent phenomena, from terahertz sources to quantum communication. Here we demonstrate a maser that is driven by single-electron tunneling events. Semiconductor double quantum dots (DQDs) serve as a gain medium and are placed inside a high-quality factor microwave cavity. We verify maser action by comparing the statistics of the emitted microwave field above and below the maser threshold.

A conventional laser uses an ensemble of atoms that are pumped into the excited state to achieve population inversion (*1*, *2*). Enabled by advances in semiconductor device technology, semiconductor lasers quickly evolved from p-i-n junctions (*3*, *4*), to quantum well structures (*5*) and quantum cascade lasers (QCLs) (*6*). In QCLs, an electrical bias is applied across exquisitely engineered multiple quantum well structures, resulting in cascaded intraband transitions between confined two-dimensional electronic states that lead to photon emission (*7*). However, QCL emission frequencies are set by heterostructure growth profiles and cannot be easily tuned in situ. At the same time, in atomic physics, researchers demonstrated a single-atom maser, where atoms prepared in the excited state transit through a microwave cavity for a precisely controlled period of time, such that the atom "swaps" its excitation to the microwave cavity, generating a large photon field (*8*). These early experiments were extended to a single atom trapped in a high-finesse optical cavity (*9*), as well as condensed-matter systems, where artificial atoms were strongly coupled to cavities (*10*–*14*).

Here we demonstrate a maser that is driven by single-electron tunneling events. The gain medium consists of semiconductor double quantum dots (DQDs) that support zero-dimensional electronic states (*15*). Electronic tunneling through the DQDs generates photons that are coupled to a cavity mode (*16*). In contrast to optically pumped systems, population inversion is generated in the DQD system through the application of a bias voltage that results in sequential single-electron tunneling.

The maser consists of two semiconductor DQDs (referred to as the left DQD and right DQD, Fig. 1), which are electric-dipole coupled to a microwave cavity. The cavity is formed from a half-wavelength ($\lambda/2$) superconducting Nb transmission line resonator with a center frequency f_c = 7880.55 MHz and a loaded quality factor $Q_c \approx 3000$ (*17*, *18*). Two lithographically defined InAs nanowire DQDs serve as the maser gain medium (*16*, *19*). Each DQD is fabricated by placing a single InAs nanowire over five Ti/Au bottom gate electrodes (Fig. 1C) (*20*, *21*). The bottom gates create a tunable DQD confinement potential in the nanowire (*21*). Electrostatically defined DQDs, often regarded as artificial molecules (*15*), are a unique gain medium. They are fully reconfigurable, with electronic transitions that can be tuned from gigahertz to terahertz frequencies.

A source-drain bias voltage V_{SD} = 2 mV is applied across the DQDs to drive a current. The energy levels of each DQD can be separately tuned and are described by the left (right) DQD detuning ε_L (ε_R). Current will flow in a nanowire DQD through a series of downhill (in energy) single-electron tunneling events (see level diagrams in Fig. 1B). In contrast with quantum well structures, current results from single-electron tunneling events between electrically tunable zero-dimensional states in the DQD (*15*, *22*). Electron tunneling results in microwave gain, which is accessed by measuring the transmission through the cavity (*16*).

To measure the gain, the cavity is driven with a coherent field at frequency $f_{in} = f_c$ with a power P_{in}. Measurements of the output power P_{out} yield the power gain $G = CP_{out}/P_{in}$, where C is a normalization constant set such that $G = 1$ when both DQDs are in Coulomb blockade (no current flow) (*16*, *23*). With V_{SD} = 0, charge dynamics within the DQD result in an effective microwave admittance that damps the electromagnetic field inside the cavity, yielding $G < 1$ (*18*, *24*, *25*). Application of a source-drain bias that drives sequential tunneling through the DQD can lead to gain in the cavity transmission, $G > 1$ (*16*). In Fig. 1D, we plot G as a function of ε_L for V_{SD} = 2 mV and $f_{in} = f_c$. For downhill electron tunneling ($\varepsilon_L > 0$), we measure a maximum gain $G \approx 7$ (*23*). In contrast, for $\varepsilon_L < 0$, the left DQD can absorb a photon from the cavity, leading to loss $G \approx 0.2$ (*18*, *25*). These data are acquired with the right DQD configured in Coulomb blockade such that the current is zero (*15*). For simplicity, we refer to a DQD as "on" when its detuning is set to achieve maximum gain and "off" when the DQD is configured in Coulomb blockade with $G = 1$.

We investigate the cavity response by measuring G as a function of f_{in} with P_{in} = –120 dBm (Fig. 2). The black curve is the "cold cavity transmission" obtained with both DQDs configured in the off state, where the maximum $G = 1$. Here the gain curve is a Lorentzian with a width set by the cavity decay rate $\kappa/2\pi$ = 2.6 MHz. When ε_L is set to the gain peak shown in Fig. 1D, we observe a maximum $G \approx 16$ at f_{in} = 7880.30 MHz. Similarly, with the right DQD on and the left DQD off, we observe a maximum $G \approx 6$ at f_{in} = 7880.41 MHz. In both configurations, the observed gain rate is too small to reach the maser threshold. In contrast, the red curve in Fig. 2 shows the gain curve with both DQDs in the on state. Here the cavity response is sharply peaked at f_{in} = 7880.25 MHz, yielding a maximum gain $G \approx 1000$, which is much larger than the product of the individual gains.

We next examine the characteristics of the device in free-running mode (with no cavity drive tone). Figure 3 shows the power spectral density $S(f)$ of microwave radiation emitted from the cavity in the on/on configuration. The spectrum is

[1]Department of Physics, Princeton University, Princeton, NJ 08544, USA. [2]Joint Quantum Institute, University of Maryland–National Institute of Standards and Technology, College Park, MD 20742, USA. [3]Joint Center for Quantum Information and Computer Science, University of Maryland and NIST, College Park, MD 20742, USA. [4]Department of Physics, University of California, Santa Barbara, CA 93106, USA.
*Corresponding author. E-mail: petta@princeton.edu

Fig. 1. Double quantum dot micromaser. (**A**) Optical micrograph of the DQD micromaser. Cavity photons are coupled to input and output ports with rates κ_{in} and κ_{out}. (**B**) Schematic illustration of the DQD micromaser. Two DQDs are electric-dipole coupled to the microwave cavity. Single-electron tunneling through the DQDs leads to photon emission into the cavity mode. Left (right) DQD detunings ε_L (ε_R) are independently tunable. (**C**) Scanning electron microscope image of an InAs nanowire DQD. (**D**) G as a function of ε_L (measured at frequency f_c) with V_{SD} = 2 mV and the right DQD configured in Coulomb blockade. Insets: For $\varepsilon_L > 0$, electron transport proceeds downhill in energy, resulting in a gain exceeding 7. With $\varepsilon_L < 0$, an electron will be trapped in the right dot until a photon is absorbed, resulting in cavity loss, $G < 1$.
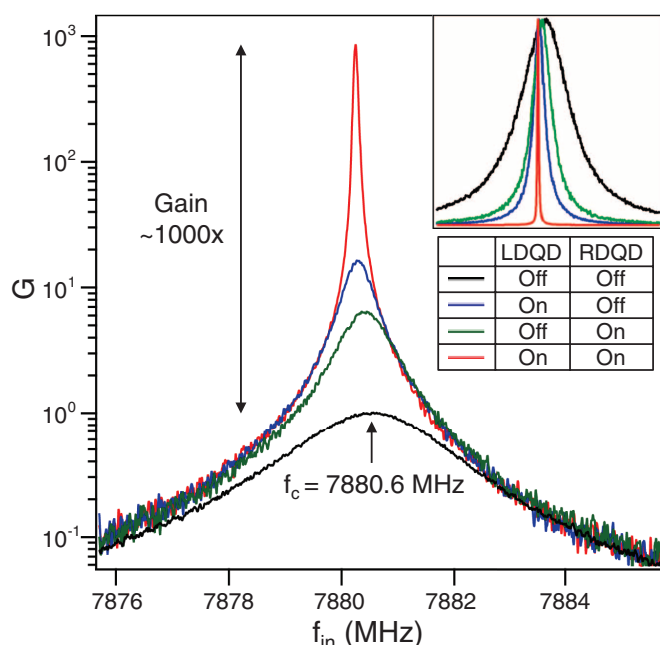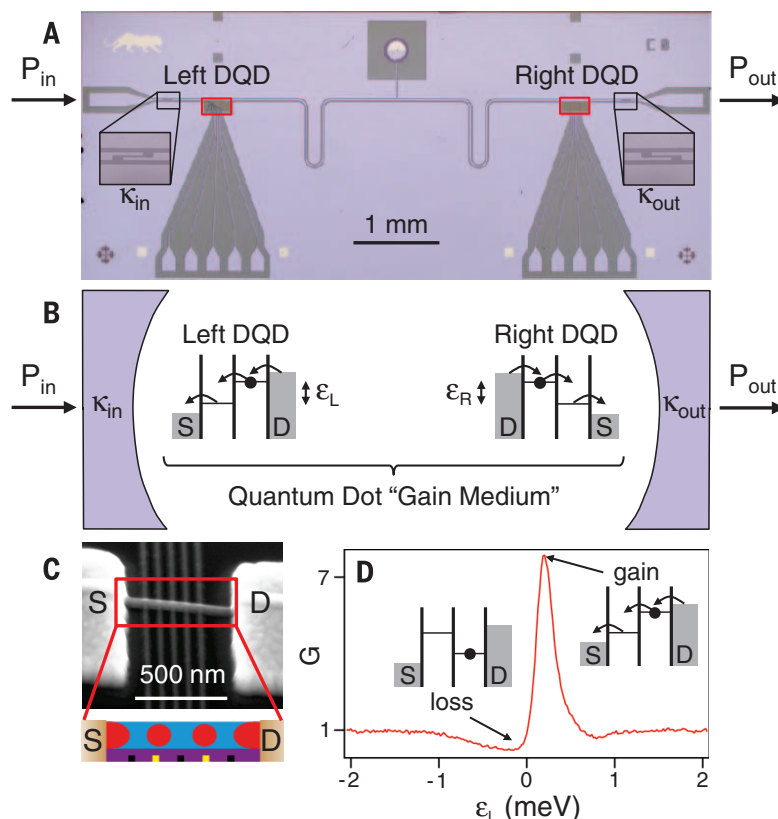
Fig. 2. Microwave gain induced by single-electron tunneling. G as a function of f_{in} with P_{in} = −120 dBm. The black curve is obtained with both DQDs in Coulomb blockade (in the off/off state). With the left DQD set at a detuning that results in gain (see Fig. 1D) and the right DQD in Coulomb blockade (on/off state), we measure a maximum $G \approx 16$. Similarly, in the off/on state, we observe a gain of ≈ 6. Maser action occurs when both DQDs are tuned to produce gain, resulting in $G \approx 1000$. (**Inset**) Data plotted on a linear scale and normalized to the same height.

Fig. 3. Maser coherence time. Power spectral density $S(f)$ measured in free-running maser mode (on/on state with no cavity drive applied). The maser emission peak width Δ_f = 34 kHz (FWHM) yields a coherence length l_{coh} = 2.8 km.

sharply peaked around f = 7880.8 MHz and has a full-width at half-maximum (FWHM) Δ_f = 34 kHz, which corresponds to a coherence time $\tau_{coh} = 1/\pi\Delta_f$ = 9.4 µs and a coherence length $l_{coh} = \tau_{coh}c$ = 2.8 km, where c is the speed of light. The measured linewidth is roughly a fac-

tor of 100 larger than the Schawlow-Townes prediction, but it is not uncommon for conventional semiconductor lasers to have broad emission linewidths (23, 26, 27). Time domain measurements of τ_{coh} are shown in (23).

The most notable evidence of above-threshold maser action is obtained by comparing the statistics of the radiation emitted from the device in the off/on and on/on configurations (23). For this purpose, we have sampled the voltages of the down-converted cavity output field to heterodyne detect the in-phase and quadrature phase components I and Q with a rate of 1 MHz
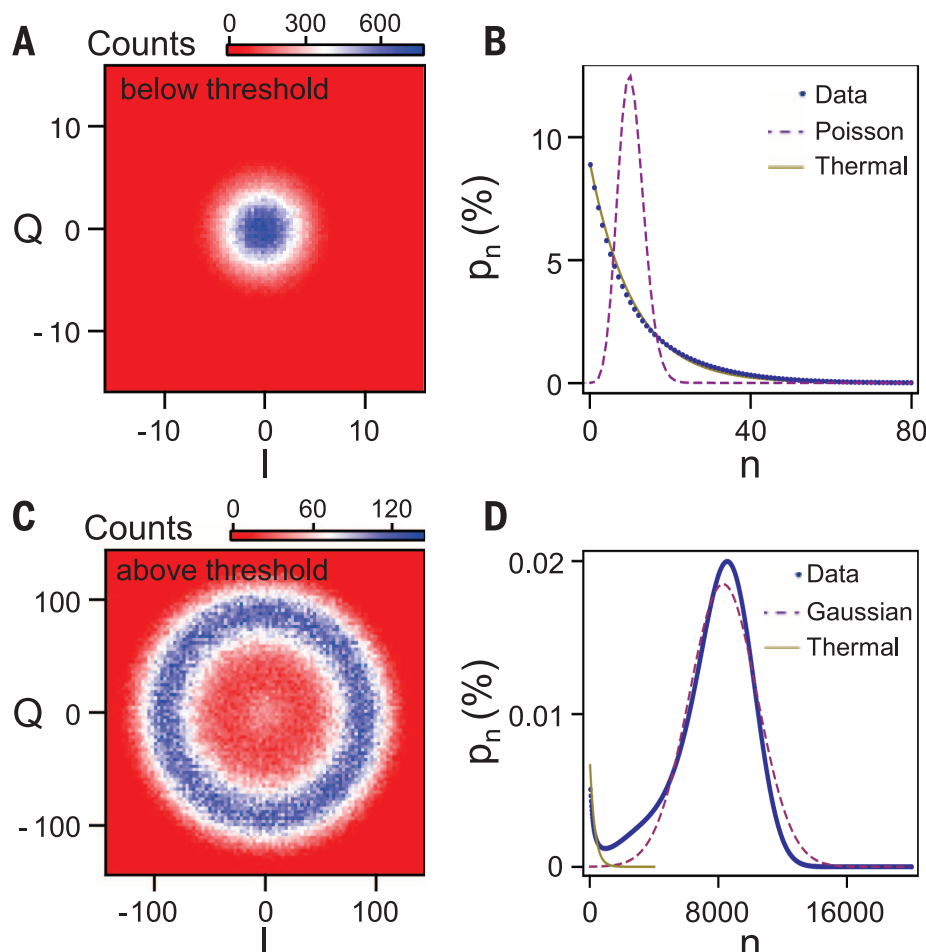
Fig. 4. Photon statistics. (**A**) *IQ* histogram acquired below threshold (off/on configuration). (**B**) The photon number distribution, p_n, extracted from the data in (A) is consistent with a thermal distribution (solid line). A Poisson distribution (dashed line) with $\overline{N} = 11.4$ is shown for comparison. (**C**) *IQ* histogram measured above threshold (on/on configuration). Here the extracted photon number distribution (**D**) is peaked around $n = 8000$ and is compared with a Gaussian distribution (dashed line). A small thermal component (solid line) is attributed to charge fluctuations, which shift the device below threshold.

after applying a 1-MHz digital filter. We store 4×10^5 individual (I,Q) measurements in two-dimensional histograms $D(I,Q)$ to analyze their statistical properties. The measured *IQ* histogram for the off/on configuration (Fig. 4A) is centered near the origin, and the extracted photon number distribution (Fig. 4B) is consistent with a thermal source (*23*). In contrast, Fig. 4C shows the *IQ* histogram for the on/on configuration. Here the *IQ* histogram has a donut shape, consistent with an above-threshold maser (*2*). The extracted photon number distribution is peaked around a photon number $n = 8000$, giving strong evidence for above-threshold behavior. The peak in the photon number distribution is well fit with a Gaussian lineshape, but its width is considerably larger than that of an ideal coherent state $\sqrt{\overline{N}} \approx 90$, where $\overline{N}$ is the average photon number (*23*). Time domain measurements of the maser emission indicate that charge noise fluctuations, which shift the detuning of the DQD gain medium, are most likely re-

sponsible for the broadening. Charge noise also occasionally shifts the system below threshold, leading to the small thermal component observed in Fig. 4D (*23*).

We have demonstrated a maser whose gain medium consists of electrically tunable semiconductor DQDs. Single-electron tunneling in the DQDs provides the energy source for maser action, and a maximum power gain of 1000 is observed. Above-threshold maser action is verified by measuring the statistics of the emitted photon field. Through further improvements in the cavity quality factor (*28*), it may be possible to exceed the lasing threshold with a single DQD emitter. In this case, theory predicts "thresholdless lasing" (*29*). Lastly, the large single-particle level spacings allow for an operation frequency that is purely set by the cavity resonance frequency. This will enable maser operation across a very wide frequency range, spanning gigahertz to terahertz frequencies, a feature that is specific to gate-defined

quantum dots, where electron tunneling takes place between confined zero-dimensional electronic states.

REFERENCES AND NOTES

1. H. Haken, *Laser Theory* (Springer, Berlin, 1983).
2. A. E. Siegman, *Lasers*. (University Science Books, Mill Valley, CA, 1986).
3. R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, R. O. Carlson, *Phys. Rev. Lett.* **9**, 366–368 (1962).
4. M. I. Nathan, W. P. Dumke, G. Burns, F. H. Dill, G. Lasher, *Appl. Phys. Lett.* **1**, 62–64 (1962).
5. R. Dingle, W. Wiegmann, C. H. Henry, *Phys. Rev. Lett.* **33**, 827–830 (1974).
6. J. Faist *et al.*, *Science* **264**, 553–556 (1994).
7. Y. Yao, A. J. Hoffman, C. F. Gmachl, *Nat. Photonics* **6**, 432–439 (2012).
8. D. Meschede, H. Walther, G. Müller, *Phys. Rev. Lett.* **54**, 551–554 (1985).
9. J. McKeever, A. Boca, A. D. Boozer, J. R. Buck, H. J. Kimble, *Nature* **425**, 268–271 (2003).
10. T. Yoshie *et al.*, *Nature* **432**, 200–203 (2004).
11. J. P. Reithmaier *et al.*, *Nature* **432**, 197–200 (2004).
12. Z. G. Xie, S. Götzinger, W. Fang, H. Cao, G. S. Solomon, *Phys. Rev. Lett.* **98**, 117401 (2007).
13. M. Nomura, N. Kumagai, S. Iwamoto, Y. Ota, Y. Arakawa, *Nat. Phys.* **6**, 279–283 (2010).
14. O. Astafiev *et al.*, *Nature* **449**, 588–590 (2007).
15. W. G. van der Wiel *et al.*, *Rev. Mod. Phys.* **75**, 1–22 (2002).
16. Y.-Y. Liu, K. D. Petersson, J. Stehlik, J. M. Taylor, J. R. Petta, *Phys. Rev. Lett.* **113**, 036801 (2014).
17. A. Wallraff *et al.*, *Nature* **431**, 162–167 (2004).
18. T. Frey *et al.*, *Phys. Rev. Lett.* **108**, 046807 (2012).
19. P.-Q. Jin, M. Marthaler, J. H. Cole, A. Shnirman, G. Schön, *Phys. Rev. B* **84**, 035322 (2011).
20. C. Fasth, A. Fuhrer, L. Samuelson, V. N. Golovach, D. Loss, *Phys. Rev. Lett.* **98**, 266801 (2007).
21. S. Nadj-Perge, S. M. Frolov, E. P. A. M. Bakkers, L. P. Kouwenhoven, *Nature* **468**, 1084–1087 (2010).
22. R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha, L. M. K. Vandersypen, *Rev. Mod. Phys.* **79**, 1217–1265 (2007).
23. Supplementary materials are available on *Science* Online.
24. M. R. Delbecq *et al.*, *Phys. Rev. Lett.* **107**, 256804 (2011).
25. K. D. Petersson *et al.*, *Nature* **490**, 380–383 (2012).
26. A. L. Schawlow, C. H. Townes, *Phys. Rev.* **112**, 1940–1949 (1958).
27. P. W. Milonni, J. H. Eberly, *Laser Physics* (Wiley, Hoboken, NJ, 2010).
28. A. Megrant *et al.*, *Appl. Phys. Lett.* **100**, 113510 (2012).
29. I. Protsenko *et al.*, *Phys. Rev. A* **59**, 1667–1682 (1999).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6219/285/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S6
References (*30–39*)

QUANTUM GASES

An Aharonov-Bohm interferometer for determining Bloch band topology

L. Duca,[1,2] T. Li,[1,2] M. Reitter,[1,2] I. Bloch,[1,2] M. Schleier-Smith,[3] U. Schneider[1,2]*

The geometric structure of a single-particle energy band in a solid is fundamental for a wide range of many-body phenomena and is uniquely characterized by the distribution of Berry curvature over the Brillouin zone. We realize an atomic interferometer to measure Berry flux in momentum space, in analogy to an Aharonov-Bohm interferometer that measures magnetic flux in real space. We demonstrate the interferometer for a graphene-type hexagonal optical lattice loaded with bosonic atoms. By detecting the singular π Berry flux localized at each Dirac point, we establish the high momentum resolution of this interferometric technique. Our work forms the basis for a general framework to fully characterize topological band structures.

More than 30 years ago, Berry (1) delineated the effects of the geometric structure of Hilbert space on the adiabatic evolution of quantum mechanical systems. These ideas have found widespread applications in physics (2) and are routinely used to calculate the geometric phase shift acquired by a particle moving along a closed path—a phase shift that is determined only by the geometry of the path and is independent of the time spent en route. Geometric phases provide an elegant description of the celebrated Aharonov-Bohm effect (3), in which a magnetic flux in a confined region of space influences the eigenstates everywhere via the magnetic vector potential. In condensed-matter physics, an analogous Berry flux in momentum space is responsible for various anomalous velocities and Hall responses (4) and lies at the heart of many-body phenomena associated with quantum Hall physics (5) and topological insulators (6). The Berry flux density (Berry curvature) is essential to the characterization of an energy band and determines its topological invariants. However, fully mapping out the geometric structure of an energy band (7–10) remains a major challenge for experiments.

Here, we demonstrate a versatile interferometric technique (9, 11) for mapping the Berry curvature of synthetic materials composed of ultracold atoms in optical lattices. In contrast to typical solid-state experiments, in which geometric effects are either averaged over the Fermi sea or largely constrained to the Fermi surface, the use of a Bose-Einstein condensate (BEC) enables measuring geometric phases along arbitrary closed paths in reciprocal space with high momentum resolution. We exploit this resolution to directly detect the topological properties of an individual Dirac cone (12) in a graphene-type hexagonal lat-

tice (Fig. 1). Concentrated at the Dirac point is a π Berry flux, which is analogous to a magnetic flux generated by an infinitely narrow solenoid (14). Signatures of this localized flux have been observed in graphene through measurements of a half-integer shift in the positions of quantum Hall plateaus (15, 16), the phase of Shubnikov–de Haas oscillations (15, 16), and the polarization

dependence in photoemission spectra (17, 18). A similar π flux also plays a crucial role in the nuclear dynamics of molecules featuring conical intersections of energy surfaces (2). Our direct detection of the singular π flux demonstrates the capability of atom interferometry to detect Berry flux features that are challenging to observe by alternative techniques based on transport measurements (7, 8, 19–21), paving the way to full topological characterization of optical lattice systems (20–27).

The effect of Berry curvature in our interferometer is analogous to the Aharonov-Bohm effect, in which an electron wave packet is split into two parts that encircle a given area in real space (Fig. 1A). Any magnetic flux through the enclosed area gives rise to a measurable phase difference between the two components. For a single Bloch band in the reciprocal space of a lattice system, an analog of the magnetic field is the Berry curvature Ω_n (Eq. 1), which we probe by forming an interferometer on a closed path in reciprocal space (Fig. 1B). The geometric phase acquired along the path can be calculated from the Berry connection A_n, the analog of the magnetic vector potential. For a lattice system with Bloch waves $\psi^n_{\boldsymbol{k}}(\boldsymbol{r}) = e^{i\boldsymbol{k}\boldsymbol{r}} u^n_{\boldsymbol{k}}(\boldsymbol{r})$ with quasimomentum $\boldsymbol{k}$ in the nth band and the cell-periodic part of the wave function $u^n_{\boldsymbol{k}}(\boldsymbol{r})$, the Berry
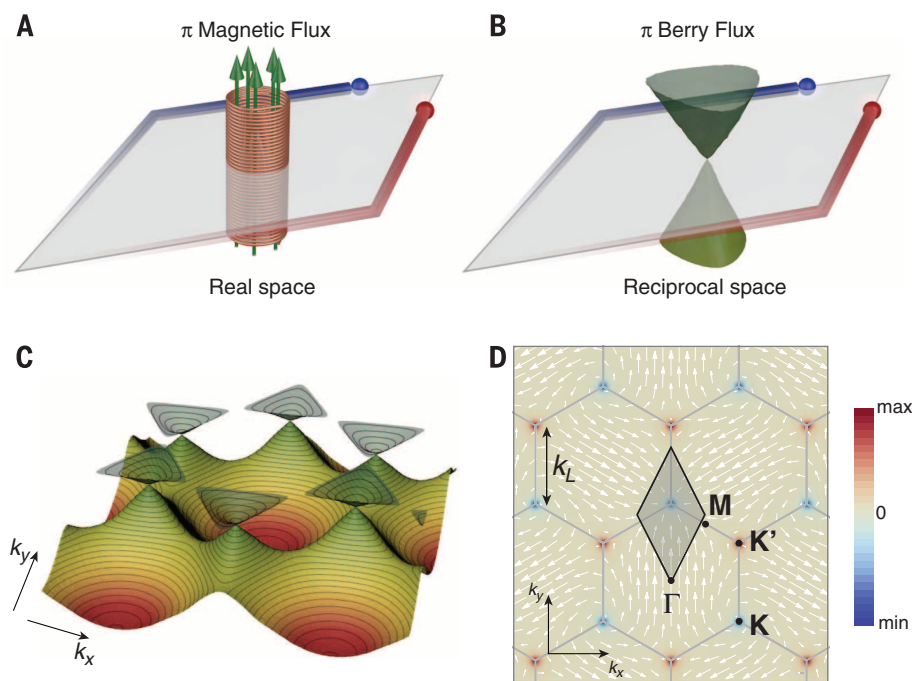
Fig. 1. Aharonov-Bohm analogy and geometric properties of the hexagonal lattice. In the Aharonov-Bohm effect (**A**), electrons encircle a magnetic flux in real space, whereas in our interferometer (**B**), the particles encircle the π Berry flux of a Dirac point in reciprocal space. In both cases, the flux through the interferometer loop gives rise to a measurable phase. (**C**) Dispersion relation of the hexagonal lattice, showing the conical intersection between the first and second band at the Dirac points. (**D**) Berry curvature of the first band calculated in the tight-binding regime with a gap of $\Delta = 0.5\,J$ for visualization purposes. Dirac points are located at the corners (K and K' points) of the BZ (gray hexagons). White arrows are a pseudospin representation of the Bloch states, with orientation indicating the phase of the coupling between sublattices; lengths of the arrows indicate the energy gap in the two-band model. Also shown is a typical interferometer path (black diamond).

[1]Fakultät für Physik, Ludwig-Maximilians-Universität München, Schellingstrasse 4, 80799 Munich, Germany. [2]Max-Planck-Institut für Quantenoptik, Hans-Kopfermann-Strasse 1, 85748 Garching, Germany. [3]Department of Physics, Stanford University, Stanford, CA 94305, USA.
*Corresponding author. E-mail: ulrich.schneider@lmu.de

connection is given by $\boldsymbol{A}_n(\boldsymbol{k}) = i\langle u_{\boldsymbol{k}}^n | \nabla_{\boldsymbol{k}} | u_{\boldsymbol{k}}^n \rangle$. Accordingly, the phase along a closed loop in reciprocal space is (1, 2)

$$\varphi_{Berry} = \oint_C \boldsymbol{A}_n(\boldsymbol{k}) d\boldsymbol{k} = \int_S \Omega_n(\boldsymbol{k}) d^2\boldsymbol{k} \quad (1)$$

where S is the area enclosed by the path $C = \partial_S$, and $\Omega_n = \nabla_{\boldsymbol{k}} \times \boldsymbol{A}_n(\boldsymbol{k})$ is the Berry curvature (color shading in Fig. 1D) (4). Although neither the magnetic vector potential nor the Berry connection is uniquely defined, the geometric phase acquired along a closed loop is gauge independent (1) and is therefore a measurable observable that encodes information on the geometrical properties of a Bloch band.

We implemented the graphene-like hexagonal optical lattice for ultracold ^{87}Rb atoms by superimposing three linearly polarized blue-detuned running waves at 120(1)° angles (Fig. 2A). The resulting dispersion relation includes two nonequivalent Dirac points with opposite Berry flux located at **K** and **K′**, which are repeated in every Brillouin zone (BZ) (Fig. 1D). The origin of the π Berry flux lies in the bipartite structure of the hexagonal lattice (12): Because the unit cell contains two nonequivalent lattice sites A and B (Fig. 2A), the Bloch wave of the lowest band has the form of a two-component spinor. This spinor is constrained to the equatorial plane of its Bloch sphere by a combination of time-reversal invar-

iance and the inversion symmetry of the lattice, and its phase winds around each Dirac point as shown in Fig. 1D. The spinor Bloch wave, just as a real spin-1/2 particle in a slowly rotating magnetic field, therefore acquires a geometric phase of π along any trajectory enclosing a single Dirac point. The Berry curvature is thus confined to a perfectly localized π Berry flux, $\Omega_n = \pm\pi\delta(\boldsymbol{k} - \boldsymbol{K}^{(\prime)})$, provided the aforementioned symmetries hold (13). Generically, the inversion symmetry may be broken by a slight ellipticity of the lattice beam polarizations, which introduces a small energy offset Δ between the A and B sites (6). Such an offset opens a small gap at the Dirac points and spreads the Berry curvature over a finite range
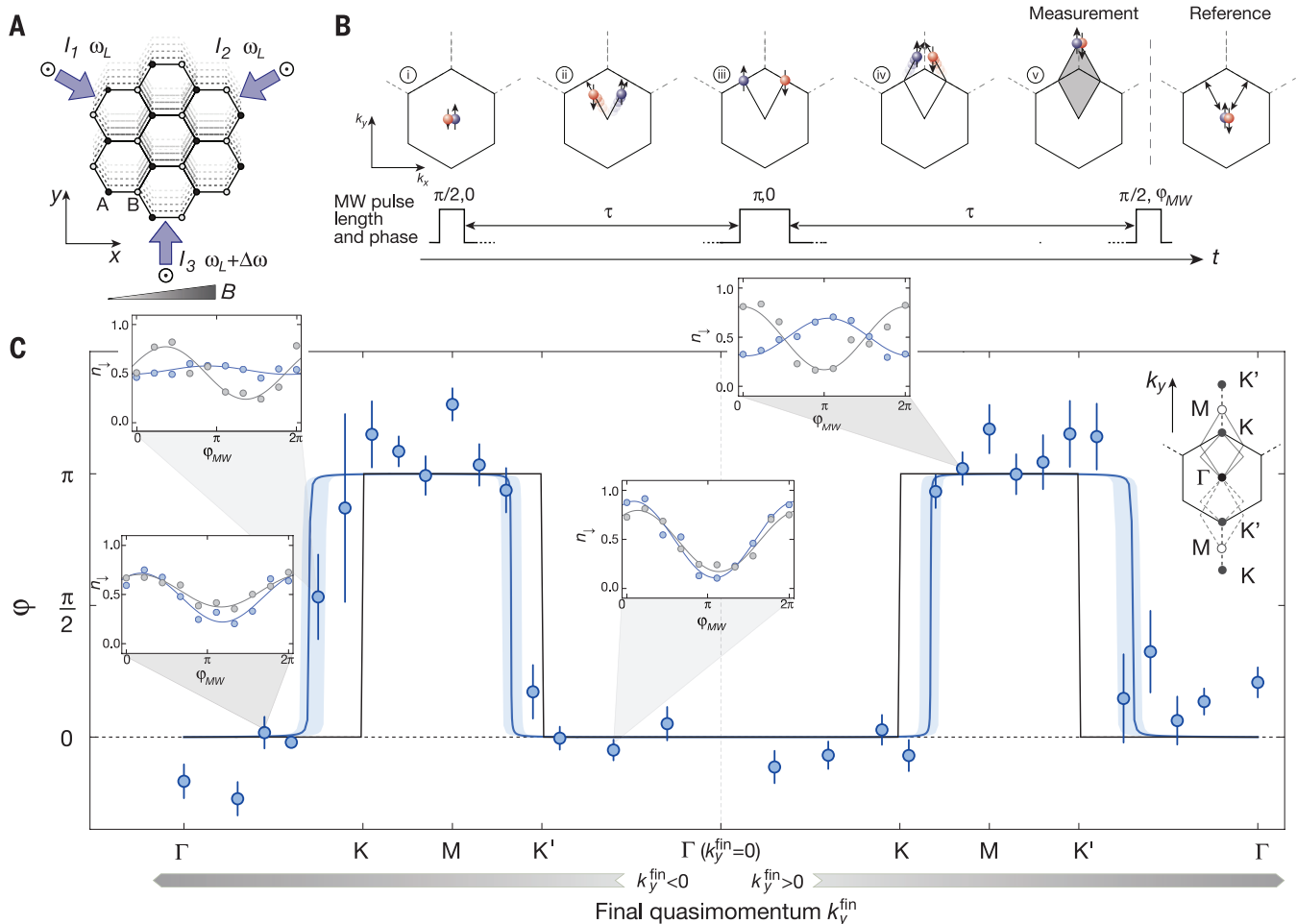
Fig. 2. Momentum-resolved detection of Berry flux at the Dirac points.
(**A**) Sketch of the hexagonal lattice in real space with A (B) sites denoted by solid (open) circles. The lattice is realized by interfering three laser beams (arrows) of wavelength λ_L, intensity I_i, and frequency ω_L, with linear out-of-plane polarizations. A linear frequency sweep of the third lattice beam creates a uniform lattice acceleration along the y direction. A magnetic field gradient $B' = 9.0(1)$ G/cm along the x axis creates an additional spin-dependent force. (**B**) Interferometer sequence. Hexagons indicate the first BZ, and red (blue) spheres are atoms in the $|\downarrow\rangle$ ($|\uparrow\rangle$) state. The duration of the interferometer sequence is $2\tau = 1.6$ ms for all measurements. (**C**) Summary of phase shifts measured relative to the zero-area reference interferometer for different final quasimomenta k_y^{fin}. Error bars denote fit uncertainties or standard deviations

in case of averages. Lines are ab initio theory using a full band structure calculation with no momentum spread $\sigma_k = 0$ and perfectly localized Berry curvature $\delta k_\Omega = 0$ (black) or $\sigma_k = 0.21$ k_L and $\delta k_\Omega \simeq 10^{-4}$ k_L (blue). Here, the shift in the phase jump results from the momentum spread σ_k (13), the broadening of the edges is caused by δk_Ω, and the shaded area accounts for a variation in $\sigma_k = 0.14 - 0.28k_L$. Insets show the fraction of atoms $n_\downarrow$ detected in state $|\downarrow\rangle$ as a function of the phase φ_{MW} for selected quasimomenta. Measurement loop data are shown in blue, and zero-area reference data are shown in gray, with corresponding sinusoidal fits. The contrast is limited by inhomogeneous broadening of the microwave transition, the finite momentum spread of the condensate, and, for large final quasimomenta, the dynamical instability of the Gross-Pitaevskii equation (32).

of quasimomenta (Fig. 1D). By probing for a spread in Berry curvature, we can place a bound on imperfections in the lattice, while simultaneously benchmarking the resolution of our interferometer.

The interferometer sequence (Fig. 2B) begins with the preparation of an almost pure ^{87}Rb BEC in the state $|\uparrow\rangle = |F = 2, m_F = 1\rangle$ at quasimomentum $\mathbf{k}$ = 0 in a V_0 = 1 E_r deep lattice, where $E_r = h^2/(2m\lambda_L^2) \approx h \times 4$ kHz is the recoil energy and h is Planck's constant. A resonant $\pi/2$-microwave pulse creates a coherent superposition of $|\uparrow\rangle$ and $|\downarrow\rangle = |F = 1, m_F = 1\rangle$ states (i). Next, a spin-dependent force from a magnetic field gradient and an orthogonal spin-independent force from lattice acceleration (Fig. 2A) move the atoms adiabatically along spin-dependent paths in reciprocal space (ii) (28). The two spin components move symmetrically about a symmetry axis of the dispersion relation. After an evolution time τ, a microwave π pulse swaps the states $|\downarrow\rangle$ and $|\uparrow\rangle$ (iii). The two atomic wave packets now experience opposite magnetic forces in the x direction, such that both spin components arrive at the same quasimomentum k^{fin} after an additional evolution time τ (iv). At this point, the state of the atoms is given by $|\psi^{fin}\rangle \propto |\uparrow, k^{fin}\rangle + e^{i\varphi}, k^{fin}\rangle$ with relative phase φ. A second $\pi/2$-microwave pulse with a variable phase φ_{MW} closes the interferometer (v) and converts the phase information into spin population fractions $n_{\uparrow,\downarrow} \propto 1 \pm \cos(\varphi + \varphi_{MW})$, which are measured by standard absorption imaging after a Stern-Gerlach pulse and time of flight.

The phase difference φ at the end of the interferometer sequence consists of the geometric phase and any difference in dynamical phases between the two paths of the interferometer. Ideally, the dynamical contribution should vanish because of the symmetry of the paths and the use of the spin-echo sequence (13). To ascertain that the measured phase is truly of geometric origin, we additionally employ a "zero-area" reference interferometer, comprising a V-shaped path (Fig. 2B) produced by reversing the lattice acceleration after the π-microwave pulse of Fig. 2B (iii).

We locate the Berry flux of the Dirac cone by performing a sequence of measurements in which we vary the region enclosed by the interferometer. This is achieved by varying the lattice acceleration at constant magnetic field gradient to control the final quasimomentum k_y^{fin} (k_x^{fin} = 0) of the diamond-shaped measurement loop. The resulting phase differences between measurement and reference loops are shown in Fig. 2C. When one Dirac point is enclosed in the measurement loop, we observe a phase difference of $\varphi \simeq \pi$. In contrast, we find the phase difference to vanish when enclosing zero or two Dirac points. We find very good agreement between our data and a theoretical model that includes the finite spread σ_k in the initial momentum of the weakly interacting BEC (blue curve in Fig. 2C) (13). Because of this spread, each atom has sampled a slightly different path in momentum space and may therefore have acquired a different geometric phase. Once the
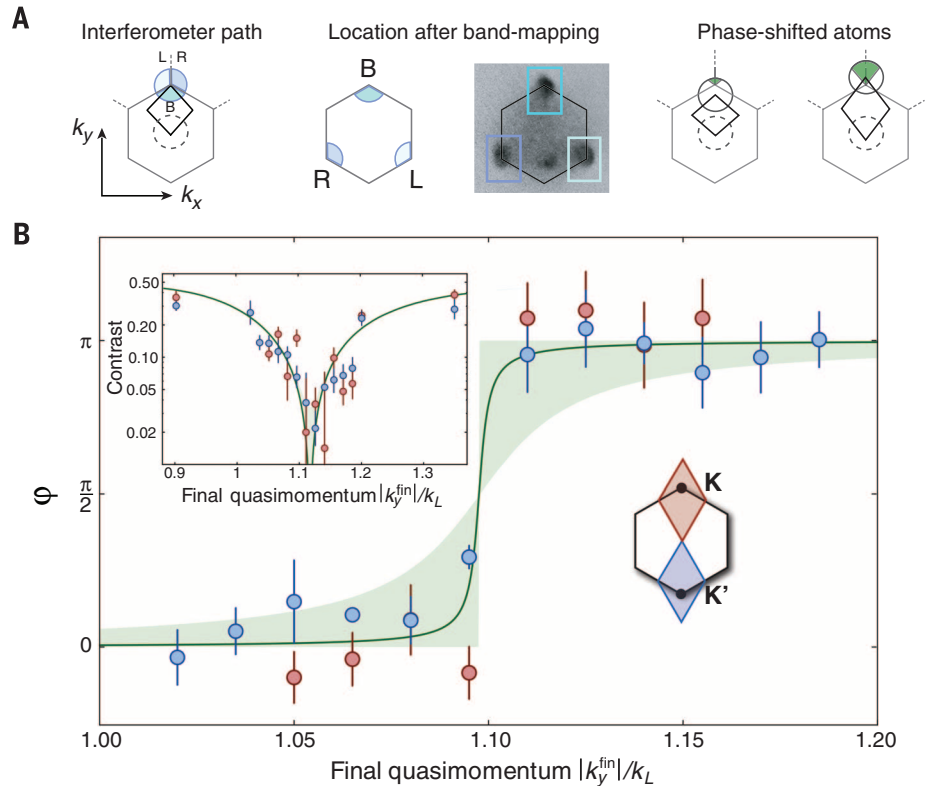
Fig. 3. Self-referenced interferometry at the Dirac point. (**A**) (Left) Interferometer path closing at the K point. Because of the initial momentum spread, the cloud (circle with colored sectors, not to scale) is split by the edges of the BZ. (Middle) Band mapping spatially separates the three different parts of the cloud onto three corners of the first BZ (schematic and image, where cloud sizes are dominated by in situ size). (Right) The fraction of atoms for which the Dirac point lies within the interferometer loop (green sectors) increases with final quasimomentum k^{fin}. (**B**) Phase differences between atoms that have crossed the band edge (sectors L and R) and those that have not (sector B) versus final quasimomentum k_y^{fin} for paths close to the K (K') point in red (blue). The shaded region indicates a range $\delta k_\Omega = 0 - 12 \times 10^{-4} k_L$ for the spread in Berry curvature, whereas the line is calculated for $\delta k_\Omega \simeq 10^{-4} k_L$ using the model described in (13), corresponding to an A-B offset of $\Delta \simeq h \times 3$ Hz. The inset shows the contrast $(n_\downarrow^{max} - n_\downarrow^{min})/(n_\downarrow^{max} + n_\downarrow^{min})$ of the interference fringes of the full cloud. Theory line and shading are for the same parameters as in the main graph and include only geometrical phases (13). All calculations assume $\sigma_k = 0.15 k_L$.

Dirac point lies within the interferometer area for exactly half of the atoms, the first phase jump occurs. Because of the small opening angle of the chosen interferometer path (~70°), this happens slightly later than in the ideal case of $\sigma_k = 0$ (black curve in Fig. 2C). Although σ_k thereby affects the positions of the π phase jumps, it does not limit their sharpness. Indeed, the data are fully consistent with the behavior expected for an inversion-symmetric lattice, where it is impossible to identify the sign of the singular Berry flux ($\pm\pi$). Small deviations of the phases from 0 or π can be attributed to an imperfect alignment of the magnetic field gradient, magnetic field fluctuations, or an imperfect lattice geometry (13). These systematic effects are particularly relevant close to the phase jump, where the contrast is minimal and can influence the perceived direction of the phase jump.

To minimize systematic errors and improve our measurement precision, we performed self-referenced interferometry close to the Dirac points. As illustrated in Fig. 3A, a standard band-mapping technique (29) projects those sectors of the cloud that have (left and right) or have not (bottom) crossed the edge of the BZ onto three different corners of the first BZ, such that we can measure their acquired phases independently. Combining these measured phases to $\varphi = (\varphi_L + \varphi_R)/2 - \varphi_B$, where φ_L, φ_R, and φ_B refer to the phases of the three sectors, eliminates the need for a separate reference measurement and significantly reduces sensitivity to drifts in the experiment. The resulting phase again shows a sudden jump from 0 to π (Fig. 3B). The position of the phase jump is in excellent agreement with a simple single-band model (13) that includes an initial momentum spread of $\sigma_k = 0.15(1) k_L$, consistent with an independent time-of-flight measurement. Notably, the phase jump occurs within a very small quasimomentum range of $<0.01\ k_L$, and an arctangent fit to the
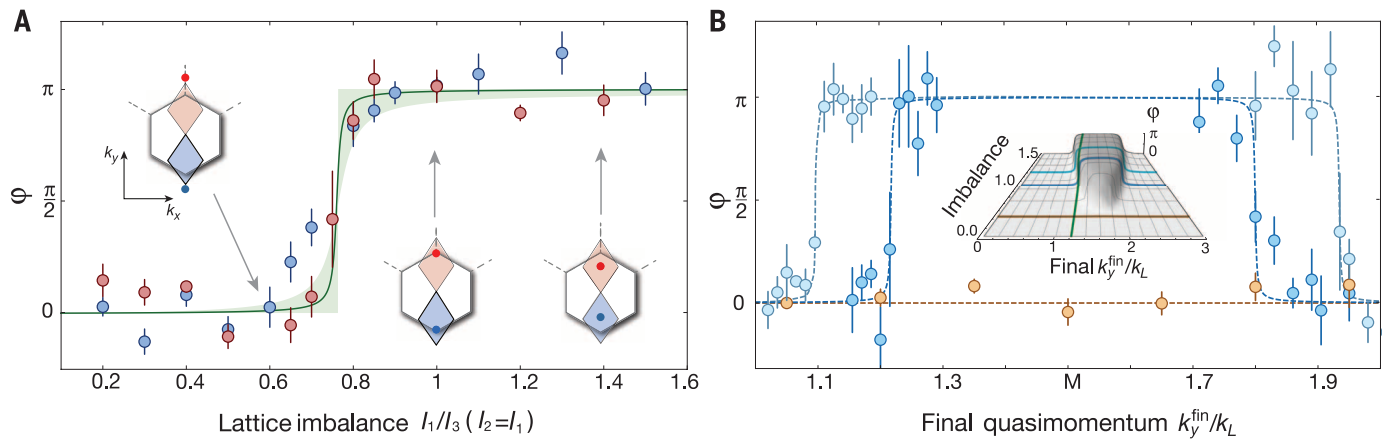
Fig. 4. Mapping the movement of Berry flux under distortion of the lattice. (**A**) Phase difference between the zero-area reference loop and measurement loop versus lattice imbalance for a fixed final quasimomentum $k_y^{fin} = \pm 1.2 k_L$ near **K** (red) or **K'** (blue). Red and blue dots in the insets give the location of Dirac points for the indicated imbalances. Theory curve is calculated for lattice depth $V_0 = 1E_r$, momentum spread $\sigma_k = 0.15 k_L$, and $\delta k_\Omega \simeq 10^{-4} k_L$. Shaded area corresponds to $\delta k_\Omega = 0 - 12 \times 10^{-4} k_L$. (**B**) Self-referenced phase near **K** and **K'** for an imbalance $I_{1,2}/I_3 = 1.0$ and $I_{1,2}/I_3 = 0.7$ in light and dark blue, highlighting the shift in the location of Berry flux. Phases are measured as in Fig. 3. Orange data are phase differences between the measurement and reference loops for an imbalance of $I_{1,2}/I_3 = 0.2$, where no phase shift is observed. Curves are guides to the eye. The inset shows the calculated Berry phase for loops with various final k_y^{fin} and lattice imbalances using the same σ_k and δk_Ω as above. Colored lines indicate parameters explored in the measurements.

experimental data gives a phase step of $\varphi = 0.95(10)\pi$. Both results are compatible with a perfectly localized and quantized π Berry flux.

To constrain the possible spread in Berry curvature, we analyze not only the phase (Fig. 3B) but also the contrast of the interference fringes (Fig. 3B, inset). The location of the Dirac cone manifests itself through a pronounced minimum in the interference contrast. The sharpness of the phase jump and the strong reduction of contrast down to our detection limit demonstrate that the interferometric protocol can map the Berry curvature with very high resolution. By comparing the experimental contrast with our theoretical model (13), we find an upper bound for the spread of the Berry curvature around the Dirac cone of $\delta k_\Omega \leq 6 \times 10^{-4} k_L$ (half-width at half maximum). This corresponds to a maximal A-B site offset of $\Delta \leq h \times 12$ Hz and a ratio of energy gap at the Dirac cone to bandwidth of $\leq 1 \times 10^{-3}$. The steepness of the phase jump in Fig. 3B suggests an even stronger localization of the Berry curvature on the order of $\delta k_W \simeq 10^{-4} k_L$ ($\Delta \simeq h \times 3$ Hz). Although the vanishing band gap precludes performing a perfectly adiabatic measurement in the immediate proximity of the Dirac point, the population in the second band is constrained by independent measurements to be $\leq 20\%$ of the total atom number (13).

To verify the method's sensitivity to changes in Berry flux, we performed interferometry in a modified lattice potential. Changing the power of two lattice beams (I_1 and I_2) relative to the third (I_3) deforms the lattice structure but preserves time-reversal and inversion symmetry. With decreasing $I_{1,2}/I_3 < 1$, the Dirac points and the associated fluxes move along the symmetry axis of the interferometer loop (12) (insets of Fig. 4A). Nonetheless, the Berry flux singularities remain protected by symmetry until the Dirac points merge and annihilate (23, 30, 31). By using a fixed

measurement loop that encloses one Dirac point in the intensity-balanced case, we can measure the change of the geometric phase as we imbalance the lattice beam intensities. The measured Berry phases drop from π to 0 as the Dirac point moves out of the loop, in very good agreement with ab initio calculations (see Fig. 4A). To map the location of the Berry flux in the imbalanced lattice, we again use the self-referenced interferometry of Fig. 3. As shown in Fig. 4B, imbalancing the lattice by decreasing $I_{1,2}/I_3$ narrows the range of final quasimomenta for which the interferometer encloses a single π flux, thereby shifting both the upward and downward phase jumps toward the M point. For strong imbalance ($I_{1,2}/I_3 = 0.2$), the two Dirac points have annihilated, and hence no phase jump is observed for any loop size. At intermediate imbalance ($I_{1,2}/I_3 = 0.7$), the position of the phase jump at $k_y^{fin} = 1.2 k_L$ is in very good agreement with theory, whereas deviations of $\simeq 10\%$ from the calculated value in the position of the second phase jump can likely be attributed to a combination of geometric imperfections, nonadiabaticity of motion at the Dirac point, and the dynamical instability of the Gross-Pitaevskii equation (32), which results in an additional broadening of the quasimomentum distribution. The latter effects can be suppressed by combining slower ramps with the use of a Feshbach resonance in an atomic species such as ^{39}K (33).

Our Aharonov-Bohm–type interferometer enabled us to detect a localization of Berry flux to better than 10^{-6} of the Brillouin zone area. This method allows one to fully resolve the Berry curvature distribution of a single Bloch band by combining local measurements of the geometric phases along small paths, thereby enabling the full reconstruction of topological invariants such as Chern numbers. The method can readily be applied to a variety of optical lattices and other

physical settings such as polariton condensates (34). Multiband extensions of this work can enable measurements of Wilson loops and off-diagonal (non-Abelian) Berry connections and thus provide a framework for determining the complete geometric tensor of Bloch bands in periodic structures (35). Controlled application of non-Abelian Berry phases would furthermore constitute a key step toward holonomic quantum computation (36). Even within a single topologically trivial band, the possibility of preparing a BEC or Fermi sea at finite quasimomentum should enable the observation of transient Hall responses due to local Berry curvature and, combined with the possibility of performing quantum quenches and the control of interactions, is expected to lead to novel many-body phenomena (37). Finally, the highly nonlinear phase jump we have observed at the Dirac point may find application in precision force sensing (38).

REFERENCES AND NOTES

1. M. V. Berry, *Proc. R. Soc. London Ser. A* **392**, 45–57 (1984).
2. A. Shapere, F. Wilczek, *Advanced Series in Mathematical Physics: Volume 5, Geometric Phases in Physics* (World Scientific, Singapore, 1989).
3. Y. Aharonov, D. Bohm, *Phys. Rev.* **115**, 485–491 (1959).
4. D. Xiao, M.-C. Chang, Q. Niu, *Rev. Mod. Phys.* **82**, 1959–2007 (2010).
5. D. J. Thouless, M. Kohmoto, M. P. Nightingale, M. den Nijs, *Phys. Rev. Lett.* **49**, 405–408 (1982).
6. M. Z. Hasan, C. L. Kane, *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
7. H. Price, N. Cooper, *Phys. Rev. A* **85**, 033620 (2012).
8. A. Dauphin, N. Goldman, *Phys. Rev. Lett.* **111**, 135302 (2013).
9. D. A. Abanin, T. Kitagawa, I. Bloch, E. Demler, *Phys. Rev. Lett.* **110**, 165304 (2013).
10. P. Hauke, M. Lewenstein, A. Eckardt, *Phys. Rev. Lett.* **113**, 045303 (2014).
11. M. Atala *et al.*, *Nat. Phys.* **9**, 795–800 (2013).
12. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109–162 (2009).
13. See the supplementary materials on *Science* Online.
14. G. P. Mikitik, Y. V. Sharlai, *Phys. Rev. Lett.* **82**, 2147–2150 (1999).

15. Y. Zhang, Y.-W. Tan, H. L. Stormer, P. Kim, *Nature* **438**, 201–204 (2005).
16. K. S. Novoselov *et al.*, *Nature* **438**, 197–200 (2005).
17. Y. Liu, G. Bian, T. Miller, T.-C. Chiang, *Phys. Rev. Lett.* **107**, 166803 (2011).
18. C. Hwang *et al.*, *Phys. Rev. B* **84**, 125422 (2011).
19. M. Aidelsburger, *et al.*, *Nat. Phys.* **2014**, 10.1038/nphys3171 (2014).
20. G. Jotzu *et al.*, *Nature* **515**, 237–240 (2014).
21. I. B. Spielman, *Annalen der Physik* **525**, 797–807 (2013).
22. M. Aidelsburger *et al.*, *Phys. Rev. Lett.* **107**, 255301 (2011).
23. L. Tarruell, D. Greif, T. Uehlinger, G. Jotzu, T. Esslinger, *Nature* **483**, 302–305 (2012).
24. J. Struck *et al.*, *Nat. Phys.* **9**, 738–743 (2013).
25. M. Aidelsburger *et al.*, *Phys. Rev. Lett.* **111**, 185301 (2013).
26. H. Miyake, G. A. Siviloglou, C. J. Kennedy, W. C. Burton, W. Ketterle, *Phys. Rev. Lett.* **111**, 185302 (2013).
27. N. Goldman, G. Juzeliunas, P. Ohberg, I. B. Spielman, *Rep. Prog. Phys.* **77**, 126401 (2014).
28. M. Ben Dahan, E. Peik, J. Reichel, Y. Castin, C. Salomon, *Phys. Rev. Lett.* **76**, 4508–4511 (1996).
29. M. Greiner, I. Bloch, O. Mandel, T. W. Hänsch, T. Esslinger, *Phys. Rev. Lett.* **87**, 160405 (2001).
30. S.-L. Zhu, B. Wang, L.-M. Duan, *Phys. Rev. Lett.* **98**, 260402 (2007).
31. P. Dietl, F. Piéchon, G. Montambaux, *Phys. Rev. Lett.* **100**, 236405 (2008).
32. L. Fallani *et al.*, *Phys. Rev. Lett.* **93**, 140406 (2004).
33. C. D'Errico *et al.*, *New J. Phys.* **9**, 223 (2007).
34. I. Carusotto, C. Ciuti, *Rev. Mod. Phys.* **85**, 299–366 (2013).
35. F. Grusdt, D. Abanin, E. Demler, *Phys. Rev. A* **89**, 043621 (2014).
36. P. Zanardi, M. Rasetti, *Phys. Lett. A* **264**, 94–99 (1999).
37. S. A. Parameswaran, R. Roy, S. L. Sondhi, *C. R. Phys.* **14**, 816–839 (2013).
38. A. K. Tuchman, M. A. Kasevich, *Phys. Rev. Lett.* **103**, 130403 (2009).

INTERFACIAL SOLVENTS

Universal solvent restructuring induced by colloidal nanoparticles

Mirijam Zobel,[1]* Reinhard B. Neder,[1] Simon A. J. Kimber[2]*

Colloidal nanoparticles, used for applications from catalysis and energy applications to cosmetics, are typically embedded in matrixes or dispersed in solutions. The entire particle surface, which is where reactions are expected to occur, is thus exposed. Here, we show with x-ray pair distribution function analysis that polar and nonpolar solvents universally restructure around nanoparticles. Layers of enhanced order exist with a thickness influenced by the molecule size and up to 2 nanometers beyond the nanoparticle surface. These results show that the enhanced reactivity of solvated nanoparticles includes a contribution from a solvation shell of the size of the particle itself.

Bulk liquids have long been known to show short-range order. The original method of choice was x-ray scattering, which was used by Zachariasen in 1935 to study the short-range order between solvent molecules (*1*). Alcohol molecules were shown to form a hydrogen-bonded network within the bulk solvent (*2–4*), and short alkanes were found to align in parallel within domains of ~2 nm (*5*). More recently, the influence of hard planar walls on bulk liquids has been investigated (*6–8*). Together with force measurements (*7, 8*), x-ray scattering confirmed that an exponentially decaying oscillatory density profile is established near the interface (*9*). With the advent of synchrotron radiation sources, Magnussen *et al.* showed by use of x-ray reflectivity that even liquid mercury orders at a flat solid interface in exactly the same way (*9*). More recent examples of such ordering phenomena at interfaces include the restructuring of nonpolar n-hexane (*10*), the assembly of fluorinated ionic liquids at sapphire surfaces (*11*),

and the exponentially decaying surface segregation profiles in Cu_3Au alloy interfaces (*12*). These restructuring phenomena—in particular, the interlayer spacings and decay lengths—are closely related to the local ordering in the bulk liquid (*9, 12*).

Although the reorganization of solvent molecules around isolated cations in solution has also been explored (*13, 14*), comparable studies on solvated nanoparticles (NPs) are rare, in par-

ticular for nonaqueous solvent (*15–17*). Solvent molecules are expected and have theoretically been modeled (*15, 16*) to rearrange at the liquid-NP interface, although no definitive experimental proof exists so far. For bulk planar surfaces (*6–8*) and ions (*13, 14*), such enhanced order is well understood. Here, we report a synchrotron x-ray scattering study of a variety of as-synthesized and commercial NPs in polar and nonpolar solvents. We show enhanced ordering of solvent molecules at the NP surface that extends several layers into the bulk liquid. This effect is largely independent of the capping agent, solvent polarity, and particle size.

We systematically redispersed different types of metal and metal oxide NPs in polar and nonpolar organic solvents (*18*). We studied the influence of the solvent molecule size within the series of primary alcohols (methanol, ethanol, and 1-propanol) and the effect on nonpolar solvents with hexane. Fourier transformation of high-energy x-ray scattering patterns yielded the pair distribution functions (PDFs), histograms of all interatomic distances within a sample. Data treatment includes the subtraction of diffraction data of the respective pure solvent so that the signal from the bulk solvent is subtracted before the Fourier transformation. The

[1]Department of Physics, Lehrstuhl für Kristallographie und Strukturphysik, Friedrich-Alexander University Erlangen-Nürnberg, Staudtstrasse 3, 91058 Erlangen, Germany. [2]European Synchrotron Radiation Facility, 71 Avenue des Martyrs, 38000 Grenoble, France.
*Corresponding author. E-mail: mirijam.zobel@fau.de (M.Z.); kimber@esrf.fr (S.A.J.K.)

Fig. 1. Fit to the PDF of redispersed ZnO NPs with citrate ligands in propanol. Experimental d-PDF of ZnO NPs (black) and their fit (red), showing the overall difference of the fit (gray), the contribution of the NP (green), and the contribution and fit of the restructured solvent (blue) in offset for means of clarity. The contribution of the restructured solvent (blue) is the dd-PDF of the experimental, background-corrected d-PDF (black) and the NP (green).

Fig. 2. Comparison of short-range order in restructured and bulk solvents. (**A**) PDF of pure bulk propanol (black) and dd-PDFs of restructured propanol around ZnO NPs with ligand acetate (red), citrate (blue), dmlt (green), and pent (yellow). The amplitudes of the dd-PDFs are scaled with respect to the contribution of ZnO NPs to the data. (**B**) Maxima positions of dd-PDFs of the redispersed ZnO NPs in all solvents plotted over the chain length of the solvent molecules. The positions follow the trend of the pure solvents (black stars), yet the dd-PDF positions are distinctly offset with respect to the error bars that resulted from least-squares fits.

resulting difference PDF (d-PDF) should only contain peaks from intraparticular distances within the NPs and hence be identical to the PDF of a NP powder sample. We observed, however, a distinct, exponentially damped sinusoidal oscillation, which vanished in dry samples (Fig. 1; further examples are shown in figs. S7 and S8) (*18*). A least-squares fit (red) of the linear combination of the respective NP powder (green) and an exponentially decaying sinusoidal oscillation (blue) describes all data sets (black) adequately (*18*) and gives access to four parameters that describe the restructuring: (i) the amplitude, (ii) the position of the first maximum of the oscillation, (iii) the wavelength, and (iv) the modulation depth [calculations are provided in (*18*)]. The amplitude contains information about how many solvent molecules restructure, and the modulation depth describes how far this enhanced ordering extents into the bulk liquid. The position and wavelength describe the arrangement and layer spacing of the molecules.

The PDF of bulk propanol is shown in Fig. 2A together with the double-difference PDFs (dd-PDFs) of the restructured solvent and their fits for four different ZnO NPs in propanol. Here, the dd-PDF is the difference between the d-PDF of the redispersed sample and the PDF of the NP powder. The resulting dd-PDFs differ from the PDF of the pure solvent in that they exhibit a distinctively larger modulation depth. In contrast to bulk propanol, the dd-PDFs do not show sharp intramolecular distances of the propanol molecules in the range 0 to 3 Å because they are entirely corrected for in the background subtraction. The restructuring does not change the internal molecular structure but only changes the relative orientation between different solvent molecules. Thus, the enhanced short-range order around the NPs becomes visible as the oscillation in the dd-PDF. We observed the same kind of re-
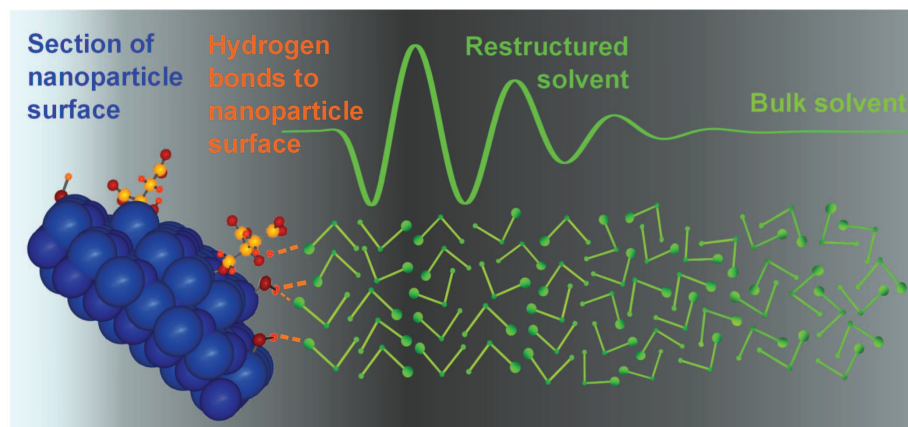
Fig. 3. Enhanced short-range order of solvent molecules at ZnO NP surfaces. The ethanol molecules (hydrogen atoms omitted) form hydrogen bonds with surface hydroxyl groups and citrate molecules. The surface coverage of these groups is reduced for means of clarity. The enhanced short-range order extends a few molecular layers into the bulk liquid before bulk properties are recovered. In bulk ethanol, molecules do not arrange in pairs, but form winding-chains or hexamers (*2–4*), and the enhanced short-range order around the NPs is not as crystalline as suggested in the scheme, which serves as simplified illustrative representation.

structuring for a wide range of metal oxide and metal NPs (figs. S5 and S6) (*18*).

In essence, the PDF represents the distribution of all interatomic distances *r*. Because we work in a highly diluted system, we can expect that the structure of the bulk solvent does not change. Likewise, the contribution of the NPs to the PDF does not differ from that of the PDF of the pure NP. Because all changes to the experimental dd-PDF occur upon redispersing the NP in the pure solvent, these changes must be induced by the NP surfaces (*18*). Thus, our dd-PDFs represent a change of the short-range order within the solvent near the NP surface. With increasing distance, the signal in the dd-PDF ex-

ponentially decays. The orientational averaging intrinsic to the PDF technique prevents us from determining the direction of rearrangement; however, the signal bears a striking similarity to those seen by using specular x-ray surface scattering, which probes only the direction normal to the interface.

A comparison of the maximum position, wavelength, and modulation depth of the oscillation for our ZnO NPs is shown in Fig. 2B and fig. S5. A linear trend in the position of the maximum amplitude is found with alcohol chain length. Furthermore, this trend occurs for four different capping agents, showing that surface composition plays a minor role. A larger spread in these

parameters is found for other types of NPs (Ag, ZrO_2, TiO_2, and In_2O_3) (fig. S6). These results show that to first approximation, and in agreement with MD simulations (15), the microscopic details of the NP surface only weakly influence the solvent restructuring. However, it is possible that the spread of data points in Fig. 2B, which increases for increasing alkyl chain length, indicates an additional role for particle shape, surface restructuring, or differences in capping agent. Our model of restructuring for ethanol at the surface of a ZnO NP decorated with citrate and hydroxyl groups of capping ligands is illustrated in Fig. 3. The surface coverage of NPs with organic ligand molecules is sufficient to prevent agglomeration but in fact is unexpectedly small according to neutron PDF data (19) and nuclear magnetic resonance (NMR) studies (20). MD simulations (15, 21, 22), in agreement with experimental evidence from NMR (20), suggest that the vast majority of ZnO surface sites are terminated by hydroxyl groups. These could form a hydrogen-bonded network with adjacent solvent molecules.

Because of the shift in the oscillation with the solvent size, we conclude that the alcohol molecules tend to align perpendicular to the NP surface. Their hydroxyl groups form hydrogen bonds with the ligand molecules and hydroxyl groups. The alkyl chains of the solvent point away from the NP surface. The next- and second-next–neighboring molecules align so that hydrogen bonds can be formed within the solvent, which results in alternating layers of methyl groups and hydroxyl groups, building layers of decreased and enhanced electron density, as depicted in Fig. 3. Adjacent molecules within such a layer would orient in parallel, as observed for liquid films (23). The extent of restructuring depends on the solvent size and packing ability. Here, the packing ability is comparable for our alcohols because the hydroxyl group is always in a terminal position and the alkane chain is not branched. A second-harmonic generation (SHG) study on the interaction of organic solvent and solute molecules with hydroxylated silica surfaces supports our hydrogen bonding model (24). This study also showed that nonpolar solvents rearrange at hydroxylated surfaces, which supports our observation that the nonpolar n-hexane restructures at the NP surfaces. However, SHG is only sensitive to broken symmetry at interfaces, whereas it cannot provide information on the decay of the restructuring into the bulk liquid, as evidenced by our dd-PDFs.

REFERENCES AND NOTES

1. W. H. Zachariasen, *J. Chem. Phys.* **3**, 158–161 (1935).
2. D. L. Wertz, R. K. Kruh, *J. Chem. Phys.* **47**, 388–390 (1967).
3. C. J. Benmore, Y. L. Loh, *J. Chem. Phys.* **112**, 5877–5883 (2000).
4. A. Vrhovšek, O. Gereben, A. Jamnik, L. Pusztai, *J. Phys. Chem. B* **115**, 13473–13488 (2011).
5. G. W. Longman, G. D. Wignall, R. P. Sheldon, *Polymer (Guildf.)* **20**, 1063–1069 (1979).
6. B. Götzelmann, A. Haase, S. Dietrich, *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **53**, 3456–3467 (1996).
7. J. Israelachvili, *Acc. Chem. Res.* **20**, 415–421 (1987).
8. I. K. Snook, W. van Megen, *J. Chem. Phys.* **72**, 2907 (1980).
9. O. M. Magnussen *et al.*, *Phys. Rev. Lett.* **74**, 4444–4447 (1995).
10. D. Doerr, M. Tolan, T. Seydel, W. Press, *Physica B* **248**, 263–268 (1998).
11. M. Mezger *et al.*, *Science* **322**, 424–428 (2008).
12. H. Reichert, P. J. Eng, H. Dosch, I. K. Robinson, *Phys. Rev. Lett.* **74**, 2006–2009 (1995).
13. D. T. Bowron, S. Díaz-Moreno, *J. Phys. Chem. B* **113**, 11858–11864 (2009).
14. R. Mancinelli, A. Botti, F. Bruni, M. A. Ricci, A. K. Soper, *Phys. Chem. Chem. Phys.* **9**, 2959–2967 (2007).
15. D. Spagnoli, J. P. Allen, S. C. Parker, *Langmuir* **27**, 1821–1829 (2011).
16. B. H. G. Hertz, I. An, *Angew. Chem. Int. Ed. Engl.* **9**, 124–138 (1970).
17. R. S. Cataliotti, F. Aliotta, R. Ponterio, *Phys. Chem. Chem. Phys.* **11**, 11258–11263 (2009).
18. Materials and methods are available as supplementary materials on Science Online.
19. K. Page, T. C. Hood, T. Proffen, R. B. Neder, *J. Appl. Cryst.* **44**, 327–336 (2011).
20. C. N. Valdez, A. M. Schimpf, D. R. Gamelin, J. M. Mayer, *ACS Nano* **8**, 9463–9470 (2014).
21. A. Kawska, P. Duchstein, O. Hochrein, D. Zahn, *Nano Lett.* **8**, 2336–2340 (2008).
22. D. Raymand, A. C. T. Duin, W. A. Goddard III, K. Hermansson, D. Spångberg, *J. Phys. Chem. C* **115**, 8573–8579 (2011).
23. H. K. Christenson, D. W. R. Gruen, R. G. Horn, J. N. Israelachvili, *J. Chem. Phys.* **87**, 1834 (1987).
24. X. Zhang, M. M. Cunningham, R. A. Walker, *J. Phys. Chem. B* **107**, 3183–3195 (2003).

PHYSICS

Observation of Fermi arc surface states in a topological metal

Su-Yang Xu,[1,2]* Chang Liu,[1]* Satya K. Kushwaha,[3] Raman Sankar,[4] Jason W. Krizan,[3] Ilya Belopolski,[1] Madhab Neupane,[1] Guang Bian,[1] Nasser Alidoust,[1] Tay-Rong Chang,[5] Horng-Tay Jeng,[5,6] Cheng-Yi Huang,[7] Wei-Feng Tsai,[7] Hsin Lin,[8] Pavel P. Shibayev,[1] Fang-Cheng Chou,[4,9] Robert J. Cava,[3] M. Zahid Hasan[1,2]†

The topology of the electronic structure of a crystal is manifested in its surface states. Recently, a distinct topological state has been proposed in metals or semimetals whose spin-orbit band structure features three-dimensional Dirac quasiparticles. We used angle-resolved photoemission spectroscopy to experimentally observe a pair of spin-polarized Fermi arc surface states on the surface of the Dirac semimetal Na_3Bi at its native chemical potential. Our systematic results collectively identify a topological phase in a gapless material. The observed Fermi arc surface states open research frontiers in fundamental physics and possibly in spintronics.

The realization of topological states of matter beyond topological insulators has become an important goal in condensed-matter and materials physics (1–15). In the topological insulators $Bi_{1-x}Sb_x$ and Bi_2Se_3 or topological crystalline insulators such as $Pb_{1-x}Sn_xTe(Se)$, the bulk has a full insulating energy gap, whereas the surface possesses an odd or even number of spin-polarized surface or edge states (3, 14–18). These are symmetry-protected topological states (19). Very recently, the possibility of realizing new topological states in materials beyond insulators, such as metals or semimetals, has attracted much attention (1–13). Semimetals are materials whose bulk conduction and valence bands have small but finite overlap; the lack of a full band-gap implies that any topological states that might exist in a semimetal should be distinct from the topological states studied in insulating materials. Theory has proposed two kinds of topological semimetals: the topological Dirac and Weyl semimetals (2, 7–12). Their low-energy bulk excitations are described by the Dirac and Weyl equations, respectively. For both types, the bulk conduction and valence bands are predicted to touch at multiple discrete points in the bulk Brillouin zone

[1]Joseph Henry Laboratory, Department of Physics, Princeton University, Princeton, NJ 08544, USA. [2]Princeton Center for Complex Materials, Princeton Institute for Science and Technology of Materials, Princeton University, Princeton, NJ 08544, USA. [3]Department of Chemistry, Princeton University, Princeton, NJ 08544, USA. [4]Center for Condensed Matter Sciences, National Taiwan University, Taipei 10617, Taiwan. [5]Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan. [6]Institute of Physics, Academia Sinica, Taipei 11529, Taiwan. [7]Department of Physics, National Sun Yat-Sen University, Kaohsiung 804, Taiwan. [8]Graphene Research Centre and Department of Physics, National University of Singapore 117542, Singapore. [9]National Synchrotron Radiation Research Center, Taiwan.
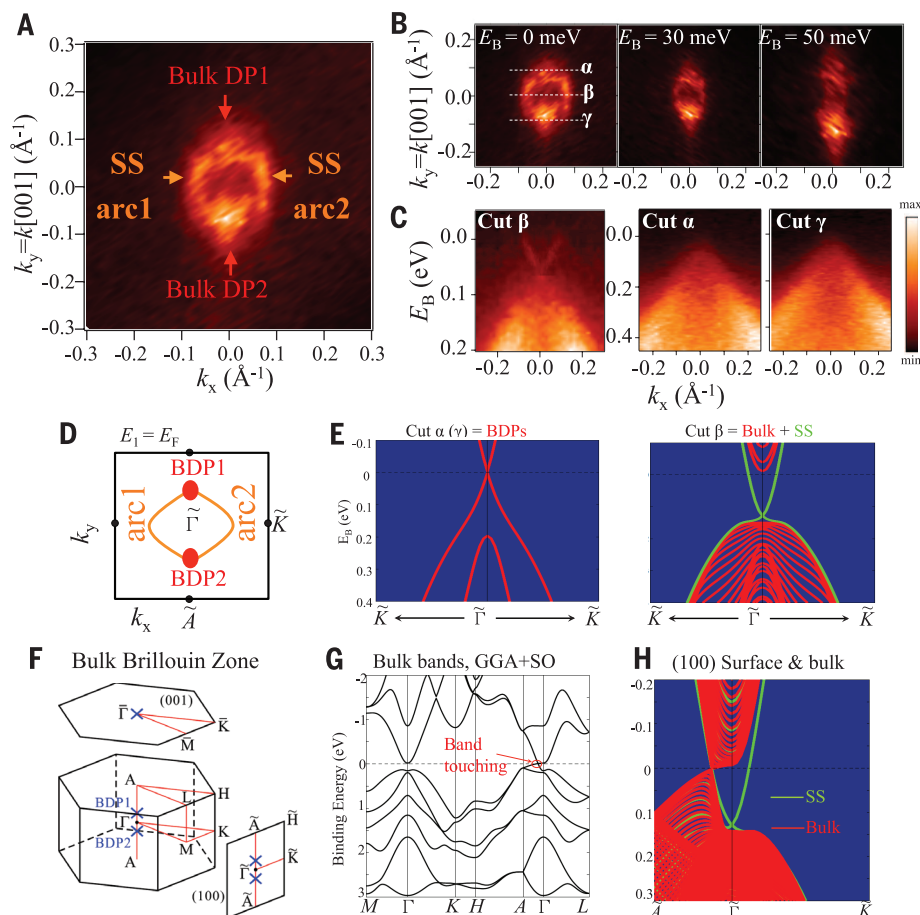*These authors contributed equally to this work. †Corresponding author. E-mail: mzhasan@princeton.edu

Fig. 1. Observation of Fermi arc surface states.
(**A**) Fermi surface map of the Na_3Bi sample at photon energy 55 eV. BDP1 and BDP2 denote the two bulk Dirac points. (**B**) ARPES constant energy contours as a function of binding energy at photon energy 55 eV. The dotted lines indicate the momentum space cuts shown in (C). (**C**) ARPES dispersion cuts α, β, and γ as defined in (B) at photon energy 55 eV. (**D**) Schematic Fermi surface of Na_3Bi. The red shaded areas and the orange lines represent the bulk and surface states, respectively. (**E**) Calculated band structure along cut β and cut α(γ). (**F**) The bulk BZ and surface BZs of the (001) and (100) surfaces. Bulk Dirac nodes are marked by blue crosses. (**G**) First-principles bulk band calculation for Na_3Bi. GGA+SO means that the calculations are based on the generalized gradient approximation (GGA) method in the presence of spin-orbit (SO) coupling. (**H**) First-principles calculation of the (100) surface electronic structure.

(BZ) and show linear behavior in the vicinity of those points (2, 7–12). For topological Dirac semimetal material candidates (9, 10), theoretical calculations (9–11) show that the surface states are a pair of Fermi arcs (Fermi arc surface states) that connect the bulk bands at the energy of the bulk Dirac nodes, which can be realized at the sample's native chemical potential. These double Fermi arc surface states are disjoint on the surface, and they connect to each other through the bulk bands. Moreover, they are predicted to show a wide range of exotic properties, such as unusual polarization (13), Friedel oscillations in tunneling experiments (6), and novel quantum oscillations in transport (12). However, they have not yet been observed in real materials.

Experimentally, a number of compounds have been identified to be three-dimensional (3D) Dirac semimetals (or 3D analogs of graphene), such as $BiTl(S_{0.5}Se_{0.5})_2$, $(Bi_{0.94}In_{0.06})_2Se_3$, $Bi_{1-x}Sb_x$, Na_3Bi, and Cd_3As_2 (20–25). Only the latter two are theoretically believed to be topologically nontrivial (2), but existing experimental data (20–25) are insufficient to prove their topological nature. We experimentally uncovered the nontrivial topological nature of the semimetal Na_3Bi. We achieved this by observing a pair of spin-polarized Fermi arc surface states at the native Fermi level on the surface of our samples using high-resolution angle-resolved photoemission spectroscopy (ARPES). Our observation of Fermi arc surface states lays the foundation for studying fundamentally new physics in nontrivial metals and may also have potential for device applications (1, 2, 4–15).

Na_3Bi is a semimetal that crystalizes in the hexagonal $P6_3/mmc$ crystal structure with a = 5.448 Å and c = 9.655 Å (26). First-principles bulk band calculations (9) show that its lowest bulk conduction (Na 3s) and valence (Bi $6p_{x,y,z}$) bands possess a band inversion of about 0.3 eV at the bulk BZ center Γ (9). The strong spin-orbit coupling in the system can open up energy gaps between the inverted bulk bands, but because of the protection of an additional threefold rotational symmetry along the [001] crystalline direction, the bulk bands are predicted to touch in two locations (Dirac nodes), even after spin-orbit coupling is considered (Fig. 1F, blue crosses). At the (001) surface, because the two bulk Dirac nodes project onto the same point in the surface BZ (Fig. 1F), Fermi arc surface states are not possible at the (001) surface. On the other hand, at the (100) surface, the two bulk Dirac nodes are separated on the opposite sides of the (100) surface BZ center $\widetilde{\Gamma}$ (Fig. 1F). Consequently, the Fermi arc surface states connecting the bulk Dirac nodes are found in the (100) surface electronic structure calculation (Fig. 1H). In order to experimentally search for the Fermi arc surface states and to probe the topological number for the Dirac semimetal state in Na_3Bi, we systematically studied its electronic structure and spin polarization at the (100) side-surface.

Figure 1A shows the ARPES-measured Fermi surface of our Na_3Bi sample at its native Fermi level. The measured Fermi surface consists of two Fermi "points" along the $k_{y[001]}$ direction and two arcs that connect the two Fermi points. We then studied the evolution of constant energy contour as a function of binding energy E_B (Fig. 1B). As we moved E_B below the Fermi level, the two bulk Dirac points were found to enlarge into contours (hole-like behavior), whereas the two surface Fermi arcs shrank (electron-like behavior). In Fig. 1C, we show the energy dispersion for important momentum space cut directions (Fig. 1B). Surface states with a surface Dirac crossing are clearly observed near the Fermi level in cut β. On the other hand, for cuts α and γ, no surface states are observed but the bulk linear band is seen to cross the Fermi level (Fig. 1E). All these ARPES results regarding the band behavior (Fermi surface, electron- or hole-like behavior for each band, and energy dispersion cuts) are in qualitative agreement with the theoretical prediction (9). Therefore, our ARPES data demonstrate that the Fermi surface of Na_3Bi is described by two Fermi arc
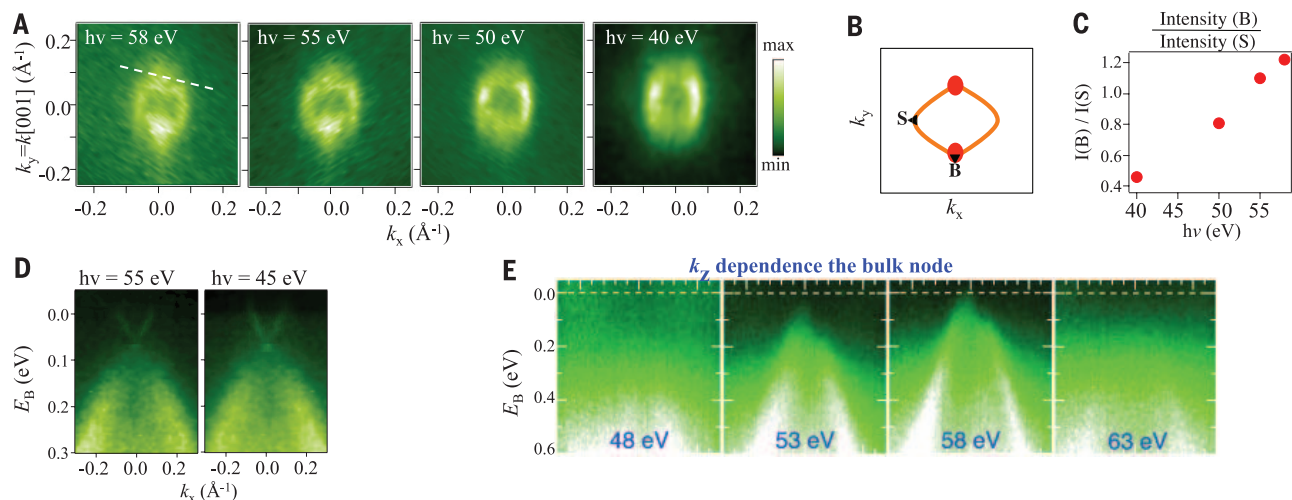
Fig. 2. Systematic studies of the double Fermi arc surface states. (**A**) ARPES Fermi surface maps at different photon energies. (**B** and **C**) We denote the middle point of the left Fermi arc as "S" and the bottom bulk Dirac point as "B." (C) shows the relative ARPES intensity between "S" and "B." (**D**) ARPES dispersion maps of the surface states at two different photon energies. (**E**) ARPES dispersion maps of the bulk Dirac band at different photon energies, along the cut indicated by the white dotted lines in (A).
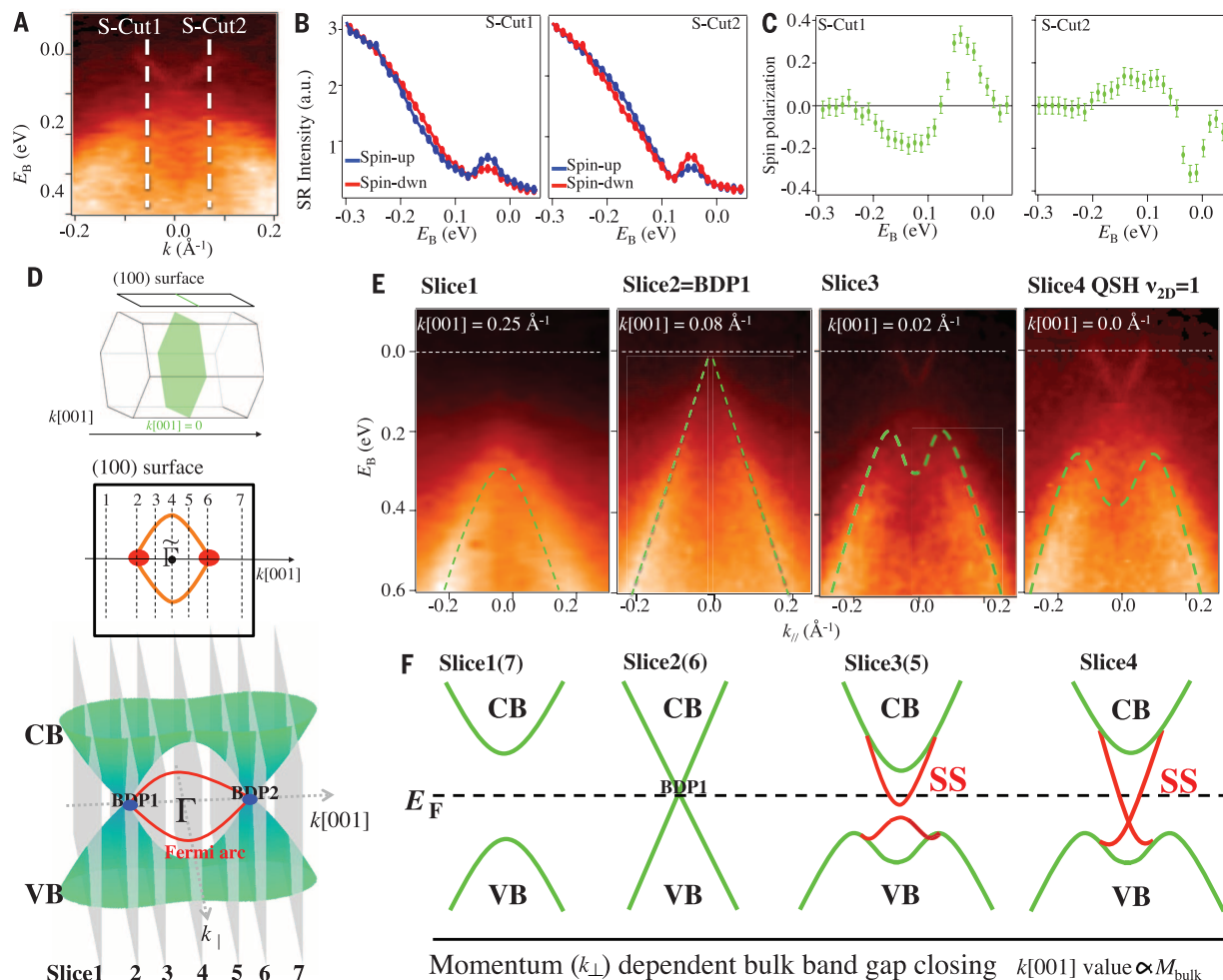
Fig. 3. Surface spin polarization and the 2D topological number. (**A**) The white dotted lines indicate the two momenta chosen for spin-resolved measurements. (**B** and **C**) Spin-resolved ARPES intensity and net spin polarization along the in-plane tangential direction for S-cuts 1 and 2 at photon energy 55 eV. The error bars in (C) show the experimental uncertainty in determining the magnitude of the spin polarization. (**D**) A schematic view of the band structure of the topological Dirac semimetal phase. Seven 2D-k slices that are taken perpendicular to the $k_{[001]}$ axis are noted. (**E** and **F**) ARPES-measured [(E)] and schematic [(F)] band structure for these slices are shown.

A Fine-tuning (topo . trivial) Dirac semimetal

Trivial insulator   Dirac at critical point   Topological insulator   SOC

C Topological Dirac semimetal

3D Dirac multiplet

Trivial insulator   Topo. semimetal SOC

B BiTl(S$_{0.5}$Se$_{0.5}$)$_2$

Γ $k_\perp$ $k_\parallel$

D Na$_3$Bi

$k_\perp$ $k_\parallel$
BDP1 Γ BDP2
Fermi arcs
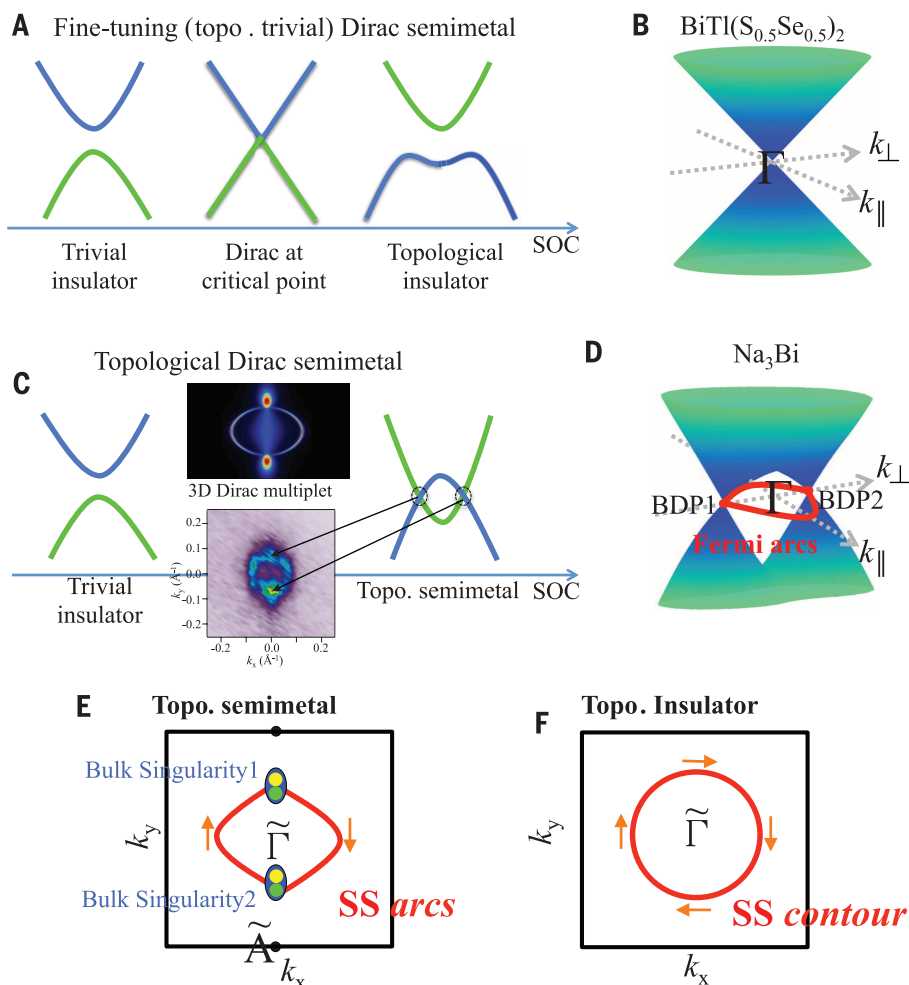
Fig. 4. Comparison between a trivial 3D Dirac semimetal achieved by fine tuning and a topological Dirac semimetal. (**A**) A 3D Dirac semimetal state with only one bulk Dirac cone can be realized by tuning a system to the critical point of a topological phase transition between a band insulator and a topological insulator. (**B**) Schematic band structure of BiTl(S$_{0.5}$Se$_{0.5}$)$_2$, which is at the critical composition of the topological phase transition system (*20*). (**C**) A topological Dirac semimetal phase is realized in an inverted band structure and is stable without fine tuning because of additional protection by crystalline symmetries. (**D**) Schematic band structure from the (100) surface of a topological Dirac semimetal. In this schematic, only surface states at the bulk Dirac point energy are plotted. Surface states at other energies are not shown. (**E** and **F**) Schematic Fermi surface of Na$_3$Bi and Bi$_2$Se$_3$. In Na$_3$Bi, each Dirac node (the blue shaded area) can be viewed as a composite of two degenerate Weyl nodes (yellow and green areas). The orange arrows note the spin polarization according to our spin-resolved ARPES measurements. The inset of (C) shows the ARPES-measured Fermi surface.

E Topo. semimetal

Bulk Singularity1
k_y
$\widetilde{\Gamma}$
Bulk Singularity2
SS *arcs*
$\widetilde{A}$
k_x

F Topo. Insulator

k_y
$\widetilde{\Gamma}$
SS *contour*
k_x

surface states that connect two bulk Dirac nodes at our samples' native chemical potential.

In order to further confirm the 2D surface nature for the double Fermi arc surface states and the 3D bulk nature for the two bulk Dirac bands, we present Fermi surface maps at different incident photon energies. The double Fermi arcs are observed at all photon energies, whereas the bulk Dirac nodes are pronounced only at 55 to 58 eV (Fig. 2, A to C). Furthermore, we show the dispersion along cut β, which also confirms that the surface state spectra do not depend on photon energy (Fig. 2D). In contrast, the bulk Dirac bands disperse strongly in photon energy (Fig. 2E). The systematic photon energy (k_z)–dependent data in Fig. 2 clearly demonstrate that the electronic states at the Fermi level are localized at the surface (2D surface state nature) everywhere around the Fermi surface contour, except at the locations of the two bulk Dirac nodes, where the states disperse strongly in all three dimensions (3D bulk band nature).

Next, we studied the surface spin polarization along cut β, where a surface Dirac crossing within the bulk bandgap is observed. Figure 3, B and C, show the in-plane spin-resolved intensity and net spin polarization. The direction of spin polarization is reversed upon going from S-cut

1 to S-cut 2 (defined in Fig. 3A), which shows the spin-momentum locking property and the singly degenerate nature of the Fermi arc surface states along the cut β direction.

We used the observed electronic and spin structure for the (100) Fermi arc surface states to obtain the topological number for the Dirac semimetal state in Na$_3$Bi. Because of the existence of the two bulk nodes (slices 2 and 6), our ARPES data (Fig. 3, D and E) show that the bulk bandgap closes and reopens as one goes across each bulk node (slice 2 or 6). We carefully examined the ARPES electronic structure for all 2D-k slices, where the bulk is gapped (slices 1, 3 to 5, and 7). It can be seen from Fig. 3, E and F, that slices 1 and 7 are gapped without any surface states; slices 3 and 5 do have surface states, but the surface states are gapped and do not connect across the bulk bandgap; however, for slice 4 ($k_{[001]} = 0$), our data show gapless Dirac surface states that span across the bulk gap. Therefore, our data indicate that slice 4 has a nontrivial 2D topological number. Because slice 4 (the $k_{[001]} = 0$ plane) is invariant under the time-reversal operation, it is possible for slice 4 to have a nonzero 2D $\mathbb{Z}_2$ number $\nu_{2D} = 1$ (*2, 14, 15*). If that is the case, then slice 4 can be viewed as a 2D quantum spin Hall–like system, and one would expect

a Dirac surface state with a helical spin texture connecting the bulk conduction and valence bands. Indeed, this is what we observe in our ARPES (Fig. 3E) and spin-ARPES (Fig. 3, A to C) data. Therefore, our data demonstrate that Na$_3$Bi is a topological Dirac semimetal that features a 2D topological number $\nu_{2D} = 1$ for the $k_{[001]} = 0$ plane, which is consistent with the prediction in (*2*). Further elaborations regarding the topological properties of Na$_3$Bi are presented in the supplementary materials (figs. S1 and S2). Figure 4, A to D, shows schematically the difference between a trivial and topological Dirac semimetal.

Our observations of topological surface states, the way in which they connect to the bulk Dirac cones, and their spin momentum locking demonstrate the topology of a Dirac semimetal in line with the theoretical prediction (*2*). The observed Fermi arc surface states represent a type of 2D electron gas that is distinct from that of the surface states in a $\mathbb{Z}_2$ topological insulator (*27*). In a typical topological insulator such as Bi$_2$Se$_3$, the surface states' Fermi surface is a closed contour. In sharp contrast, the Fermi surface in Na$_3$Bi consists of two arcs, which are bridged by the two bulk nodes (Fig. 4, E and F). Therefore, as one goes along the surface Fermi arc and reaches a bulk node, the wave function of the

surface state gradually loses its surface nature and becomes a bulk band. Such exotic behavior does not exist in surface states of Bi_2Se_3 or any known topological insulator. In surface electrical transport experiments at high magnetic field, it would be interesting to study how the surface electrons wind along the arc and enter the bulk singularity as considered by a recent theoretical work (5), if the chemical potential can be placed precisely at the bulk Dirac node so that the bulk carrier density is vanishingly small (5). Finally, it would be interesting to introduce superconductivity or magnetism in these new topological metals (9, 10, 28). In particular, breaking time-reversal symmetry via magnetism can split the Dirac nodes into Weyl nodes (Fig. 4E).

REFERENCES AND NOTES

1. A. M. Turner, A. Vishwanath, http://arxiv.org/abs/1301.0330 (2013).
2. B.-J. Yang, N. Nagaosa, *Nat. Commun.* **5**, 4898 (2014).
3. M. Z. Hasan, S.-Y. Xu, M. Neupane, http://arxiv.org/abs/1406.1040 (2014).
4. F. D. M. Haldane, http://arxiv.org/abs/1401.0529 (2014).
5. A. C. Potter, I. Kimchi, A. Vishwanath, *Nat. Commun.* **5**, 5161 (2014).
6. P. Hosur, *Phys. Rev. B* **86**, 195102 (2012).
7. S. Murakami, *New J. Phys.* **9**, 356 (2007).
8. S. M. Young et al., *Phys. Rev. Lett.* **108**, 140405 (2012).
9. Z. Wang et al., *Phys. Rev. B* **85**, 195320 (2012).
10. Z. Wang, H. Weng, Q. Wu, X. Dai, Z. Fang, *Phys. Rev. B* **88**, 125427 (2013).
11. A. Narayan, D. Di Sante, S. Picozzi, S. Sanvito, http://arxiv.org/abs/1408.3509 (2014).
12. X. Wan, A. M. Turner, A. Vishwanath, S. Y. Savrasov, *Phys. Rev. B* **83**, 205101 (2011).
13. T. Ojanen, *Phys. Rev. B* **87**, 245112 (2013).
14. M. Z. Hasan, C. L. Kane, *Rev. Mod. Phys.* **82**, 3045–3067 (2010).
15. J. E. Moore, *Nat. Phys.* **5**, 378–380 (2009).
16. D. Hsieh et al., *Nature* **452**, 970–974 (2008).
17. D. Hsieh et al., *Science* **323**, 919–922 (2009).
18. Y. Xia et al., *Nat. Phys.* **5**, 398–402 (2009).
19. X.-G. Wen, *Phys. Rev. B* **89**, 035147 (2014).
20. S.-Y. Xu et al., *Science* **332**, 560–564 (2011).
21. M. Brahlek et al., *Phys. Rev. Lett.* **109**, 186403 (2012).
22. S.-Y. Xu et al., http://arxiv.org/abs/1312.7624 (2013).
23. Z. K. Liu et al., *Science* **343**, 864–867 (2014).
24. M. Neupane et al., *Nat. Commun.* **5**, 3786 (2014).
25. S. Borisenko et al., *Phys. Rev. Lett.* **113**, 027603 (2014).
26. T. B. Massalski, *Binary Alloy Phase Diagrams* (ASM, Materials Park, OH, ed. 2, 1990).
27. M. Z. Hasan, www.fysik.su.se/~ardonne/nobel/ns156-program.pdf (2014).
28. S. A. Yang, H. Pan, F. Zhang, *Phys. Rev. Lett.* **113**, 046401 (2014).

~ 30TH ANNUAL ~
MOLECULAR BIOLOGY SUMMER WORKSHOP

Learn Molecular Biology in 2 Weeks!

We are pleased to announce the thirtieth annual Molecular Biology Summer Workshop, sponsored by New England Biolabs in conjunction with Smith College. The workshop is held at Ford Hall, Smith College, Northampton, MA, USA. Well over 4,000 people have graduated from this intensive training program in the past twenty-nine years. This intensive course emphasizes hands-on molecular biology laboratory work and covers a wide variety of topics and techniques. This is the largest and longest running molecular biology course for professionals in the world. No prior experience in molecular biology is required.

when:
two-week session:
July 19 -August 1, 2015

where:
Ford Hall
Smith College
Northampton, MA 01063
USA

Dr. Steven A. Williams, Director

to apply:
website:
http://www.science.smith.edu/neb

email:
biolabs@smith.edu

phone:
(413) 585-3915

Two-week Molecular Biology Session Available

This two-week long course covers in-depth DNA cloning, PCR, DNA sequencing, genomics, next-gen sequencing and bioinformatics. Learn hands-on techniques used in gene expression analysis including microarray analysis, RNAi and quantitative RT-PCR, bioinformatics and genomics and proteomics.

application information: No previous experience in molecular biology is required or expected. Forty participants will be selected from a variety of disciplines and academic backgrounds, including principal investigators, directors of programs, medical doctors, postdoctoral fellows, graduate students, research assistants, sales associates, equipment engineers, etc.

fee: $4300 per participant. This fee includes lab manual, use of all equipment and supplies, and room and board (all rooms are singles).

application deadlines: July 1, 2015.
First come, first served (apply now!). Late applications will be accepted on a space available basis.

payment deadline: Three weeks following receipt of your email acceptance letter.

topics / techniques:

- **gene cloning
  (and library construction)**
- **gene expression analysis**
- **PCR and quantitative RT-PCR**
- **genomics and bioinformatics**
- **DNA sequencing, including
  NextGen**
- **RNAi and siRNA**
- **RNA-Seq**
- **and much more — visit our
  website for a complete list!**

NEW ENGLAND BioLabs Inc.

www.neb.com

Unique.

Over 1000 IHC validated antibodies

Trusted | **Worry-Free**® | **Customizable**

With over 20 years of experience manufacturing antibodies, Enzo offers highly characterized antibodies that have been rigorously validated for multiple applications (IHC, ICC, and IF). But don't take our word for it, their use is documented in 1000s of peer-reviewed publications proving their reliable performance time and again. As a comprehensive provider of detection technologies, Enzo also offers high definition chromogens and detection reagents to enable accurate, reproducible multiplex analysis of antigens which increase efficiency and conserve precious tissue samples.

scientists *enabling* scientists.™

www.enzolifesciences.com

Enzo®

This is the start of something big.

Science Advances | AAAS

SIGNIFICANT RESEARCH, GLOBAL IMPACT

Introducing *Science Advances* – the new, online-only, open-access journal from *Science* and AAAS. Find out how you can be among the first authors published at **scienceadvances.org.**

AAAS 2015 ANNUAL MEETING

Join Us in San Jose

Learn about science and technology addressing current and future global challenges.

- Seminars on computing; imaging Earth; infectious disease; and communicating science
- 150+ symposia in 13 disciplinary tracks covering the latest research advances
- Network with colleagues and attend career development workshops

Connect with us

🐦 @AAASmeetings   #AAASmtg

f facebook.com/AAAS.Science

www.aaas.org/meetings

Reporters: The EurekAlert! website hosts the AAAS Annual Meeting Newsroom. For details please visit eurekalert.org/aaasnewsroom/2015

Dear Colleagues:

On behalf of the AAAS Board of Directors, it is my honor to invite you to join us in San Jose, CA for the 2015 AAAS Annual Meeting, 12-16 February. This annual event is one of the most widely recognized global science gatherings, with hundreds of diverse scientific sessions and communication opportunities with broad U.S. and international media coverage.

This year's theme—*Innovations, Information, and Imaging*—will focus on transformation across all disciplines of science and technology brought about by rapid progress in organizing, visualizing, and analyzing data.

Everyone is welcome at the AAAS Annual Meeting. We hope you will join us in San Jose.

Gerald R. Fink

Gerald R. Fink, Ph.D.
AAAS President and Program Chair
Margaret and Herman Sokol Professor of Genetics
Massachusetts Institute of Technology

PRESIDENT'S ADDRESS

Gerald R. Fink
AAAS President and Program Chair
Margaret and Herman Sokol
Professor of Genetics
Massachusetts Institute of Technology

Thursday, 12 February
6:00–7:00 p.m.
San Jose Convention Center, Room 220A

Dr. Gerald Fink's work in genetics, biochemistry, and molecular biology has advanced our understanding of gene regulation, mutation, and recombination. He developed a technique for transforming yeast that allowed researchers to introduce a foreign piece of genetic material into yeast cells and study the inheritance and expression of that DNA. The technique, fundamental to genetic engineering, laid the groundwork for the commercial use of yeast as biological factories for manufacturing vaccines and other drugs, and set the stage for genetic engineering in all organisms. Fink chaired a National Research Council Committee that produced the 2003 report *Biotechnology Research in an Age of Terrorism: Confronting the Dual Use Dilemma,* recommending practices to prevent the potentially destructive application of biotechnology research while enabling legitimate research. A founding member and past director of the Whitehead Institute, he received a Ph.D. in genetics from Yale University and a bachelor's degree in biology from Amherst College. He is a member of the National Academy of Sciences and Institute of Medicine and a fellow of the American Academy of Arts and Sciences and the American Philosophical Society. He has received the National Academy of Sciences Award in Molecular Biology, the Genetics Society of America Medal, the Emil Christian Hansen Award for Microbiology, the George W. Beadle Award, the Gruber Prize in Genetics, and a Guggenheim Fellowship.

PLENARY LECTURES

All plenary lectures will be held in the San Jose Convention Center, Room 220A

Daphne Koller
President and Co-Founder
Coursera

***The Online Revolution:
Learning Without Limits***

Friday, 13 February
5:00–6:00 p.m.

Karl Deisseroth
D.H. Chen Professor of Bioengineering
Stanford University

***Optical Desconstruction of Fully-Assembled
Biological Systems***

Sunday, 15 February
5:00–6:00 p.m.

David Baker
Professor of Biochemistry
University of Washington

***Post-Evolutionary Biology: Design of Novel
Protein Structures, Functions, and Assemblies***

Saturday, 14 February
5:00–6:00 p.m.

Neil Shubin
Professor, Organismal Biology and Anatomy
University of Chicago

Finding Your Inner Fish

Monday, 16 February
8:30–9:30 a.m.

TOPICAL LECTURES

David Altshuler
Professor of Genetics and Medicine, Harvard Medical School
Human Genome Sequence Variation and Disease

Gerbrand Ceder
R.P. Simmons Professor of Materials Science and
Engineering, Massachusetts Institute of Technology
*The Materials Genome and Its Role in Accelerating
Materials Development*

Sally Davies
Chief Medical Officer and Chief Scientific Adviser,
Department of Health, United Kingdom
Antimicrobial Resistance: A Rising Global Threat

Ann McKee
Professor of Neurology and Pathology,
Boston University School of Medicine
Emerging Concepts in Chronic Traumatic Encephalopathy

Geoffrey Nunberg
Adjunct Professor of Information,
University of California, Berkeley
The Science of Grammar and Vice Versa

Naomi Oreskes
Professor of the History of Science, Harvard University
Science: Why Should They Believe Us?

Naledi Pandor
Minister of Science and Technology, South Africa
Why Science In, With, and For Africa Matters

GEORGE SARTON MEMORIAL LECTURE IN THE HISTORY AND
PHILOSOPHY OF SCIENCE
Paul Farber
Distinguished Professor Emeritus, Modern Life Sciences,
Intellectual History, Oregon State University
Darwinian Evolution and Human Race

JOHN P. MCGOVERN LECTURE IN THE BEHAVIORAL SCIENCES
Susan Fiske
Eugene Higgins Professor of Psychology and Public Affairs,
Princeton University
Humans are Intent Detectors: Implications for Society

SEMINARS

Thursday, 12 February

Communicating Science

Scientific and technological issues may trigger societal conflict when they intersect with personal or political views. Today's scientists and engineers are challenged to communicate and engage with the public, particularly amid pressure on research budgets and related concerns about transparency and accountability. This seminar will share expertise for scientists participating in science communication and public engagement.

Organized by: AAAS Center for Public Engagement with Science and Technology

Scientists Communicating Challenging Issues

MODERATOR: Susanne Moser, Susanne Moser Research and Consulting, Santa Cruz, CA

SPEAKERS:
Noah Diffenbaugh, Stanford University, CA
Kathleen Hall Jamieson, University of Pennsylvania, Philadelphia
Lisa Krieger, *San Jose Mercury News,* CA

Public Engagement for Scientists: Realities, Risks, and Rewards

MODERATOR: Bruce Lewenstein, Cornell University, Ithaca, NY

SPEAKERS:
Elizabeth Babcock, California Academy of Sciences, San Francisco
Heidi Ballard, University of California, Davis
Anthony Dudo, University of Texas, Austin
Nalini Nadkarni, University of Utah, Salt Lake City

Friday, 13 February

The Future of Computing

New ways of collecting and using information are transforming science, technology, and the fabric of society. This seminar will address some of the most important advances and needs in computing and the internet, including information-centric networking; high-performance computing; artificial intelligence; mobile and wearable devices; and cybersecurity.

Our Computational Foundation Crisis and Life Beyond

Organized by: Jon Candelaria, Semiconductor Research Corporation, Durham, NC; Larry A. Nagahara, National Cancer Institute, Bethesda, MD

SPEAKERS:
Tilak Agerwala, IBM T.J. Watson Research Center, Yorktown Heights, NY
Peter Norvig, Google Inc., Mountain View, CA
Charles Bergan, Qualcomm Inc., San Diego, CA

The Future of the Internet: Meaning and Names or Numbers?

Organized by: Glenn T. Edens, PARC, Palo Alto, CA; J.J. Garcia-Luna-Aceves, University of California, Santa Cruz

SPEAKERS:
Vinton Cerf, Google Inc., Mountain View, CA
David Oran, Cisco Systems, Cambridge, MA
Glenn T. Edens, PARC, Palo Alto, CA

Engineering Information: Adapting Risk and Resilience Frameworks to Cybersecurity

Organized by: Igor Linkov, U.S. Army Engineer Research and Development Center, Concord, MA; Sankar Basu, National Science Foundation, Arlington, VA

SPEAKERS:
Shankar Sastry, University of California, Berkeley
John E. Savage, Brown University, Providence, RI
Ahmad-Reza Sadeghi Sadeghi, Technical University of Darmstadt, Germany
Stephanie Forrest, University of New Mexico, Albuquerque
Ken Heffner, Honeywell International Inc., Clearwater, FL
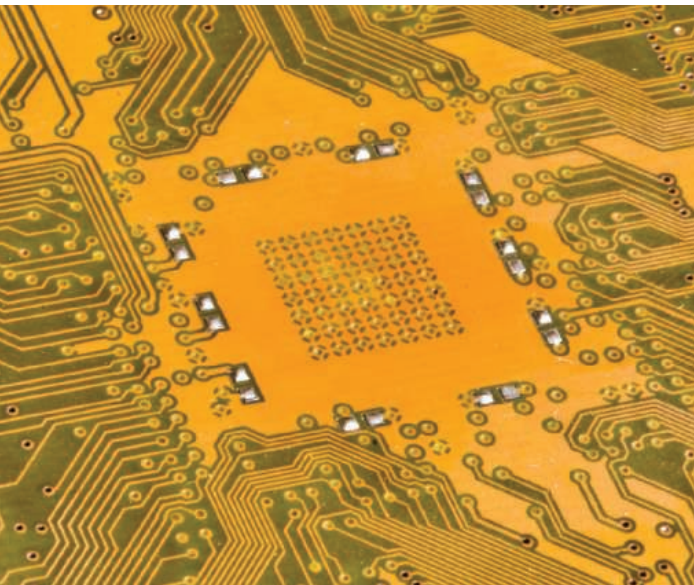
Saturday, 14 February

Infectious Disease: Monitoring and Response

This seminar will consider how infectious disease can be effectively monitored to inform public health response. A discussion about the Ebola virus outbreak in West Africa will feature perspectives from individuals and organizations involved in monitoring and responding and provide information on current and future challenges. Another session will review vaccine development and advances in genomics, computational and structural biology, and systems biology that address genetic diversity and immune evasion, providing greater probability for success in the next generation of vaccines. The last session will discuss how Earth observation data can be used to predict the spatial and/or temporal distribution of disease in association with environmental factors, resulting in diverse health applications such as support for development of static spatial predictions of disease risk, outbreak early warning systems, and forecasts of climate change health impacts.

Lessons from the Ebola Outbreak: Response and Responsibility

MODERATOR: Jon Cohen, AAAS/*Science*

SPEAKERS:

Anthony Fauci, National Institute of Allergy and Infectious Diseases, Bethesda, MD

Keiji Fukuda, World Health Organization (WHO), Geneva, Switzerland

Stephen Gire, Harvard University and Broad Institute, Cambridge, MA

The Human Vaccines Project: Transforming the Future of Vaccine Development

Organized by: Wayne Koff, International AIDS Vaccine Initiative, New York City

SPEAKERS:

Wayne Koff, International AIDS Vaccine Initiative, New York City

Bali Pulendran, Emory Vaccine Center, Atlanta, GA

Peter Kwong, Vaccine Research Center, National Institute of Allergy and Infectious Diseases, Washington, DC

Earth Observation Approaches to Spatial Disease Prediction, Surveillance, and Control

Organized by: Archie C.A. Clements, Australian National University, Canberra

SPEAKERS:

Uriel Kitron, Emory University, Atlanta, GA

Kenneth J. Linthicum, U.S. Department of Agriculture, Gainesville, FL

Archie C.A. Clements, Australian National University, Canberra

Sunday, 15 February

Innovations in Imaging Earth

From Global Positioning System (GPS) mapping devices to internet-connected location tagging, geospatial information permeates society in the modern digital world. This seminar will highlight advances in remote sensing technology, numerical simulations, and computation that enable monitoring of the Earth at high resolution. Big data analysis techniques combined with low cost, high quality data on environmental patterns and processes from satellites, unmanned aircraft, and remote sensor networks provide insights into global change dynamics and how they impact food security, environmental sustainability, and human welfare.

Advances in Earth Observation: Enabling New Insights into Global Environmental Change

Organized by: Lyndon Estes, Princeton University, NJ; Tom Evans, Indiana University, Bloomington

SPEAKERS:

Matt Hansen, University of Maryland, College Park

Kelly Caylor, Princeton University, NJ

Maggi Kelly, University of California, Berkeley

Imaging Earth

Organized by: Ana Barros, Duke University, Durham, NC; Sally McFarlane, U.S. Department of Energy, Washington, DC

SPEAKERS:

Piers Sellers, NASA Goddard Space Flight Center, Greenbelt, MD

Melba Crawford, Purdue University, West Lafayette, IN

William Collins, Lawrence Berkeley National Laboratory, CA

Geospatial Innovations in Imaging Information Intelligently

Organized by: Steven Whitmeyer, James Madison University, Harrisonburg, VA; Declan De Paor, Old Dominion University, Norfolk, VA

SPEAKERS:

David Thau, Google Inc., Mountain View, CA

Barbara Tewksbury, Hamilton College, Clinton, NY

Richard Treves, University of Southampton, United Kingdom

Robert Kolvoord, James Madison University, Harrisonburg, VA

Jennifer Piatek, Central Connecticut State University, New Britain

Declan De Paor, Old Dominion University, Norfolk, VA

SYMPOSIA TRACKS

Organizers are listed under symposia titles.

ANTHROPOLOGY, CULTURE, AND LANGUAGE

Discovering Lost Horizons: People, Land, and Society in Prehistory
George R. Milner, Pennsylvania State University, University Park; Michael Frachetti, Washington University, St. Louis, MO

History Written in Skeletons: Intersections of History, Archaeology, and Biology
Anita Guerrini, Oregon State University, Corvallis

Imaging the Past: Using New Information Technologies To Nurture Historical Analysis
Cecile Menetrey-Monchau and Annekathrin Jaeger, European Research Council, Brussels, Belgium

Radiography of the Past: Revealing the Invisible at Archaeological Sites
Louise Byrne, European Commission, Brussels, Belgium; Gabriela Chira, Marie Curie International Fellowships, Brussels, Belgium

The Linguistics of Status, Influence, and Innovation: A Computational Perspective
Jacob Eisenstein, Georgia Institute of Technology, Atlanta

The Sense of Smell as a Novel Means to Explore Language, Culture, and Biology
Asifa Majid, Radboud University Nijmegen, Netherlands

Visualizing Verbal Culture: Seeing Language Diversity
John Nerbonne, University of Groningen, Netherlands

Worth More Than a Thousand Words: State-of-the-Art Visualization in Cultural Heritage
Francesca Casadio, The Art Institute of Chicago, IL; Marc Walton, Northwestern University, Evanston, IL

BEHAVIORAL AND SOCIAL SCIENCES

Human Mathematical Abilities: From Intuition to the Classroom and Back
James L. McClelland, Stanford University, CA

Human-Made Noise and Nighttime Lighting
Clinton D. Francis, California Polytechnic State University, San Luis Obispo

Innovations in the Family: New Structures, New Challenges
Toni Antonucci, University of Michigan Institute for Social Research, Ann Arbor

Measuring Research Integrity: Survey of Organizational Research Climate
C.K. Gunsalus, National Center for Professional and Research Ethics, Urbana, IL

Social Influences on Health Service Use Following Disasters
Eric Jones, University of North Carolina, Greensboro

Social, Emotional, and Cognitive Bases of Communication: New Analytic Approaches
Judith F. Kroll, Pennsylvania State University, University Park; Cecilia Aragon, University of Washington, Seattle; Laurie Beth Feldman, University at Albany, NY

Virtual Labs: Transforming the Social, Behavioral, and Information Sciences
Michael W. Macy, Cornell University, Ithaca, NY

Visualizing the Experience and Use of Space in the Built Environment
Annekathrin Jaeger and Katja Meinke, European Research Council, Brussels, Belgium

Watching the Brain Think: Naturalistic Approaches To Studying Human Brain Function
Susan Hagen and Jessica F. Cantlon, University of Rochester, NY

BIOLOGY AND NEUROSCIENCE

3D Chemical Imaging: New Frontier Across Disciplines
Barbara Illman, U.S. Forest Service, Madison, WI; Carol Hirschmugl, University of Wisconsin, Milwaukee

Emerging Trends in Visualizing Physical Models and Rapid Prototyping for Biological Systems
Promita Chakraborty and Ronald N. Zuckermann, Lawrence Berkeley National Laboratory, CA

Extracting Evidence From Biological Data: Multiple Disciplines Get In On the Act
Veronica Vieland, The Research Institute at Nationwide Children's Hospital, Columbus, OH

Insights from Social Networks: Visualizing Big Data from Cells to Cell Phones to Societies
Barbara Illman, U.S. Forest Service, Madison, WI; Daniel Rubenstein, Princeton University, NJ

Multisensory Development, Plasticity and Learning: From Basic to Clinical Science
Mark Wallace, Vanderbilt University, Nashville, TN

Neuroimaging: Discoveries from Fetus to Adult
Pat Levitt, University of Southern California, Los Angeles

Neuroscience of Hearing: Brain Implants in the Fight Against Childhood Deafness
Nan Ratner, University of Maryland, College Park

Novel Technologies for Exploring the Uncultivated Microbial Majority
Susannah Tringe and Tanja Woyke, U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA

Proteomics: How Big Data Opens New Vistas in Personalized Medicine
Barbara Wankerl and Patrick Regan, Technical University of Munich, Garching, Germany

Reproducibility of Science: A Roadmap Forward
Pat Levitt, University of Southern California, Los Angeles; Allan R. Sampson, University of Pittsburgh, PA

Searching for Alternative Chemistries of Life on Earth and Throughout the Universe
David G. Lynn, Emory University, Atlanta, GA; Jay T. Goodwin, National Science Foundation, Arlington, VA

To Bug or Not to Bug the Immune System: Benefits and Costs of Altering the Microbiome
Victor J. Johnson, Burleson Research Technologies Inc., Morrisville, NC; Berran Yucesoy, University of Cincinnati College of Medicine, OH

Visualizing Biomedical Data and Processes Across Space and Time Scales
Sean E. Hanlon, National Cancer Institute, Rockville, MD

What Are Race and Sex Doing in Our Genomes? Perspectives on Human "Types"
Sally Lehrman, University of California, Santa Cruz

When Experts Collide: Driving Cross-Cutting Innovation in Biological Imaging and Informatics
Carol Lynn Alpert, Museum of Science, Boston, MA

CLIMATE CHANGE, ENVIRONMENT, AND ECOLOGY

Avoiding Collapse: Human Impacts on the Biosphere
Anthony D. Barnosky, University of California, Berkeley; Paul R. Ehrlich, Stanford University, CA

Climate Change and Big Data
So-Min Cheong, University of Kansas, Lawrence

Climate Intervention and Geoengineering: Albedo Modification
Marcia McNutt, AAAS/Science, Washington, DC

Earth History: Innovative Approaches to Studying Critical Transitions
Dena M. Smith, Geological Society of America, Boulder, CO

Going Negative: Removing Carbon Dioxide From the Atmosphere
Jennifer Milne and Sally Benson, Stanford University, CA

Marine Ecosystems in Hot Water: Some Like It Hot (But Some Do Not)
Frank E. Muller-Karger, University of South Florida, St. Petersburg

Modeling Earth's Interior from Atomic to Global Scale
Renata M. Wentzcovitch, University of Minnesota, Minneapolis; Rob van de Hilst, Massachusetts Institute of Technology, Cambridge; David Bercovici, Yale University, New Haven, CT

Ocean Acidification and Hypoxia: Planning For Regional Action
Michael J. O'Donnell, California Ocean Science Trust, Oakland

Seeing the Invisible: How Sequencing Diverse Eukaryotes Transforms Ocean Science
Jian Guo and Alexandra Z. Worden, Monterey Bay Aquarium Research Institute, Landing Moss, CA; Adam Monier, University of Exeter, United Kingdom

Severe Weather in a Changing Climate: Informing Risk
Michael E. Mann, Pennsylvania State University, University Park; Michael Wehner, Lawrence Berkeley National Laboratory, CA; Donald J. Wuebbles, University of Illinois, Urbana-Champaign

The Atlantic Ocean: Our Unknown Treasure
Marco Weydert and John Bell, European Commission, Research and Innovation, Brussels, Belgium

Visual Cultures of Prediction: Imaging Climate Change Data
Lynda Walsh, University of Nevada, Reno; Birgit Schneider, University of Potsdam, Germany

◣AAAS 2015 ANNUAL MEETING
INNOVATIONS, INFORMATION, AND IMAGING

Advance Registration Rates until 22 January

	AAAS Member Rates for members in good standing	New Member Reduced rates if you join AAAS today	Non-Member Rates for all other attendees
Professional	$295	$380	$399
Postdoc	$235	$320	$335
K-12 Teacher	$235	$320	$335
Emeritus	$235	$320	$335
Student	$60	$70	$90

One-day rates are also available. Visit **www.aaas.org/meetings/registration.**

COMMUNICATION AND PUBLIC PROGRAMS

Citizen Science from the Zooniverse: Cutting-Edge Research with 1 Million Scientists
Ramin A. Skibba, University of California, San Diego

Citizen Science: Advancing Innovations for Science, Information, and Engagement
Meg Domroese, Citizen Science Association/Schoodic Institute, Winter Harbor, ME; Jennifer Shirk, Cornell University, Ithaca, NY

Comics, Zombies, and Hip-Hop: Leveraging Pop Culture for Science Engagement
Kishore Hari and Rebecca L. Smith, University of California, San Francisco

Engagement with Intent? Scientists' Views of Communication and Why It Matters
John C. Besley, Michigan State University, East Lansing

From Art to Mathematics: A Visual Mode of Communication
George W. Hart, Stony Brook University, NY

Going Public: Investing in Science Communication for Scientists
Keegan Sawyer, National Academy of Sciences, Washington, DC; Brooke Smith, COMPASS, Portland, OR; Russ Campbell, Burroughs Wellcome Fund, Research Triangle Park, NC

National Climate Assessment: Resource for Climate Literacy and Making Decisions
Emily Therese Cloyd, U.S. Global Change Research Program, Washington, DC

Public Engagement with Science: What's In It for Scientists?
Katherine Nielsen, University of California, San Francisco

Scientists Engaging with Reporters, the Public, and Social Media: Survey Findings
Lee Rainie, Pew Research Center, Washington, DC

Science Visualization: The Art of Making Data Beautiful
Andy Freeberg, SLAC National Accelerator Laboratory, Menlo Park, CA

Scientific Visualization: Collaborations Between Museums and Scientists
Erika C. Shugart, American Society for Microbiology, Washington, DC

Strategies for Effective Broader Impacts Work
Justin Lawrence and Elise Lipkovitz, AAAS Science and Technology Policy Fellows, National Science Foundation, Arlington, VA

Using Cartoons To Convey Science
Yoram Bauman, StandUpEconomist.com, Seattle, WA

EDUCATION AND HUMAN RESOURCES

Assessment in Support of K–12 Science Learning
Natalie Nielsen and Heidi Schweingruber, National Academy of Sciences, Washington, DC

Diversity in the Academic STEM Workforce: Understanding Career Experiences
Julia E. Melkers, Georgia Institute of Technology, Atlanta

Gender in STEM Policy, Practice, and Research: Advances in North America and Europe
Wanda Ward, National Science Foundation, Arlington, VA

Graduate Science Education in Flux: Alternate Pathways to Science Careers
Daryl E. Chubin, Independent Consultant, Savannah, GA; Marilyn J. Suiter, National Science Foundation, Arlington, VA

Inclusive STEM High Schools: Innovative Pathways and Partnerships for Success
Edith Gummer, National Science Foundation, Arlington, VA

Innovations in Broadening Participation and Diversifying the Science Workforce
Jessi L. Smith, Montana State University, Bozeman

Learning from Visualization: Insights from STEM and Cognitive Science Collaboration
Mary Hegarty, University of California, Santa Barbara; Thomas F. Shipley, Temple University, Philadelphia, PA

Learning Science from Scientific Data
Cathryn A. Manduca, Carleton College, Northfield, MN

Next-Generation Learning and Assessment Environments for Science Inquiry Practices
Michael A. Sao Pedro, Worcester Polytechnic Institute, MA

Paving Smoother Pathways for All Students: Systemic Undergraduate STEM Reform
Martin Storksdieck, Oregon State University, Corvallis; Michael Feder, National Research Council, Washington, DC

Preparing Researchers for the Quantitative Biology of the Future
Frederick R. Adler, University of Utah, Salt Lake City; M. Gregory Forest, University of North Carolina, Chapel Hill

The Future of Graduate Education in STEM: Thinking Beyond Disciplines
Jay B. Labov, National Academy of Sciences, Washington, DC

Undergraduate Research at the Community College: Models and Sustainability
James Hewlett, Finger Lakes Community College, Canandaigua, NY

ENGINEERING, INDUSTRY, AND TECHNOLOGY

A New Paradigm for Electron Microscopy: Fast Detectors and Extreme Data Experimentation
Frances M. Ross, IBM T.J. Watson Research Center, Yorktown Heights, NY; Jim Ciston and Andrew Minor, University of California, Berkeley

Computer-Aided Design of Catalysts for Sustainable Energy Conversion
Jens Norskov, Stanford University, CA

Correlating Properties of Nano-Building Blocks Via Hyperspectral Nano-Optical Imaging
P. James Schuck and Alexander Weber-Bargioni, Lawrence Berkeley National Laboratory, CA

I See, Therefore I Can
Rene Martins, European Commission, Brussels, Belgium

Integrated Computational Materials Engineering Principles for Additive Manufacturing
Sudarsanam Babu, University of Tennessee/Oak Ridge National Laboratory, Knoxville

Next-Generation Batteries for Mobile Devices and the Grid
Johanna Nelson Weker and Glennda Chui, SLAC National Accelerator Laboratory, Menlo Park, CA

Revolutionary Vision: Implants, Prosthetics, Smart Glasses, and the Telescopic Contact Lens
Megan Williams, Christian Simm, and Melanie Picard, swissnex, San Francisco, CA

The Business of Innovation: Great Science Is Only the First Step
Fred Ledley, Bentley University, Waltham, MA

The Road to Autonomous Cars
Chris Gerdes, Stanford University, CA

Whole Lotta Shakin': Man-Made Earthquakes and Energy Development
William Savage, Seismological Society of America, Las Vegas, NV; Rex Buchanan, Kansas Geological Survey, Lawrence

GLOBAL PERSPECTIVES AND ISSUES

21st Century Global Food Security and the Environment: Improving or Deteriorating?
Felix Kogan and Alfred M. Powell, National Oceanic and Atmospheric Administration, College Park, MD

Access to Scientific Expertise in Fast Growing African Countries
Max Goldman, Sense About Science, London, United Kingdom; Sarah Evanega, Cornell University, Ithaca, NY

Astrobiology: Expanding Views of Life and Encountering New Societal Questions
Margaret S. Race, SETI Institute, Mountain View, CA; Brian Patrick Green, Santa Clara University, CA

Bounded Gaps Between Prime Numbers: Individual Research Versus Crowdsourcing
Carl Pomerance, Dartmouth College, Hanover, NH; Daniel A. Goldston, San Jose State University, CA

Creationism in Europe
Stefaan Blancke, Ghent University, Belgium; Peter C. Kjærgaard, Aarhus University, Denmark

Earth Observation Systems and Citizen Scientists Tackle Urban Environmental Hazards
Gilles Ollier, European Commission, Directorate General for Research and Innovation, Brussels, Belgium

Gender Equality in the Knowledge Society: Better Analysis Through Better Information
Sophia Huyer, Women in Global Science and Technology, Brighton, Canada

Historical and Philosophical Perspectives on Innovation in Science and Society
Manfred Laubichler, Arizona State University, Tempe

Novel Government Partnerships: Integrating Developing Countries in Global Research
Katherine E. Himes and Cameron D. Bess, U.S. Agency for International Development, Washington, DC

Science Diplomats Tackling Our Lifestyle Killers
Aidan Gilligan, SciCom–Making Sense of Science, Brussels, Belgium; Michel Kazatchkine, United Nations, Geneva, Switzerland

Science for Haiti Reconstruction: Advances in Science and Science Education Capacity
Jorge L. Colón, University of Puerto Rico, San Juan

Solutions for Achieving Gender Equity: International Perspectives
Lynnette D. Madsen, National Science Foundation, Arlington, VA; Catherine Didion, National Academy of Engineering, Washington, DC

Strategies for Pandemic Emergency Response
Eva K. Lee, Georgia Institute of Technology, Atlanta

Ukraine's Scientific Future: International Cooperation and Science Diplomacy
Michela Greco, CRDF Global, Arlington, VA

Unlocking Natural History Collections To Model the Biosphere
Ian Owens, Natural History Museum, London, United Kingdom; Kirk Johnson, National Museum of Natural History, Washington, DC

INFORMATION AND DATA TECHNOLOGY

Advancing University Career Paths in Interdisciplinary Data-Intensive Science
Cecilia Aragon and William Howe, University of Washington, Seattle

Beyond Silicon: New Materials for 21st Century Electronics
Glennda Chui, SLAC National Accelerator Laboratory, Menlo Park, CA

Differential Privacy: Analyzing Sensitive Data and Implications
Salil Vadhan, Harvard University, Cambridge, MA; Cynthia Dwork, Microsoft Research, Mountain View, CA

From the Grid to the Cloud: Computing for Big (and Small) Science
Terry O'Connor, Science and Technology Facilities Council, Swindon, United Kingdom; Vincenzo Napolano, National Institute for Nuclear Physics, Rome, Italy

Holistic Computing Risk Assessment: Privacy, Security, and Trust
L. Jean Camp and Diane Henshel, Indiana University, Bloomington

Improving Quality of Life By Transforming Images to Health Care Decisions
Ram D. Sriram, National Institute of Standards and Technology, Gaithersburg, MD; Ramesh Jain, University of California, Irvine; Donald Henson, Uniformed Services University, Bethesda, MD

Privacy in an Era of Big Data: Directions, Advances, and Reflections
Ersin Uzun, Palo Alto Research Center, CA

Seeing Earth in the "Light" of Gravity: New Views Through Satellite Geodesy
Barbara Wankerl and Patrick Regan, Technical University of Munich, Germany

The Convergence of Massive Data and Complex Control in the Future U.S. Power Grid
Jeffrey Taft, Pacific Northwest National Laboratory, Richland, WA

Visualization Insights from Big Data: Envisioning Science, Engineering, and Innovation
Katy Borner, Indiana University, Bloomington; Joseph E. Sabol, Chemical Consultant, Racine, WI

Wise Computing: Collaboration Between People and Machines
Kazuo Iwano and Yosuke Takashima, Japan Science and Technology Agency, Tokyo; Tateo Arimoto, National Graduate School for Policy Studies, Tokyo, Japan

MEDICAL SCIENCES AND PUBLIC HEALTH

Affordable Diagnostics for All: High-Resolution Medical Imaging for Saving Lives
Comert Kural, Ohio State University, Columbus; Ahmet Yildiz, University of California, Berkeley

Antibiotic Resistance: An Environmental Problem Threatening Global Health Care
Anneli Waara, Uppsala University, Sweden

Bugs Without Borders: A Data-Driven Approach to Tackling Antimicrobial Resistance
Stefania Di Mauro-Nava, British Consulate-General, San Francisco, CA; Lindsay R. Chura, British Embassy, Washington, DC

Cannabis and Medicine: A New Frontier in Therapeutics
Mark Ware and Julie Robert, McGill University Health Center, Montreal, Canada

Challenges in Communicating about Stem Cells
Judy Illes and Julie M. Robillard, University of British Columbia, Vancouver, Canada

Chronic Pain: No Longer an Invisible Disease
Linda Koffmar, Uppsala University, Sweden

E-Cigarettes: Killing Me Softly or Our Greatest Public Health Opportunity?
Aidan Gilligan, SciCom—Making Sense of Science, Brussels, Belgium; Julian Kinderlerer, European Commission, Brussels, Belgium

Imaging the Future of Cancer Research
Eleni Zika and Sally Donaldson, European Research Council, Brussels, Belgium

Informatics and Bioimaging: New Ways to Better Medicines
Deanna L. Kroetz, University of California, San Francisco

Innovations in Clinical Trial Registers
Victoria Murphy, Sense About Science, London, United Kingdom

Obesity and Microbiome: Concepts and Contradictions
Mary C. Dinauer, Washington University School of Medicine, St. Louis, MO; Shili Lin, Ohio State University, Columbus

Responsible Data Sharing for Clinical Trials
David L. DeMets, University of Wisconsin, Madison

The Digital Human Head: 3D Imaging and the Restoration of Craniofacial Function
Luisa Ann DiPietro, University of Illinois, Chicago

The Human Vaccines Project: Transforming the Future of Vaccine Development
Wayne Koff, International AIDS Vaccine Initiative, New York City

Worldwide Epidemic of Senile Dementias: Challenges of Pre-Clinical Treatment
Edward J. Goetzl, University of California, San Francisco

PHYSICS AND ASTRONOMY

A Science Cycle: From Novel Imaging to Novel Physics, and Vice Versa
Norman Chonacky, Yale University, New Haven, CT

Advancing Solar and Space Physics: Combining Observations, Models, and Analysis
Michael Wiltberger, National Center for Atmospheric Research, Boulder, CO

Building Galaxies: Some Assembly Required
Mark T. Adams, National Radio Astronomy Observatory, Charlottesville, VA

Celebration of 2015: The International Year of Light
Anne Matsuura, The Optical Society, Washington, DC; Anthony Johnson, University of Maryland, Baltimore County; Phil Bucksbaum, Stanford University, CA

Cosmic Neutrinos
Marcela S. Carena, University of Chicago/Fermi National Accelerator Laboratory, Batavia, IL; Graciela Gelmini, University of California, Los Angeles

Exoplanets: New Worlds Aplenty
William Joseph Borucki and Steve B. Howell, NASA Ames Research Center, Moffett Field, CA

General Relativity at 100: Looking Forward and Looking Back
Beverly Berger, International Society on General Relativity and Gravitation, Livermore, CA; Stanley Whitcomb, California Institute of Technology, Pasadena

Imaging the Unseen: Advanced Adaptive Optics Enabling Scientific Discovery
Lisa A. Poyneer, Lawrence Livermore National Laboratory, CA

Innovations in Accelerator Science
Maria Spiropulu, California Institute of Technology, Pasadena; Saul Gonzalez Martirena, U.S. Office of Science and Technology Policy, Washington, DC

Scanning the Southern Skies: Harnessing Big Astronomy Data with the Square Kilometer Array
William Garnier, Square Kilometer Array Organization, Macclesfield, United Kingdom; Daan du Toit, South African Science Ministry, Brussels, Belgium

Solar System Exploration by Remote Imaging
Rolf M. Sinclair, University of Maryland, College Park; Fuk K. Li, California Institute of Technology, Pasadena

The Cosmic Microwave Background: Window into New Physics
Glenn E. Roberts, SLAC National Accelerator Laboratory, Menlo Park, CA; Jay Marx, California Institute of Technology, Pasadena

The Hubble Space Telescope: 25 Years of Imaging the Cosmos
Jennifer Wiseman, NASA Goddard Space Flight Center, Greenbelt, MD; Ray Villard, Space Telescope Science Institute, Baltimore, MD

Transformational Opportunities of Quantum Information Technologies
Martin Laforest, University of Waterloo, Canada; Jenny Hogan, National University of Singapore, Malaysia

Wave-Particle Duality of Neutrons, Atoms, and Molecules
Charles W. Clark, Joint Quantum Institute, Gaithersburg, MD

What Can We Expect from the Second Run of LHC in 2015?
James Gillies, European Organization for Nuclear Research (CERN), Geneva, Switzerland; Jon Weiner, Lawrence Berkeley National Laboratory, CA

PUBLIC POLICY

Active SETI: Is It Time To Start Transmitting to the Cosmos?
Jill C. Tarter and David C. Black, SETI Institute, Mountain View, CA

Big Data: Challenges and Social Impacts
Chris Tyler, U.K. Parliamentary Office of Science and Technology, London; Timothy M. Persons, U.S. Government Accountability Office, Washington, DC

Creating and Using New Information To Promote Innovations in Crime and Justice Policy
William Alex Pridemore, Georgia State University, Atlanta

Crisis in Quantitative Training for Biomedical Science
David L. DeMets, University of Wisconsin, Madison

Dementia: Research Milestones and Policy Priorities
Lindsay R. Chura, British Embassy Washington, DC

Developing Data-Driven Policies To Cure and Prevent Age-Related Diseases
Sharotka Simon, AAAS Science and Technology Fellow, Diplomacy Security and Development Program, Washington, DC

Integrity of Science
Thomas Arrison, National Academy of Sciences, Washington, DC

Making Sense of Chaos: A Path to Understanding Environmental Chemicals and Health
Louis J. D'Amico, Kacee Deener, and Samantha J. Jones, U.S. Environmental Protection Agency, Washington, DC

Memories Imaged and Imagined: How the Science of Memory Will Challenge the Law
Henry Greely, Stanford University Law School, CA; Nita Farahany, Duke University, Durham, NC

Metrics for Science Policy and Policy for Metrics
Ismael Rafols, University of Sussex, Brighton, United Kingdom; Michele S. Garfinkel, European Molecular Biology Organization, Heidelberg, Germany

Open Access To Your Emails: Tensions Between Academic Freedom and Open Records Laws
Michael Halpern, Union of Concerned Scientists, Washington, DC

Science During Crisis: Amidst Oil Spills, Hurricanes, and Other Disasters
Gary Machlis, U.S. Department of the Interior Strategic Sciences Group, Clemson, SC; Kristin Ludwig, U.S. Geological Survey, Reston, VA

Science Policy Professionals in the Era of Big Data: In Honor of Stephen D. Nelson
Susan E. Cozzens, Georgia Institute of Technology, Atlanta

The Human Microbiome: Implications of the Microcosm Within Us
Se Y. Kim and Jennifer Wiseman, AAAS Center for Science, Policy, and Society Programs, Washington, DC

SUSTAINABILITY AND RESOURCE MANAGEMENT

Beyond Intensification: Measuring the "Sustainable" in Sustainable Intensification
Jerry Glover, U.S. Agency for International Development, Washington, DC

Can Our Ocean Commons Be Sustainably Managed? Innovative Strategies for the High Seas
Cassandra M. Brooks and Larry Crowder, Stanford University, CA

Data Infrastructure for Sustainability
Thomas Dietz, Michigan State University, East Lansing; Adam Douglas Henry, University of Arizona, Tucson

Drivers of Tropical Forest Regeneration
Ashwini Chhatre, University of Illinois, Urbana-Champaign

Dynamic Ocean Management: Supporting Ecological and Economic Sustainability
Sara M. Maxwell, Stanford University, CA; Rebecca Lewison, San Diego State University, CA

Ecology, Economics, and Engineering of Nature-Based Coastal Defenses
Jane Carter Ingram, Wildlife Conservation Society, New York City; Michael W. Beck, The Nature Conservancy, Santa Cruz, CA

Feeding 9 Billion+: Information and Imaging for Innovation in Next-Generation Agriculture
Vijaya Gopal Kakani, Oklahoma State University, Stillwater

Information Accelerators: Using Online Tools To Address Sustainability Challenges
Richard Sharp and Rebecca Chaplin-Kramer, The Natural Capital Project, Stanford, CA

Sustainable Intensification in Agriculture: New Scientific Approaches
Karl Zimmerer, Pennsylvania State University, State College

AAAS 2015 ANNUAL MEETING
INNOVATIONS, INFORMATION, AND IMAGING

AAAS, publisher of *Science*, thanks the sponsors and supporters of the 2015 Annual Meeting

The Art and Science of Traditional Medicine
Part 2: Multidisciplinary Approaches for Studying Traditional Medicine

Sponsored by

北京中医药大学
BEIJING UNIVERSITY OF CHINESE MEDICINE

香港浸會大學
HONG KONG BAPTIST UNIVERSITY

Contents

In this second of three special supplements, herbal genomics as a novel approach for revolutionizing research on, and ultimately use of, traditional herbal medicines and other materia medica, as well as advances in their quality control and standardization, is highlighted. A prominent focus is the U.S. Food and Drug Administration's practical framework for developing botanicals (including traditional medicines) into new drugs based on the same standards as small molecule drugs. The application of mechanistic studies to drug discovery and development from traditional therapies is discussed, with an emphasis on preclinical toxicology assessments, pharmacovigilance, comparative effectiveness research, and the practice of "P4" medicine, particularly in the context of influenza, ischemic heart disease, stroke, and cancer.

Editorial Team

Tai-Ping Fan, Ph.D. (Guest project editor)
University of Cambridge, UK

Josephine Briggs, M.D.
National Center for Complementary & Alternative Medicine, NIH, USA

Liang Liu, M.D., Ph.D.
Macau University of Science & Technology, Macau SAR, China

Aiping Lu, M.D., Ph.D.
Hong Kong Baptist University, Hong Kong SAR, China

Jan van der Greef, Ph.D.
University of Leiden and TNO, The Netherlands

Anlong Xu, Ph.D.
Beijing University of Chinese Medicine, China

Editor: Sean Sanders, Ph.D.
Assistant Editor: Tianna Hicklin, Ph.D.
Proofreader/Copyeditor: Yuse Lajiminmuhip
Designer: Amy Hardcastle

Bill Moran, Global Director
Custom Publishing
bmoran@aaas.org
+1-202-326-6438

Ruolei Wu, Associate Director, Asia
Custom Publishing
rwu@aaas.org
+86-186-0082-9345

Articles

ILLUSTRATION (FRONT) CHARLOTTE LOKIN

Herbal genomics: Examining the biology of traditional medicines

Authors:
Shilin Chen[1*],
Jingyuan Song[2]
Chao Sun[2],
Jiang Xu[1],
Yingjie Zhu[1],
Rob Verpoorte[3],
Tai-Ping Fan[4*]

Traditional herbal medicines, such as plant- and fungi-based remedies, have been used for more than 5,000 years. However, the genetic background, the agricultural traits, and the medicinal quality of most traditional herbs are poorly understood. With rapid advances in high throughput sequencing technologies and greatly reduced costs, a new discipline called "herbal genomics" has emerged. Researchers are now systematically categorizing medicinal herbs by sequencing, assembling, and annotating their genomes, and by analyzing their genes' functions. The genomes of some commonly used herbs have already been sequenced, such as Lingzhi (*Ganoderma lucidum* or "mushroom of immortality"). This species has provided an effective model system that has facilitated the study of the biosynthetic pathways of secondary metabolites in medicinal fungal species (*1*). Genomic information, together with transcriptomic, proteomic, and metabolomic data, can therefore be used to predict secondary metabolite biosynthetic pathways and their regulation, triggering a revolution in discovery-based research aiming to understand the genetics and biology of herbs.

Herbal genomics provides an effective platform to support the chemical and biological analyses of complex herbal products that may contain more than one active component. Therefore, it is now being applied to many areas of herb-related biological research to help understand the quality of traditional medicines and for molecular herb identification through the establishment of an herbal gene bank. Moreover, functional herbal genomics can contribute to model herb research platforms, geoherbal research, genomics-assisted herb breeding, and herbal synthetic biology, all of which are important for securing the sourcing of the medicinal plants and their active compounds in the future.

Creating model herbs

With the recent developments in biotechnology and genomics, several species including *Ganoderma lucidum*, *Salvia miltiorrhiza*, and *Catharanthus roseus* have emerged as valuable models for studying the genetics and metabolic activities of herbs. These species have been shown to synthesize active pharmaceutical components, including triterpenes, diterpene quinone, and indole alkaloids. Although the core biosynthetic pathways of secondary metabolites in herbs are conserved, downstream pathways have evolved and became highly diverse (*2*). Therefore genes from different cultivars of medicinal herbs or evolutionarily related species can be evaluated using these herbal models to understand the mechanisms underlying natural variation. These model systems can also be used to identify novel biosynthetic pathways for convergent secondary metabolites in closely related herb species. Recent advances in genome editing have provided feasible approaches by introducing or altering specific alleles; hence, genetic control over metabolites can be investigated (*3*). Although the elucidation of biosynthetic pathways is one of their most appealing features, model herbs can also provide information on perennial habits, development patterns, cultivation requirements, and resistance to environmental or biological insult (Figure 1).

Biological basis of geoherbalism

The Chinese concept of geoherbalism encompasses the use of "authentic" or "superior" herbs, which are produced in a specific region or environment, to generate remedial products that have a high efficacy. Through the application of new 'omics technologies to geoherbalism, information can be obtained concerning the optimal growth conditions of medicinal products and specific herb genotypes, allowing both genetic and environmental factors to be taken into account when considering herbal growth. 'Omics provide new and powerful tools to elucidate the molecular basis underlying geoherbalism and to select elite varieties. Creating herb pangenomes—the entire genetic code for all of the strains within a given species—can provide insights into identifying the "core genomes" and "dispensable genomes" of the species as well as the individual genetic variations that exist in different regions or ecological circumstances. Environmental stressors can lead to epigenetic modifications; techniques such as DNA methylome analysis, chromatin immunoprecipitation (ChIP)-sequencing, and small RNA sequencing are useful for investigating the influence of epigenetic factors. In addition, soil microbes can affect an herb's environment, and metagenomic analysis of soil microbial populations can point to important interactions between microbes and herbs that may alter growth conditions (*4*).

Targeted herb breeding

Molecular breeding requires the availability of polymorphic markers and/or information about trait-associated genes. Since they are considered minor crops, herbs have been limited in genomics-assisted improvements due to the high cost; however, next generation sequencing and its increasing affordability have dramatically accelerated marker selection breeding programs through the sequencing of wild varieties and different cultivars of herbs that represent a valuable reservoir of genetic diversity. Herbal 'omics research has

[1]WHO Collaborating Centre for Traditional Medicine, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing, China
[2]Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China
[3]Natural Products Laboratory, IBL, Leiden University, Leiden, The Netherlands
[4]Department of Pharmacology, University of Cambridge, Cambridge, UK
*Corresponding Authors: slchen@implad.ac.cn (S.C.) and tpf1000@cam.ac.uk (T.F.)

accelerated the identification of many functional genes in model species and has also allowed the development of functional markers specific to the production of desired compounds, information that can be used for targeted molecular breeding (*5*).

Researching herbal synthetic biology

Although herbs are sources of novel and known therapeutic compounds, problems in sourcing are common. Biotechnology and genetic engineering offer approaches to alternative production methods. Metabolic engineering of medicinal plants has been studied extensively, resulting in, for example, *Atropa belladonna* plants producing scopolamine instead of atropine. However, it is clear that in order to improve the overall production of a plant compound, the overexpression of the primary coding genes or even regulatory genes in the pathway are not sufficient, since compartmentalization, transport, storage, and co-factor availability may be important rate limiting factors. A better understanding of pathways involved would build a foundation for a more comprehensive approach to

FIGURE 1. Model herb systems and their applications.

metabolic engineering. This is the goal of herbal synthetic biology, which involves the alteration or de novo synthesis of genomes, with the potential to address resource and purity issues. Furthermore, natural products for drug discovery can be structurally diversified by combining and introducing plant metabolic pathways into other organisms, such as bacteria or yeast (*6*). The conventional practice in herbal synthetic biology is to introduce the heterologous biosynthetic pathway into an expression system able to produce the products. However, a different approach for the large-scale production of a pure compound could be the engineering of an entirely novel synthetic genome, as described for *Mycoplasma* (*7*).

Defining a molecular identity

DNA barcoding is revolutionizing the practice of herbal identification, utilizing the concept of "one sequence, one species." Standardized DNA barcoding identification systems are available, but the process can be tedious. Analysis of a plastid genome as a superbarcode is a promising alternative for closely related species or cultivars that cannot be unambiguously distinguished by traditional DNA barcoding (*8*, *9*). With the increasing availability of DNA barcodes, current market issues with herbal medicines that result from the use of inferior substitutes, adulterants, and counterfeits could be resolved. Overall, a standardized identification system based on DNA barcoding can play an important role in controlling the quality of traditional

medicines through the accurate identification of herbal materials.

Constructing an herbal gene bank

Herbal genetic information is being accumulated with increasing speed, making the need for a common platform for integrated and consolidated access to all 'omics data paramount. Several herb-related databases have already been developed to categorize current resources, including genomic information (http://herbalgenomics.org), transcriptomic information (http://medicinalplantgenomics.msu.edu), a DNA barcode database (http://tcmbarcode.cn/en), and a metabolic pathway database (http://cathacyc.org). However, these distributed resources are not subjected to long-term maintenance and require bioinformatics skills to use. A comprehensive and easily accessible database is required that stores molecular and biological data for herbal medicines. The DNA/protein sequences and metabolomes of herbs can be integrated into such a database (*10*, *11*). With improved bioinformatics approaches, genomic and chemical information can be used to identify the biosynthetic pathways of secondary metabolites leading to the design of more efficient and targeted searches for plant- and fungi-based remedies (*12*).

Despite its success thus far, herbal genomics still faces significant technological and ethical challenges. For instance, there have been only a few well-assembled herbal genomes released to date, partly because of their complexity. High
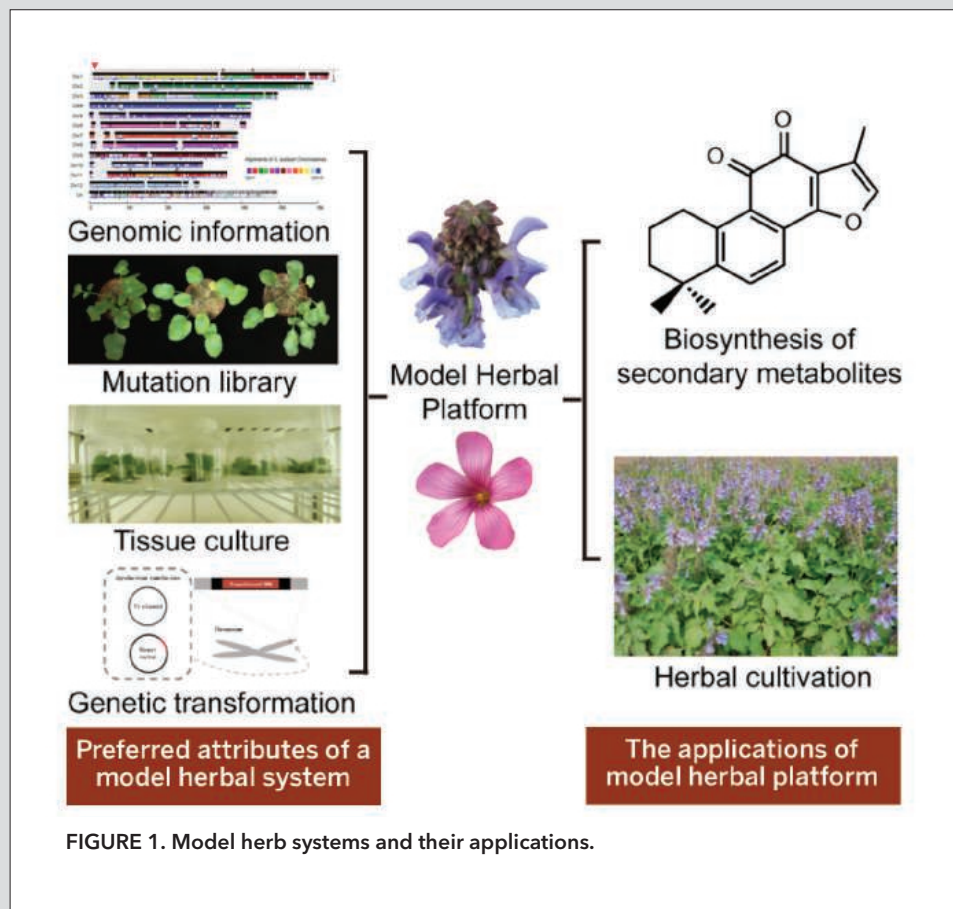
heterozygosity, repetition-rich DNA sequences, and polyploidy are factors that impede data assembly from short-read, whole-genome shotgun sequencing. Furthermore, the lack of high throughput methods reduces the efficiency of identifying enzymes and pathways involved in the biosynthesis of secondary metabolites. There have also been ethical and biosecurity concerns regarding synthetic biology expressed by the scientific community and the public. Nevertheless, herbal genomics provides an unprecedented opportunity to revolutionize the use and acceptance of traditional herbal medicines, while contributing to the knowledge base essential for further proteomic and metabolomic studies.

References

1. S. Chen *et al.*, *Nat. Commun.* **3**, 913 (2012).
2. S. J. Gagne *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 12811 (2012).
3. Q. Shan, Y. Wang, J. Li, C. Gao, *Nat. Protoc.* **9**, 2395 (2014).
4. P. R. Hirsch, T. H. Mauchline, *Nat. Biotechnol.* **30**, 961 (2012).
5. R. K. Varshney, A. Graner, M. E. Sorrells, *Trends Plant Sci.* **10**, 621 (2005).
6. F. Geu-Flores, N. H. Sherden, V. Courdavault, *Nature* **492**, 138 (2012).
7. D. G. Gibson *et al.*, *Science* **329**, 52 (2010).
8. X. Li *et al.*, *Biol. Rev. Camb. Philos. Soc.* doi:10.1111/brv.12104 (2014).
9. S. Chen *et al.*, *Biotechnol. Adv.* **32**, 1237 (2014).
10. B. Berger, J. Peng, M. Singh, *Nat. Rev. Genet.* **14**, 333 (2013).
11. M. Yandell, D. Ence, *Nat. Rev. Genet.* **13**, 329 (2012).
12. L. Chae, T. Kim, R. Nilo-Poyanco, S. Y. Rhee, *Science* **344**, 510 (2014).

A holistic approach to the quality control of traditional Chinese medicines

Authors:
De-an Guo[1*],
Wan-Ying Wu[1],
Min Ye[2],
Xuan Liu[1],
Geoffrey A. Cordell[3]

Plants and traditional medicines (TMs) are used around the world for the prevention and treatment of diseases as well as the sources of numerous prescription and over-the-counter therapeutics ([1]). In many cases, these TMs have been used for thousands of years and are still largely harvested from the wild. However, the quality control for the growth and isolation of most TMs is poor or nonexistent. Patients and practitioners alike need to be confident of the quality, safety, efficacy, and consistency (QSEC) of TMs, which requires standardization of all aspects of the plant preparation. This begins with the identification of the correct plant, and includes isolation and characterization of all bioactive constituents. The evidence-based criteria to determine the QSEC of TMs differs considerably across the globe. As a result, the evidentiary standards required for the marketing of TM products can vary greatly from country to country.

As acknowledged by the famous Chicago architect, Louis Sullivan, "form forever follows function." For patients and practitioners, the expectation for the TMs is that they must be effective (function). Strategies (form) to assure this outcome require a strong evidence base and necessitate overcoming certain barriers to success. In many societies, the acceptance of TMs precedes integration into the health care system, causing them to face philosophical and regulatory barriers due, at least in part, to the absence of an evidence base.

The World Health Organization has recognized some of these issues ([1, 2]) and, through the Beijing Declaration, has encouraged the integration of evidence-based TMs into national health care systems and promoted regulations and standards that "ensure appropriate, safe, and effective use of traditional medicine" ([3]).

In order to fully integrate TMs into national health care systems, and provide an evidence-based justification for their use, certain entrenched customs and beliefs need to be addressed. These include the ideas that a traditional medicine used for thousands of years must be safe and effective; that simply using the "correct" plant is adequate; that the biological

[1]National Engineering Laboratory for TCM Standardization Technology, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, Shanghai, China
[2]School of Pharmaceutical Sciences, Peking University, Beijing, China
[3]Natural Products Inc., Evanston, IL 60203, U.S.A. and Department of Pharmaceutical Sciences, College of Pharmacy, University of Florida, Gainesville, FL
*Corresponding Author: daguo@simm.ac.cn

effects will be consistent, irrespective of the geographic origin of the plant, the plant part used, or the method of plant preparation; that older or cultivated plant material is less effective than fresh, wild plants; that complex plant mixtures are necessary for effectiveness, but cannot be standardized; and that the traditional knowledge and the particular medicinal plant will always be available. The application of contemporary information, together with botanical, chemical, biological, and clinical research, provides an evidence-based research agenda for traditional medicine that serves both the practitioner and the patient (*4*). The next step, standardization (*5*), includes proper plant authentication through DNA barcoding (*6, 7*), and the chemical profiling and quantification of all bioactive constituents in the material (*8, 9*).

Traditional Chinese medicine (TCM) is an ancient, holistic treatment system established through empirical evaluation, and exists in many related forms in Greater China, Japan, Korea, Vietnam, Malaysia, and Singapore. It seeks to restore energy (*qi*) and balance (*yin* and *yang*) through the use of medicinal plants, fungi, animal products, and minerals, and, superficially, appears quite different from the reductionist approach of Western medicine.

However, modern biomedical science is now embracing the concept of systems biology, which views human diseases as the result of a multifactorial instability in homeostasis (*10*). Treatment of cancer or HIV-AIDS now involves a cocktail of drugs targeting different mechanisms of action. At the same time, TCM is embracing network pharmacology, which investigates how the major constituents in a plant (or plants) act on various biological pathways to produce multiple, synergetic actions (*11, 12*).

The quality control (QC) of TCMs should begin in the field and continue throughout the production process. Developing a QC system for a TCM preparation is a critical, foundational step for the manufacture of a standardized product suitable for biological and clinical studies. A typical TCM preparation, often consisting of an admixture of multiple plants, represents a vast array of chemical constituents that work synergistically to bring about the observed therapeutic effects. Establishing a chemical and biological quality standard for such a complex TCM preparation represents a daunting analytical challenge. A comprehensive analytical approach, integrating chemical, metabolic, and biological methods, was therefore developed to serve as a paradigm for establishing quality standards for TCMs.

Development of systemic analytical methods

A comprehensive analytical assay that can provide the chemical fingerprint of each individual component of a complex preparation is necessary to monitor quality and biological consistency of TCMs (Figure 1) (*13, 14*).

In our laboratory, liquid chromatography/multiple-stage mass spectrometry (LC-MS^n) techniques have been employed in order to explore the chemical profile of various TCM plant materials. Ginseng, notoginseng, and American ginseng (*Panax* species, Araliaceae) are commonly used in TCM formulae and contain ginsenosides like triterpene and steroid saponins as their active ingredients. These three similar plants have differing clinical efficacies and are easily confused, particularly in their post-processing forms.
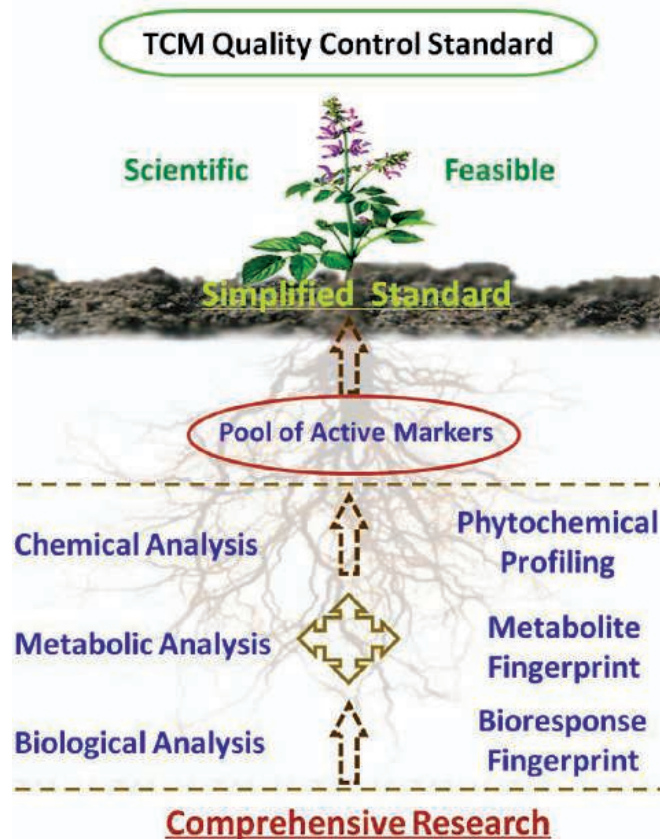
FIGURE 1. A systemic traditional Chinese medicine (TCM) quality research approach: from comprehensive research to simplified standard.

A total of 623 ginsenosides, including 437 potential new ginsenosides, were characterized from the three plants, allowing specific biomarkers to be developed that can unequivocally differentiate between them (*15*). This technique was similarly applied to a number of TCM herbs, including *Salvia miltiorrhiza*, *Ganoderma lucidum*, *Glycyrrhiza uralensis*, and *Rheum palmatum*.

In one case, a metabolic fingerprinting technique identified seventy metabolites of the major tanshinones and salvianolic acids of *S. miltiorrhiza* in rats after oral administration of the plant, enabling the determination of metabolic pathways and excretion routes (*16, 17*). Combining both chemical and biological analyses provides an effective strategy for revealing active components. Given the complex metabolic matrix of *S. miltiorrhiza,* a three-tier strategy involving analysis of single compounds, extracted fractions, and the whole herb was adopted. A multi-level biological approach was used that integrated pharmacology, molecular biology, and systems biology. The target proteins and mechanisms of action were found to be impacted at the molecular, cellular, tissue, and whole animal levels, supporting the contention that TCMs work on multiple targets through multiple pathways. In the case of *S. miltiorrhiza*, salvianolic acid B was determined to modulate

FIGURE 2. Holistic quality control model for traditional Chinese medicines. TLC, thin layer chromatography; LC-MS^n, liquid chromatography/multiple-stage mass spectrometry.
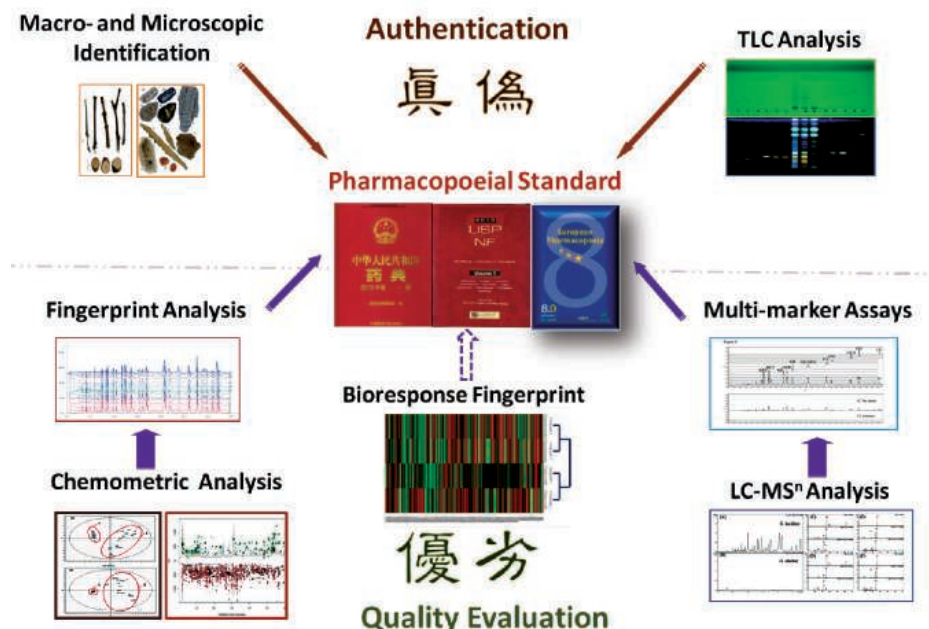
several molecular targets, including matrix metalloproteinase 9, epidermal growth factor receptor, and integrin (*18*), and the major tanshinone derivatives were found to be cardioprotective and antioxidant agents (*19*).

Elaboration of an overall quality standard

A new quality control model that combines analytical fingerprinting to monitor batch-to-batch consistency, and a multicomponent assay to assure authenticity and quality, has been developed and effectively applied to several Chinese herbal materials and their preparations (Figure 2). Based on the salvianolic acids and the tanshinones, an overall quality standard for *S. miltiorrhiza* was established using the single standard to determine multicomponents (SSDMC) method, which uses a single reference standard to quantify the content of many related compounds in a mixture, and has been adopted by both the Chinese and United States pharmacopoeias. It is recommended that only pharmacopoeial quality material be used in clinically evaluations in order to ensure product consistency and safety.

Conclusions

A multitude of intrinsic and extrinsic factors affect the quality, consistency, and stability of medicinal plants and their metabolites (*12, 13*). The regulated application of Good Agricultural and Collection Practices, Good Laboratory Practices, Good Manufacturing Practices, and Good Clinical Practices, are necessary to assure high quality, safe, effective, and consistent TCM products for practitioners and patients (*5*). In order for TCM products to be accepted into a global, evidence-based health care system, it is imperative for robust international standards and procedures to be developed that govern their growth, collection, processing, and administration. Furthermore, the extreme complexity of TCMs necessitates the integration of new technologies

and strategies for proper analysis of bioactive constituents, allowing the targets and pathways impacted to be fully understood (*20*).

References

1. World Health Organization, The Regional Strategy for Traditional Medicine in the Western Pacific (2011–2020). WHO, Regional Office for the Western Pacific, Manila, Republic of the Philippines. 71 pp. (2012).
2. World Health Organization, WHO Traditional Medicine Strategy 2014–2023. WHO, Geneva, Switzerland. 71 pp. (2014).
3. World Health Organization, http://www.who.int/medicines/areas/traditional/congress/beijing_declaration/en/ (2008). Accessed on August 15, 2014.
4. G.A. Cordell, *Phytochem. Lett.*, Epub before print (2014), doi: 10.1016/j.phytol.2014.06.002.
5. H. Uzuner *et al.*, *J. Ethnopharmacol.* **140**, 458 (2012).
6. N. Techen, I. Parveen, Z. Q. Pan, I.A. Khan, *Curr. Opin Biotech*. **25**, 103 (2014).
7. Ž. F. Pečnikar, E. V. Buzan, *J. Appl. Genet.* **55**, 43 (2014).
8. J. Liang *et al.*, *J. Chromatogr. A* **1294**, 58 (2013).
9. J. Jing, W. C. Ren, S. B. Chen, M. Wei, H. S. Parekh, *Trends Anal. Chem.* Vol. 44 (2013).
10. J. van der Greef, *Nature* **480**, S87 (2011).
11. G. B. Zhang, Q. Y. Li, Q. L. Chen, S. B. Su, *Evid. Based Complement. Alternat. Med.*, id: 621423 (2013), doi: 10.1155/2013/621423.
12. S. Li, B. Zhang, *Chin. J. Nat. Med*. **11**, 110 (2013).
13. G. A. Cordell, *Planta Med*. **77**, 1129 (2011).
14. G. A. Cordell, M.D. Colvard, *J. Nat. Prod*. **75**, 514 (2012).
15. W. Z. Yang *et al.*, *Anal. Chim. Acta* **739**, 56 (2012).
16. M. Yang, C. R. Cheng, J. L. Yang, D. A. Guo, *Curr. Drug Metab*. **13**, 535 (2012).
17. X. Qiao *et al.*, *AAPS J.* **16**, 101 (2014).
18. L. X. Feng *et al.*, *Proteomics* **11**, 1473 (2011).
19. X. Liu, W. Y. Wu, B. H. Jiang, M. Yang, D. A. Guo, *Trends Pharmacol. Sci*. **34**, 620 (2013).
20. R. Tilton *et al.*, *Chin. Med.*, **5**, 30 (2010).

Evolution of traditional medicines to botanical drugs

Authors:
Sau L. Lee[*],
Jinhui Dou[*],
Rajiv Agarwal,
Robert Temple,
Julie Beitz,
Charles Wu,
Andrew Mulberg,
Lawrence X. Yu,
Janet Woodcock

Botanicals constitute an important source for new drugs (*1*, *2*). To facilitate botanical drug development, the Center for Drug Evaluation and Research (CDER) of the U.S. Food and Drug Administration (FDA) established the Botanical Review Team in 2003 and published its first *Guidance for Industry: Botanical Drug Products* in 2004 (*3*). This guidance represents FDA's thinking and provides recommendations on quality, nonclinical, clinical, and other unique aspects associated with botanical new drug development through the investigational new drug (IND) and new drug application (NDA) processes. From 2004 to 2013, CDER received over 400 botanical IND applications and pre-IND meeting requests (Table 1). Most of the INDs were allowed to enter phase 2 clinical trials for evaluation of preliminary safety and efficacy of the investigational botanical products in patients. FDA approved the first botanical NDA for Veregen (sinecatechins) in 2006 (*4*, *5*) and the second botanical NDA for Fulyzaq (crofelemer) in 2012 (*6*, *7*). These two NDA approvals show that new therapies derived from natural complex mixtures can be developed to meet modern FDA standards of quality, safety, and efficacy. A small number of INDs are currently in phase 3 clinical trials, which may lead to more NDAs in the future.

"Totality-of-evidence" approach

For new botanical products intended to be marketed as drugs in the United States, applicants need to provide evidence of safety and efficacy at the same level as is expected for small-molecule products. Botanical products also need to meet standards for product quality, so the marketed product batches deliver a therapeutic effect consistent with that observed for product batches tested in the clinical studies (i.e., therapeutic consistency). However, quality control of botanical products is challenging because these products contain complex mixtures in which the active components may not be known and may also have considerable batch-to-batch variations (e.g., in chemical composition). The conventional chemistry, manufacturing, and controls (CMC) data (primarily from chemical testing) used to ensure the quality of small-molecule products may be insufficient for botanical products. To address this challenge, FDA has developed a "totality-of-evidence" approach based on knowledge and experience acquired from the review of botanical IND and NDA submissions. In addition to conventional CMC data, this integrated approach considers other evidence including raw material control, clinically relevant bioassay(s), and other non-CMC data (including clinical data on the dose-response

generated based on multiple batches of the drug product). The degree of reliance on these other data for ensuring consistency of quality depends on the extent to which the botanical mixture can be characterized and quantified. Such an integrated approach is best explained and illustrated through the FDA's experience with the first two botanical NDAs.

Veregen

Veregen (sinecatechins) ointment was approved for topical treatment of external genital and perianal warts (*4*, *5*). It contains 15% (w/w) sinecatechins, a botanical substance that is a partially purified fraction of the water extract of green tea leaves from *Camellia sinensis* (L.) Kuntze (Theaceae). Sinecatechins is a mixture of catechins (85% to 95% by weight of the total drug substance) and other green tea components. These catechins include more than 55% of epigallocatechin gallate and other catechin derivatives (Figure 1A) (*4*).

For FDA approval, the safety and efficacy of Veregen were established in two randomized, double-blind, vehicle-controlled clinical studies. The study results showed that the treatment group receiving Veregen demonstrated a significantly higher response rate compared with the vehicle-control group (53.6% vs. 35.3%) (*4*). The response rate was defined as the proportion of patients with complete clinical (visual) clearance of all external genital and perianal warts (baseline and new) by week 16. Although extensive human consumption of green tea and literature research data on catechins provided considerable reassurance of the safety of topical use of catechins, this information was not considered sufficient to support the NDA approval for Veregen. Standard nonclinical studies were conducted to support the safety of Veregen.

As a naturally occurring mixture in which the active components are not well defined, the identified individual major and minor chemical components in Veregen need to be monitored and controlled for each marketed product batch. In the absence of data correlating chemical properties and clinical response, the acceptance ranges for these components were primarily established based on their levels observed in the multiple batches tested in the clinical studies. As significant variations of catechins and other chemical components have been identified from the tea leaves of different cultivars (*8*), this botanical product can have considerable batch-to-batch variability (e.g., in total catechins and ratios of different catechins). In addition, although the majority of components can be adequately characterized and quantified, there may still be some residual uncertainties about the chemical nature of minor components in sinecatechins. Therefore, to ensure consistent quality for Veregen, FDA considered two other important pieces of information from the application, summarized below:
- The botanical raw materials for future marketed product

Center for Drug Evaluation and Research, Food and Drug Administration, Silver Spring, MD

[*]Corresponding Authors: sau.lee@fda.hhs.gov (S.L.L.) and jinhui.dou@fda.hhs.gov (J.D.)

TABLE 1. Number of botanical pre-INDs and INDs submitted to CDER (2004–2013).

Submissions	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	Total
Pre-IND	12	6	11	7	14	10	10	8	5	8	91
IND	21	38	22	39	27	27	27	33	33	51	318
Pre-IND and IND	33	44	33	46	41	37	37	41	38	59	409

batches were limited to the cultivars used in clinical studies of Veregen, which were identified by botanists and tea experts based on the characteristics and history of the specimens collected at the tea farms. It was emphasized that cultivation sites or farms should follow the principles of good agricultural and collection practices (GACP). These control measures reduce the variability in chemical composition at the plant and raw material levels, so the marketed product batches will exhibit a natural variability similar to the range observed in batches used in clinical studies.

• The efficacy results of clinical studies showed no significant difference in clinical response between two doses (i.e., 10% and 15%). This observation suggests that subtle variations in any uncharacterized fractions may not have an impact on the therapeutic effect.

Fulyzaq

Fulyzaq (crofelemer) is the first FDA-approved drug for symptomatic relief of noninfectious diarrhea in patients with HIV/AIDS on antiretroviral therapy (6, 7). It is available as a delayed-release tablet containing 125 mg crofelemer, a botanical drug substance derived from the red latex of *Croton lechleri* Müll.Arg. (Euphorbiaceae). Crofelemer is an oligomeric proanthocyanidin mixture primarily composed of (+)-catechin, (-)-epicatechin, (+)-gallocatechin, and (-)-epi-gallocatechin monomer units linked in random sequence (Figure 1B) (6). The red latex, commonly known as Dragon's Blood, has been commonly used in South America as an herbal medicine for treating diarrhea (9, 10). Results from the randomized, double-blind, placebo-controlled and placebo-free clinical study showed that a proportion of HIV-positive patients (who were on stable antiretroviral therapy with a history of diarrhea) in the crofelemer 125 mg bid group was significantly higher than that in the placebo group in experiencing clinical response (17.6% vs. 8.0%) (6). The clinical response was defined as less than or equal to two watery bowel movements per week during at least two of the four weeks of the placebo-controlled phase.

Compared with sinecatechins, the chemical characterization of crofelemer presented an even greater challenge. This drug consists of a mixture of oligomers that vary in composition, sequence, and length, which precluded adequate separation and quantitation of proanthocyanidin oligomers based on multiple conventional chromatographic methods. Additional chromatographic, spectroscopic, spectrometric, and acid hydrolysis methods, which likely represented the state of the art of analytical technologies at the time, were needed to provide a comprehensive characterization of crofelemer. These analytical methods collectively revealed extensive information on the structural signatures of crofelemer (e.g., the composition of proanthocyanidin oligomers), but they were ultimately considered insufficient to support the character-

ization and quality control of this complex botanical mixture (e.g., difficulty of quantification of proanthocyanidins at the molecular level and limitations of the hydrolysis method that provided less than optimal yields) (11).

Because the degree of uncertainty regarding chemical characterization of crofelemer was much greater compared with sinecatechins, FDA concluded that in addition to raw material control and clinical data from multiple doses and product batches, a clinically relevant bioassay to assess the drug product activity was warranted for the approval of Fulyzaq. The non-CMC data that FDA relied on to ensure consistency of quality for Fulyzaq are summarized below:

• Similar to Veregen, botanical raw material controls included implementation of GACP and restricted harvesting of botanical raw material to specific eco-geographic regions (EGRs). These control measures help to reduce the variability at the plant and raw material levels.

• Crofelemer's mechanism of action (MOA) was reported in the literature as a potent inhibitor of both the cyclic adenosine monophosphate-stimulated cystic fibrosis transmembrane conductance regulator chloride ion channel and the calcium-activated chloride ion channel at the luminal membrane of enterocytes (12). Knowledge of crofelemer's MOA led to an important development of a clinically relevant bioassay, which enabled the establishment of clinically relevant specifications. The existence of a clinically relevant bioassay could potentially provide more flexibility for the manufacturer to make postapproval changes (e.g., expansion of EGRs to increase and diversify the botanical raw material supply).

• The clinical data from multiple doses (125–500 mg bid) showed that the drug's effects were not sensitive to the tested doses. This is consistent with in vitro observations and clinical pharmacology data, suggesting that the estimated drug concentrations in the gastrointestinal tract after oral dosing (125 mg bid) are several-fold higher than the concentrations used for maximum chloride ion channel inhibition, which leads to drug saturation at the action sites (7, 12, 13). The clinical data from multiple batches of crofelemer also did not reveal any noticeable differences among drug product batches manufactured by using different drug substance batches. The above information suggested that the natural variations observed in crofelemer were unlikely to have significant impact on the efficacy of Fulyzaq.

Conclusion

The approvals of Veregen and Fulyzaq demonstrated that FDA's science-based "totality-of-evidence" regulatory approach can ensure the consistency of product quality and thereby therapeutic effect of botanical drugs. For quality control purposes, relying on other evidence including raw material control, clinically relevant bioassay(s), and/or clinical

Component	Abbrev.	R1	R2	R3	R4	R5
(–)-Epigallocatechin Gallate	(–)-EGCg	G	OH	–	–	–
(–)-Epicatechin Gallate	(–)-ECg	G	H	–	–	–
(–)-Epigallocatechin	(–)-EGC	H	OH	–	–	–
(–)-Epigcatechin	(–)-EC	H	H	–	–	–
(–)-Gallocatechin Gallate	(–)-GCg	–	–	G	OH	–
(–)-Gallocatechin	(–)-GC	–	–	H	OH	–
(–)-Catechin Gallate	(–)-Cg	–	–	G	H	–
(+)-Catechin	(+)-C	–	–	–	–	H

R = H or OH
range n = 5.5
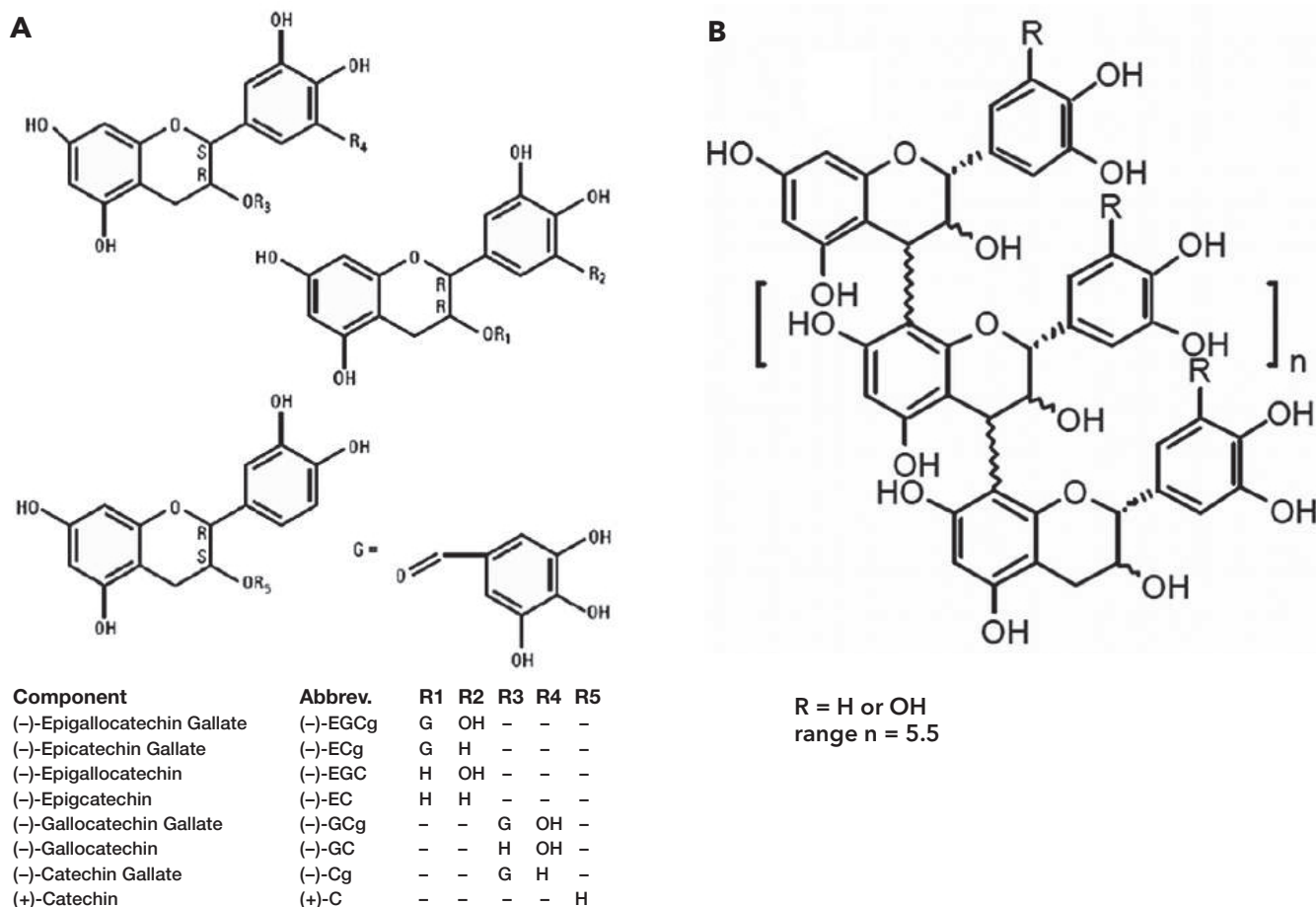
FIGURE 1. (A) Known catechins in sinecatechins (*4*). (B) Proanthocyanidins in Crofelemer (*6*).

data can overcome the limited ability to characterize the entire botanical mixture or its active components, based on the analytical technology available. The approval of these two botanical NDAs demonstrates the success of an integrated approach and provides the industry with a practical framework for developing botanicals (including traditional medicines) to new drugs that are held to the same FDA standards as small-molecule drugs.

References
1. S.T. Chen *et al.*, *Nat. Biotechnol.* **26**, 1077 (2008).
2. D.J. Newman, G.M. Cragg, *J. Nat. Prod.* **75**, 311 (2012).
3. FDA Guidance for Industry-Botanical Drug Products, http://www.fda.gov/downloads/Drugs/GuidanceCompliance RegulatoryInformation/Guidances/ucm070491.pdf, accessed on June 28, 2014.
4. Approved Labeling for Veregen (NDA 021902) http://www.accessdata.fda.gov/drugsatfda_docs/nda/2006/021902s000_prntlbl.pdf, accessed on June 28, 2014.
5. Drugs@FDA FDA: Veregen (NDA 021902) http://www.accessdata.fda.gov/drugsatfda_docs/nda/ 2006/021902s000TOC.cfm, accessed on June 28, 2014.
6. Approved Labeling for Fulyzaq (NDA 202292) http://www. accessdata.fda.gov/drugsatfda_docs/nda/ 2012/202292Orig1s000Lbl. pdf, accessed on June 28, 2014.
7. Drugs@FDA FDA: Fulyzaq Delayed-Release Tablets (NDA 202292) http://www.accessdata.fda.gov/drugsatfda_docs/nda/2012/202292Orig1s000TOC.cfm, accessed on June 28, 2014.
8. L. Chen, Z.X. Zhou, *Plant Foods Hum. Nutr.* **60**, 31 (2005).
9. K. Jones, *J. Altern. Complement. Med.* **9**, 877 (2003).
10. D. Gupta, B. Bleakley, R. K. Gupta, *J. Ethnopharmacol.* **115**, 361 (2008).
11. J. A. Kennedy, G. P. Jones, *J. Agric. Food Chem.* **49**, 1740 (2001).
12. L. Tradtrantip, W. Namkung, A. S. Verkman, *Mol. Pharmacol.* **77**, 69 (2010).
13. J. Cottreau, A. Tucker, R. Crutchley, K. W. Garey, *Expert Rev. Gastroenterol. Hepatol.* **6**, 17 (2012).

Developing influenza treatments using traditional Chinese medicine

Authors:
Zi-Feng Yang[1‡],
Elaine Lai-Han Leung[2‡],
Liang Liu[1],
Zhi-Hong Jiang[1],
Nanshan Zhong[2*]

Humans have been faced the threat of epidemics such as influenza throughout their existence. Traditional Chinese medicine (TCM) practitioners began documenting their diagnostic and treatment principles related to epidemic diseases in the classic Chinese medical book, "Emperor Internal Medical Classic" (*1*). The unique treatments and herbal formulas used to combat influenza may serve as a source of information and inspiration for the development of new drugs (*2*).

Chinese herbal medicines and influenza

A major difference between Western and Chinese influenza treatments is the mode and targets of their actions. The first antiviral chemical drugs appeared in the West in the mid-1960s. Since then, many single-target therapeutics have been designed, but drug resistance is common. To circumvent this, Western medicine has incorporated multiple molecular targets into a single treatment using combination therapies, a practice now well accepted in the West.

Chinese herbal formulas (CHFs), on the other hand, often act via multiple modes that target not only the virus, but also various components of the host's immune response (Table 1), creating a synergistic effect. For example, *jinchai* capsules blunt viral replication by blocking adsorption of virions and preventing virus hyperalgesia-induced membrane fusion (*3*), while evodiamine blocks viral action by interfering with the AMPK/TSC2/mTOR signaling pathway, which is associated with virus-induced autophagy (*4*). Figure 1 summarizes the points of action of CHFs when treating influenza.

Isatis indigotica roots and influenza

Isatis indigotica roots (IIR) (*Banlangen*) have long been used to treat seasonal influenza in China. Currently, more than 100 chemical constituents of IIR have been identified. Among them, the compounds of epigoitrin; 2,4(1H,3H)-quinazolinedione; 4(3H)-quinazolinone; and clemastanin B have been demonstrated to kill or significantly inhibit the influenza virus. Studies from our laboratory have shown that polysaccharides extracted from IIR can prevent the influenza virus from attaching to host cell surfaces through a process involving hemagglutinins (*5*). Moreover, an indole alkaloid has been found to play a major role in preventing viral infection of host cells (*6*), while compounds derived from IIR can block translocation of the nucleocapsid protein at the early stage of replication, primarily through modulation of NF-κB signaling, thus inhibiting viral replication (*7*). In addition, IIR has been shown to exert immune modulatory effects in vitro and in vivo. In lipopolysaccharide (LPS)-stimulated RAW264.7 murine macrophages, the methanolic extracts of IIR inhibited degradation of IκBα and production of nitric oxide, prostaglandin E_2, and interleukin (IL) 6 (*8*). The polysaccharides from IIR could promote proliferation of lymphocytes and macrophages, as well as production of IL-2 and interferon (IFN) γ in mouse models (*9*). Indirubin and its derivatives can suppress a number of pro-inflammatory cytokines/chemokines in infected human bronchial epithelial cells, human peripheral blood-derived macrophages, and alveolar epithelial cells (Table 1) (*10*, *11*). Taken together, these data imply that IIRs play a variety of roles protecting against viral infection by targeting both the virus and the host—a markedly different effect than that of marketed chemically synthesized drugs.

Drug development strategies using TCM

High-quality consistency, treatment effectiveness, safety assurance, and patient affordability are the key factors for drug development. TCM can inform research into these areas in the following ways.

Firstly, the strategies and principles underpinning the translational research used in TCM-based influenza treatments could be applied more broadly. Two possible approaches can be taken: the standard, bottom up bench-to-bedside strategy, or a more innovative approach that transitions empirical medical knowledge from TCM into an evidence-based research strategy. We proffer that the latter better reflects the real-world interaction between basic science and the TCM clinical experience.

Secondly, basic research and clinical studies on CHFs could be conducted in parallel. For example, the effects of extracts and/or combinations of the active compounds from commonly prescribed CHFs could be investigated concurrently with standardized clinical trials based on documented clinical experience.

Thirdly, well-defined methodologies for standardized assessment of the quality, efficacy, and safety of CHFs are still lacking. It is important to standardize the composition and level of active components of herbs in CHFs before including them in a basic research project or clinical trial so as to maintain the data integrity.

Finally, TCM research is complex. It therefore behooves all researchers to develop interdisciplinary, innovative, and collaborative research projects, through which the scientific foundation of TCMs can be elucidated and a new framework that incorporates modern medical science can be built.

We have been pioneers in an attempt to implement the abovementioned strategies using IIR, launching the first randomized control trial in China in 2010. Various 'omics

[1]State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, China
[2]State Key Laboratory of Quality Research in Chinese Medicine, Macau University of Science and Technology, Macau (SAR), China.
‡Contributed equally to this work
*Corresponding Author: nanshan@vip.163.com

TABLE 1. Examples of TCM and Western anti-influenza drugs.

Antivirals				Target	Target subject	Mechanism of action or therapeutic effect	Year Documented	Reference
Western anti-influenza compounds	Cyanovirin-N			Surface glycoproteins of enveloped viruses	Virus	Inhibits entry of virus	1997	*12*
	Nitazoxanide			Influenza hemagglutinin	Virus	Impairs hemagglutinin intracellular trafficking	2009	*13*
	DAS181			Influenza viral receptor	Host	Removes viral receptor	2006	*14*
	Amantadine/Rimantadine			Influenza M2 protein	Virus	Inhibits proton conductance of M2	1965	*15, 16*
	Favipiravir(T-705)			Influenza RNA polymerase	Virus	Inhibits viral RNA polymerase activity	2002	*17, 18*
	Oseltamivir/Peramivir/Zanamivir/Laninamivir			Influenza neuraminidase	Virus	Blocks release of virus	1996/2000/1993/2005	*19–22*
Chinese anti-influenza herbal formulas, compounds, and constituents	Single herb	Ban Lan Gen (*Isatis indigotica* root)	Methanol extract	NF-κB signaling	Host	Inhibits nitric oxide and prostaglandin E_2 production, and NF-κB signaling in macrophages	1260s	*8*
			Polysaccharides	–	Host	Promotes transformation of lymphocytes and production of IL-2 and IFN-γ		*9*
			Clemastanin B	–	–	Blocks influenza ribonucleoprotein nuclear export, prolongs mean lifespan of infected mice		*7, 23*
			Indirubin	NF-κB signaling	Host	Interrupts virus-induced p38 MAP kinase activation and NF-κB translocation, and reduces expression of CCL5 in human bronchial epithelial cells		*10*
			Indirubin derivatives	–	Host	Suppresses pro-inflammatory cytokines and chemokines		*11*
		Da Qing Ye (*Folium isatidis*)	Monomer	–	–	Reduces mortality rate of influenza virus-infected mice	202–220 B.C.E.	*24*
		Jin Yin Hua (*Lonicera japonica*)	Ethanol extracts	Antiviral, immune-modulatory, and anti-inflammatory protein in mouse serum	Host	Reduces lung index and alleviates lung lesions in influenza virus-infected mice	1400s	*25*
		LianQiao (*Forsythia suspense*)	Ethanol and water extracts	–	Host	Regulates CCL5 and MCP-1 secretion in H1N1 virus-infected A549 cells		*26*
	Complex herbal formula	Ma Xing Shi Gan Tang + Yin Qiao San		–	–	Reduces time to fever resolution in patients with H1N1 influenza virus infection	2011	*27*
		Lian Hua Qing Wen capsule		–	–	Reduces time to fever resolution in patients with H1N1 influenza virus infection	2004	*28*
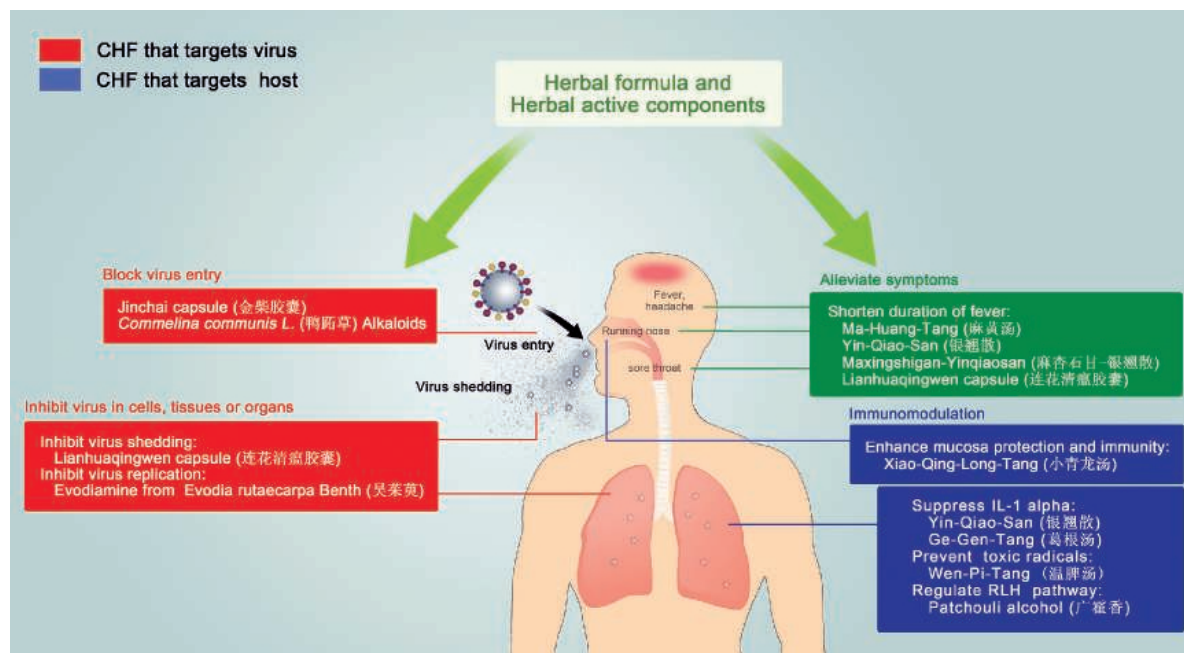
FIGURE 1. Holistic intervention to treat influenza.

technologies have been concurrently used to search for bioactive compounds, and we expect that additional active constituents with unique pharmaceutical activities will be found in the future. We have also combined the application of modern technologies with TCM clinical experience. For example, practitioners have noted that IIR appears to display beneficial clinical effects if administered during early onset of the disease (*9*). These studies suggest that further investigation of the mechanisms of IIR action is warranted. Importantly, using treatments with multiple sites of action may prevent or delay the generation of resistant viral strains.

References

1. J. Curran, G. P. Locum, Glasgow, *Brit. Med. J.* **336**, 777 (2008).
2. L. Jiang, L. Deng, T. Wu, *Cochrane Database Syst. Rev.* **3**, CD004559 (2013).
3. J. Zhong, X. Cui, Y. Shi, Y. Gao, H. Cao, *J. Tradit. Chin. Med.* **33**, 200 (2013).
4. J. P. Dai *et al.*, *PLOS ONE* **7**, e42706 (2012).
5. Z. Yang *et al.*, *Mol. Med. Rep.* **5**, 793 (2012).
6. Q. Zhu *et al.*, *Chinese Patent* CN 102836168 A (2012).
7. Z. Yang *et al.*, *Int. J. Mol. Med.* **31**, 867 (2013).
8. E. K. Shin, D. H. Kim, H. Lim, H. K. Shin, J. K. Kim, *Inflammation* **33**, 110 (2010).
9. Y. L. Zhao *et al.*, *Chin. J. Integr. Med.* **14**, 207 (2008).
10. N. K. Mak *et al.*, *Biochem. Pharmacol.* **67**, 167 (2004).
11. C. K. Mok *et al.*, *Antiviral Res.* **106**, 95 (2014).
12. B. R. O'Keefe *et al.*, *Antimicrob. Agents Chemother.* **47**, 2518 (2003).
13. J. F. Rossignol, S. La Frazia, L. Chiappa, A. Ciucci, M. G. Santoro, *J. Biol. Chem.* **284**, 29798 (2009).
14. M. P. Malakhov *et al.*, *Antimicrob. Agents Chemother.* **50**, 1470 (2006).
15. R. R. Grunert, J. W. McGahen, W. L. Davies, *Virology* **26**, 262 (1965).
16. J. R. Schnell, J. J. Chou, *Nature* **451**, 591 (2008).
17. Y. Furuta *et al.*, *Antimicrob. Agents Chemother.* **46**, 977 (2002).
18. Y. Furuta *et al.*, *Antimicrob. Agents Chemother.* **49**, 981 (2005).
19. C. U. Kim *et al.*, *J. Am. Chem. Soc.* **119**, 681 (1997).
20. Y. S. Babu *et al.*, *J. Med. Chem.* **43**, 3482 (2000).
21. M. von Itzstein *et al.*, *Nature* **363**, 418 (1993).
22. M. Yamashita *et al.*, *Antimicrob. Agents Chemother.* **53**, 186 (2009).
23. D. D. Sun, S. H. Yan, J. W. Chen, X. Li, L. W. He, *J. Nanjing Univ. TCM* **29**, 53 (2013).
24. Z. Liu, Z. Q. Yang, H. Xiao, *Virol. Sin.* **25**, 445 (2010).
25. H. Zhang, J. Song, J. Shi, *Chin. J. Chin. Materia Medica* **36**, 1071 (2011).
26. H. C. Ko, B. L. Wei, W. F. Chiou, *J. Ethnopharmacol.* **102**, 418 (2005).
27. C. Wang *et al.*, *Ann. Intern. Med.* **155**, 217 (2011).
28. H. X. Ouyang, Q. Y. Tang, Y. Z. Chen, Y. Wei, G. S. Li, *Chin. Med. Herald* **30**, 7 (2010).

Acknowledgments

A novel drug discovery strategy inspired by traditional medicine philosophies

Authors:
Xinfeng Zhao[1†],
Xiaohui Zheng[1†],
Tai-Ping Fan[2*],
Zijian Li[3],
Youyi Zhang[3],
Jianbin Zheng[1*]

For thousands of years, traditional medicines in China (traditional Chinese medicine, TCM), Japan (Kampō medicine), Korea (traditional Korean medicine), Indonesia (Jamu), India (Ayurvedic medicine), North America (phytotherapy), and Europe (herbalism) have been the primary means for maintaining health as well as preventing and treating human diseases (1). Over time, experiential knowledge derived from the medical application of natural products led to their incorporation into complex medical knowledge bodies—materia medica—characterized by the understanding of nature unique to each culture. For almost 200 years, the traditional use of natural products has also represented a source of effective drugs (2–5). This strategy represents a successful approach to novel drug identification and development through isolation and purification of active ingredients from natural products, high-throughput and high-content screening, and subsequent analysis and testing according to the guidelines of the U.S. Food and Drug Administration and other regulatory agencies (6–8). However, the pharmacologically active ingredients of a phytocomplex are not always the original natural molecules, but may be their host-specific metabolites or molecular complexes formed following co-administration with other herbs. This complexity has generated significant scientific challenges in the study of natural products (9, 10). The multicomponent nature of traditional medicines leading to multiple potential molecular interactions, multiple targets, and numerous metabolic byproducts, suggests that a conventional reductionist approach will have limitations in identifying active ingredients, making a more network-oriented, holistic approach preferable (12).

The *Jun-Shi* medicinal compatibility model

To address the challenge innate in the complex composition of natural products, inspiration was taken from the theoretical principles underlying TCM (13). This includes meticulous documentation of clinical observations, which can then inform the practice of traditional healing and help to develop guidelines and principles. The principles of these practices will be validated when successfully applied to practical clinical problems (13). We applied this approach to analyzing the *Jun-Chen-Zuo-Shi* principle of combining different materia medica in a specific manner when creating TCM compound formulations (*Fufang*). Additionally, as a strategy for drug discovery, we propose a simplified *Jun-Shi* model to identify the ingredients from TCM formulations that reach the bloodstream and the pharmacological effects they may have in the body.

Jun-Chen-Zuo-Shi and *Qiqing* (seven ways of pairing compatibile herbs) are the basic theories behind the formulation of TCM treatments. *Jun-Chen-Zuo-Shi* theory guides the combination of different herbal medicines in *Fufang*, based on the healing/pharmacological properties and constituents of each herb. The *Jun* (emperor) component is the principal phytocomplex targeting the major symptom of the disease. There are only a few varieties of *Jun* medicinals that are administered as a single formula, usually in large doses. The *Chen* (minister) herbs synergize with *Jun* to strengthen its therapeutic effects, and may also treat secondary symptoms. The *Zuo* (assistant) medicinal reduces or eliminates possible adverse or toxic effects of the *Jun* and/or *Chen* components, while also enhancing their effects and sometimes treating secondary symptoms. Finally, the *Shi* (courier) herbs facilitate delivery of the principal components to the lesion sites, or facilitate the overall action of the other components (14–15). The principles of *Qiqing* describe how herbs can be used independently, to reinforce (when both herbs have similar properties) or enhance (when the efficacy of the primary medicinal is improved) the effects of other herbs, or to antagonize certain unwanted, negative side-effects. In practice, *Qiqing* helps to determine the optimal pairing and proportions of two medicinals in a formulation.

We propose combining the principles of *Jun-Chen-Zuo-Shi* and *Qiqing* to create the notion of a *Jun-Shi* medicinal pair in order to provide better therapeutic efficacy when compared with a single medicinal (16). A *Jun-Shi* medicinal pair also has the synergistic characteristics of a *Fufang*, targeting the active phytochemicals to their designated sites of action. At the same time, the simpler composition of the *Jun-Shi* medicinal pair provides a less complex formulation for scientific analysis, which should reduce the complexity and difficulty when seeking new drugs from TCM-related sources. Based on these theories, we propose an innovative strategy for new drug discovery through the screening of in vivo effector compounds from *Jun-Shi* medicinal pairs. The *Jun* herb performs the primary action while the *Shi* herb potentiates this activity either by modifying the physicochemical properties of the *Jun* herb or facilitating its interaction with the pharmacological target. The identification of pharmacologically meaningful differences could therefore be made by comparing action and efficacy of the *Jun* herb with or without the *Shi* herb.

[1]Northwest University, Xi'an, China
[2]Department of Pharmacology, University of Cambridge, Cambridge, United Kingdom
[3]Institute of Vascular Medicine, Peking University Third Hospital; Key Laboratory of Cardiovascular Molecular Biology and Regulatory Peptides, Ministry of Health; Key Laboratory of Molecular Cardiovascular Sciences, Ministry of Education; and Beijing Key Laboratory of Cardiovascular Receptors Research, Beijing, China
[†]These authors contributed equally to this work
[*]Corresponding Authors: tpf1000@cam.ac.uk (T.F.) and zhengjb@nwu.edu.cn (J.Z.)

FIGURE 1. Schematic of the strategy for drug discovery based on the identification of an in vivo effector compound using the *Jun-Shi* medicinal compatibility model. We firstly selected *Jun* and *Shi* herbs as a research target from a complex *Fufang* using the principles of the *Jun-Shi* medicinal compatibility model. In vivo and in vitro compounds of the *Jun-Shi* medicinal pair were analyzed by high-performance liquid chromatography-mass spectrometry (LC-MS) and nuclear magnetic resonance. The data obtained by LC-MS was extracted and further analyzed by principal component analysis (a variance-focused procedure seeking to reproduce the total parameter variance, in which components reflect both common and unique variance of each parameter) to uncover the differential compounds. The resulting compounds were then examined using causal relationship analysis such as the Granger causality test. The bioactivity of those effector compounds identified was screened by receptor affinity chromatography and confirmed by classical functional pharmacological assays to identify leads of interest. One lead, isopropyl 3-(3,4-dihydroxyphenyl)-2-hydroxypropanoate (IDHP), was evaluated further according to guidelines established by the U.S. Food and Drug Administration and other regulatory agencies for the development of an investigational new drug. We hope that the mechanism of action of IDHP will further confirm the broad applicability of the *Jun-Shi* medicinal compatibility model to drug discovery.
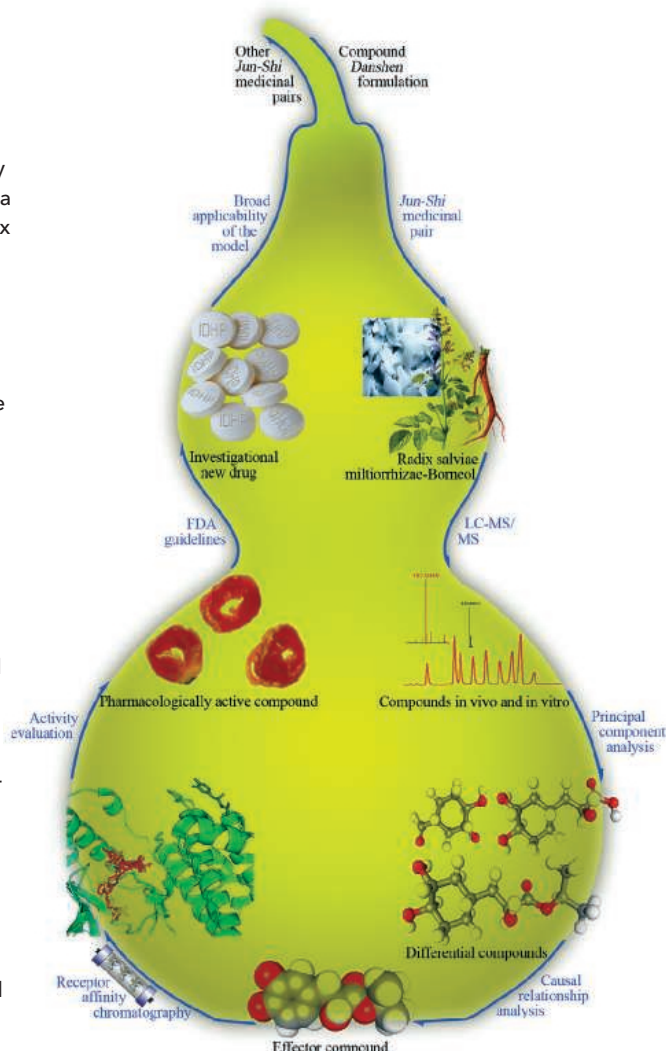
(*Danshensu*) were found to be increased in rabbit heart by co-administration of *Bingpian*. Additionally, isopropyl 3-(3,4-dihydroxyphenyl)-2-hydroxypropanoate (IDHP), a novel metabolite of *Danshensu* in *Fufang Danshen Diwan* was identified (*18*). Pharmacokinetic studies showed that this new compound was preferentially found in heart and brain tissues, in agreement with the lesion sites expected to be targeted by this formulation according to TCM principles. Furthermore, synthetic IDHP was generated and shown to have a protective function against myocardial and cerebral ischemia injury (*19*, *20*), strongly suggesting that IDHP is the effector compound in the *Danshen-Bingpian* medicinal formulation in *Fufang Danshen Diwan*.

Conclusions

The studies described here propose a novel strategy for drug discovery based on the identification of active substances in TCM herbal formulations. The strategy consists of the selection of *Jun-Shi* medicinal pairs, the establishment of a quantitative model for analysis of causal relationships, and a system for the identification of active substances by receptor affinity chromatography (Figure 1) (*21*, *22*). Compared with the conventional procedure of drug discovery using a reductionist approach, the current strategy employs the principles of *Jun-Chen-Zuo-Shi* and *Qiqing* in devising a *Jun-Shi* medicinal compatibility model to reduce the number of targets for the identification of active ingredients. This can reduce the arbitrary nature of the drug discovery process and also improve its efficiency. Significantly, it can further enrich TCM theory and offer new perspectives in the research of complex biomedical questions.

Testing the *Jun-Shi* model

The *Fufang Danshen Diwan* (Dantonic pill) has successfully completed phase 2 clinical trials in the United States and is currently undergoing phase 3 trials (*17*), which made this a perfect candidate on which to test our strategy of drug discovery. We choose to focus on the *Danshen* (*Radix Salviae miltiorrhizae*) plus *Bingpian* (Borneol) pair of medicinals present in this pill to investigate our *Jun-Shi* compatibility model. The process is described in Figure 1. The levels of *Danshen*-derived phenolic acids such as 3-(3,4-dihydroxyphenyl)-2-hydroxylpropionic acid

Future perspectives

Based on the successful work with IDHP, we have further synthesized a series of *Jun-Shi* compounds including 1,7,7-tri-methylbicyclo[2.2.1]heptan-2-yl 3-(3,4-dihydroxyphenyl)-2-hydroxypropanoate by incorporating borneol in the structures of active compounds of *Danshen* such as *Danshensu*, caffeic acid, and rosmarinic acid, according to our drug discovery strategy. Patents on these compounds have been approved in 34 countries or regions, including the United States (No. 8017786), Canada (No. 2652299), Russia (No. 2421443), and the European Union (No. 2019090). Preliminary pharmacological investigations have shown that some of these compounds are effective against the formation of atherosclerotic plaques in ApoE$^{-/-}$ mice (Yonggong Zhai, unpublished observations). Results indicate that these *Jun-Shi* compounds may be potential candidates for constructing a database of combinational TCM-derived molecules for drug discovery.

References

1. H. Itokawa, S. L. Morris-Natschke, T. Akiyama, K. H Lee, *J. Nat. Med.* **62**, 263 (2008).
2. W. Lam *et al.*, *Sci. Transl. Med.* **45**, 1 (2010).
3. W. H. J. Li, J. C. Vederas, *Science* **325**, 161 (2009).
4. D. Normile, *Science* **299**, 188 (2003).
5. V. D. Luca, V. Salim, S. M. Atsumi, F. Yu, *Science* **336**, 1658 (2012).
6. F. E. Koehn, G. T. Carter, *Nat. Rev. Drug Discov*. **4**, 207 (2005).
7. Editorial, *Nat. Rev. Drug Discov*. **6**, 3 (2007).
8. L. Paterson, E. A. Anderson, *Science* **310**, 451 (2005).
9. G. Le, L. Zhang, S. B. Su, *Mode. Tradit. Chin. Med. Mater. Med.* **12**. 15 (2010).
10. J. P. Zhu, R. Ken, *Clin. Acupunct. Orient. Med*. **3**, 77 (2002).
11. S. Prizker, *Clin. Acupunct. Orient. Med*. **3**, 99 (2002).
12. S. E. Phelan, *Emergence* **3**, 120 (2001).
13. Y. Jiang, *Inform. Tech*. **11**, 20 (2001).
14. L. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 4826 (2008).
15. T. P. Fan, J. C. Yeh, K. W. Leung, P. Y. K. Yue, R. N. S. Wong, *Trends Pharmacol. Sci*. **27**, 297 (2006).
16. S. P. Wang *et al.*, *J. Ethnopharmacol*. **143**, 412 (2012).
17. http://www.clinicaltrials.gov/ct2/show/study/NCT01659580.
18. X. H. Zheng *et al.*, *J. Sep. Sci.* **30**, 851 (2007).
19. S. P. Wang *et al.*, *Eur. J. Pharmacol*. **579**, 283 (2008).
20. J. W. Tian *et al.*, *Neurosci. Lett*. **442**, 279 (2008).
21. X. H. Zheng *et al.*, *Chinese. Sci. Bull*. **53**, 842 (2008).
22. X. F. Zhao *et al.*, *J. Chromatogr. B*. **878**, 2029 (2010).

Deciphering ancient combinatorial formulas: The *Shexiang Baoxin* pill

Authors:
Runhui Liu[1‡],
R. Scott Runyon[2‡],
Yuchong Wang[3],
Stephen G. Oliver[3],
Tai-Ping Fan[2*],
Weidong Zhang[1*]

An ancient combinatorial formula (*Fufang*) carries similarities to the polypill used in Western medicine. A polypill is a single medication that contains multiple pharmaceutical therapies, and may be intended for the treatment of a single or even multiple diseases. An example of a polypill is Polycap—comprising aspirin, simvastatin, ramipril, atenolol, and hydrochlorothiazide—used to treat hypertension and prevent heart attack and stroke (*1*). In traditional Chinese medicine (TCM), *Fufang* refers to a group of therapeutic compounds derived from multiple plant, mineral, or occasionally animal sources. For example, the formulation Realgar-Indigo naturalis—containing tetra-arsenic tetra-sulfide (As_4S_4), indirubin, and tanshione IIA—has been demonstrated to be an effective treatment for promyelocytic leukemia (*2*). In Europe, a five-herb formulation of cowslip (*Primula veris/elatior*), yellow gentian (*Gentiana lutea*), black elder (*Sambucus nigra*), common sorrel (*Rumex* species), and vervain (*Verbena officinalis*) has been approved by the European Medicines Agency Committee on Herbal Medicinal Products for sinusitis and bronchitis.

To study the efficacy and bioactivity of combinatorial formulas, the *Shexiang Baoxin* Pill (SBP) was chosen as a test case. SBP originates from a classical ancient prescription *Suhexiang* Pill recorded in *Prescriptions of the Bureau of Taiping People's Welfare Pharmacy*, which was used for the treatment of chest pain with dyspnea in the Southern Song Dynasty (10th to 12th centuries C.E.). The modern form of SBP was developed in 1981 under Professor Rui-Hong Dai's leadership (*3*), and consists of seven Chinese materia medica (medical materials) (Figure 1). Today in China, SBP has become a widely used *Fufang* for the treatment of stable angina pectoris, chest pain or discomfort caused by coronary heart disease (CHD). Several randomized, prospective clinical studies strongly suggest SBP can decrease the frequency of angina due to CHD and daily nitroglycerine use (*4, 5*) and may even reduce ischemic myocardial changes as measured by electrocardiogram and perfusion imaging (*6*).

However, as with many *Fufang*, numerous questions must

[1]College of Pharmacy, Second Military Medical University, Shanghai, China
[2]Department of Pharmacology, University of Cambridge, Cambridge, UK
[3]Cambridge Systems Biology Centre & Dept. Biochemistry, University of Cambridge, Cambridge, UK
‡Contributed equally to this work
*Corresponding Authors: tpf1000@cam.ac.uk (T.F.) and wdzhangy@hotmail.com (W.Z.)

FIGURE 1. Composition of *Shexiang Baoxin* Pill (SBP). Fifteen of the more abundant compounds that comprise SBP and the materia medica from which they are derived are shown. Pure compounds were identified by high-performance liquid chromatography-mass spectrometry of SBP (*7*, *11*). (Images courtesy of www.tcmwiki.com.)

be answered if SBP is to reach patient populations in countries with stringent food and drug regulations. In contrast to modern Western pharmaceuticals, where the active ingredients are synthesized and examined independently in clinical trials before being co-administered, the process in TCM is reversed. The current challenge is to define SBP's composition, pharmacodynamics, pharmacokinetics, and, since it is a formulation, any pharmacodynamic synergies that may be present.

A reductionist approach: Breaking down SBP

Advanced separation and analysis techniques such as gas and liquid chromatography, coupled with mass spectrometry, have made it possible to identify the exact chemical species that comprise *Fufang*-based therapies. To date, over 70 non-volatile and over 40 volatile chemical species have been identified in SBP (*7, 8*). Following oral administration of SBP in rats, as many as 22 of these pure compounds and eight metabolites could be observed in blood plasma (*9, 10*). These analytical techniques help to establish a "chemical fingerprint" for SBP, which allows for batch-to-batch comparison of individual materia medica and the resultant drug, an important step towards quality control in manufacturing (*11, 12*), and reproducible safety and efficacy.

Such a chemical fingerprint is also vital for identifying active ingredients. For instance, several of the major classes of compounds that have been identified in SBP include bufadienolides, ginsenosides, and bile acids (Figure 1). Two bile acids identified, ursodeoxycholic acid and chenodeoxycholic acid, have been approved by the U.S. Food and Drug Administration (FDA) and marketed as chenodiol and ursodiol, respectively. Prescribed for the prevention and management of gallstones, these bile

acids decrease the production of cholesterol and reduce hypertriglyceridemia, a strong risk factor for CHD (*13*). The bufadienolides that are abundant in *V. bufonis* and SBP (*11*) are cardiac glycosides, and carry similarities to digoxin, a plant-derived cardiac glycoside, FDA-approved class V anti-arrhythmic, and positive inotropic agent (*14*). The main bufadienolide in *V. bufonis*, bufalin, is also a Na^+,K^+-ATPase inhibitor that increases cardiac contractility (*15*) and may protectively downregulate the renin-angiotensin system during heart failure (*16*). The small molecule cinnamaldehyde, abundant in *Cortex cinnamomi* (*7*), has been shown to be a strong vasodilator (*17*) and activates transient receptor potential channels (TRPV1 and TRPA1) involved in nociception (*18, 19*). Active reduction of pain perception, cholesterol and triglyceride levels, and effects on contractility may represent a few of the ways the bioactive ingredients in SBP exert a cardioprotective effect. As part of a reductionist approach for any *Fufang*, a deeper examination of how identified bioactive ingredients affect disease processes may prove vital in characterizing the formulas' mechanism of action.

A systems biology approach: Keeping SBP intact

Systems biology provides an additional strategy for exploring *Fufang* mechanisms of action and already has contributed to the successful development of two antianginal therapeutics, ranolazine and ivabradine (*20*). In contrast to a reductionist approach, which studies complex systems by investigating its individual components, systems biology focuses on understanding biological networks in a broader context by integrating transcriptomics, proteomics, metabolomics, and bioinformatics data and analyses (*20,*

FIGURE 2.
Chemogenomic profiling.
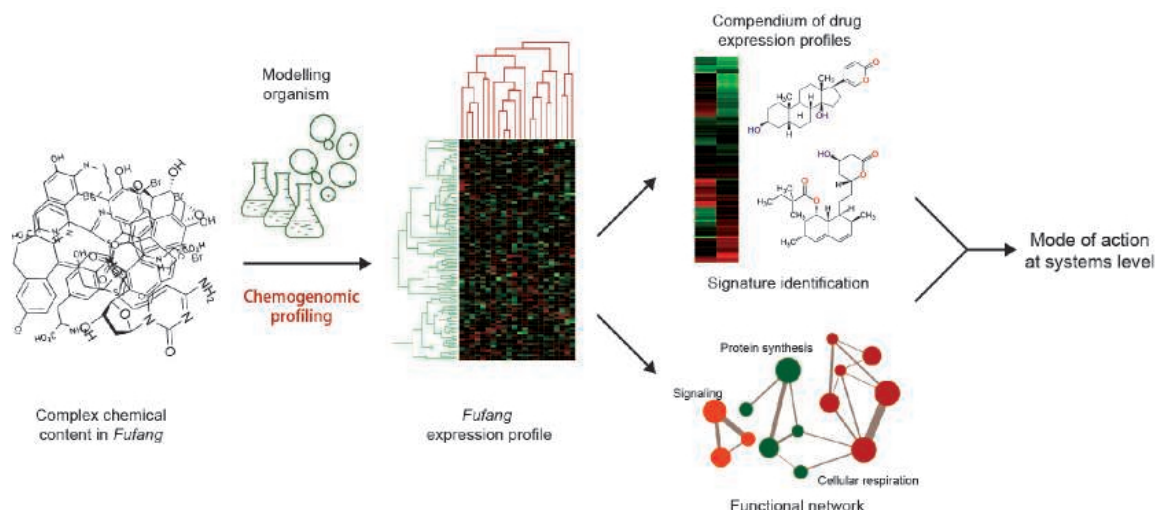
21). Recently, we initiated a more comprehensive study of the in vivo mechanisms of *Fufang* using a chemogenomic approach (Figure 2). For example, using drug-induced haploinsufficiency profiling in *Saccharomyces cerevisiae* (*22*), functional information can be obtained from loss-of-function assays to systematically investigate the cellular response to either individual bioactive entities or combined subsets of SBP. Heterozygous strains that show hypersensitivity to SBP can reveal pathways and targets that respond to the drug, thereby providing clues about its mode of action in a cellular context. A previous compendium of cellular responses to small molecules allows mechanisms of novel compounds to be inferred on the basis of profile similarity to established drugs (*23*).

Metabolomic methods have already proven useful in characterizing SBP. In a rat model of acute myocardial infarction (MI), numerous plasma and urinary biomarkers involved in oxidative injury, dysfunction of energy and amino acid metabolism, and inflammation have been identified using partial least squares discriminant analysis plots (*24, 25*). SBP given orally before MI can significantly reverse the changes in a number of these identified biomarkers—including lactic acid, homocysteine, and prostaglandin E_2—nearly returning their levels to normal (*25*). Analysis of the chemogenomic and metabolomic response signatures of SBP and other *Fufang* can clarify their impact on broader cellular processes and identify potential targets.

From complexity to simplicity—A new development strategy for *Fufang*

As not all components in *Fufang* are active, a combination of its active components may provide a simplified *Fufang* that facilitates easier identification of therapeutic targets and mechanisms of action. Early attempts to develop a simplified formulation of SBP (Figure 1, molecules in bold) have shown some promising results in rat models of MI (*26*), illustrating the potential of integrating reductionist and systems biology approaches in the development of *Fufang*. Ultimately, by

conducting rigorous quality control and purification, and removing extraneous compounds, such an approach would be equally applicable to create a new generation of polypills for other diseases.

References
1. N. J. Wald, D. S. Wald, *Postgrad. Med. J.* **86**, 257 (2010).
2. L. Wang *et al., Proc. Natl. Acad. Sci. U.S.A.* **105**, 4826 (2008).
3. X. P. Luo, H. M. Shi, E. H. Fan, R. H. Dai, *Chin. J. Integ. Trad. Wester. Med.* **22**, 718 (2002).
4. S. Y. Wang *et al., Chin. J. Integ. Trad. Wester. Med.* **16**, 717 (1996).
5. H. Zhu *et al., Chin. J. Integ. Trad. Wester. Med.* **30**, 474 (2010).
6. L. J. Wang *et al., Chin. J. Integ. Med. Cardio-/Cerebrovascular Disease* **6**, 129 (2008).
7. P. Jiang *et al., Chromatogr.* **70**, 133 (2009).
8. R. H. Liu *et al., Chem. Nat. Compd.* **45**, 599 (2009).
9. L. A. Guo *et al., J. Pharm. Practice* **30**, 207 (2012).
10. P. Jiang *et al., Biomed. Chromatogr.* **23**, 1333 (2009).
11. S. K. Yan, W. D. Zhang, R. H. Liu, Y. C. Zhan, *Chem. Pharm. Bull.* **54**, 1058 (2006).
12. S. K. Yan, Y. Yang, Y. W. Wu, R. H. Liu, W. D. Zhang, *J. Anal. Chem.* **64**, 141 (2009).
13. M. Watanabe *et al., J. Clin. Invest.* **113**, 1408 (2004).
14. A. P. Ambrosy *et al., J. Am. Coll. Cardiol.* **63,** 1823 (2014).
15. A. R. Patel, T. Kurashina, J. P. Granger, K. A. Kirchner, *Hypertension* **27**, 668 (1996).
16. M. M. Kau, J. R. Wang, S. C. Tsai, C. H. Yu, P. S. Wang, *Br. J. Pharmacol.* **165**, 1868 (2012).
17. A. Yanaga *et al., Biol. Pharm. Bull.* **29**, 2415 (2006).
18. A. Szallasi, D. N. Cortright, C. A. Blum, S. R. Eid, *Nat. Rev. Drug Discov.* **6**, 357 (2007).
19. F. Sui *et al., J. Asian Nat. Prod. Res.* **12**, 76 (2010).
20. C. Macilwain, *Cell* **144**, 839 (2011).
21. E. Holmes, I. D. Wilson, J. K. Nicholson, *Cell* **134**, 714 (2008).
22. G. Giaever *et al., Nature* **418**, 387 (2002).
23. A. Y. Lee *et al., Science* **344**, 208 (2014).
24. P. Jiang *et al., Mol. Biosyst.* **7**, 824 (2011).
25. L. Xiang *et al., Mol. Biosyst.* **8**, 2434 (2012).
26. L. Xiang *et al., Alt. Integr. Med.* **6**, doi:10.4172/2327-5162. 1000127 (2013).

Lessons from the development of the traditional Chinese medicine formula PHY906

Authors:
Wing Lam,
Shwu-Huey Liu,
Zaoli Jiang,
Yung-Chi Cheng*

Traditional Chinese medicine (TCM) has been practiced for thousands of years. While the historical usage of TCM is well documented, it is not broadly accepted by mainstream physicians who question the quality and consistency of TCM products, the scientific basis for usage, and the lack of evidence-based clinical studies. Nonetheless, TCM formulas are currently being used to relieve the side effects of nonhematological toxicities caused by chemotherapy, including diarrhea, nausea, vomiting, and fatigue. We decided to further explore the mechanisms of action of TCM in chemotherapy, laying the groundwork for its potential use as an adjuvant treatment.

We selected several TCM formulas composed of common herbs for the treatment of the abovementioned symptoms. Considering their long history of usage, these formulas should be relatively safe. Simple formulas consisting of a limited number of herbs were used to facilitate quality control and simplify analysis of the mechanisms of action. Among these formulas, we found that *Huang-Qin Tang* could enhance the therapeutic index of irinotecan, a chemotherapeutic agent for the treatment of metastatic colon and rectal cancer. This four-herb formula (*Glycyrrhiza uralensis* Fisch, *Paeonia lactiflora* Pall, *Scutellaria baicalensis* Georgi, and *Ziziphus jujube* Mill) has been used to treat gastrointestinal disorders for approximately 1,800 years. The formula, named PHY906, was manufactured using standard operation procedures and following current good manufacturing practice standards.

Quality control for TCM

Since the sites of action and bioactive compounds found in TCMs are not always known, it is not sufficient to rely on either chemistry (*1, 2*) or biological analysis (*3*) alone for quality control (QC) purposes. We therefore developed Phytomics QC, an analysis system that integrates both chemical and biological data (from in vitro and in vivo studies) to assess the consistency of TCM using a novel statistical methodology (*4*). Using this system, we demonstrated that batches of PHY906 formulations spanning over a decade had a Phytomics Similarity Index (PSI) > 0.9 (1, perfectly identical; 0, no similarities). Interestingly, commercial *Huang-Qin Tang* products had a wider PSI range and showed inconsistent in vivo biological activity when compared to PHY906.

Are all herbs in a TCM formula required?

Based on TCM principles, a formulation should have the proper ratio of the "imperial" herb (the main ingredient), the

"ministerial" herb (ancillary to the imperial herb), the "assistant" herb (reduces side effects of the main herb), and the "servant" herb (aids in harmonizing the other herbs) to achieve the best therapeutic effect (*5*). By comparing formulations in which one herb has been removed, as well as single-herb preparations, a formulation for PHY906 was found that achieved the optimal therapeutic effect in combination with irinotecan (*6*). Different herbs appear to play different roles in the enhancement of antitumor activity and in protection against weight loss and mortality (*6*). These findings support the theory that each herb might play a specific role in a TCM formulation and suggests that multiple targets, and multiple active compounds for each target, are involved in the PHY906's action. TJ14, a seven-herb TCM formula, has the potential to reduce diarrhea and oral mucositis due to chemotherapy (*7–9*). While TJ14 shares three herbs with PHY906, it did not increase the therapeutic index of irinotecan in our preliminary studies (unpublished data).

Scientific validation of TCM

In preclinical studies, PHY906 was found to reduce diarrhea and intestinal damage following irinotecan or irradiation treatment by inhibiting multiple inflammatory processes and/or by promoting intestinal recovery (*10, 11*). These preclinical studies of PHY906 may help to explain why *Huang-Qin Tang* is effective in treating diarrhea. Based on these findings, the use of PHY906 in treating inflammatory bowel disease and stimulating stem/progenitor cell growth is being explored.

Preclinical studies indicate that PHY906 could potentiate the antitumor activity of a broad spectrum of anticancer agents in vivo (*12*). PHY906 plus irinotecan may increase tumor cell apoptosis associated with strong macrophage infiltration and induce an acute inflammation in tumors (*13*). Chemicals and metabolites of PHY906 varied based on the particular tissue being studied, which could partly explain why intestine and tumor tissue had different inflammation responses to PHY906 (*13*).

Like PHY906, many other herbal formulas have been reported to have multiple effects. TJ14 is thought to inhibit multiple targets of the cyclooxygenase pathway through a range of bioactive compounds (*14*). TJ48 improves patient host-immunity and quality of life following chemotherapy (*15*), and was shown to have anti-angiogenic activity in tumors (*16*) as well as stimulating hematopoietic stem cell proliferation (*17*).

TCM provides an opportunity to treat patients in a holistic manner. The polychemical nature of TCMs allows them to target multiple organs in which absorption and metabolism could be quite different. Additionally, the in vivo impact of PHY906 on RNA expression across different tissues (liver, spleen, and tumor) was variable (*13*). A more detailed analysis of the responses of different organs and tissues to herbal

Department of Pharmacology, Yale University School of Medicine, New Haven, CT
*Corresponding Author: yccheng@yale.edu

formulations could support the TCM theory that there may be a preferential response in certain organs to certain herbs. This work could lead to new applications for TCM.

Challenges of TCM clinical trials

Traditionally, TCM is prescribed as a decoction based on the diagnosis of a practitioner and the individual patient's needs. It is therefore challenging to design randomized, placebo-controlled, dose-escalation studies. Herbal formulas have been administered in capsule form for many years, so it made sense to initiate clinical studies using encapsulated products. Phase 1/2 or phase 2 clinical trials in the United States have suggested that an encapsulated form of PHY906 could have beneficial effects for cancer patients treated with irinotecan or capecitabine for advanced colorectal cancer, hepatocellular carcinoma, and pancreatic cancer (*18–22*). Recently, clinical trials of PHY906 with irinotecan, sorafenib, or radiation for colorectal cancer, liver cancer, or rectal cancer have been initiated in the United States. They are using a comprehensive systems biology approach to identify predictive pharmacodynamics biomarkers associated with PHY906 treatment, including immunocytokines, metabolomic profiles, herbal metabolites, and circulating tumor DNA. These results may aid in the identification of active compounds and in stratifying patient populations prior to treatment.

A number of herbal products are at various clinical de-velopment stages, including Dantonic (T89) (*23*), Selected Vegetable and Herb Mix (NCT00246727), *Rhodiola rosea* extract (NCT01098318), *Fuzheng Huayu* (NCT00854087), and HMPL-004 (NCT01805791). Thus far, the U.S. Food and Drug Administration (FDA) has approved only two highly purified botanical drugs with defined polychemicals: Veregen Oint-ment, a green tea extract for the topical treatment of genital warts, in 2006, and Fulyzaq (crofelemer), purified oligomeric proanthocyanidin from *Croton lechleri,* for the treatment of diarrhea in HIV patients, in 2012 (see page S32). FDA does not provide specific quality control guidance for orally admin-istrated herbal mixtures. However, they advise that multiple batches of an herbal product go through phase 3 clinical trials to demonstrate consistent efficacy. Although it is currently not required that the active ingredient from the herbal mixture be identified, it is anticipated that FDA will require information from more in-depth studies in the future.

In summary, although TCM formulas often vary, it is possible to make consistent preparations, as exemplified by PHY906. TCMs often have multiple sites of action and the active compounds acting at each site may be different. Systems biology and modern bioinformatics technologies are needed to fully explore the value of TCM for future medical applications.

References
1. J. Wang *et al., J. Pharm. Biomed. Anal.* **83**, 57 (2013).
2. X. Cheng *et al., Sci. Rep.* **4**, 5147 (2014).
3. A. Buriani *et al., J. Ethnopharmacol.* **140**, 535 (2012).
4. R. Tilton *et al., Chin. Med.* **5**, 30 (2010).
5. H. Y. H. Hsu, C. S. Hsu, *Commonly used Chinese herb formulas with illustrations*. (Oriental Healing Art Institute, Long Beach, CA, 1980).
6. S. H. Liu, Y. C. Cheng, *J. Ethnopharmacol.* **140**, 614 (2012).
7. Y. Sakata, H. Suzuki, T. Kamataki, *Cancer & Chemotherapy* **21**, 1241 (1994).
8. T. Aoyama *et al., Cancer Chemother. Pharmacol.* **73**, 1047 (2014).
9. K. Mori *et al., Cancer Chemother. Pharmacol.* **51**, 403 (2003).
10. W. Lam *et al., Sci. Transl. Med.* **2**, 45 (2010).
11. S. Rockwell *et al., Int. J. Radiat. Biol.* **89**, 16 (2013).
12. S. H. Liu *et al., Proc. Am. Assoc. Cancer Res.* **45**, 128 (2004).
13. E. Wang *et al., BMC Med. Genomics* **4**, 38 (2011).
14. T. Kono *et al., Integr. Cancer Ther.* **13**, 435 (2014).
15. J. J. Gao *et al., Drug Discov. Ther.* **6**, 1 (2012).
16. H. Kamiyama *et al., Biol. Pharm. Bull.* **28**, 2111 (2005).
17. H. Hisha *et al., Blood* **90**, 1022 (1997).
18. M. P. Farrell, S. Kummar, *Clin. Colorectal Cancer* **2**, 253 (2003).
19. S. Kummar *et al., Clin. Colorectal Cancer* **10**, 85 (2011).
20. M. W. Saif *et al., Phytomedicine* **17**, 161 (2010).
21. Y. Yen *et al., Anticancer Res.* **29**, 4083 (2009).
22. M. W. Saif *et al., Cancer Chemother. Pharmacol.* **73**, 373 (2014).
23. S. Ling *et al., J. Cardiovasc. Pharmacol.* **60**, 513 (2012).

The potential role of Chinese herbal medicines in cancer management

Authors:
Clara Bik-San Lau[1,2*], Grace Gar-Lee Yue[1,2], Kwok-Pui Fung[1,2,3], Ning-Hua Tan[4], Ping Chung Leung[1,2]

The management of cancer involves multiple disciplines, including surgery, chemotherapy, radiation therapy, targeted therapy, biological therapy, and systemic therapy. In spite of scientific advances and the evidence-based practice of these treatments, limitations in their benefits still exist, resulting in the increasing use of complementary and alternative medicine (CAM) by cancer patients and survivors (*1*). Numerous preclinical and clinical studies of CAM have been documented over the past decade (*1, 2*). Recent surveys revealed that the overall prevalence of CAM use among cancer patients in Germany and Ireland was 77% (*3*) and 32.5% (*4*), respectively. Other prospective and multicenter studies in the United States have shown that CAM usage was reported in 52–54% of cancer patients (*5, 6*). One of the modalities commonly used in Chinese cancer patients is Chinese herbal medicines (CHMs). A similar proportion (53%) of cancer patients in southwestern China using CHMs was reported (*7*). However, the potential benefits of CHMs as a cancer therapy have been less well studied. This article aims to illustrate the potential role of CHMs in cancer management and their adjuvant value in conventional cancer therapy.

In the last decade, scientists (including our research group) have carried out a series of investigations using different experimental systems examining four possible targets of CHM in cancer treatment, namely cytotoxicity (including reversal of multidrug resistance), immunomodulation, anti-angiogenesis, and antimetastasis (Table 1) (*11–24*).

Despite of the fact that hundreds of CHMs have been used in clinics to treat cancer patients in China (*25*), only a handful are undergoing clinical trials that meet international quality standards (double-blinded and placebo-controlled). One example is our own clinical trial demonstrating that oral consumption of the *Yunzhi-Danshen* capsule was beneficial in promoting immunological function after conventional
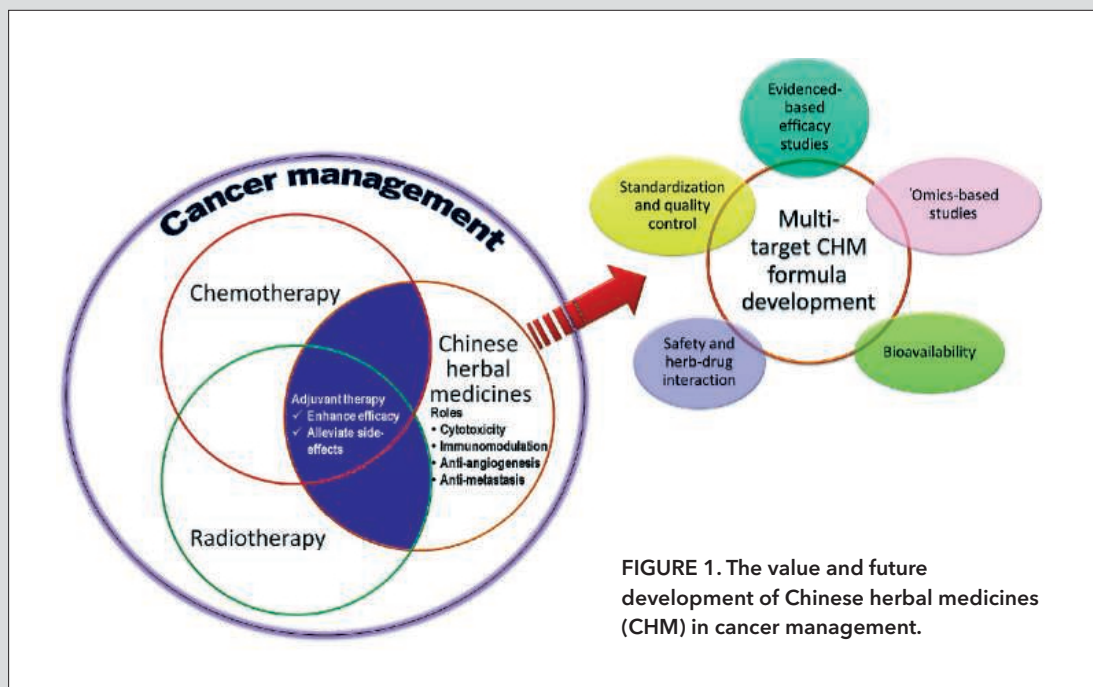
FIGURE 1. The value and future development of Chinese herbal medicines (CHM) in cancer management.

Possible CHM targets in cancer management

Although various active antitumor compounds have been isolated from CHMs (*8*), the therapeutic rationale for the treatment of cancer using CHMs is not limited to only cytotoxicity. Other therapeutic principles include boosting the natural host immune response, improving quality of life, and preventing relapse after surgery (*8–10*).

treatment of nasopharyngeal carcinoma patients (*26*). In the United States, PHY906 (*Huang Qin Tang*) has recently passed a phase 1 trial in colorectal cancer patients (*27*) and a phase 2 trial in pancreatic cancer patients (*28*). Furthermore, a Kampō formula (TJ14, *hangeshashinto*) is undergoing a phase 2 trial for gastric cancer patients in Japan (*29*).

Single mode of action vs. multitarget approach

An herb can serve as a CAM for the treatment of cancer when it is used alone or in combination with other herbs.

[1]Institute of Chinese Medicine, [2]State Key Laboratory of Phytochemistry and Plant Resources in West China (CUHK), and [3]School of Biomedical Sciences, The Chinese University of Hong Kong, Hong Kong
[4]State Key Laboratory of Phytochemistry and Plant Resources in West China, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming, China.
*Corresponding Author: claralau@cuhk.edu.hk

TABLE 1. Possible targets of Chinese herbal medicines (CHMs) in experimental systems and their implications.

Possible targets	Examples of CHM or isolated compounds	Experimental systems	Principal parameters measured	Implications
Cytotoxicity	• Andrographolide from *Andrographis paniculata* (11) • Curcumin from *Curcuma longa* (12) • Eriocalyxin B from *Isodon eriocalyx* (13)	➢ Cancer cell lines ➢ Tumor-bearing animals	◊ Inhibition of proliferation ◊ Induction of apoptosis	Inhibit tumor growth at primary site
	• *Hedyotis corymbosa* (14) • *Scutellaria barbata* (15)	➢ Multidrug resistance hepatoma cells	◊ Reverse P-glycoprotein-mediated multidrug resistance ◊ Induction of apoptosis	Improve chemotherapy efficacy
Immunomodulation	• *Astragalus* species (16) • *Coriolus versicolor* (17) • *Ganoderma sinense* (18)	➢ Immune cells (e.g., lymphocytes, dendritic cells) ➢ Tumor-bearing animals	◊ Production of cytokines and/or chemokines ◊ Population changes of certain immune cell types	Improve immune response; strengthen attack system of immune cells against cancer cells
Anti-angiogenesis	• Bigelovin from *Inula helianthus-aquatica* (19) • n-Butylidenephthalide from *Angelica sinensis* (20) • Cyclopeptide RA-V from *Rubia yunnanensis* (21)	➢ Endothelial cells	◊ Inhibition of √ proliferation √ tube formation √ migration	Inhibit new blood vessels formation towards and inside the tumor
		➢ Zebrafish embryos ➢ Tumor-bearing animals	◊ Inhibition of blood vessels growth	
Antimetastasis	• *Andrographis paniculata* (22) • *Camellia sinensis* (23) • *Ganoderma lucidum* (24)	➢ Invasive cancer cell lines	◊ Inhibition of √ migration √ invasion	Prevent migration of tumor cells from primary site to other organs
		➢ Tumor-bearing animals	◊ Inhibition of metastasis	

Active compounds isolated from herbs can also be developed as anticancer drugs, often aimed at specific targets (8). Certain extracts from single herbs, containing a complex array of constituent molecules, have been shown to exhibit direct or indirect antitumor activities (30, 31). In this case, the extract could be considered to be a multitarget combination therapy. CHMs are generally administered as combinations of multiple herbs, emphasizing one of the CAM principles of treating multiple targets while protecting internal harmony. The multicomponent, multitarget, and synergistic nature of CHMs can thus be fully utilized, particularly in multifactor metastasis management, an active area of research in our laboratory.

Adjuvant therapy

CHM is increasingly being used in conjunction with chemotherapy and radiotherapy, with the hope that it can alleviate or even eliminate the adverse effects of the treatments, as well as improve their overall efficacy (1, 32). Previous meta-analyses have suggested that *Astragalus*-based CHM may increase the effectiveness of platinum-based chemotherapy (33). Recent studies showed enhanced antitumor effects (27, 28), improved prognosis (34), and beneficial effects on the survival of cancer patients (35) receiving chemotherapy in combination with CHM, emphasizing the potentially positive adjuvant role of CHMs in cancer therapy.

The way forward

The potential roles of CHM in cancer management are summarized in Figure 1. Future development of CHMs will require extensive further research such as evidenced-based efficacy studies, standardization and quality control, 'omics-

based studies into the mechanisms of action, bioavailability studies, and careful examination of safety and herb-drug interactions. We suggest that future research should focus on developing CHM formulae for multitarget therapy approaches to improve prognosis and survival outcomes. Further research on interactions between CHM and chemotherapeutics can provide more information on the safety and/or potential benefit of adjuvant therapies. In this way, the full holistic benefits of Chinese herbal medicines for cancer management can be realized.

References
1. G. Deng, B. Cassileth, *Nat. Rev. Clin. Oncol.* **10**, 656 (2013).
2. U. Werneke, D. Ladenheim, T. McCarthy, *Cancer Therapy* **2**, 475 (2004).
3. J. Huebner *et al.*, *Oncol. Res. Treat.* **37**, 304 (2014).
4. K. H. Chang *et al.*, *BMC Cancer* **11**, 196 (2011).
5. M. S. Moran *et al.*, *Int. J. Radiat. Oncol. Biol. Phys*. **85**, 40 (2013).
6. A. Naing *et al.*, *Cancer* **117**, 5142 (2011).
7. T. G. Liu *et al.*, *Evid. Based Complement. Alternat. Med.* **2012**, 769042 (2012).
8. D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **75**, 311 (2012).
9. Z. Wang *et al.*, *Evid. Based Complement. Alternat. Med.* **2013**, 268963 (2013).
10. L. Y. Wong *et al.*, *Patient Prefer. Adherence* **4**, 223 (2010).
11. J. C. Lim *et al.*, *Clin. Exp. Pharmacol. Physiol.* **39**, 300 (2012).
12. S. Prasad *et al.*, *Biotechnol. Adv.* **32**, 1053 (2014).
13. L. Li *et al.*, *Curr. Mol. Med.* **14**, 673 (2014).
14. G. G. Yue *et al.*, *Xenobiotica* **42**, 562 (2012).
15. P. M. Tang *et al.*, *Mol. Cancer* **8**, 56 (2009).
16. F. Hong *et al.*, *Planta Med.* **77**, 817 (2011).
17. G. Ragupathi *et al.*, *Vaccine* **26**, 4860 (2008).
18. G. G. Yue *et al.*, *Nutr. Cancer* **65**, 765 (2013).
19. G. G. Yue *et al.*, *Eur. J. Med. Chem.* **59**, 243 (2013).
20. J. C. Yeh *et al.*, *Angiogenesis* **14**, 187 (2011).
21. G. G. Yue *et al.*, *Br. J. Pharmacol.* **164**, 1883 (2011).
22. G. G. Yue *et al.*, *Planta Medica* **78**, 1183 (2012).
23. K. W. Luo *et al.*, *J. Nutr. Biochem.* **25**, 395 (2014).
24. J. Loganathan *et al.*, *Int. J. Oncol.* **44**, 2009 (2014).
25. X. Li *et al.*, *PLOS ONE* **8**, e60338 (2013).
26. Y. X. Bao *et al.*, *J. Altern. Complement. Med.* **12**, 771 (2006).
27. S. Kummar *et al.*, *Clin. Colorectal Cancer* **10**, 85 (2011).
28. M. W. Saif *et al.*, *Cancer Chemother. Pharmacol.* **73**, 373 (2014).
29. T. Aoyama *et al.*, *Cancer Chemother. Pharmacol.* **73**, 1047 (2014).
30. Y. Niu, Q. X. Meng, *J. Asian Nat. Prod. Res*. **15**, 550 (2013).
31. C. C. Lin *et al.*, *In Vivo* **25**, 633 (2011).
32. J. Chiu, T. Yau, R. Epstein, *Support Care Cancer* **17**, 231 (2009).
33. M. McCulloch *et al.*, *J. Clin. Oncol.* **24**, 419 (2006).
34. H. Guo *et al.*, *Integr. Cancer Ther.* **10**, 127 (2011).
35. L. Zhao *et al.*, *Evid. Based Complement. Alternat. Med.* **2014**, 625493 (2014).

Evaluating the safety of herbal medicines: Integrated toxicological approaches‡

Authors:
Elizabeth M. Williamson[1*],
Kelvin Chan[2],
Qihe Xu[3], Amandine Nachtergael[4],
Valérian Bunel[5,6],
Li Zhang[7],
Moustapha Ouedraogo[5,8], Joëlle Nortier[6], Fan Qu[9],
Debbie Shaw[10],
Xinmin Liu[11],
Caroline Stévigny[5],
Joseph Kahumba[12],
Olavi Pelkonen[13],
Pierre Duez[4,5*]

Many complex herbal mixtures are already commonly used worldwide, either for primary health care or as complementary or alternative medicines (*1*). Ancient traditional remedies—notably traditional Chinese medicine (TCM) and Ayurveda—have been passed down and refined over their long history of clinical use. Often perceived as innocuous, some herbs exhibit delayed or cumulative toxicity that may not be obviously attributed to the herbs, but instead identified by serendipity or unfortunate clinical findings. Given the large number of herbal products on the market and the relatively low budgets available for research to date, safety assessment in accordance with modern guidelines has been carried out on relatively few herbs (*2*). Despite these concerns, a recent survey of practitioners in Europe and China, although limited in scope, provides some reassurance that the vast majority of herbs in regular use are known to be relatively safe (*3*). Reports of serious adverse events regarding TCM mainly concern those that are used very rarely in Europe and extremely carefully in China (*4*). It is also important to differentiate between intrinsic herbal toxicity and malpractice: A recent Hong Kong study found that, of 52

[1]The School of Pharmacy, University of Reading, UK
[2]University of Western Sydney and Faculty of Pharmacy, Australia
[3]Department of Renal Sciences, King's College London, UK
[4]Department of Therapeutical Chemistry and Pharmacognosy, Université de Mons, Belgium
[5]Laboratory of Pharmacognosy, Bromatology and Human Nutrition, Université Libre de Bruxelles (ULB), Belgium
[6]Laboratory of Experimental Nephrology, Université Libre de Bruxelles, Belgium
[7]Center for Drug Re-evaluation, China Food and Drug Administration, Beijing, China
[8]Laboratory of Pharmacology and Toxicology, University of Ouagadougou, Burkina Faso
[9]Women's Hospital, School of Medicine, Zhejiang University, China
[10]Guy's and St Thomas NHS Trust, London, UK
[11]Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences, Beijing, China
[12]Department of Pharmacognosy, Université de Lubumbashi, D.R. Congo
[13]Department of Pharmacology and Toxicology, University of Oulu, Finland
*Corresponding Authors: e.m.williamson@reading.ac.uk (E.M.W.) and pierre.duez@umons.ac.be (P.D.)
‡This article is dedicated to our close colleague and co-author, Moustapha Ouedraogo, who died in January 2014.

clinical case reports of aconite poisoning, the majority were actually related to poor-quality herbs, poor prescribing practices, or dispensing errors (5). In Europe, adverse events have mainly resulted from contaminated products and a practitioner's incompetence, rather than herbal medicines being inherently risky. The reasons why the safety of herbal products for clinical use may be compromised are summarized in Figure 1. Pharmacovigilance systems have only recently been established for herbal medicines, thus the true incidence of adverse events may be under-reported; nevertheless the available data indicate that their overall safety is better than would be suggested by widely publicized incidents involving adulterated products and herbs already known to be toxic.

Preclinical and clinical toxicology research on herbal medicines is in most cases inadequate or insufficient to fulfil official medicine registration requirements. Safety has primarily relied on rational clinical use, avoiding drug interactions, ensuring correct botanical identification and labeling, quality control for adulterants

FIGURE 1. Factors influencing toxicity of herbal medicines.

and contaminants (mycotoxins and heavy metals, among others), mastery of sometimes complex processing methods, detection of new structural alerts, and avoidance of known toxicophore-bearing species. These measures are, however, inadequate for evaluating complex mixtures of incompletely known composition and where adverse events may take months or even years to manifest (2, 4).

Preclinical toxicology assessments for medicinal herbs

There is a need for realistic evaluation of the toxicity of herbal drugs as they are prepared traditionally and used clinically. Special attention must be paid to insidious toxicities which are not easily detected by pharmacovigilance, including genotoxicity, carcinogenicity, and developmental toxicity (6).

The complexity of herbal medicines is a major issue for safety assessments since the often large number of components have variable potencies and affinities for various targets. Toxicity, as with efficacy, may be the result of a mixture of active compounds rather than a single chemical entity, therefore different methods of testing are required to replace or supplement those routinely used in classical medicine risk assessment. Toxicological and pharmacokinetic studies of individual constituents of herbal materials must be compared with qualitative

and quantitative profiling data for the total extract, using appropriate analytical tools (7), in vitro and in silico methods (2), and modern screening techniques (8). In vitro models using subcellular organelles and human-derived cellular systems form the basis of pharmacological and toxicological screenings and are generally low-cost, efficient, and easy-to-handle. However, they are reductionist techniques, which focus on nonspecific phenomena at the cellular level and neglect pharmacokinetic and pharmacodynamic aspects relevant to clinical conditions that affect the nature and concentration of active moieties. Therefore, the recently developed in vitro 3-D tissue and organ-on-a-chip models—which simulate the in vivo cell microenvironment—should also be considered for toxicological and pharmacological herbal assessment. Herbal matrix effects must also be considered in the experimental design of toxicity screens. In vitro and in silico models should therefore be used only for comparison of toxicity profiles or deciphering toxic activities at the molecular level rather than for direct estimation of toxicological risks. In silico methods linking components with potential targets can provide an indication of adverse effects that may result from having multiple constituents and signal the need for a closer investigation of relevant toxicological and pharmacological issues. Measuring the kinetic characteristics of metabolism, transport, disposition of active components, and matrix effects should provide

TABLE 1. Integrated strategies for the toxicological assessment of herbal drugs.

Strategy	Information yielded
Analytical chemistry	• Chemical profiles, identification of components of interest or concern
In vitro, in silico methods and nonmammalian animal models, eventually combined with 'omics-based methods	• Identification of main pharmacological and toxicological issues • Comparison of toxicity profiles • Deciphering toxic activities at a molecular level
Mammalian models, eventually combined with 'omics-based methods	• Acute and repeated dose toxicities (animal models only) • Pharmacokinetic data, metabolism • Identification of pharmacological and toxicological issues
Clinical studies, eventually combined with 'omics-based, personalized methods	• Identification of exposure or toxicity biomarkers • Genotoxicity, carcinogenicity; reproductive and developmental toxicity
Pharmacoepidemiology and pharmacovigilance	• Evaluation of safety in clinical use • Identification of clinically relevant adverse effects

the necessary information to ensure that in vivo studies effectively integrate these critical parameters, and offer the opportunity to consider the realistic exposure levels in various organs and the biological activities at concentrations relevant to human consumption (*9, 10*). In vivo models such as zebrafish, *Drosophila*, and *Caenorhabditis elegans* offer convenient methods for high throughput toxicity assessment and inform further toxicological considerations, but must be cautiously interpreted due to their differences from human physiology.

Preclinical pharmacological and toxicological assessment of herbal medicines therefore needs to combine 'omics, bioinformatics, and network-centred systems biology with in vitro and in vivo assays, targeting studies on biomarkers identified from animal experiments (*11–13*), as summarized in Table 1. At all stages, it is imperative to remember the multicomponent nature and variability of herbal remedies, and that studies using wrongly identified or processed, contaminated, or adulterated preparations are not only useless but misleading for toxicological evaluation (*14*).

Pharmacovigilance and the clinical assessment of herbal medicines

To compensate for the lack of safety evaluations of herbal drugs, empirical clinical assessment must be carried out through pharmacovigilance. This can only happen if herbal products and practitioners are regulated in some way so that adverse events can be traced to a particular medicine or practice. Continuing professional development should be mandatory to ensure that practitioners and other healthcare workers are able to provide safe and appropriate advice to patients.

New techniques are being developed for evaluating the toxicological risk of herbal drugs in a clinical context. Urinary metabolomics appears to be a promising noninvasive tool for the prediction of specific organ toxicity (*15*). This type of active surveillance allows better assessment of causality and detection of possibly harmful components of the treatment, and can be adapted for other herbal medicines.

In most herbal traditions, certain combinations of herbal drugs are believed to reduce toxicity and others to enhance it. To rationalize formulations and enhance safety, it is crucial to take into account the complex chemical interactions taking place in herbal mixtures (and mixtures of mixtures) using systems-based approaches.

Conclusions

Herbal safety is compromised when any element of the herbal medicine-practitioner-patient triangle is flawed. To meet the challenge, integrating emerging systems-based technologies with conventional means is essential. There remains a clear and urgent need for novel methods able to rapidly pinpoint indicators of major mid-term and long-term toxicities, to yield warning signals, and identify those herbal drugs and formulae that need further toxicological investigation. Recent advances in 'omics and bioinformatics techniques have made it possible to investigate efficacy and toxicity at the organism level and in an individual manner. When further developed and validated, these methods should enhance the detection of insidious toxicities, provide the necessary background information for effective pharmacovigilance, and aid mechanistic studies of specific herbal medicines.

References
1. Q. Xu *et al.*, *BMC Complement. Altern. Med.* **13**, 132 (2013).
2. M. Ouedraogo *et al.*, *J. Ethnopharmacol.* **140**, 492 (2012).
3. E. M. Williamson, A. Lorenc, A. Booker, N. Robinson, *J. Ethnopharmacol.* **149**, 453 (2013).
4. D. Shaw, *Planta Med.* **76**, 2012 (2010).
5. S. P. Chen *et al.*, *Drug Safety* **35**, 575 (2012).
6. J. Zhou, M. Ouedraogo, F. Qu, P. Duez, *Phytother. Res.* **27**, 1745 (2013).
7. Y. Z. Liang, P. S. Xie, K. Chan, *Comb. Chem. High Throughput Screen.* **13**, 943 (2010).
8. V. Bunel, M. Ouedraogo, A. T. Nguyen, C. Stevigny, P. Duez, *Planta Med.* **80**, 1210 (2014).
9. O. Pelkonen *et al.*, *J. Ethnopharmacol.* **140**, 587 (2012).
10. I. C. Munro, A. G. Renwick, B. Danielewska-Nikiel, *Toxicol. Lett.* **180**, 151 (2008).
11. M. L. Coghlan *et al.*, *PLOS Genet.* **8**, e1002657 (2012).
12. K. M. Wu, J. G. Farrelly, R. Upton, J. Chen, *Phytomed.* **14**, 273 (2007).
13. S. G. Newmaster, M. Grguric, D. Shanmughanandhan, S. Ramalingam, S. Ragupathy, *BMC Med.* **11**, 222 (2013).
14. K. Chan *et al.*, *J. Ethnopharmacol.* **140**, 469 (2012).
15. S. C. Hsieh, J. N. Lai, P. C. Chen, H. J. Chen, J. D. Wang, *Pharmacoepidemiol. Drug Saf.* **15**, 889 (2006).

Combining 'omics and comparative effectiveness research: Evidence-based clinical research decision-making for Chinese medicine

Authors:
Claudia M. Witt[1,2,3*],
Jianping Liu[4],
Nicola Robinson[4]

Systematic reviews and meta-analyses of Chinese medicine trials have demonstrated issues with consistent quality, and an evidence gap between the practice of, and research on, traditional treatments. Clinical practice is built on knowledge, clinical experience, and patient preferences, all of which can be influenced by values and belief systems. A current movement in clinical medicine research, known as comparative effectiveness research, supports the development of evidence-based recommendations to enable more informed decision-making in the clinic and more valid health policies that also meet the criteria for practicing "P4" (predictive, preventive, personalized, and participatory) medicine. Creating a modern, strategic research framework for Chinese medicine that takes into account the stakeholders' perspectives, follows a patient-centered approach, uses mixed methods research methodologies, and combines modern scientific techniques such as systems-biology–based 'omics technologies would be beneficial for bridging the gap between Chinese medicine theory and modern clinical research methodologies.

Background

The most prominent medical research model compares "one disease, one treatment;" however, this strategy often does not have comparable clinical practices. Further, clinical trials based on such research models are usually performed in a standardized setting with a carefully selected patient group, and often produce results that are neither generalizable nor able to guide and inform clinical care. Systematic reviews and meta-analyses summarizing such trials might even be misleading for various chronic diseases, but especially for complex conditions such as diabetes, cardiovascular disease, and pain, which often occur in patients with multiple comorbid diseases who are receiving a number of different treatments. Decision makers—clinicians, patients, and funders—require studies that are comparable with actual treatment options in real life settings (1). Comparative effectiveness research (CER) is intended to provide real-world evidence that helps clinicians and patients choose the options that best fit the individual's needs and preferences (2). CER involves the stakeholders' needs at all relevant steps and includes a number of different types of research designs, clinical trials being one of them. These so called pragmatic trials are characterized by including more "real-life" patients presenting in routine clinical care including

those that have comorbidities and use comedication, providing more individualized treatments, using patient-relevant outcomes, and being performed in a setting that is "in line" with routine clinical care (3).

Chinese medicine has been historically based on a descriptive and phenomenological approach and has relied on complex mixtures of herbal medicines as well as nonpharmacological interventions such as acupuncture and lifestyle advice. Research on some of the individual treatment components of Chinese medicine (e.g., acupuncture) have already made relevant contributions to CER evidence (4) and provided guidance for the design of further acupuncture CER (5). However, in clinical practice Chinese medicine is a complex intervention that focuses on the whole system's organization, and not on physiological pathways or single targets. Treatments are built on knowledge accumulated from ancient texts, experts, clinical experiences, and patient preferences, which are influenced by values and belief systems (6). Chinese medicine has a fundamental patient participation element, including general lifestyle aspects (e.g., diet and exercise) in the complex intervention strategies. Traditional Chinese diagnoses (or "syndrome differentiation"), a comprehensive analysis of clinical information from a Chinese medicine perspective (e.g., information derived from case taking, examining the patient's pulse and tongue), is used to guide personalized treatment options (7). Each syndrome consists of symptoms that determine their own unique treatment protocol.

Integrating syndrome differentiation with the biomedical techniques of modern clinical practice would be helpful for determining personalized treatments (8). The beta version of the International Classification of Diseases (ICD) 11 already allows Chinese syndrome coding in addition to Western diagnoses (9). Currently, these Chinese syndromes are considered important tools for predicting disease (10, 11) and ongoing efforts are correlating them with measurable biomarkers. Recently, a systems-biology–based approach has been utilized for Chinese medicine syndrome differentiation studies enabling the stratification of patient populations (8). This strategy may help researchers optimize their clinical trial design by having the ability to determine which patients are most appropriate for a specific intervention. One advantage of a systems biology approach is that it aims to understand both the connectivity and interdependence of individual components within a dynamic and nonlinear system, such

[1]Institute for Complementary and Integrative Medicine, University Hospital Zurich and University of Zurich, Zurich, Switzerland
[2]University of Maryland School of Medicine, Center for Integrative Medicine, Baltimore, MD
[3]Institute for Social Medicine, Epidemiology and Health Economics, Charité – Universitätsmedizin, Berlin, Germany
[4]Centre for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China
[5]Faculty of Health and Social Care, London South Bank University, London, UK
*Corresponding Author: claudia.witt@uzh.ch

as traditional medicine, as well as the properties that emerge at different organizational levels. In addition, the use of 'omics techniques (including genomics, proteomics, metabolomics, transcriptomics, and lipidomics) align very closely with the concepts and practices of Chinese medicine (8). Using 'omics-based techniques in Chinese medicine presents the unique opportunity to better understand both personalized diagnoses and the systems-based interventions of Chinese medicine (12). Two recent studies on rheumatoid arthritis have demonstrated the potential of this approach for conventional medicine. By combining Chinese syndrome diagnoses with the identification of biomarkers and the use of genomics to determine the diagnostic subgroups, opportunities for better treatment outcomes were provided (13, 14).

In the future, it is hoped that P4 medicine will enable the prediction and prevention of diseases rather than reactive health care (15). Understanding how genomic differences in individuals, along with an individual's environmental exposures, influence biological systems has the potential to enable medical professionals to make patient-specific predictions followed by personalized treatments or, even better, preventive interventions. In the future, health care consumers will be increasingly equipped with their personal health information, including genome sequences, molecular profiles of diseased tissues, and biomarker panels (16). Participation from all major stakeholders will be needed to provide clinical and health policy guidance for this new medical era.

Moreover, one strategy CER could benefit from is incorporating at least genomics as part of a future research approach (17). Including 'omics techniques into CER would be a new area for Chinese medicine—one yet to be incorporated into methodological CER guidance (18). It could directly bridge the gap between the personalized approach of Chinese medicine theory and Western science. The trend toward P4 medicine in CER also creates an ideal setting to provide information from large, real-life populations. Furthermore, the characteristics of P4 medicine dovetails well with the foundations of Chinese medicine.

Below, we have proposed some recommendations that combine the underlying concepts of CER with systems biology based 'omics technologies in order to collect scientific evidence for Chinese medicine that can be used to broaden the use of traditional medicine and optimize clinical decision-making.

Recommendations

When there is insufficient evidence for a treatment, a combination of both existing data from trials and systematic reviews of the literature should be used to inform future research and support clinical decision-making. This requires:

• Tools for systematic reviews and meta-analyses that provide comprehensive information about both the context of the studies and the extent to which the results are generalizable

• Secondary data-analyses of existing studies that have utilized a systems biology approach to identify possible associations between syndrome differentiation in Chinese medicine, other patient characteristics, and disease progression.

Future clinical research on Chinese medicine would benefit from combining the evolving CER methodology, modern systems biology 'omics approaches, and patients' needs during routine care. In practice, this would require:

1. Strategic clinical trial designs:

• Trials that include heterogeneous and "realistic" patient samples, are performed in settings reflective of a patient's routine care, and have sample sizes that facilitate further subgroup analyses

• Trial designs that balance multiple factors, including the type of study and context of the relevant diagnostic scenarios with both qualitative information (e.g., Chinese syndrome differentiation, patient preferences, and expectations) and quantitative parameters (e.g., systems biology including 'omics analysis)

• The development of guidance on the appropriate outcome measures for future research

• Realistic treatment protocols that shift patient treatments toward a personalized care model that allows the use of complex interventions, including lifestyle factors, and reflects the changes patients experience during routine clinical practice; 'omics-based analyses should be used to answer open questions about the complex pharmacological networks that are activated by complex herbal preparations (12)

• More studies with a diagnostic focus are needed to learn how predictive factors from Chinese medicine correlate with modern quantitative parameters.

2. Stakeholder involvement:

• Each stakeholders' needs (patients, clinicians, government, and payers) should be taken into account when identifying a study's focus, planning clinical trial designs, and interpreting results; to allow a high level and systematic inclusion of stakeholder viewpoints, both qualitative and quantitative research methodologies would need to be applied (19).

References
1. M. Helfand *et al.*, *Clin. Transl. Sci.* **4**, 188 (2011).
2. P. H. Conway, C. Clancy, *N. Engl. J. Med.* **361**, 330 (2009).
3. M. Zwarenstein *et al.*, *BMJ* **337**, a2390 (2008).
4. C. M. Witt *et al., PLOS ONE* **7**, e32399 (2012).
5. C. M. Witt *et al., BMC Complement. Altern. Med.* **12**, 148 (2012).
6. N. Robinson *et al.*, *J. Ethnopharmacol.* **140**, 604 (2012).
7. M. Jiang *et al.*, *J. Ethnopharmacol.* **140**, 634 (2012).
8. J. van der Greef, *Nature* **480**, S87 (2011).
9. ICD-11 Beta. http://apps.who.int/classifications/icd11/browse/f/en (2012).
10. J. Dai *et al.*, *Evid. Based Complement. Alternat. Med.* **706762** (2013).
11. S. Sun *et al.*, *Evid. Based Complement. Alternat. Med.* **453503** (2012).
12. A. Buriani *et al.*, *J. Ethnopharmacol.* **140**, 535 (2012).
13. H. A. van Wietmarschen *et al.*, *PLOS ONE* **7**, e44331 (2012).
14. H. A. van Wietmarschen *et al.*, *J. Clin. Rheumatol.* **15**, 330 (2009).
15. C. Auffray, D. Charron, L. Hood, *Genome Med.* **2**, 57 (2010).
16. L. Hood, Q. Tian, *Genomics Proteomics Bioinformatics* **10**, 181 (2012).
17. N. I. Simonds *et al.*, *J. Natl. Cancer Inst.* **105**, 929 (2013).
18. C. M. Witt *et al.*, *Trials.* **15,** 169 (2014).
19. P. A. Deverka *et al.*, *J. Comp. Eff. Res.* **1**, 181 (2012).
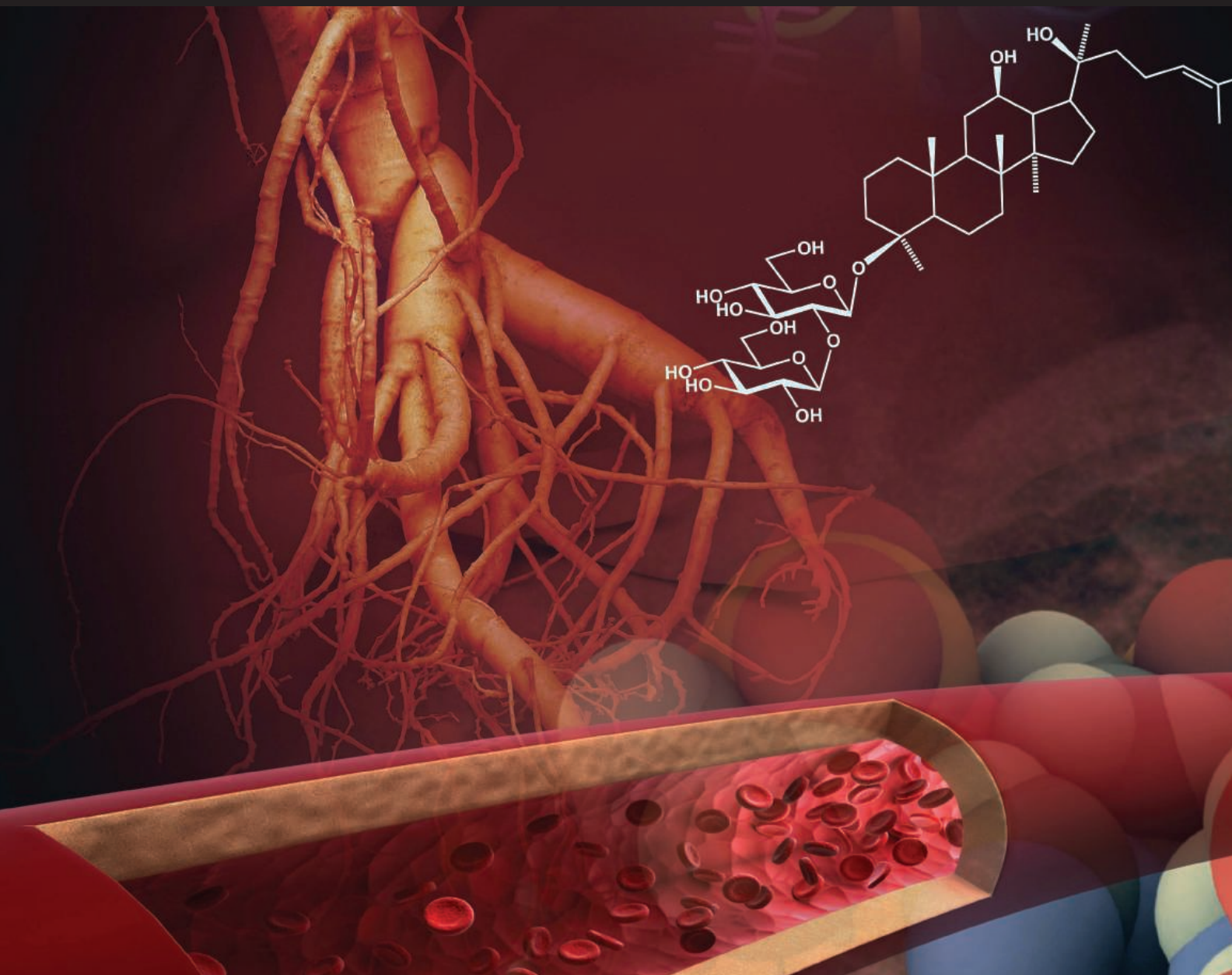
The American Association for the Advancement of Science
1200 New York Avenue NW
Washington, DC 20005

Spectrophotometers

Qualoupe Lite is a new low-cost, intuitive laboratory information management system (LIMS). Compatible with Jenway 7300, 7305, 67 series, Genova Plus, and Genova Nano spectrophotometers, Jenway 4510 bench conductivity meter, and the Stuart SMP40 automatic melting point, the system provides a convenient solution for laboratories in need of simple, auditable data storage. Qualoupe Lite is ideal for small and medium laboratories in the academic and commercial sectors, delivering powerful but straightforward functions for storing and retrieving laboratory results and data. This affordable LIMS eliminates the need to spend large amounts of money on overcomplicated software—a common problem faced by laboratories looking to upgrade from a traditional paper or multiple PC storage system. Using Qualoupe Lite, analysis results and method information can be transferred easily from the instruments to the database, either by a direct connection or via removable media, ready to be recalled for review at any point.

Bibby Scientific
For info: +44-(0)-1785-812121
www.bibby-scientific.com

Fluorescence Spectroscopy System

The new FluoroMax Plus is a convenient, affordable, and easy-to-use benchtop spectrofluorometer. The new FluoroMax Plus expands on the already class-leading capabilities of our FluoroMax series with the option of a second detector and two position grating turret, providing an extended range out to 1,650 nm, and TCSPC performance down to 25 ps. The higher sensitivity offered by FluoroMax Plus enables faster measurement of weaker and smaller samples. Horiba Scientific offers a broad range of gratings, detectors, and accessories for the FluoroMax series to fit any research application needs. These include measurement of solid and liquid samples, high throughput screening, cryogenic or elevated temperatures, absolute quantum yields, microliter volumes, and even micron-scale measurements with a fiber optic coupling to a microscope.

Horiba Scientific
For info: 877-546-7422
www.fluorsolutions.com

Microtube Homogenizer

The BeadBlaster 24 Microtube Homogenizer can grind, lyse, or homogenize up to 24 samples simultaneously. Tubes in the stainless steel tube carrier are subjected to an intense 3-D motion, causing high-energy impacts between the samples and microbeads to release their cellular contents. After homogenization, samples are centrifuged and the supernatant collected for additional processing. Operation of the BeadBlaster 24 is programmable, allowing for short bursts of homogenizing with rests in between to protect sensitive samples. Up to 50 programs can be stored in memory and recalled quickly. The transparent lid prevents access to the samples when the unit is in operation while a brushless motor provides cool and quiet operation. Benchmark offers a variety of prefilled tube kits to meet your sample needs.

Benchmark Scientific
For info: 908-769-5555
www.benchmarkscientific.com

Chromatography Medium

Capto S ImpAct cation exchange chromatography medium is designed for efficient and flexible purification of monoclonal antibodies (MAbs). Capto S ImpAct is a strong cation exchanger for the intermediate and polishing steps when purifying MAbs and a wide range of other biomolecules. Capto S ImpAct offers both high binding capacity and high resolution at high sample loads. The high-flow characteristics of the Capto base matrix enables excellent pressure-flow properties. The small bead size for high resolution, in combination with the novel extended-surface ligands, supports high dynamic binding capacity at high sample loads. The ability to run at both high flow rates and bed heights combines with these attributes to increase productivity and flexibility in process design. Key benefits of Capto S ImpAct include: high binding capacity, high resolution for efficient aggregate removal at high sample loads, large operational window of flow rates and bed heights for easy optimization and scale up, and high productivity.

GE Healthcare Life Sciences
For info: 800-526-3593
www.gelifesciences.com/captosimpact

Biological Sample Concentrator

The miVac Quattro is the perfect workhorse sample concentrator for biological laboratories tasked with handling larger numbers of samples. The miVac Quattro centrifugal vacuum concentrator rapidly removes a wide range of solvents, from volatile organic solvents through to DMF, including water, making it the ideal system for concentrating precious samples in applications including oligonucleotide synthesis, RNA/DNA preparation, peptide preparation, sequencing, molecular biology, and ADME/toxicology research. A wide variety of high-capacity rotors means that the miVac Quattro is capable of removing water and organic solvents from biological samples in a range of formats including tubes, microplates, and vials. Fitted with a four-swing position rotor the miVac Quattro can simultaneously accommodate up to twenty shallow-well microplates or eight deep-well plates enabling significant increases in throughput and slashing drying times. The intuitive controls allow inexperienced users to get first-class results first time with most samples, while providing more sophisticated programming capability for experienced workers.

Genevac
For info: +44-(0)-1473-240000
www.genevac.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/products/newproducts.dtl for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.

Warwick Integrative Synthetic Biology
WISB

THE UNIVERSITY OF
WARWICK

We are offering exciting new research employment opportunities in multiple areas of synthetic biology. The aim of this burgeoning area of science is to construct and test new biological systems utilising either modified versions of naturally evolved biological components or entirely synthetic components.

The newly established Warwick Integrative Synthetic Biology Centre (WISB) develops cutting-edge technologies to generate enhanced understanding of biological systems and to deliver innovative solutions to important challenges in human and animal health, food security and the environment. General information about WISB can be obtained from the Director (Professor John McCarthy, John.McCarthy@warwick.ac.uk) or the Deputy Director (Professor Declan Bates, D.Bates@warwick.ac.uk).

Fifteen Postdoctoral Research Positions

(WISB; http://www.bbsrc.ac.uk/pa/grants/AwardDetails.aspx?Funding Reference=BB/M017982/1).

£28,695 - £34,233 pa

This first round of appointments will contribute to the development of the four WISB research themes. Candidates for the PDRA posts should have demonstrated the capacity to apply cutting-edge computational and/or experimental methods to research areas relevant to the projects outlined below. Further details about the respective posts can be obtained informally via the contacts indicated below.

Theme 1, Predictive Biosystems Engineering, is the underpinning research theme, integrating computational and experimental methods to inform our fundamental understanding of synthetic systems and thereby driving innovative approaches in our applied research areas. There are seven posts available:

- **Modelling and control of novel synthetic systems**
  (3 years; Professor Declan Bates) Ref: 75429-124

- **A modelling framework for synthetic biology**
  (2 years; Dr Sara Kalvala, Sara.Kalvala@warwick.ac.uk)
  Ref: 75430-124

- **Noise-aware design tools for synthetic biology**
  (2 years; Professor Nigel Stocks, N.G.Stocks@warwick.ac.uk)
  Ref: 75431-124

- **Novel regulatory components for robust SGCs and Synthetic protein circuitry** (3 years each; Professor John McCarthy)
  Ref: 75432-124

- **RNA-based circuitry in multiple hosts**
  (3 years; Professor Alfonso Jaramillo) Ref: 75433-124

- **Noise optimisation for SGCs in eukaryotic hosts**
  (3 years; Dr Daniel Hebenstreit, D.Hebenstreit@warwick.ac.uk)
  Ref: 75434-124

Theme 2, Engineering Biosynthetic Pathways, uses novel combinatorial assembly methods to build biosynthetic pathways *de novo* in actinobacteria. There are two posts available:

- **Biosynthetic pathway elucidation**
  (3 years; Professor Greg Challis, G.L.Challis@warwick.ac.uk)
  Ref: 75435-124

- **Orthogonal inducible expression systems for actinobacteria**
  (3 years; Dr Chris Corre, C.Corre@warwick.ac.uk) Ref: 75436-124

Theme 3, Engineering Microbial Communities, utilises novel methods to build defined (bespoke) microbial communities that can be utilized in a wide range of applications. There are three posts available:

- **Synthetic microbial communities for conversion of sunlight**
  (2 years; Professor Orkun Soyer) Ref: 75437-124

- **Synthetic gut microbial communities**
  (2 years; Dr Yin Chen, Y.Chen.25@warwick.ac.uk) Ref: 75438-124

- **Engineering microbial communities for bioremediation**
  (2 years; Dr Munehiro Asally, m.asally@warwick.ac.uk) Ref: 75439-124

Specific details of the posts and the postdoctoral roles as well as information on how to apply are available via the HR website below.

Closing date: 15 February 2015

Theme 4, Engineering Microbial Effector Systems in Plants, constructs new types of genetic regulatory pathway in plants with a view to introducing enhanced properties related to stress- and pathogen- resistance. Three posts are available:

- **SynEffector modules to re-engineer signalling pathways**
  (3 years; Dr Vardis Ntoukakis, V.Ntoukakis@warwick.ac.uk)
  Ref: 75440-124

- **Novel activation of signalling pathways**
  (3 years; Dr Patrick Schäfer, P.Schafer@warwick.ac.uk) Ref: 75441-124

- **Designing synthetic signalling pathways**
  (2 years; Dr Katherine Denby, K.J.Denby@warwick.ac.uk)
  Ref: 75442-124

We pursue a broad programme of research into the ethical, legal and societal aspects of synthetic biology (for more information contact Dr Nick Lee, N.M.Lee@warwick.ac.uk), offer undergraduate and postgraduate training in synthetic biology, and run an iGEM team (Professor Alfonso Jaramillo, Alfonso.Jaramillo@warwick.ac.uk). WISB is a partner in the RCUK-funded Centre of Doctoral Training in Synthetic Biology (http://www2.warwick.ac.uk/fac/sci/sbcdt/ Warwick contact: Professor Orkun Soyer, O.Soyer@warwick.ac.uk).

WISB Administrative Manager

£38,511 - £45,594 pa (subject to job evaluation)

The appointee will work closely with the WISB management team to support operational and financial aspects of the centre's activities. The post holder will have multiple responsibilities, including preparing and delivering reports, monitoring expenditure and progress in WISB activities, arranging workshops and meetings, and promoting external activities (including outreach to the wider public). This challenging position will be suited to an individual with a strong background in both research and project management.
Ref: 75443-124

WISB Research Support Facility Manager

£28,695 - £37,394 pa (subject to job evaluation)

The responsibility of this role will be to manage the core equipment facilities of WISB, which will feature, among other technologies: flow cytometry, high resolution microscopy, robotics, microfluidics, microbioreactors, mass spectrometry, RT-PCR and biophysical instrumentation, including maintenance requirements and associated expenditure. In addition, the appointee will provide training to WISB personnel and to WISB partners and support the generation of research data. Candidates should have experience in the relevant technical areas. Ref: 75444-124

WISB Data Manager

£28,695 - £37,394 pa (subject to job evaluation)

The appointee will help develop and manage the computational infrastructure required to support the research activities of WISB, and oversee the procedures for storage and utilisation of data within the centre. The role will also include establishing and supporting data analysis pipelines, maintaining the WISB high performance computing cluster and supporting the WISB website. Candidates should be proficient in computational/mathematical skills and have technical management experience. Ref: 75445-124

The three appointments above will be for up to five years.

BBSRC

EPSRC
Pioneering research
and skills

The University Values Diversity   **http://www2.warwick.ac.uk/services/humanresources/jobsintro/**

Auburn University
Harrison School of Pharmacy
Department of Drug Discovery and Development
Assistant/Associate/Full Professor
Pharmaceutical and Biomedical Sciences

The Department of Drug Discovery and Development at Auburn University's Harrison School of Pharmacy invites applications for multiple 12-month tenure-track faculty positions. Individuals with expertise in cardiovascular disease, cancer, neuroscience, infectious disease, cellular signaling, synthetic medicinal chemistry, pharmacognosy/natural products chemistry, pharmacokinetics/drug metabolism, or pharmaceutics/drug formulation/drug delivery are of particular interest. These new positions offer the opportunity to join an expanding enterprise in pharmaceutical and biomedical research that features new and enhanced research facilities and infrastructure, as well as a partnership with the Auburn campus of the Edward Via College of Osteopathic Medicine. Moreover, Auburn University's Research Park offers opportunities for pharmaceutical and biomedical technology transfer and entrepreneurial incubation. Opportunities for collaboration are present within the School of Pharmacy as well as with faculty throughout the University, including those in the College of Veterinary Medicine, the College of Engineering and the College of Sciences and Mathematics.

Successful candidates are expected to:
- Develop and/or maintain a sustainable, extramurally-funded research program;
- Lead collaborative, interdisciplinary research programs;
- Demonstrate a high level of scholarly activity, as evidenced by quality publications in peer-reviewed scientific journals and active participation in professional societies;
- Actively participate in the teaching mission of the Harrison School of Pharmacy at the graduate (M.S., Ph.D.) and professional (Pharm.D.) levels; and
- Contribute to Department, School and University service activities.

Applicants must have a Ph.D. or equivalent degree, postdoctoral training and demonstrated research abilities, strong creative problem solving characteristics with a strong desire and track record for establishing and maintaining a competitive extramurally funded research program. The successful candidate will be provided with competitive start-up funding and salary commensurate with experience. Excellent written and interpersonal communication skills are required. All applicants must meet eligibility requirements to work in the U.S. at the time the appointment is scheduled to begin and to working legally for the proposed term of employment. Candidates should submit a letter of application, curriculum vitae, statements of immediate and long term research aspirations and teaching interests, as well as the names and addresses (include e-mail address and phone number) of 3 references. Review of applications will begin **February 1, 2015**. Interested candidates should apply on line at the following link: **http://aufacultypositions.peopleadmin.com/postings/833**

Auburn, AL offers a unique professional and living environment, and is located approximately 100 miles south of Atlanta, GA, 50 miles from Montgomery, AL, 40 miles from Columbus, GA, and 120 miles south of Birmingham, AL. The Auburn and Opelika communities consist of approximately 100,000 people and offer a high quality of life with excellent schools and recreational opportunities. Auburn is a true multi-cultural college town with all the amenities associated with a full-service university. Auburn University is one of the nation's premier land, sea, and space grant institutions. In the 2014 edition of U.S. News & World Report, it was ranked 48th among public universities. Auburn is a high research activity institution, offering Bachelor's, First Professional, Master's, Educational Specialist, and Doctor's degrees. Auburn University has 1,184 full-time faculty, and in Fall 2014 enrolled 25,912 students from all fifty states, the District of Columbia, Puerto Rico and approximately 85 countries, including 20,629 undergraduates, 4,198 graduate students, and 1,085 first professional students.

Auburn University is an EEO/Vet/Disability Employer.

Linköping University

Post-Doctoral Fellows / Research Associates

The Center for Social and Affective Neuroscience at Linköping University will launch soon, established with long term funding from the Swedish Research Council and other sources. CSAN aims to become a vibrant international research center combining the strengths of its constituent research groups. Main areas of focus are: motivation and emotion as they relate to addiction (**Prof. Markus Heilig**); the social neuroscience and physiology of affective touch (**Assoc. Prof. India Morrison / Prof. Håkan Olausson**, **www.grasplab.se**); and the interaction of affect and decision making (**Prof. Daniel Västfjäll / Assoc. Prof. Gustav Tinghög**). The Center also has a preclinical / translational component (**Assoc. Profs. David Engblom and Annika Thorsell**).

The center is now seeking to fill several positions for projects involving functional brain imaging. We are seeking post-doctoral fellows or research associates with a strong background in human neuroscience, and a documented publication track record of probing cognitive and affective processes using fMRI.

For its imaging needs, the center relies on the Center for Medical Image Science and Visualization at Linköping University which provides state of the art technology incl. a research dedicated 3T magnet and extensive computational infrastructure.

To inquire about these positions or to submit an expression of interest (CV plus a 2 page maximum research statement), please e-mail research coordinator **Ann-Charlotte Johansson** at **ann-charlotte.johansson@liu.se**.

Submission deadline February 28, 2015.

MICHIGAN STATE UNIVERSITY

Open Position in Immunology
Department of Microbiology and Molecular Genetics

The Department of Microbiology and Molecular Genetics at Michigan State University seeks candidates at the Assistant, Associate, or Full Professor level in the broad area of immunology, a research focus currently expanding at MSU. The Department is highly collegial and collaborative, with diverse research interests, **including molecular and cellular immunology**, bacterial and viral pathogenesis, prokaryotic and eukaryotic genetics, systems biology, and evolution at the molecular level. State of the art research support facilities (including high performance computing, flow cytometry, bio-informatics, genomics, proteomics, etc.) are available to the successful applicant. MSU is home to Colleges of Human Medicine, Osteopathic Medicine, and Veterinary Medicine. We invite applications from outstanding researchers who desire a highly collaborative and diverse environment. The successful applicant is expected to develop (or maintain) a rigorous, externally-funded research program with national visibility. Academic rank will be commensurate with experience. Teaching in graduate, professional, or undergraduate programs is expected.

Applications should be submitted electronically at **https://jobs.msu.edu** (posting **#0667**) with requisite material (letter of interest, curriculum vitae listing past and current funding if appropriate, a statement of future research goals, and names of three potential references (not to be contacted until approval is received from the applicant)) uploaded as a single PDF. Review of applications will begin immediately and remain open until filled.

Additional information could be obtained via email to the Search Committee Executive Assistant at **mmgchair@msu.edu**; **http://www.mmg.msu.edu**

MSU is an Affirmative-Action, Equal-Opportunity Employer and is committed to achieving excellence through diversity. The University actively encourages applications of women, persons of color, veterans, and persons with disabilities, and we endeavor to facilitate employment assistance to spouses or partners of candidates for faculty and academic staff positions.

EMBL
40 YEARS | 1974 2014

EMBL offers a highly collaborative, uniquely international culture. It fosters top quality, interdisciplinary research by promoting a vibrant environment consisting of young, independent researchers with access to outstanding graduate students and postdoctoral fellows.

Group Leader / Laboratory Head Biological X-ray Free Electron Laser Research at EMBL Hamburg, Germany

The most powerful X-ray Free Electron Laser (European XFEL, www.xfel.eu) is presently under construction in Hamburg/Schenefeld and is expected to start operation in 2017. EMBL will provide its expertise as the coordinator of the XFEL-based Biology Infrastructure (XBI) user consortium to establish a state-of-the-art sample preparation facility for life science applications using the future XFEL. The goal of this facility is to serve both the members of the XBI user consortium and the international research community interested in X-ray FEL based life science applications.

We are seeking an outstanding candidate to lead this future facility and to carry out excellent research in biological X-ray FEL applications. The future Group Leader will have the following tasks:

- To lead the construction and subsequent operation of the future X-ray FEL-based infrastructures for preparation, characterisation and delivery of biological samples, in consultation with the members of the XBI consortium and in close cooperation with other relevant European XFEL user consortia and the European XFEL GmbH in Hamburg, Germany. The Group Leader will represent EMBL in the XBI user consortium and other relevant committees.

- To carry out research in life science applications of X-ray FEL radiation, possibly combined with synchrotron radiation. This may include research & development of relevant methods, approaches and protocols; or focus on a biological theme of high scientific relevance, using X-ray scattering and diffraction experiments, including time-resolved studies, and ranging from crystalline samples to single particles. Collaborations with other group leaders from EMBL with complementary research interests are encouraged.

Qualifications and Experience

- Demonstrated expertise and track record of challenging structural biology and/or imaging projects, making use of synchrotron radiation and preferably X-ray free electron laser radiation;

- Proven expertise in biological sample production, biophysical and biochemical characterisation & crystallisation;

- At least three years of postgraduate (PhD or similar) expertise and an excellent research track record (publications, fellowships, awards).

State-of-the-art infrastructures for applications in life sciences on DESY campus, Hamburg, Germany, provide a world-class environment for the most challenging experiments in structural biology with the highest scientific significance. The EMBL Hamburg Unit has built an integrated facility in structural biology, which includes three brand-new synchrotron radiation beamlines at the PETRA III storage ring and a user facility for sample preparation, characterisation and crystallisation. EMBL is a partner of the new Centre for Structural Systems Biology (CSSB) in Hamburg focusing on research in structural biology, systems biology and infection biology.

Application Instructions

Please apply online through **www.embl.org/jobs** and include a cover letter, CV and a concise description of research interests and future research plans.

Please also arrange for 3 letters of recommendation to be emailed to **references@embl.de** at the latest by 28 February 2015.

Interviews are planned for March and April 2015.

Further details on Group Leader appointments can be found under **www.embl.org/gl_faq**.

EMBL is an inclusive, equal opportunity employer offering attractive conditions and benefits appropriate to an international research organisation.

www.embl.org

Learn more and bring your job search in for a smooth landing.

- Search thousands of job postings
- Create job alerts based on your criteria
- Get career advice from our Career Forum experts
- Download career advice articles and webinars
- Complete an individual development plan at "myIDP"

Target your job search using relevant resources on **ScienceCareers.org**.

ScienceCareers

FROM THE JOURNAL SCIENCE   AAAS

A European postdoc for the family

In the spring of 2011, I started a 2-year Whitaker International postdoctoral scholarship in bioengineering. My husband, an experienced mechanical engineer who is fluent in German, left his job in the United States and found a new one in Vienna. The next year, I learned I was pregnant. We didn't go to Europe to start a family—our goal was to experience a new culture and remain competitive in our fields—but we recognized almost immediately the advantages of Austria's support for young professionals pursuing careers and a family.

Austria is far ahead of the United States in providing options for both women *and* men to take time off from work for birth and child care. First there is *Mutterschutz*, a mandatory paid maternity leave and job protection period of 8 weeks before and 8 weeks after the birth of a child. Then there is *Elternkarenz*, up to 36 months of financial support with allowances of as much as 80% of your monthly wage. Parents frequently opt for 14 months, with the father taking 2 months.

Other European countries offer similar benefits. Dutch employment law, for example, entitles working women to 16 weeks of paid *zwangerschapsverlof*, men to 2 days of paid *vaderschapsverlof*, and both parents to 6 months of additional unpaid leave with job protection. In the United Kingdom, statutory maternity pay (SMP) is granted for up to 39 weeks, and employers often supplement these benefits with occupational maternity pay. Under the SMP program, fathers can receive 2 weeks of paid leave, with an additional 26 weeks if the mother returns to work. Specific restrictions vary, but non-E.U. nationals can qualify as long as they have an employment contract and an appropriate visa. A science Ph.D. makes it easier to secure the required credentials.

According to a Pew Research Center report covering 38 countries around the world, the United States is the only one that does not mandate paid leave for new mothers. California is one of a few states to offer paid family leave, available to workers who have contributed to the state disability insurance program. But in most of the country, the postdoc parenting experience is institution dependent.

After our son was born, I took 5 months off and my husband took 3 months. We wanted to share parenting responsibilities equally, as much as nature allowed. He planned to take off 5 months, too, but job opportunities in the United States brought us back sooner than expected. Still, my husband was able to experience caring for our son full time while I worked; they bonded in a way that is hardly possible for working fathers in the United States, where paid paternity leave is rare. Being the primary caregiver for 3 months allowed my husband time to find his own way of parenting while I got back to work.

I wanted to stay connected during my leave, so I worked remotely when I could and often brought my son to lab meetings, which my colleagues were kind enough to schedule around naps. My immediate team of male colleagues did not seem to mind the infant intrusion— it helped that he was a calm baby— and meetings were productive. This flexibility, and the absence of pressure to return to work, allowed ample time to connect with my son as I learned to be a mom. Work was secondary when it needed to be.

In the United States, the pool of qualified postdocs has grown and postdocs have gotten longer. There's now a greater likelihood than ever that training will overlap with starting a family. The decision when and where to have children is personal and depends on many factors; there's something to be said, for example, for having your mother nearby. But in deciding where to train, postdocs should consider the whole experience of working and living, not just time spent in the lab. Add to the mix Europe's ample opportunities for professional enrichment, and the parental-leave advantages that Europe can offer postdocs are worth considering. ∎

> *"The parental-leave advantages that Europe can offer postdocs are worth considering."*

Michelle Gabriele Sandrian is now an assistant professor of ophthalmology and bioengineering at the University of Pittsburgh in Pennsylvania. For more on life and careers, visit www.sciencecareers.org. Send your story to SciCareerEditor@aaas.org.