Science

Choosing a climate future

Beyond the Paris talks

CONTENTS

1030 & 1061

Reflections on water splitting

ON THE COVER

Symbolizing the choices facing climate negotiators next week in Paris, a windmill and the smokestack of a coal-burning power plant stand less than a kilometer apart in a coal-mining area of North Rhine-Westphalia, Germany. See pages 1007, 1016, and 1034. *Photo: © Jochen Tack/imageBROKER/Corbis*

NEWS

INSIGHTS

1038 & 1077
Cichlid extinction

1069
Flow paths through salt

Published by AAAS

Turning the focus to solutions

Next week, the United Nations Climate Change Conference, COP21, will be held in Paris. The goal is to achieve an international agreement to stem climate change—in particular, an agreement on how to keep global warming below a 2°C rise, or less, over preindustrial levels. As the newly elected chair of the Intergovernmental Panel on Climate Change (IPCC), I am hopeful that an agreement will be reached that builds a more sustainable, prosperous world.

There will always be more to know about climate change, but we know more than enough to act. So while the IPCC will continue to advance the world's understanding about the science of climate change, it will better serve global policy-makers by providing a more in-depth, and clear, understanding of the solutions. The focus on solutions will be a major component of my tenure at the IPCC, along with enhancing the participation of developing-country experts and improving the way in which the organization communicates.

We will need to deploy a wide range of technological measures and behavioral changes, and only major institutional and technological change will give us a better-than-even chance that global warming will not exceed a dangerous threshold. Here the IPCC's Fifth Assessment Report (AR5), released last year, provided a wide range of potential solutions to this challenge. Adaptation and mitigation options from AR5 are numerous and encompass not just technological but social and institutional questions and economic considerations. For example, AR5 showed that ambitious mitigation would delay expected global economic growth by 1.5 years by 2050. This does not include the numerous co-benefits of mitigation, such as improved ecosystems and human health from cleaner air, biodiversity conservation, water availability and food security, and a more secure and resilient energy supply. The options indicated by the IPCC also involve thoughtful evaluations of expected risks and benefits, such as recognizing the importance of governance, ethical dimensions, equity, value judgments,

"The goal is to achieve an international agreement to stem climate change..."

and other factors. These options also require additional knowledge about the co-benefits and adverse side effects of mitigation.

The scientific community needs to more fully understand how the private and public sectors can work together to finance adaptation and mitigation measures. And we need to better assess the costs of adaptation measures, funding, and investment. There is also a shortage of research on the economic costs of unmitigated climate change. We tend to talk about the cost of action, but what about the cost of inaction? We also need greater knowledge about the social cost of carbon emissions, which would aid in the efficient and effective pricing of carbon.

We need better information about the impact of climate change, and the solutions to it, for developing countries. That's why I am calling for training and assistance for developing-country scientists who have firsthand experience with their regions. Their knowledge will help determine the most appropriate adaptation and mitigation measures. The IPCC has already done a great deal to recruit more developing-country experts, but we can do more by identifying and networking with local centers of excellence in areas of climate science, adaptation, and mitigation. The IPCC can do the same with issues related to economic development and poverty reduction.

As the IPCC begins to plan its next assessment cycle, I am encouraged that many of my IPCC colleagues share a desire to give more prominence to what needs to be done to build a world of opportunity by tackling climate change. There is also a growing recognition at the IPCC that our reports are too difficult to understand and that something needs to be done during the next assessment cycle to ensure that AR6 can be read and understood not just by policy-makers but the average citizen. It is never easy to translate complex scientific literature into simple language, but we will have failed in our mandate to inform policy-makers if we do not reach this goal.

– **Hoesung Lee**

Hoesung Lee is chair of the Intergovernmental Panel on Climate Change, Geneva, Switzerland.

PHOTOS: (RIGHT) IPCC; (INSET) EZPERK/LAUZEN

NEWS

1,000,000,000

The upper limit, in tons, of carbon in soils around the world that is being removed by erosion. It's a previously unsuspected contributor to climate change, equivalent to as much as 55% of the greenhouse gas flux each year due to agriculture (*Nature Climate Change*).

IN BRIEF
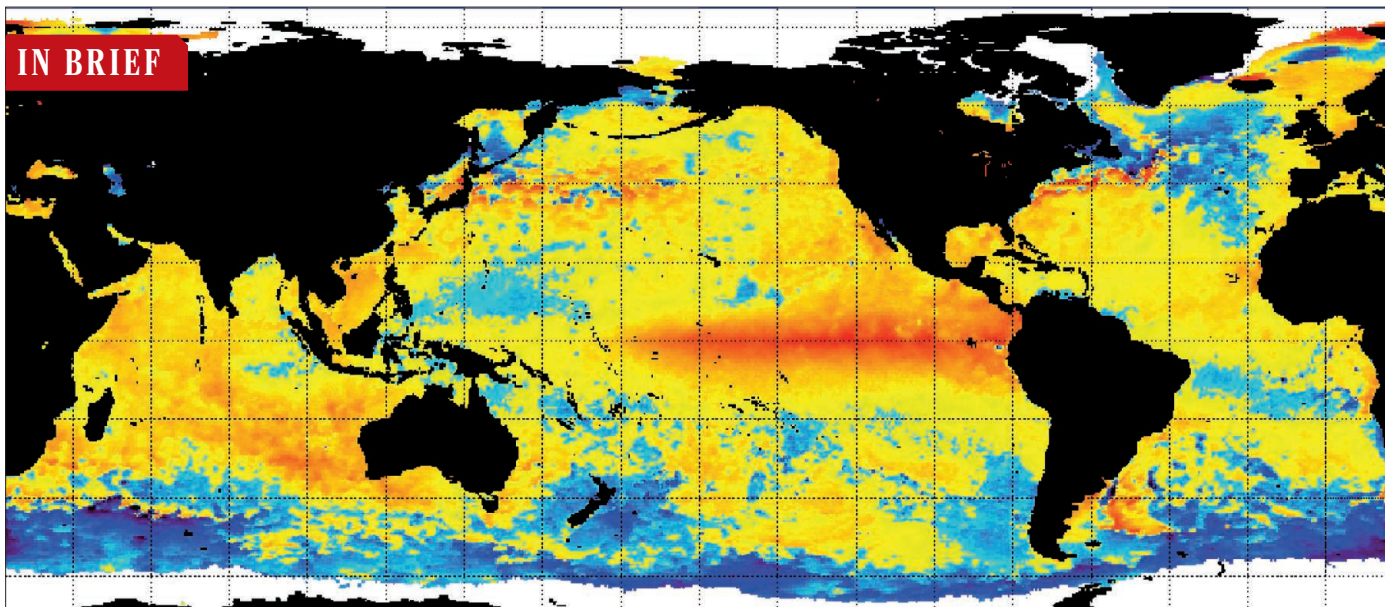
The equatorial Pacific Ocean was unusually warm (anomalous temperatures in red) on 19 November—a hallmark of an El Niño event.

Intensifying El Niño among biggest on record

Rising sea surface temperatures in the eastern tropical Pacific Ocean show that El Niño continues to strengthen and is set to be among the three biggest on record, the World Meteorological Organization said last week. The global weather pattern, driven by the warm tropical waters, has already brought drought to Indonesia, where forest fires are currently raging, and it is set to bring extra rain to the southwest United States this winter. Ocean temperatures, 2°C above normal, place this year's El Niño alongside the most severe El Niño events previously on record: in 1972–73, 1982–83, and 1997–98. And models are predicting it will stick around. There is an 80% chance that El Niño conditions will persist through spring of 2016, according to model results released on 19 November by Columbia University and the National Oceanic and Atmospheric Administration.

AROUND THE WORLD

Streamlining U.K. science funds

LONDON | As the U.K. government's looming budget cuts threaten science funds, a major review for the Department for Business, Innovation & Skills, issued 19 November, suggests ways to increase the efficiency of the seven research councils, which spend £3 billion a year. The report suggests creating an umbrella organization that would take on administrative tasks and ideally standardize and accelerate grant reviews. Paul Nurse, who leads the Francis Crick Institute, conducted the review and emphasizes the importance of having a prominent leader who can provide "a stronger strategic voice for research." Already, U.K. scientists are welcoming Nurse's review, hoping that the suggestions can reshape the research landscape of the U.K. http://scim.ag/FundingStreamline

U.S., Cuba team up to protect seas

HAVANA | Two U.S. agencies and Cuba's science ministry agreed to work together this week to manage and study marine protected areas. The new

A Florida Keys marine sanctuary is one collaboration site.

memorandum of understanding, signed by the National Oceanic and Atmospheric Administration (NOAA) and the National Park Service, covers the U.S. Flower Garden Banks and Florida Keys marine sanctuaries and two U.S. national parks, as well as Cuba's Guanahacabibes National Park and an offshore reef area known as the Banco de San Antonio. Efforts will include inventories and baseline studies of fish, coral reefs, and seagrass beds. Practically, that should translate into more NOAA ships, equipment, and technology going to Cuba, which has a dearth of research resources. "This particular agreement is the highest profile [attempt] to truly remove the barriers to scientific [collaboration]," says Daniel Whittle, who runs the Environmental Defense Fund's Cuba program. http://scim.ag/USCubaMarine

Scientists: Take wolves off ESA

WASHINGTON, D.C. | The gray wolves (*Canis lupus*) of the western Great Lakes states—Minnesota, Wisconsin, and Michigan—should be removed from the federal Endangered Species Act (ESA), say 26 scientists and wildlife managers in a 18 November letter to the secretary of the interior, Sally Jewell. By 2014, the letter states, the wolves' numbers were more than 3700, greatly surpassing the original criteria of 300. Some conservation and animal advocacy groups dispute the notion that the wolves are fully recovered, and have used lawsuits to successfully block previous attempts to down-list them. But letter signatory L. David Mech, a wolf biologist with the U.S. Geological Survey in St. Paul, and his co-authors argue that keeping the animals on the federal list only harms them by creating ill will. "When an animal is recovered but not delisted, it impugns the ESA," Mech says, "and gives ammo to those who dislike the act." Other scientists argue, however, that the Great Lakes wolves' greatest threat persists: People hate them. http://scim.ag/GreatLakeswolves

Europe's first *T. rex*

BERLIN | Tristan the *T. rex*, one of the best-preserved *Tyrannosaurus rex* skeletons yet discovered, is being prepped to go on display next month at Berlin's Museum of Natural History. The fossil, more fully known as "Tristan Otto," was named for the sons of the two private collectors who own the fossil and have loaned it to the museum for at least 3 years. Researchers at the museum, in cooperation with several other institutes, have already begun studying the bones. Broken ribs, a deformed jaw,

Tadpole's-eye view wins new photo competition

The grand prize winner of a new nature photography competition invites viewers to imagine themselves as tadpoles. In the upward-facing shot (shown), toad larvae seem to fly against a blue sky and fluffy clouds in the clear waters of a Belgian canal. It's a side of the ecosystem we usually miss, but it's one "that is very much part of the tadpoles' own view—the clouds, the trees, and the sky," said contest judge and evolutionary biologist Alex Badyaev of the University of Arizona in Tucson in a statement. The photo was one of more than 1000 submissions to the Royal Society Publishing photography competition, an event launched this year by two of the society's biology journals. The competition celebrates the 350th anniversary of the world's oldest scientific journal, the *Philosophical Transactions of the Royal Society*. Belgian biologist Bert Willaert, who took the winning shot while snorkeling, said in a statement that he hopes his photo will help people appreciate the ecological wonders in their own backyards.

and damaged teeth suggest that the dinosaur suffered several injuries and illnesses before it died on the verge of adulthood. Researchers hope further studies will help them answer still-controversial questions about the species, such as how fast it could run and how powerful a bite it had. The museum will unveil Tristan to the public on 17 December—the first full *T. rex* skeleton to go on display in Europe.

U.S. gives GM salmon green light

WASHINGTON, D.C. | A fast-growing salmon developed by AquaBounty Technologies has become the first genetically modified (GM) animal to get the blessing of the U.S. Food and Drug Administration (FDA). On 19 November, the biotech company received FDA's approval for the animal, which grows twice as fast as non-GM Atlantic salmon thanks to the overexpression of a growth hormone. FDA's green light marks the end of a nearly 20-year-long struggle for the right to sell the fish in grocery stores. Although a 2012 FDA draft assessment found that the salmon is unlikely to have any harmful impact on the environment, some environmental groups fear that the fish could become invasive in the unlikely event that they are released into the wild. Meanwhile, many grocery stores, including Target, Trader Joe's, and Whole Foods, have said they won't sell the GM salmon. http://scim.ag/_GMsalmon

FINDINGS

Pathogen infects snakes' skin

The fungus suspected in a mysterious skin disease afflicting snakes across the eastern and midwestern United States has been confirmed as infectious and pathogenic. First described in 2006, telltale lesions have been seen in seven species and reports of sick snakes are becoming more frequent. Severity varies, but the lesions can be deadly; a population of imperiled timber rattlesnakes in New Hampshire fell by half from 2006 to 2007. The fungus *Ophidiomyces ophiodiicola*, found in the lesions, was a top suspect. To see whether it is the culprit, biologists with the U.S. Geological Survey cultured the fungus from a water snake (*Nerodia sipedon*) and then infected eight corn snakes.

Crusty scales cover skin blisters on this infected northern water snake.

Within days, swelling and lesions appeared, they reported in *mBio* last week. Although the snakes rid themselves of the fungus by shedding their skin, two refused food and others left their hiding places—behaviors the researchers say could indirectly contribute to deaths.

NEWSMAKERS

Salk names new president

Elizabeth Blackburn, a 2009 Nobel laureate in medicine for her discovery of telomeres and their role in the aging process, last week was named president of the Salk Institute for Biological Studies. She will succeed William Brody, who led a $350 million capital campaign that has put the 55-year-old San Diego, California, institute on firmer financial footing. A professor at the University of California, San Francisco, the 66-year-old Blackburn calls the move to Salk a "natural transition" in her continuing quest to explore "how life works." http://scim.ag/SalkPresident

Female space pioneer honored

Before **Katherine Johnson** could even get a job as a "woman computer" at Langley Research Center in the 1950s, she had to break down two formidable barriers: Working as a woman in science, and being an African-American. But Johnson, a research mathematician who this week received the Presidential Medal of Freedom, never let others define her place. Johnson plotted the trajectory for the first American into space, Alan Shepherd. And even after computers had taken over the task, Johnson says, John Glenn "wanted to see what she comes up with" before he made his historic orbital flight in 1962. At 97, Johnson has been a role model for countless African-Americans and women.

A grooving chemist in the video that won this year's Dance Your Ph.D. contest.
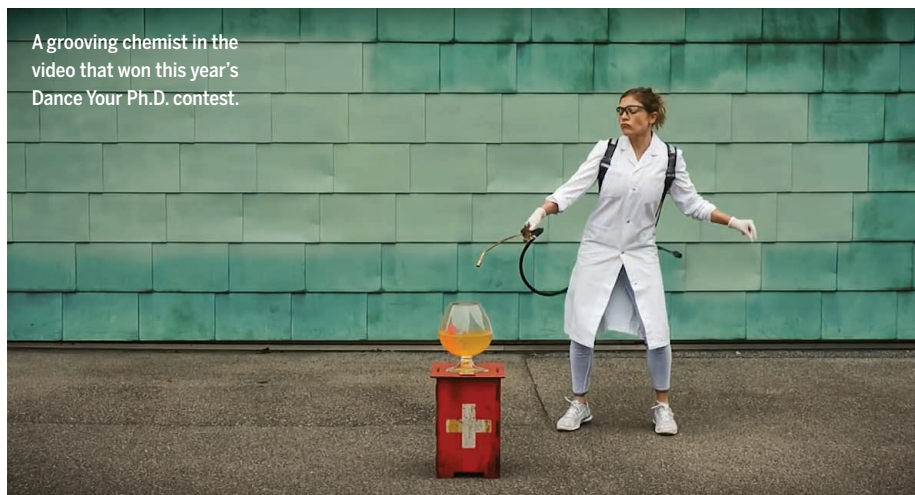
Water resource protection triumphs in dance-off

When Florence Metz turned in her Ph.D. thesis on water protection policy this year at the University of Bern in Switzerland, she thought her work was done. But then a friend sent her an email with congratulations—and an order: "Dance your Ph.D.!" So Metz recruited a small army of friends to help her create an interpretive dance describing her thesis research. Her video, one of 32 submissions this year to *Science*'s annual contest, uses various dance styles—including hip hop, salsa, and even acro-yoga—to represent the different interest groups that shape the evolution of policies around water resources. The panel of judges from the science and dance worlds awarded Metz the top prize for her efforts: $1000 and a trip to Stanford University in Palo Alto, California, in the spring, where she'll screen her dance video and give a talk. "My main aim with this video was to make people laugh," she says. But it really does help people understand her work better, she adds. "This bridge between academia and the nonacademic world is crucial." **http://scim.ag/DanceYourPhD2015**

The ITER construction site in southern France earlier this year.

FUSION ENERGY

More delays for ITER fusion project

Review notes progress but estimates "first plasma" will take 6 years longer than planned

By **Daniel Clery**,
in Saint-Paul-lès-Durance, France

Managers of the troubled ITER fusion project delivered a dose of reality last week: a new schedule that is likely to push the estimated date of completion back by 6 years, to 2025, and add roughly €2 billion to the project's ballooning cost. The project's international partners—China, the European Union, India, Japan, Russia, South Korea, and the United States—did not immediately balk at the new "baseline," presented at a meeting of the governing council here. But they plan to carry out an independent review, looking for ways to tighten the schedule and costing, and have put off approving the baseline until the next council meeting in 6 months.

In the meantime, the council approved the proposed schedule for 2016 and 2017, set out milestones, and agreed to make extra resources available. "It was a very important meeting for us and it went well," says ITER Director-General Bernard Bigot. One member of the council, Jianlin Cao, vice minister at China's Ministry of Science and Technology, acknowledged the challenges facing the project: "ITER is a new thing, and success does not come easily."

ITER aims to show that nuclear fusion is feasible as a source of energy. Although fusion powers the sun and stars, researchers have never managed to achieve a controlled fusion reaction on Earth that produces more energy than it consumes. ITER, with a doughnut-shaped "tokamak" reaction chamber able to contain 840 cubic meters of superheated hydrogen gas, or plasma, is the biggest attempt so far and should produce 500 megawatts of power from a 50 megawatt input. The project began in 2006 with an estimated cost of €5 billion and a start date—or first plasma—in 2016. The figures quickly changed to €15 billion and 2019, but confidence in those numbers has eroded over the years.

When Bigot took charge earlier this year, he ordered a full review of the project, which currently has numerous buildings springing up at the Cadarache site in southern France, and components arriving from contractors around the globe. That review resulted in the new baseline and the 6-year delay, which is not mentioned in the communiqué but is widely acknowledged. The cost of running the ITER organization and the seven "domestic agencies" that handle industrial contracts for each partner is very roughly €350 million per year, so the delay will add about €2 billion. Many factors have slowed progress, including the complexity of the project, delays in finalizing the design, and the demands of France's nuclear regulator.

The international partners "need more time to fully endorse this or another schedule and to reconcile it with the resources they have," Bigot says. Delegates agreed. "We must take the schedule home and discuss it with the finance ministry," says Anatoly Krasilnikov, head of Russia's ITER domestic agency. As an interim measure, the council will shift money already allocated for 2016 and possibly provide new money for 2017. The ITER organization will hire 150 new staff. In return, the council set 25 major milestones to be met by 2017. "If we meet the milestones, it will consolidate the trust," Bigot says.

ITER's organizational structure is almost as complex as its technology. Each partner manufactures a share of the necessary components: 45% from the European Union (as host), and 9% from each of the others. How much each partner spends to fulfill its share is its own concern and is not revealed, making the true cost of the project difficult to assess. In addition, each contributes to the costs of the central ITER organization.

The ITER organization, in turn, faces the challenges of drawing up the design, ensuring everyone sticks to it, and supervising assembly of the reactor while also satisfying French regulators. Last year, a highly critical management assessment faulted the organization for failing to establish a workable "project culture" (*Science*, 28 February 2014, p. 957). Bigot has gone to great lengths to get contractors, domestic agencies, and ITER staff working better together, as the council communiqué acknowledges.

Despite the prospect of delays, fusion researchers are "super anxious for a firm schedule," says Stewart Prager, director of the Princeton Plasma Physics Laboratory in New Jersey, America's top fusion facility. In the United States, he says, disarray in the project could give ammunition to opponents in Congress, whereas "a credible schedule could have a positive effect." Steven Cowley, director of the Culham Centre for Fusion Energy in the United Kingdom, thinks the council is right to call for an independent review. But then, he says: "Let's get on and do it. It's the best chance we've got to achieve a fusion burn." ■

INFECTIOUS DISEASES

An obscure mosquito-borne disease goes global

After racing through Oceania last year, the Zika virus is now spreading in the Americas

By **Martin Enserink**

Seven years ago, after returning from a field trip to Senegal, medical entomologists Brian Foy and Kevin Kobylinski came down with a serious disease so obscure that no one could find out what it was. The duo, both at Colorado State University, Fort Collins, had a rash, fatigue, headaches, and swollen and painful joints, but they tested negative for known infectious agents.

Both recovered, and they might never have known the cause of their illness if, more than a year later on another African trip, Foy hadn't run into medical entomologist Andrew Haddow, then at the University of Texas Medical Branch in Galveston. In 1947, Haddow's grandfather was working in Africa, where he had helped discover a mosquito-borne virus named Zika. The symptoms it caused seemed to match Foy's. Haddow helped arrange for Foy and Kobylinski to get tested for antibodies. Bingo: They had been infected with Zika.

Now, the oddity that floored Foy and Kobylinski has hit the big time. Eight years ago, after cropping up sporadically in Africa and Asia for half a century, Zika went on a rampage through the Pacific islands. In May of this year it arrived in Brazil, where it has now caused disease in 14 of the country's 27 states; in the past 2 months, cases have also been reported from neighboring Colombia and Suriname. In May and again in October, the Pan American Health Organization warned other countries in the region to prepare for the virus's arrival.

Researchers are only beginning to study the disease and its transmission in earnest. But because the *Aedes* mosquitoes that spread Zika are ubiquitous in urban areas throughout the Americas, it is almost certain that the virus will spread through South and Central America, Mexico, and the Caribbean, says Duane Gubler, who directs the Emerging Infectious Diseases program at National University of Singapore. It will make occasional inroads into the southern United States and southern Europe as well, he predicts.

Gubler feels so sure because two other diseases that are spread by the same species of *Aedes* mosquitoes—most notably *A. aegypti*, the yellow fever mosquito—have expanded in this way. One of the two, dengue, has caused vast epidemics in Latin America the past 40 years. The other, chikungunya, virtually unknown a decade ago, began conquering the Western hemisphere in 2013 (*Science*, 16 May 2014, p. 678), where to date there have been more than 600,000 suspected and confirmed cases in more than 30 countries.

The good news is that Zika has caused no known deaths. Its symptoms resemble those of dengue and chikungunya, but in most patients Zika is milder than those diseases, which both can cause excruciating pain. (Dengue also can progress to dengue hemorrhagic fever, which can be fatal.) But having so many similar viruses circulating at the same time complicates diagnoses and strains public health systems. And Zika's arrival in the Americas raises the question of how many other *Aedes*-borne viruses might emerge. "We have an unholy trinity now—but it could easily grow," Foy says.

Both Zika and dengue are flaviviruses, a genus that also includes yellow fever and the West Nile virus. During the half century after Zika's 1947 discovery—in a rhesus monkey in the Zika Forest near Entebbe, Uganda—fewer than 15 cases were reported, all from Africa and Southeast Asia.

Something changed in 2007, when Zika erupted in a big outbreak in Yap, an island group in the Western Pacific that is part of the Federated States of Micronesia. A 2009 study based on antibody surveys estimated that an astonishing 73% of the population had become infected, although no one died or even was hospitalized.

That was the start of an extended island-hopping tour. Zika infected close to 30,000 people—one-tenth of the population—during an outbreak in French Polynesia in 2013–14; this time, some people did end up in the hospital, and a few developed Guillain-Barré syndrome, a muscle weakness caused by damaged peripheral nerves. That dispelled the notion that Zika is always benign, says Didier Musso of the Institut Louis Malardé on Tahiti, French Polynesia. Infected travelers then probably took the virus to New Caledonia, the Cook Islands, Vanuatu, and Easter Island, where local mosquitoes bit them and started fresh outbreaks.

Although the combination of air travel and large mosquito populations guarantees rapid spread, there may be other transmission routes. After Foy returned home in 2008, his nurse wife, Joy Chilson Foy, contracted Zika

Aedes aegypti

Virus on the move

Zika, which occurred sporadically in Africa and Southeast Asia before 2007 (inset), has caused major outbreaks in Oceania and now has reached the Americas.

2007
Yap
(Federated States
of Micronesia)

2013
French
Polynesia

April 2015
Vanuatu

2014
New Caledonia
(France)

2014
Cook
Islands

2014
Easter Island
(Chile)

November 2015
Suriname

October 2015
Colombia

May 2015
Brazil
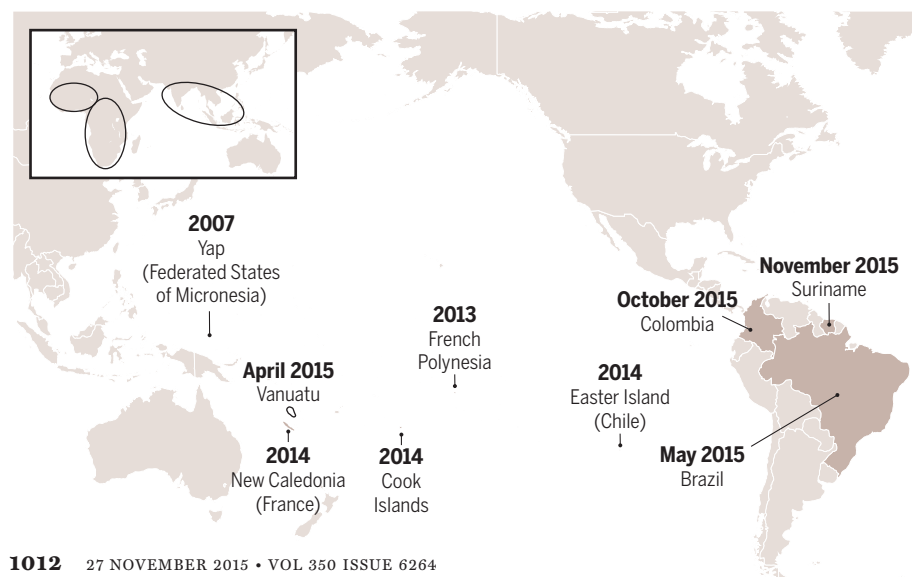
as well. Because northern Colorado hosts no mosquitoes that can act as Zika vectors, and because Foy had prostatitis and blood in his semen at the time, the most likely explanation is sexual transmission—the first documented instance ever for a mosquito-borne disease. (The couple, along with Kobylinski and others, wrote up their bit of medical history in a 2011 paper in *Emerging Infectious Diseases*, in which they dispassionately refer to themselves as patients 1, 2, and 3.)

Musso and his colleagues also found the virus in the bloody semen of a Zika patient in French Polynesia. Still, the fact that many children and elderly people fell ill suggests that sexual transmission is a "curiosity" rather than playing a major role in Zika's spread, Musso says. The team in Tahiti also documented a case of mother-to-child transmission during birth, and showed that the virus occurred in the blood of 3% of asymptomatic blood donors, a surprisingly high number that suggests a real risk of Zika transmission during transfusions.

Scientifically speaking, Zika is largely terra incognita. A PubMed search for the virus last week turned up just over 200 papers, compared with more than 2500 for chikungunya and more than 14,500 for dengue. There are no drugs or vaccines, and no animal model except monkeys, which aren't practical for most studies. After his own brush with the virus, Foy applied for a National Institutes of Health grant to study the disease but was rejected. Now that Zika appears to be exploding, he expects a wave of new research, as happened for chikungunya, then a similarly obscure agent, after it caused a massive epidemic in Indian Ocean islands in 2005.

For the moment, there is little hope of stopping Zika, Gubler says. Controlling the mosquitoes that carry it is a frustrating affair; *Aedes* breed in ubiquitous small water reservoirs, such as flower pots or abandoned car tires, and spraying insecticides has not been particularly effective.

There is hope, however, that new technologies may eventually help control mosquito populations. For instance, Oxitec, a startup in the United Kingdom that was recently bought by the U.S. synthetic biology company Intrexon, has developed genetically modified sterile male mosquitoes that can bring down a population. Other researchers have infected mosquitoes with the *Wolbachia* bacterium, which reduces their capacity to host pathogens. And long-running attempts to make mosquitoes resistant to pathogens through genetic modification have received a fresh boost from new gene-editing systems (see story, p. 1014). If one or more of these technologies works out and becomes widely adopted, Gubler says, "we could finally see some real progress." ∎

BIOMEDICAL RESEARCH

An end to U.S. chimp research

NIH announces plans to retire its last chimpanzees

By **Jocelyn Kaiser**

The U.S. National Institutes of Health (NIH) is ending its support for invasive research on chimpanzees. Agency head Francis Collins said last week that a colony of 50 chimps it had planned to keep in reserve for research—after retiring the rest—is no longer needed. NIH also made clear that it will no longer fund invasive studies on any other chimps. "I think it is the natural next step in what has been a process over the last 5 years, really, of deep thinking about the appropriateness of research on our closest relatives, the chimpanzees," Collins told *Science*.

It's "amazing and historic news," says Kathleen Conlee, vice president of animal research issues for The Humane Society of the United States in Washington, D.C. But some biomedical researchers expressed disappointment. Chimpanzees "have been a critical model for life-saving research" in infectious disease and other areas, and the colony would have been available "in the event there was a national, critical need for research in the future," said a spokesperson for the Texas Biomedical Research Institute in San Antonio, which hosts the Southwest National Primate Research Center.

NIH announced the move on 18 November, but the news broke earlier when someone at NIH leaked an internal staff email from Collins. In the 16 November memo, he wrote that several factors, including the fact that no researchers have asked to use chimpanzees in recent years, led him to conclude that the 50 animals are no longer needed. "Given this complete absence of interest in a space now approaching 3 years, I think it's fair to say the scientific community has come up with other ways to answer the kinds of questions they used to ask with chimpanzees," Collins says.

The United States, unlike many other countries, has kept the door open for invasive research on chimpanzees but has tightened restrictions. In June 2013, NIH announced it would phase out most agency-funded chimpanzee research and retire all but 50 of its research chimps. As NIH-funded chimpanzee grants ended, investigators would need to meet new and tougher standards to qualify for new grants.

A decision this past June by the U.S. Fish and Wildlife Service (FWS) to list captive chimpanzees as endangered added another hurdle by requiring a permit to use any of the 700 research chimpanzees in the country in invasive research. Researchers wanting to study chimps would also have to show that the work would somehow benefit chimpanzees in the wild.

As a 14 September deadline approached for the FWS requirement, the agency had not received a single research permit application (*Science*, 21 August, p. 777). And NIH has only received one research application—a proposal from an intramural researcher that was later withdrawn, Collins says.

In the leaked email to James Anderson, director of the NIH Division of Program Coordination, Planning, and Strategic Initiatives, Collins describes other factors behind his decision. These included the fact that in 2013, Congress lifted a cap on how much NIH can spend on supporting chimpanzee retirement, and that there is room for

> *"The scientific community has come up with other ways to answer the kinds of questions they used to ask with chimpanzees."*
>
> **Francis Collins**, NIH director

more animals at Chimp Haven in Keithville, Louisiana, the federal sanctuary for retired NIH chimps.

Collins has asked Anderson's office to prepare a retirement plan for the 20 or so chimps NIH still owns at the Southwest center and the 140 at MD Anderson Cancer Center's primate facility in Bastrop, Texas. About 150 more chimps at the Alamogordo Primate Facility in New Mexico will be moved later. "We still have work to do," Collins says. But it will be easier, he says, "without having to go through the complicated calculus of which chimps ought to be in the group of 50 to be saved for research."

Collins also asked Anderson to plan to phase out funding for about 82 research chimpanzees at the Southwest center that NIH supports but does not own. And, an NIH spokesperson wrote in an email, "NIH will not fund biomedical (i.e., invasive) research involving chimpanzees (full stop)." ∎

SCIENCE sciencemag.org

27 NOVEMBER 2015 • VOL 350 ISSUE 6264 **1013**

Published by AAAS

The blue in these mosquitoes indicates where they express engineered mouse genes that can interfere with malaria parasites.

SCIENCE AND SOCIETY

Gene drive turns mosquitoes into malaria fighters

Antiparasite genes made to spread among lab insects

By **Elizabeth Pennisi**

The war against malaria has a new ally: a controversial technology for spreading genes throughout a population of animals. In the laboratory, researchers have harnessed a so-called gene drive to efficiently endow mosquitoes with genes that make them immune to the malaria parasite—and unable to spread it. On its own, gene drive won't get rid of malaria, but if successfully applied in the wild the method could help wipe out the disease, at least in some corners of the world. The approach "can bring us to zero [cases]," says Nora Besansky, a geneticist at the University of Notre Dame in South Bend, Indiana, who specializes in malaria-carrying mosquitoes.

But testing that promise in the field may have to wait until a wider debate over gene drives is resolved. The essence of this long-discussed strategy for spreading an engineered gene is to bias inheritance so that more than the expected half of a subsequent generation inherits it. Gene drive attracted new attention earlier this year, when geneticists studying fruit flies adapted a gene editing technology called CRISPR-Cas9 to help spread a mutation—and were startled to find it worked so well that the mutation reached almost all fly progeny. Their report, published this spring in *Science* (20 March, p. 1300), came out less than a year after an *eLife* paper had warned that a CRISPR-Cas9 gene drive system could disrupt ecosystems and wipe out populations of entire species.

A firestorm quickly erupted over the risks of experimenting with gene drives, nevermind applying them in the field. The U.S. National Academy of Sciences has convened a committee to weigh the risks and propose safeguards, and the authors of the *eLife* and *Science* papers have laid out guidelines for experiments (*Science*, 28 August, p. 927).

Meanwhile, for the past 20 years, Anthony James, a geneticist at the University of California (UC), Irvine, has tried to engineer mosquitoes so they can no longer host the malaria parasite. In 2012, his team pinned down mouse genes for antibodies that make rodents immune to the human malaria pathogen and put them in a mosquito species that spreads malaria in India. The antibodies, as hoped, interrupted the parasite's life cycle within the insect. But James had no way to spread those antibody genes through countless millions of mosquitoes in nature.

Earlier this year James got an email from Ethan Bier, the geneticist at UC San Diego whose lab was doing the soon-to-be-published gene drive research in fruit flies. Bier thought he had a solution to James's dilemma. "As soon as we saw [gene drive] could work, we thought of mosquitoes," Bier says. James was thrilled, but wondered whether the system could ferry the hefty 17,000 bases of DNA containing the mouse antibody genes. "The question was, 'Would it carry a large cargo that would remain active?'" James recalls. He and Bier teamed up to see.

Valentino Gantz from Bier's lab and Nijole Jasinskiene, a molecular biologist at UC Irvine, began by engineering male and female *Anopheles stephensi* to carry the gene drive system. They had designed the system so that, along with spreading the antibody genes from one half of a chromosome pair to the other—the key to biasing inheritance—it would also cut out a piece of a gene responsible for eye color. When they mated the altered mosquitoes with normal ones, they could quickly see whether the gene drive had worked: Offspring that had inherited the antibody genes also had white eyes.

The technology was efficient, endowing about 99% of the transgenic male's offspring with the added genes, Bier and James's teams reported this week in the *Proceedings of the National Academy of Sciences*. And as hoped, those genes were active in the mosquitoes. Earlier experiments had shown that if the antibody genes were expressed, they thwarted the parasite. "We've got all the pieces," James says. "It's a question of [making] a product that people will want."

And that is the big if. James, Bier, and their colleagues adequately addressed concerns about accidental releases of the transgenic mosquitoes, say several outside researchers contacted by *Science*. The insectaries were behind five sets of doors, and they used a mosquito that doesn't survive in California, should it manage to escape.

But before such work continues, says evolutionary engineer Kevin Esvelt from Harvard University, biologists should look at the ecological effects of gene-driven changes, make sure the changes are stable over many generations, and develop a way to counter or get rid of the gene drive if problems arise.

Because the antiparasite genes should continue to spread ad infinitum among a mosquito population, national and international regulations need to be worked out before gene drives are deployed in the field, adds social scientist Kenneth Oye from the Massachusetts Institute of Technology in Cambridge. "How are we going to decide as a society whether, when, and how to use gene drive to solve a problem?" he asks. Adds Esvelt, who even before there is an appropriate gene drive technology to spread anti–Lyme disease genes in mice is discussing the idea with the public, doctors, and government officials: "At the end of the day, unless you have widespread public support, you can't do it."

James accepts that his dream may be deferred for now. "If it turns out we are too far ahead of the curve, we'll just have to wait for people to catch up," he says. "I hope I don't have to wait the rest of my productive career, but if we can't find a way to do it ethically, then it won't be done." ∎

ACADEMIC MISCONDUCT

China pursues fraudsters in science publishing

Measures may not be enough to stem the tide, some fear

By **Mara Hvistendahl**

China's main basic research agency is cracking down on scientists who used fake peer reviews to publish papers, demanding that serious offenders return research funding. The move accompanies an announcement by the China Association for Science and Technology (CAST) in Beijing, first reported by state media on 12 November, that it had investigated dozens of scientists involved in peer-review scams. The probe's findings highlighted the role of China's many unscrupulous paper brokers, which peddle ghostwritten or fraudulent papers.

"If it wasn't obvious before, it is now difficult to deny China's research community has serious underlying ethical issues," says Benjamin Shaw, China director for the English-language editing company Edanz in Beijing. Others caution that the sanctions on discredited authors are not severe enough to deter academic dishonesty. But the coordinated response by funding agencies and CAST, which links China's science and technology community with the government, suggests China is taking the publishing abuses seriously.

Since 2012 scores of authors, many of them Chinese, have been snagged in a peer-review scandal involving papers published in international journals. Journals discovered that authors or their brokers had suggested their own reviewers—many of whom were real researchers—but provided email addresses to accounts controlled by the perpetrators and then reviewed their own work. The findings, first reported by the blog Retraction Watch, prompted major publishers to retract scores of papers. In March, the London-based BioMed Central (BMC) began retracting 43 papers, and on 18 August Springer, which owns BioMed Central, said that it would retract 64 papers. Elsevier and SAGE have also retracted papers en masse.

In some cases the publishers say that authors weren't solely to blame. "Some researchers may have innocently become implicated in attempts to manipulate the peer review process by disreputable services," Elizabeth Moylan, senior editor for research integrity at BMC, wrote on the publisher's blog last March after an internal investigation. Four months later, *Diagnostic Pathology: Open Access*, a BMC journal, took the unusual step of updating a retraction notice, noting that the authors' institute in Shanghai, China, had found that the researchers "intended to purchase language editing services for their manuscript only and did not participate in influencing the peer review process."

The CAST investigation underscores the role of paper brokers, who profit from China's publish-or-perish mentality. According to *People's Daily*, the association contacted each of the 31 Chinese authors who had papers retracted by BMC. (BMC provided CAST with information when asked but did not collaborate on the investigation,

64

Number of peer-review scam papers Springer is retracting.

93.5

Percent of CAST-contacted authors of retracted papers who admitted to using a broker.

$5500

Upper range of fees that tarnished authors said they had paid to brokers.

says the publisher's spokesperson, Shane Canning.) Fully 29 authors admitted to using a broker, with many shelling out fees ranging from $600 to more than $5500.

The CAST investigation identified five companies that helped authors of the retracted papers secure fraudulent peer reviews, *People's Daily* noted. (In 2013, before the peer-review scandal came to light, *Science* published an investigation [29 November, p. 1035] into China's paper brokers, uncovering schemes in which scientists could purchase authorship of accepted papers or have papers ghostwritten. One broker singled out by *Science* was also targeted by the CAST investigation.)

The Chinese government is taking steps to prevent fraud. Earlier this month, the National Natural Science Foundation of China (NSFC) announced that it had investigated authors of 22 retracted papers whom it had supported, revoking funding in egregious cases. If a retracted paper was submitted as the basis for a grant application, "the employer of the offending researcher has to return all of the funding for the grant, regardless of how much of the money has been spent," NSFC President Yang Wei told *Science*. For fraud committed after grant approval, the foundation is revoking all money due after a paper's submission.

Because many of the retracted papers were in medical science, China's National Health and Family Planning Commission in Beijing is also reacting to the scandals. In September, it released new regulations requiring institutions to fully investigate cases of scientific misconduct and forbidding researchers from signing their names to papers they did not help research or write.

Companies that provide legitimate English-language editing services are attempting to distance themselves from less-principled brethren. Last month, six editing companies formed the Alliance for Scientific Editing in China and adopted industry standards, such as requiring members to publish ethics policies and forbidding them from manipulating the peer-review process.

Such measures may not be sufficient, says Lin Songqing, an editor with the Chinese Academy of Sciences in Wuhan. Until more institutions begin firing scientists who commit fraud, he says, "Paper trading will still exist for a long time." Science administrators and officials themselves feel pressure to rack up publications, he adds—which gives them "incentive to hide the truth" about publishing abuses. Zhang Yuehong, editor of the *Journal* of *Zhejiang University-SCIENCE* in Hangzhou, says more journals should do their own policing, adopting tools like Open Researcher and Contributor ID, which allows editors and readers to easily examine authors' academic backgrounds.

Fighting misconduct in scientific publishing is a long-term struggle, warns the NSFC's Yang. "Academic fraud in different varieties comes and goes like tidal waves," he says. "One has to watch for new forms of fraud constantly." ∎

Haze from coal and other
fossil fuels dims the sky
in Changchun, China.

CLIMATE CROSSROADS

After decades of failure, a new approach to negotiations has raised hopes that nations meeting in Paris will agree to meaningful climate steps

By **Eli Kintisch**

t's reasonable to view the upcoming Paris climate talks with skepticism. More than 2 decades have passed since nations met in Rio de Janeiro, Brazil, to create the 1992 United Nations Framework Convention on Climate Change. Since then, a succession of international meetings under the framework—most notably in Kyoto, Japan, in 1997 and Copenhagen in 2009—have done little to alter our planet's worrisome trajectory. Annual global emissions of carbon dioxide have risen steadily from 21 billion tons in 1992 to 32 billion tons in 2012. The rate of increase in the atmospheric concentration of greenhouse gases has accelerated, by roughly 30% since the 1990s. Nine of the 10 warmest years on record have occurred since 1998, and the impacts of climate change are already being felt from the tropics to the poles.

Optimists point to the growing use of solar, wind, and other renewable power sources and the success of some nations, such as Denmark (see p. 1020), in curbing emissions. But rising emissions from China, India, and other developing nations are swamping that progress. And the dismal track record of global climate talks inspires little confidence that nations can agree to make the huge changes required to stop treating the atmosphere like a carbon sewer.

Negotiators huddling in Paris next week are convinced these talks will be different. In Kyoto, nations attempted to create a legally binding agreement, which subsequently failed to deliver results in part because the United States would not ratify the treaty. This time, nations—164 of them, by the time *Science* went to press—have each prepared pledges, called Intended Nationally Determined Contributions (INDCs), which detail their promised emissions cuts and other actions through 2030.

Negotiators hope the bottom-up INDC approach will prevail where the top-down Kyoto strategy failed. Developing nations largely stuck to the sidelines in previous talks. This time almost everyone—including China and India (see p. 1024)—has pledged to limit emissions. And by arriving in Paris with pledges in hand, negotiators hope to avoid the last-minute deadlocks that have doomed past efforts.

That's not to say that negotiators aren't under pressure. There's still the question of how developing nations will raise the hundreds of billions of dollars they'll need to curb emissions and adapt to climate change. And it's not clear how national pledges will be monitored and verified.

Another thorny question is whether nations will agree to review any Paris deal every 5 years. That would create regular opportunities for countries to extend their reduction policies past 2030 and ratchet up cuts. Climate experts say such action is needed, because the Paris pledges alone won't keep planetary warming by 2100 below the 2°C ceiling that many consider safe.

The Washington, D.C.–based nonprofit Climate Interactive, for example, estimates that without further action, the pledges will allow the world to warm 3.5°C by 2100 (see page 1018). The U.N. Environment Program (UNEP), however, assumes that nations will extend their pledges, which could keep the warming to 2.7°C. The lower number is "a reason for hope," says Cassie Flynn, a UNEP official, because it puts the 2°C threshold within reach—if nations can agree to work together after Paris. ■

PHOTO: WANG ZHE · IMAGINECHINA

AFTER PARIS: THE ROCKY ROAD AHEAD

Officials call the Paris talks a beginning, but what's the destination? Below are three possible future paths for annual global greenhouse gas emissions. Models suggest each would produce very different ranges of atmospheric greenhouse gas levels and warming in 2100. And all would cause the seas to rise for centuries after 2100. — *Eli Kintisch*

Global emissions in CO$_2$ equivalent (Gtons per year)

World population

3.0 billion 6.1 9.6 11.2
1960 2000 2050 2100

2015 Paris
Nations arrive in France with pledges in hand, hoping to achieve global emissions curbs that have eluded negotiators in the past.

2030 Paris pledges end
Emissions will resume their rise unless nations extend, and strengthen, their climate measures.

1997 Kyoto
The U.S. Senate undermined talks by opposing any deal that didn't require action by developing nations; few countries subsequently fulfilled their commitments.

Atmospheric CO$_2$ (ppm)

319 367 404 451 908 674 478

EMISSIONS SCENARIOS DATA COURTESY OF CLIMATE INTERACTIVE

Alternative worlds

1. Business as usual
If nations don't fulfill their pledges, temperatures could rise by 2100 to levels not seen for millions of years. The world would be committed to centuries of sea level rise after 2100. (Range of sea levels shown for a 3.3°C rise.)

Atmospheric CO$_2$ level	Warming above preindustrial	Committed sea level rise
High prediction — 922 ppm	5.9°C	9.9 m
Middle prediction — 908 ppm	4.5°C	7.1 m
Low prediction — 889 ppm	2.7°C	4.3 m
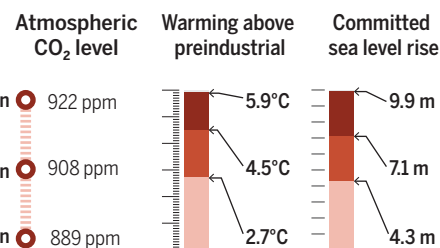
2. Paris pledges only
The Paris commitments are expected to cause global emissions to dip initially. But emissions will surge after 2030 if nations take no further action. (Range of sea levels shown for a 2.3°C rise.)

684 ppm	4.6°C	7.2 m
674 ppm	3.5°C	5 m
661 ppm	2°C	2.7 m
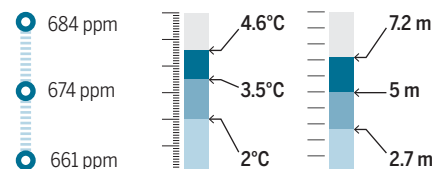
3. The path to 2°C
Avoiding dangerous warming will require no less than a global energy revolution. But even with aggressive action global coasts will see a battering from rising seas. (Range of sea levels shown for a 1.1°C rise.)

484 ppm	2.7°C	4.2 m
478 ppm	2°C	2.4 m
470 ppm	1°C	0.7 m
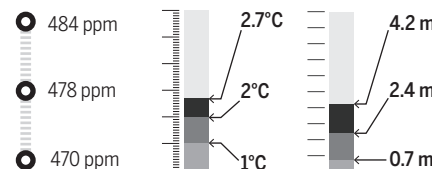
Lasting consequences
Even under the 2°C path, one measure of progress—atmospheric CO$_2$ levels (left)— will remain flat for centuries as the gas is very stable. Another benchmark, CO$_2$ equivalent (above) includes emissions of shorter lived gases such as methane, as well as emissions from agriculture and deforestation.

1. With business as usual, a world transformed

Average temperatures would rise most dramatically at the poles and over the continents (left), whereas rising seas would swamp countless communities (right).

Modeling a hotter planet

Temperature change (°C) 2071–99 v. 1970–99

0.6   1.7   2.8   3.9   5.0   6.0   7.2   8.3

Ocean engulfs the North Carolina coast

Flooding by 5 m sea level rise (after 2100)

20 km

Fairfield

Lake Mattamuskeet

Scranton

Atlantic Ocean

2. With Paris pledges only, gains could be swamped by developing nation growth

An ongoing shift to renewables will help nations meet their Paris pledges. But population and economic growth in developing nations could push up emissions after 2030.

One scenario for a changing energy mix (TWh)

Total energy 34193

Total energy 22970

Fossil fuels 15257 →

Nuclear 2432 →

Renewables 5281 →

−0.1% change

+34% change

+197% change

← 15236 Fossil fuels

← 3269 Nuclear

← 15688 Renewables

2015   2040

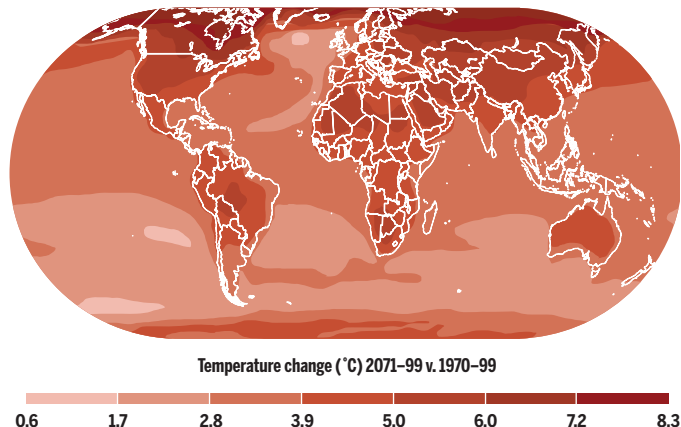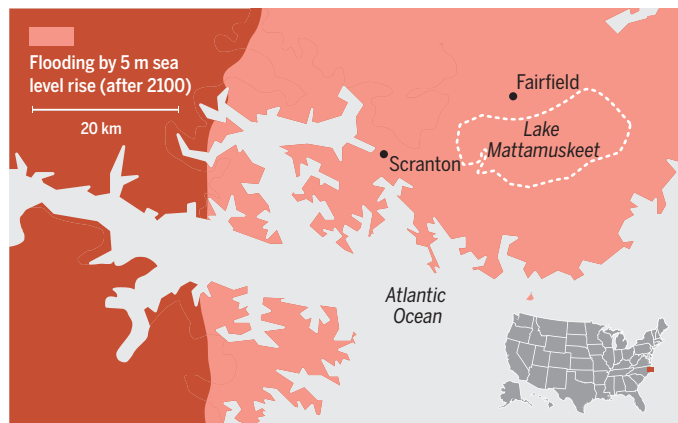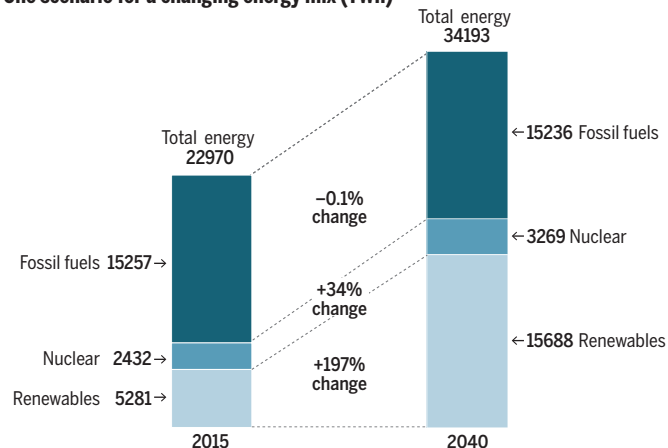## Emissions surge after 2030 with no new action

Gtons CO$_2$e/year

Other developing nations

China

Other developed nations

India

United States

Europe

2000 2020 2040 2060 2080 2100
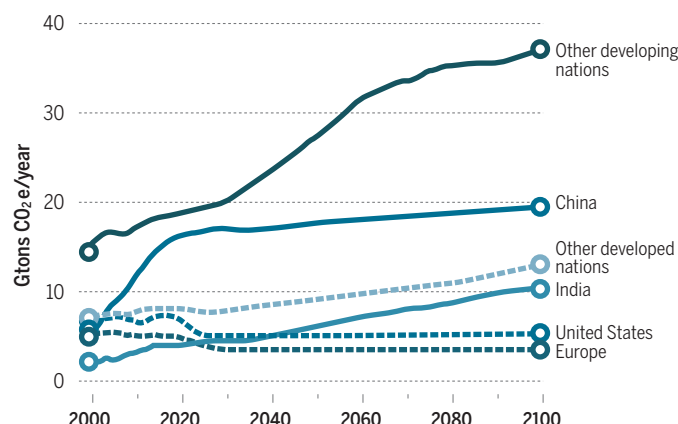
3. The path to 2°C will require new technologies and major investments

Present trends lead to massive fossil fuel use (left). Under some scenarios, slashing emissions will require carbon capture and storage (CCS) on a huge scale (middle), and a major shift in spending away from fossil fuels and to efficiency and renewables (right).

Fossil fuel dominance...

Primary energy use (EJ)

← Nuclear
← Renewables
← Bioenergy

← Fossil fuels

1980   2020   2060   2100

or an energy transformation...

← Nuclear

← Renewables

← Bioenergy +CCS

← Fossil fuels +CCS

← Bioenergy

← Fossil fuels

1980   2020   2060   2100
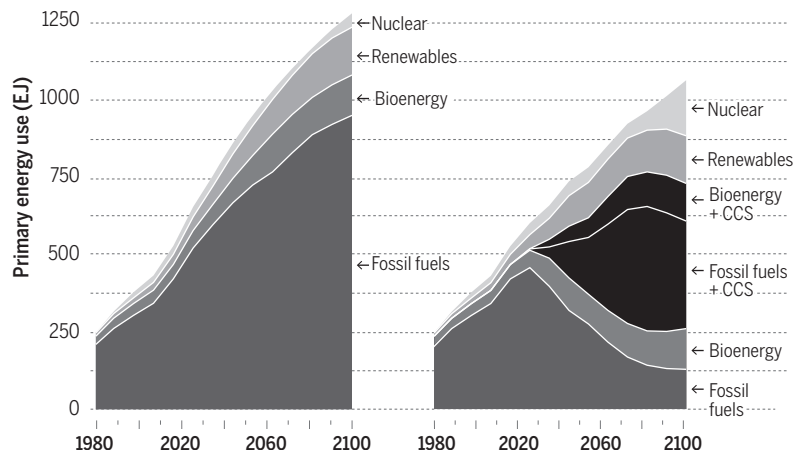
propelled by shifts to greener spending.

Reduced annual investment   Increased annual investment

Energy efficiency — 328

Renewables — 86

Fossil fuels + CCS — 41

Nuclear — 43

Fossil fuels — −66

Coal, oil, and gas extraction — −116

−200 −100   0   100   200   300

Change needed to meet 2°C goal, 2010–29 (mean from multiple studies, $ in billions/year)
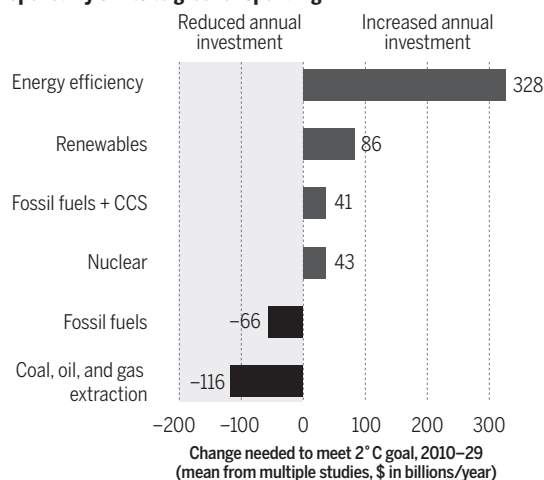
CLEAN REVOLUTION

Denmark is striving to become the world's first carbon-neutral nation

By **Robert F. Service**, *in Bornholm, Denmark*

Not long ago, Maja Felicia Bendtsen spent a fair amount of time making sure things didn't spin out of control on this small, windswept island in the Baltic Sea. Her employer, the power utility Østkraft, operates 35 wind turbines that help supply electricity to Bornholm's 41,000 residents. But sometimes there wasn't enough demand for the electrons produced by those spinning blades. So Østkraft engineers would have to act quickly—at times in the middle of the night, when the wind was howling—to put the brakes on the turbines, lest the excess electricity melt the island's power

lines. It was a frustrating routine: Østkraft was throwing away essentially free renewable energy.

The utility is thriftier now. Through a project led by Bendtsen and others called EcoGrid, Østkraft is using cutting-edge smart grid technologies to make more efficient use of its bountiful wind power. When electricity demand is low, for instance, Østkraft uses the excess power to heat water in residential hot water systems to a couple of degrees above normal, banking the energy for later use. Such strategies have meant an 80% drop in the number of times that Østkraft has had to rein in its turbines.

That may sound like a modest achievement, but it marks an important step toward achieving Bornholm's ambitious goal of completely weaning itself off fossil fuels by 2025—and toward Denmark's bigger target. This small, resource-poor country of 5.5 million people has set the most ambitious climate goal in the world: to become a carbon neutral economy by 2050. As delegates gather in Paris to hammer out a global agreement to slow climate change, many are looking to Bornholm and Denmark to understand how their nations might also rapidly transform their energy systems. "The Denmark model is really important," says Dan Kammen, an energy

On exceptionally blustery days, wind turbines such as these off the coast of Copenhagen already generate more electricity than Denmark consumes.

in droves. The island's population became smaller, older, and poorer. Looking for ways to save money for residents, in 1985 the mayor of Rønne, the island's largest city, persuaded five nearby municipalities to set up a district heating system. It used waste heat produced by the island's central power station to warm water that was piped to nearby homes. Other green energy projects soon followed, including two new straw-fueled heating plants, a biogas facility that turned fermented farm waste into natural gas, and the first wind turbines. In 2007, a committee of island residents went further, producing a plan that envisioned a "bright green island" with an economy powered completely by renewables by 2025.

It seems to be working. Today, wind and other renewables provide 43.4% of Bornholm's electricity. That number jumps when biomass burned for district heating and power generation is added. Østkraft, meanwhile, continues to add new wind turbines, solar panels, and other sources of renewable power at a brisk pace. Next year the company will replace the coal burner in its electrical plant with another that burns wood chips. But burning wood still churns out carbon dioxide (CO_2), the primary greenhouse gas. So Østkraft is trying to shift away from wood and even further toward wind and solar.

Early on, skeptics worried that because wind and solar power fluctuate depending on the weather and time of day, they couldn't provide a reliable source of electricity. "Engineers [predicted] that if we used 2% to 3% more renewables in our system, it would collapse," Aagaard recalls.

BENDTSEN'S EcoGrid research project is helping prove them wrong. For the past 3 years, the E.U.-funded project has turned Bornholm into one of the world's largest laboratories for developing smart grid technologies. These automated systems work behind the scenes to maximize the use of electricity when renewable power is abundant and slow consumption when it's not. Every 5 minutes, for example, the EcoGrid sends an electricity price update to smart controllers installed in approximately 1200 homes and 100 businesses. The controllers can be set to reduce electricity use when power is expensive, and to ramp up consumption when power is cheap. The devices don't turn off essentials, such as lights, but can postpone a refrigerator's next burst of cooling until the price declines. "Our goal is to be invisible," Bendtsen says. "Customers won't see us, but still get everything they need."

Earlier this year, Jacob Østergaard, an electrical engineer at the Technical Uni-

policy expert at the University of California, Berkeley. "It's an illustration of what can be done."

But Bornholm is also helping highlight the potential technical and political obstacles to going green. Denmark has struggled to align its bold emissions goal with tax and economic policies, and some aspects of the carbon neutral push have become politically contentious. The experience, says Lars Aagaard, managing director of the Danish Energy Association in Copenhagen, "is certainly not a walk in the park."

DENMARK LAUNCHED its energy transformation long before the recent wave of concern about climate change. In the 1970s, the nation was buffeted by the Arab oil shocks, and because it lacks extensive petroleum and coal reserves, Danish farmers and later politicians called for a push toward renewable power. Although Denmark doesn't have rushing rivers for hydropower or intense sun for photovoltaics, it does

have plenty of windy coastlines on the North and Baltic seas.

In 2012, Danish politicians committed the country to using 100% renewables in the electricity and transportation sectors. Today, renewables provide one-quarter of Denmark's energy, and the nation's use of coal, oil, and natural gas is declining. Some 2500 wind turbines provided 39.1% of Denmark's electricity in 2014, and that share is expected to jump to 50% over the next 5 years. Power stations that once burned coal or oil now use renewable straw and wood. And thanks to efficiency gains—through home renovation programs and stricter construction codes—the country has decreased its overall energy use by 12% since 2007.

All along, Bornholm has provided a model for the rest of Denmark. Historically, residents of its picturesque, cobble-paved and red-tiled towns made a living by farming and fishing. But in the 1980s, the fisheries declined and young people left

versity of Denmark (DTU), Lyngby, and his colleagues analyzed the results of the EcoGrid's efforts to shift electricity demand. The approach allowed Østkraft to increase its use of renewable energy by 8%, they concluded. The gains came despite some technical bugs and the participation of only about 6% of the island's homes. If scaled up, the approach could produce far higher gains, Bendtsen says.

The Bornholm experiment has provided other benefits as well, including helping stabilize power quality at the needed frequency of 50 hertz. Conventional power systems avoid fluctuating frequencies, which can harm electronics, by turn-

technology now accounts for nearly 7% of the nation's exports, helping support thousands of jobs.

THAT'S THE GOOD NEWS. But as Bornholm and Denmark push their energy transition beyond its formative stage, it's getting more difficult. "We have taken the low-hanging fruit," Bendtsen says. "Now, we are moving to the place where it starts to hurt."

One big challenge will be persuading the country's European neighbors to upgrade their electrical grids to allow for easier sharing of renewable energy. Analysts say that for Denmark, such robust connections

low. But Denmark could face bottlenecks as it seeks to export its growing supply of green power.

For now, excess power is often shipped to Germany (which has pledged to have all of its electricity provided by renewables by 2050). But much of Germany's population and industry is located in the southern part of the country, which has relatively few hefty grid connections with the north. Meanwhile, residents of northern Germany have resisted efforts to add new transmission lines that would benefit either the Danes or their countrymen to the south, but not them.

Unless such bottlenecks are cleared,

Researchers at the Danish Technical University's PowerLab closely track electricity flows on Bornholm, a Danish island in the Baltic Sea that has become laboratory for renewable energy.

ing on additional power plants when demand surges, causing the frequency to dip. EcoGrid's smart metering technology is able to accomplish the same thing, reducing the need for spot power generation by controlling when power is fed to appliances. "We get the exact same type of response, and even higher quality, than changing the generating equipment," Østergaard says.

On Bornholm, as in Denmark as a whole, the shift to renewable power is yielding economic and environmental benefits, supporters say. There's less air pollution, and Denmark's gross domestic product has expanded by 38% since 1990—even as overall energy use and greenhouse gas emissions dropped. Green energy

will be key to becoming carbon neutral by 2050—meaning the country will produce enough zero-carbon energy to cover its entire energy budget even if it still burns some fossil fuel. Strong power links will allow Danes to import low-carbon energy when they run short, reducing the need for fossil fuels. And they also will enable Denmark to export excess green power, helping to offset domestic carbon emissions from its transportation, manufacturing, and power sectors.

Denmark already has good electrical ties with its immediate neighbors, Sweden, Norway, and northern Germany. It imports hydropower from Norway's vast system of dams, for example, when domestic production of solar and wind power is

it won't "make sense to have a very high [renewable energy] target in Denmark after 2020," says energy analyst Anne Grete Holmsgaard, who directs the BioRefining Alliance in Copenhagen.

AN EVEN TOUGHER challenge for Danish policymakers could be transportation, still powered mainly by internal combustion, which generates about a quarter of the nation's greenhouse gas emissions. On Bornholm, the islanders have committed to eliminating gas- and diesel-powered cars and trucks just 10 years from now. But Bendtsen predicts that "won't happen."

In part, that's because Denmark doesn't have any domestic car manufacturers, and its population isn't large enough to

sway foreign automakers toward offering cleaner cars. Domestic tax policies have further complicated matters. When Danes register a new car, they pay a tax that can more than double the price of the vehicle. Those high taxes help fund the nation's welfare state and foster the widespread use of bikes and public transportation. But the tax code also represents an obstacle to jump-starting sales of electric vehicles (EVs), perhaps the best alternative to fossil-fueled vehicles.

A few years ago, Denmark's government waived the taxes on EVs in a bid to spur sales of greener cars. The transportation ministry also spent millions of dollars to create a nationwide network of charging stations, enough to power 100,000 EVs. But customers have barely budged—buying just a few thousand EVs a year. Most buyers balk at the short range of lower priced EVs, which remain relatively expensive even with the tax break. Wealthy Danes, however, who can afford pricey EVs with a longer driving range, have cashed in on the subsidy to buy high-end Tesla models. That's angered many Danish voters, who feel they're being forced to subsidize the rich. Politicians have responded by extending vehicle taxes to EVs, threatening to wipe out their sales altogether. "It's not a political success story," says Anders Jelstrup Besenbacher, who directs the Danish EV Alliance in Copenhagen, and favors keeping the tax break for lower priced EVs.

Denmark's renewables push faces other tax troubles. The country heavily taxes energy use, forcing Danes to pay electricity tariffs that are nearly seven times the wholesale cost. The upside, Aagaard says, is that such hefty fees promote conservation. The downside is that they discourage people from switching from gas to electricity for heating and cooking, and they are a further brake on electric car use. "The challenge today is not building additional renewable electricity capacity," Aagaard says. "The challenge now is moving to consumption, creating a society built for using electricity. The tax system is not appropriate anymore."

THESE GROWING PAINS have potential solutions. The European Commission, for instance, is working to create what's known as the Energy Union, which aims to upgrade grid connections between the continent's 28 energy markets. And

within Denmark, analysts say an expansion of grid-scale technologies for storing renewable energy could reduce the need to export or import power. A host of such technologies exists, including flywheels and systems that use electricity to compress air, which later drives a turbine. But so far most are expensive. The cheapest energy-storage approach is to use electricity to pump water uphill, and later release the water through a turbine. Unfortunately, Denmark is flat.

Denmark does, however, have an abundance of underground salt caverns, which the country already uses to store natural gas. The caverns could also store methane gas produced with renewable electricity, researchers proposed earlier this year in *Energy & Environmental Science*. The team showed that it could use renewable electricity to break down water to produce hydrogen, which can be reacted with carbon dioxide in devices called solid oxide fuel cells (SOFCs) to generate methane. Later, the methane could be piped through a SOFC running the reverse reaction to generate electricity at a cost of about 3 cents per kilowatt hour—about the same as pumped water systems. Such results suggest "there will be room for a lot of different technologies" as the country pursues carbon neutrality, says the head of the team, Mogens Mogensen of DTU, Risø.

DENMARK WILL NEED to move quickly if it is to meet its 2050 goal. Cost is not a major obstacle: Last year, researchers with

---

Key numbers

Population: 5.5 million

Annual emissions:* 51.1 megatons CO_2 equivalent (global rank 75)

Carbon intensity: 213 tons CO_2 equivalent/\$ million GDP (rank 167)

Per capita emissions: 9.2 tons CO_2 equivalent/person annually (rank 46)

Paris pledge: At least 40% cut from 1990 domestic greenhouse gas emissions by 2030.

....

* Without land use change and forestry. All figures from 2012.

---

the Danish Energy Agency compared the cost of abandoning the green energy commitment and sticking to fossil fuels—the cheapest option for the future—with that of four green-energy scenarios. They envisioned a mix of green energy technologies: One focused on wind power; two added in biomass for producing heat, electricity, and biofuels; and a fourth relied on biomass and wind to generate hydrogen. The bottom line: Switching to renewable energy would cost more, the researchers concluded—but not much more, because Denmark already faces the need to replace aging coal, oil, and natural gas plants. Overall, the Danish Energy Agency estimated that the greener options would cost 6 billion to 29 billion Danish kroner annually through 2050, or about \$0.9 billion to \$4.2 billion per year. That's on par with the additional 4% per year that the switch to renewables is already costing Danes.

The bigger challenge is time, the Danish Energy Agency suggests. To meet the 2050 carbon neutrality goal, Denmark's leaders will need to decide which path to pursue as soon as possible, and no later than the early 2020s, given the decades-long lifetime of newly installed energy infrastructure. An early clue could come next year, when the Danish Parliament—now controlled by the nation's right-of-center Liberal Party—is expected to begin discussing a new energy plan that covers 2020 to perhaps 2030. One question is whether Liberal Party politicians will continue to back the 2050 goal, which was set under a previous government.

In the meantime, energy experts from around the world continue to board the ferry to Bornholm, to see what others might learn from the island's experience. In the run-up to the Paris meeting, eight of the world's top 10 economies have already announced plans to collectively double their use of renewables by 2030, according to the World Resources Institute in Washington, D.C. And one lesson from Denmark is already clear, Holmsgaard says: "Danes in general take a lot of pride in green energy development … But now we are having some growing pains." ∎

Danish politicians have tried to encourage consumers to buy electric vehicles, shown here during a rally, but have had little success so far.

CAN INDIA KEEP ITS PROMISES?

India hopes that steps to limit climate change will also improve its citizens' lives. Critics say such "cobenefits" may be a pipe dream

By **Priyanka Pulla**

On 3 December 2009, a loud debate took place in India's parliament. The Copenhagen climate summit was due to begin in 4 days, and Jairam Ramesh, then the minister of environment and forests, proposed a bold commitment. The country would pledge to reduce the "emissions intensity" of its economy—the amount of carbon dioxide emitted per unit of gross domestic product—by between 20% and 25% by the year 2020.

The proposal drew withering criticism from the opposition party. Opponents said India—an immense, poor country with a pressing need for development—needed "carbon space" to grow. "India today has a dramatically lower share of carbon space compared to its fair share," said a member of the parliament from the South Indian state of Kerala. "Our policy should be based on this sober reality." Over the next few days, as the Copenhagen summit drew near, the debate rumbled into a full-blown attack on Ramesh.

In his defense, Ramesh insisted that robust climate commitments would benefit India, too. "Forget Copenhagen for the time being. Climate change is of great significance to our country," he argued. India needed to set mitigation targets for domestic reasons, he said: They would help produce not just a more benign climate, but also a better quality of life

and improved health for Indians because of reduced air pollution and cleaner water.

In the end, the Copenhagen meeting collapsed without a strong international agreement. But now, as India prepares for next week's climate summit in Paris, Ramesh's former opponents have embraced his argument: India needs to tackle climate change, they say, both for the sake of the planet and for the "cobenefits" that action will bring. The opposition party in 2009, now in power, is speaking language remarkably similar to Ramesh's. "We want our air to be clean, we want our waters to be clean, we want our environment to be clean, we want our energy to be clean, and we want more green. That is our own agenda; it is with conviction, not by compulsion," Prakash Javadekar, India's environment minister, said earlier this year after India announced its Paris climate pledge, or Intended Nationally Determined Contribution (INDC).

For a country that had resisted any climate action, the pledge, released in the early hours of 2 October in India to coincide with the birthday of freedom fighter Mohandas Karamchand Gandhi, is a major step forward. Soon to be the world's most populous country, India has Earth's third largest coal reserves, which it has not hesitated to tap as it industrializes. The country is on course to eventually vie with China as the world's top greenhouse gas

Key numbers

Population: 1.2 billion

Annual emissions:* 3013 megatons of CO_2 equivalent (global rank 3)

Carbon intensity: 491 tons CO_2 equivalent/\$ million GDP (rank 81)

Per capita emissions: 2.44 tons CO_2 equivalent/person annually (rank 129)

Paris pledge: 33% to 35% cut in 2005 emissions intensity by 2030; 40% electricity from renewables; sequester 2.5 billion to 3 billion tons CO_2 equivalent through afforestation; adopt "a climate friendly and cleaner path."

··

* Without land use change and forestry. All figures from 2012.

Coal from mines like this one in the state of Chhattisgarh dominates India's energy supply. The country has pledged a major shift to fossil-free energy by 2030.

A worker repairs tangled power lines in Noida, near New Dehli, part of India's antiquated electricity grid. It could falter under the strain of the wind and solar power that the country has committed to adding.

emitter. Even under its new pledge, India continues to insist on "carbon space" and has not committed to any future cuts in emissions. But its INDC will keep its per capita emissions well below China's for the foreseeable future.

The INDC's core targets include deriving 40% of India's electric power capacity from fossil fuel–free sources by 2030, reducing its emissions intensity by 33% to 35% by 2030, and expanding forests to create a carbon sink capable of absorbing 2.5 billion to 3 billion tons of carbon from the atmosphere. The emissions intensity goal will be relatively easy to achieve, most analysts agree, given that India is already on course to a 20% to 25% cut by 2020. But the goals for fossil-free energy and an expanded carbon sink are decidedly ambitious. "The 40% [fossil-free electricity] goal is a stretch," says Chandra Bhushan, deputy director general at Delhi-based environment and climate advocacy group Centre for Science and Environment. "I can tell you that right now."

What's more, the government's claim that these goals will improve life for India's citizens has sparked a new domestic debate. Critics say the promised cobenefits may not materialize. Indeed, they argue, some of India's climate commitments could actually threaten equity, inclusiveness, and quality of life. "The point is: Who will get the benefits of all this?" asks Sujatha Byravan, a climate policy researcher at Bangalore's Center for Study of Science, Technology and Policy and the lead author of a report that examined the cobenefits of climate mitigation in India. "One can continue to have emissions planning and development planning, and the poor will still not have energy services," she says. "They will have the same quality of life and nothing will change."

THE BIGGEST POTENTIAL FLASHPOINT is the pledge to sequester carbon by protecting or expanding forests, analysts say. About 275 million Indians subsist on resources extracted from forests, including fruits and tubers, firewood, and marketable products, the World Bank estimated in 2006. For a large part of this population, earnings from forest products make up most of their incomes. And past conservation efforts have often come at the expense of forest people, pushing them off the land and sparking unrest.

In 2004, a major conservation debacle in the Sariska Tiger Reserve of the Western Indian state of Rajasthan highlighted how alienating forest dwellers in the name of conservation can backfire. That year, reports began to emerge that no tigers, pugmarks, or even tiger carcasses had been seen in Sariska for a while. In response, officials from India's national tiger conservation mission, Project Tiger, launched an emergency tiger census. The census threw up a shocker: Local forest officials had been inflating tiger numbers all along, and the real number of tigers in the park was zero. Organized poachers had killed all of them—with the collusion of locals and corrupt officials.

The incident prompted an official investigation, which concluded that shabby treatment of the local tribal people when the sanctuary was established had contributed to the debacle. Indignant villagers had turned against forest officials, preferring instead to assist poachers.

To avoid similar debacles and protect tribal communities against injustices, India passed the 2006 Forests Rights Act, which recognizes the rights of forest-dwellers to live in and sustainably extract resources from forests. But several studies show that little has changed on the ground. "The process has moved extremely slowly for the obvious reason that the forest department doesn't favor it," says Sharachchandra Lélé, an environmental researcher at the Bangalore-based Ashoka Trust for Research in Ecology and the Environment. Tribal inhabitants of protected areas continue to be harassed by forest officials and evicted from their dwellings without being compensated.

Against that background, many are skeptical of India's ambitious afforestation target: to increase forest cover by an estimated 200,000 to 300,000 square kilometers. Social unrest and political backlash could easily derail the effort, Lélé says. "You cannot talk about carbon sequestration leading to benefits to the people who do sequestering without resolving the question of ownership," he says.

Other researchers suspect that India's rapid industrialization could make the afforestation goal unachievable in any case. When India announced last year that its forest cover had increased by 5871 square kilometers between 2010 and 2012, environmentalists were skeptical. Under Javadekar, India's current ministry of environment has adopted a no-delay approach to environmental clearances for industrial projects. "There is no way they are going to be able to meet the afforestation goal, because this government is going exactly in the opposite direction. They are diluting all the environmental laws. They are virtually going to open half the forests," says D. Raghunandan, a climate policy researcher at the New Delhi–based nonprofit Centre for Technology and Development.

One reason India is reporting a growing forest cover despite widespread deforestation could be that official figures conflate forests with commercial plantations. The figures include not just natural forest, but also monocultures such as coconut, areca nut, mango, and orange, concluded a study published last year in the Indian journal *Current Science* by a team of researchers from Bangalore's Indian Institute of Science.

Although plantations could effectively sequester carbon, they could worsen the lot of forest-dwellers if they replace forests, says Rohini Chaturvedi, who leads the Washington, D.C.–based World Resources Institute's forest restoration program in India. "If we are looking only at carbon sequestration, we might find that plantations are a better route," Chaturvedi says. "But where we need multiple uses and have large populations de-

pending on forests for food, fodder, and fuel, plantations often cannot provide the range of ecosystem services."

India's INDC does not specify how plantations and forests would contribute to the afforestation goal, Chaturvedi notes. "We will have to wait for the government to clarify." But she is heartened by the pledge, calling it "definitely an encouraging sign."

Lélé, though, thinks that the afforestation goal will either fail or exact an unacceptable social cost. "If local people are powerful enough, they will block the afforestation efforts, burn down, or uproot the trees. Conversely, if [government forestry officials are] powerful enough, they will impose their will, make money out of these international commitments, and the people will suffer."

LÉLÉ IS MORE OPTIMISTIC about the potential cobenefits of India's renewable energy commitments. To meet the target of converting 40% of total installed power capacity to fossil fuel–free sources by 2030, India will have to add between 196 and 276 gigawatts (GW) of hydropower, nuclear, solar, wind, and other renewable power. Nuclear and hydro will not grow very rapidly, especially given the political obstacles nuclear power faces right now in India, which means that much of the added capacity will have to come from solar and wind, according to the New Delhi–based Centre for Policy Research. "You are basically adding on the order of magnitude of your entire current grid in renewables, which is a large number," says Navroz Dubash, a climate policy researcher with the organization.

The cobenefits could be substantial. By displacing coal, renewables could reduce India's dependence on coal imports and also limit air pollution, fast becoming a major health hazard in metropolises such as Delhi. Yet the technological challenges are formidable. India's power grid is antiquated and suffers from massive transmission and distribution losses. Adding hundreds of GW of power from fluctuating solar and wind sources is a challenge even in countries with sophisticated and stable grids (see p. 1020). In India, a massive surge of renewables could worsen power quality and blackouts.

Avoiding that outcome will require a vast investment in the grid, including the development of automated systems that can manage the fluctuations in renewable energy, and storage systems that can bank excess green power for later use. The INDC estimates that, all told, India will spend about

$834 billion on mitigation by 2030. (It doesn't clarify how much of the money will be raised internally and how much will be sought from external donors.) Despite such challenges, "the INDC target is in that zone where you can reach it," believes Ashok Sreenivas, an energy policy researcher at Pune's Prayas Energy Group, an energy research and advocacy nonprofit. But "you have to take special, extra efforts to get there."

THE INDC GOALS might be easier to reach if India had pushed the cobenefits philosophy further, Raghunandan says. Instead of focusing exclusively on the grid and on forests, he argues that it should have pledged to clean up transportation, solid waste management, and building emissions—all sectors that promise both emission cuts and tangible payoffs such as improved quality of life and equity. "I am, frankly speaking,

A barber in Srinagar lights a candle during a power cut. A better grid would aid the shift to new energy sources and boost the economy.

disappointed," says Raghunandan, who co-wrote an article with Dubash and Sreenivas in 2013 that argued in favor of a cobenefits approach to climate change.

He notes that the INDC gives no more than a mention to transportation measures such as expanding railway freight corridors, which could shift freight away from energy-inefficient, polluting trucks. And he complains that India's domestic policies are boosting transportation emissions. For example, in May 2014, it announced a plan to

build 200 no-frills airports in smaller cities. Many are underused, Raghunandan says. "These are vanity projects." Instead of relying on climate-friendly rail links between cities, "the government is promoting regional airports, which are for the rich. This is not equity."

Another missed opportunity, he says, is boosting the energy efficiency of residential buildings—a step that could help consumers by cutting power bills. A 2010 report by McKinsey & Company consulting firm estimated that up to 80% of all buildings that will stand in India in 2030 have yet to be constructed. "A house is not like a mobile phone. When people build a house, they build it for 50 years," Raghunandan says. "If you can reduce the energy demand on domestic construction by even 15% to 20% by construction methods—hollow bricks, insulating and reflective materials—so that cooling loads reduce, that will be a huge contribution. But the INDC doesn't talk about it."

INDIA COULD ULTIMATELY take greater action in those sectors, regardless of the outcome in Paris next month. But those critical of India's current negotiating stance say that, overall, the nation continues to view its climate commitments as something it is offering to the world in return for carbon space and some unspecified amount of aid. The INDC says that "successful implementation" of its core climate targets will require foreign funding, though it does not define success. "It is not that we are not going to do anything. It is like: We may do something, but we may do more if we have funding. Or, we may do everything by ourselves. All possibilities are open," Sreenivas says.

The focus on international negotiations has kept Indian policymakers from rigorously analyzing the potential domestic benefits of stronger climate measures, critics say. And some fear a lack of stronger action by India could ultimately give more-developed nations—historically the major contributors to global emissions—an excuse to weaken their own climate commitments. In contrast, Dubash says, a more aggressive move by India could push industrialized nations to do more. It would, he says, "open up the space to also call out the industrialized countries and say: 'Listen, we are actually doing a lot despite starting from a low base. You guys need to do more.'" ∎

Priyanka Pulla is a freelance writer in Bangalore, India.
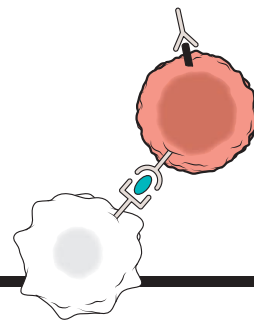
PERSPECTIVES

HUMAN GENOME

The indispensable genome

The core genes essential for life in human cells are defined

By **Charles Boone** *and* **Brenda J. Andrews**

Game-changing moments in functional genomics often reflect the development and application of powerful new reagents and methods to provide new phenotypic insight on a global scale. Three independent studies describe systematic, genome-scale approaches to defining human genes that are indispensable for viability, which collectively form the essential gene set. On pages 1092 and 1096 of this issue, Blomen *et al.* (*1*) and Wang *et al.* (*2*), respectively, report a consistent set of ~2000 genes that are indispensable for viability in human cells. Moreover, very similar results were obtained by Hart *et al.* (*3*). For the first time, we now have a firm handle on the core set of essential genes that are required for human cell division. This opens the door to studying the roles of essential genes, how gene essentiality depends on genetic and tissue contexts, and how essential genes evolve.

Scientists could anticipate identifying a core set of human cell essential genes since the first description of the essential gene set in yeast, a model organism that has been a reliable test bed for functional genomics studies. In yeast, defining the essential gene set was relatively straightforward because it has a compact genome and efficient homologous recombination, which enables precise
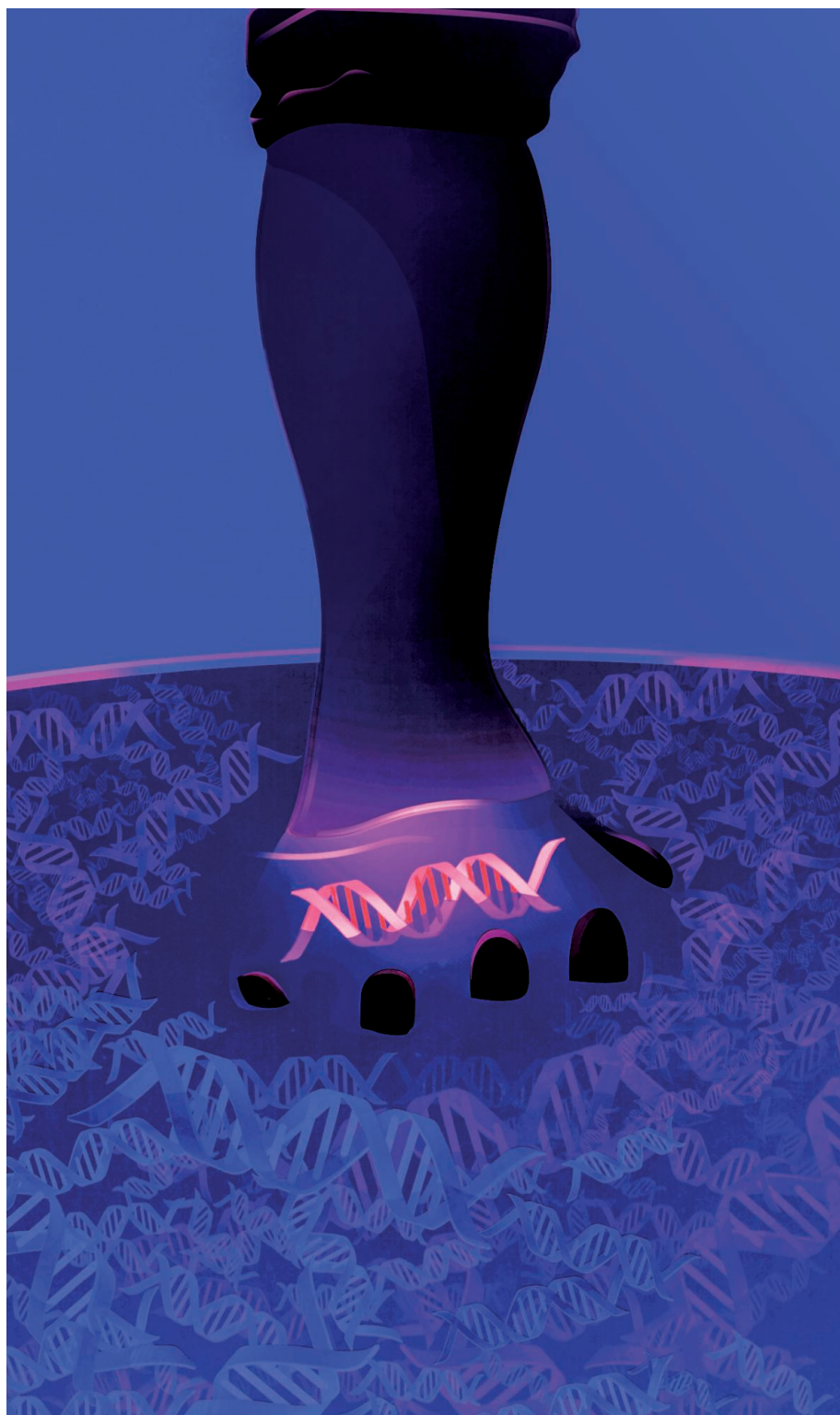
Donnelly Centre and Department of Molecular Genetics, University of Toronto, Toronto, Ontario M5S 3E1, Canada. E-mail: charlie.boone@utoronto.ca; brenda.andrews@utoronto.ca

gene deletions within their normal chromosomal context.

Yeast gene deletion can be performed in a diploid cell, which then can be induced to undergo meiosis, producing pure haploid cells carrying a precise deletion of a single gene. The failure of a haploid deletion mutant strain to grow after such manipulations identifies genes that are essential for cell division. Genome-scale application of this deletion mutant analysis revealed that ~1000 of yeast's 6000 genes are essential for viability in standard growth conditions (4, 5). The yeast essential genes encode proteins that drive basic cellular functions such as transcription, translation, DNA replication, cell division cycle control, and fundamental metabolism. Moreover, the yeast essential genes share several attributes that reflect their critical role in cellular life. For example, they are often conserved and evolutionarily constrained, are highly expressed, and encode abundant proteins that tend to form stable complexes and thus are rich in protein-protein interactions (6–8).

The landscape of essential genes in human cells can now be explored using the conceptual framework established in yeast. Blomen et al. surveyed the essential genes in a near-haploid chronic myelogenous leukemia (CML) cell line, KBM7, and its nonhematopoietic cell derivative HAP1, which is haploid for all chromosomes. These haploid cell lines enable essential genes to be identified using the gene-trap insertional mutagenesis methodology. Wang et al. also surveyed KBM7 cells using gene-trap analysis, but they further used a gene-editing strategy based on clustered regularly interspaced short palindromic repeats (CRISPR) that enabled exploration for essential genes in diploid cell lines, including another CML cell line (K562) and two Burkitt's lymphoma cell lines (Raji and Jiyoye). Similarly, Hart et al. used a CRISPR-Cas9 gene-editing strategy to explore essential genes, but they did so across a diverse series of adherent cell types including colorectal cancer cell lines, cervical cancer cells, primary patient-derived glioblastoma cells, and immortalized retinal epithelial cells. All three groups found that human essential genes are highly conserved, and much like yeast, they encode abundant proteins that engage in protein-protein interactions. The core set of human cell essential genes also tend not to be duplicated and appear to have increased evolutionary constraints, as they evolve slowly and are associated with fewer deleterious single-nucleotide polymorphisms. Although many essential genes are involved in fundamental biological processes including transcription, translation, and DNA replication, a substantial fraction remains functionally uncharacterized. Indeed, each analysis prioritized a wealth of uncharacterized genes whose essential roles are waiting to be explored.

Blomen et al. and Wang et al. identified a coherent and overlapping set of essential genes in two related haploid cell lines, which emphasizes the potential to characterize the core set of human cell essential genes that drives life in all cell types. Nonetheless, the essentiality of some genes, as demonstrated previously by large-scale RNA interference–based screens (9, 10) as well as the Hart et al. analysis, is context-dependent and affects viability in a cell type–specific manner. Indeed, by screening additional CML or Burkitt's lymphoma diploid cell lines, Wang et al. discovered 48 genes that exhibited cell type–specific essentiality. These results highlight the potential of genome-scale genetic screens to reveal the biology underlying cell differentiation and the potential to identify specific cancer cell vulnerabilities that can be exploited as targets for personalized therapeutic strategies. Large-scale genetic screens will have to be implemented in a number of different cell lines to generate a complete picture of the human cell essential gene set.

All three studies (1–3) conclude that ~10% of the ~20,000 genes in human cells are essential for cell survival. Although this frequency may change somewhat with screening of other cell lines, it is clear that most human genes are nonessential under laboratory culture conditions. This finding is consistent with the results of large-scale screens from various model organisms that have established a comparable ratio of essential to nonessential genes, highlighting the buffering against genetic and environmental insults inherent in eukaryotic genomes (11). Many nonessential genes impinge on essential functions, but they do so in the context of a sophisticated biological machine that has been wired with backup pathways (12).

The essential functions of pairs of nonessential genes can be examined in synthetic lethal double-mutant screens. Synthetic lethality occurs when two mutations, neither of which is lethal on its own, combine to generate a lethal double mutant—a genetic interaction that has been explored in yeast. Global genetic interaction screens have identified hundreds of thousands of synthetic lethal genetic interactions, revealing contextual lethality that is at least an order of magnitude more prevalent than the lethality caused by single gene perturbation (12). Blomen et al. begin to address the extent of synthetic lethal interactions in human cells by screening a set of five nonessential genes with roles in secretion for synthetic lethal negative genetic interactions. They discovered an average of ~20 synthetic lethal double-mutant interactions for a given nonessential gene, and these interactions tend to occur with functionally related genes. Even this relatively small genetic network suggests that the properties of the extensive genetic networks mapped for yeast are conserved and can now be mapped efficiently in human cells. The genetic network described by Blomen et al. ought to catalyze large-scale, collaborative efforts to map genetic interactions in human cells. Such an effort promises to enable functional annotation of the human genome, because genetic interaction profiles are rich in functional information and provide a quantitative measure of gene function.

The ability to systematically assess all genes within a genome for a particular phenotype greatly accelerated the functional analysis of yeast genes (13). The yeast deletion mutant collection also opened the door to the field of chemical genetics, enabling new approaches for linking bioactive compounds to their cellular targets (14). Genome-wide mutant collections can be analyzed in pools, but there is also the potential to generate arrays of mutants, providing a new resource for high-content screening approaches to address cell biology phenotypes. The studies of Blomen et al., Wang et al., and Hart et al. reveal the core essential gene set for human cells, setting the stage for the next wave of new genetic and chemical-genetic science that will take place directly in human cells. A future challenge will be to develop genetic tools, such as conditional alleles of essential genes, for exploring the terminal phenotypes and the various molecular mechanisms underlying the lethality associated with perturbation of different essential functions. ∎

> "...the essentiality of some genes... is context dependent..."

REFERENCES

1. V. A. Blomen et al., Science 350, 1092 (2015).
2. T. Wang et al., Science 350, 1096 (2015).
3. T. Hart et al., Cell 10.1016/j.cell.2015.11.015.
4. E. A. Winzeler et al., Science 285, 901 (1999).
5. G. Giaever et al., Nature 418, 387 (2002).
6. P. M. Kim et al., Science 314, 1938 (2006).
7. A. E. Hirsh, H. B. Fraser, Nature 411, 1046 (2001).
8. C. Pal, B. Papp, L. D. Hurst, Genetics 158, 927 (2001).
9. B. Luo et al., Proc. Natl. Acad. Sci. U.S.A. 105, 20380 (2008).
10. R. Marcotte et al., Cancer Discov. 2, 172 (2012).
11. J. L. Hartman IV et al., Science 291, 1001 (2001).
12. M. Costanzo et al., Science 327, 425 (2010).
13. T. R. Hughes et al., Curr. Opin. Microbiol. 7, 546 (2004).
14. A. Y. Lee et al., Science 344, 208 (2014).

SOLAR FUELS

A quick look at how photoelectrodes work

Transient photoreflectance spectroscopy reveals charge carrier dynamics in water splitting

By **Ole Hansen,**[1] **Brian Seger,**[2]
Peter C. K. Vesborg,[2] **Ib Chorkendorff**[2]

I t is appealing to harvest solar energy directly into chemical bonds with photo-electrochemical (PEC) cells—for example, by splitting water into hydrogen (H_2) and oxygen (O_2), as first demonstrated by Fujishima and Honda (1). Achieving the highest possible efficiency requires rapid transfer of the charge carriers generated by semiconductor photoabsorbers (2) to the catalysts for H_2 and O_2 evolution. Long-term stability requires protection layers for the semiconductors against strong acid or base. Direct experimental observation of charge carrier dynamics at these complex interfaces, which is critical for optimization, has been a major challenge. On page 1061 of this issue, Y. Yang et al. (3) show how transient photoreflectance spectroscopy can reveal information about the carrier dynamics

[1]Department of Micro- and Nanotechnology, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. [2]Center for Individual Nanoparticle Functionality, Department of Physics, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. E-mail: ole.hansen@nanotech.dtu.dk

and the electric field near the semiconductor surface.

Recently, computational screening of a wide range of materials has revealed the difficulty of finding a combination of semiconductor materials that fulfills the requirements for tandem operation unless a protection layer strategy is adopted (4). The photoabsorber requirements can be decoupled from the stability requirements because stability is provided by thin, transparent, semiconducting protection layers [for example, titanium dioxide (TiO_2) or nickel oxide (NiO)] that separate the absorber materials from the electrolyte. The use of protection layers also introduces a heterojunction interface between the semiconductor absorber and the protection layer, where the band lineup and charge carrier dynamics become critical. In other words, the photogenerated electrons (or holes) must be transported efficiently from the absorber through the protection layer to reach the catalyst and ultimately into the pertinent redox couple in the electrolyte.

Yang et al. interrogated carrier dynamics using an optical pump and probe beam setup with ~150-fs pulses. The short-wave-

length pump beam creates electron-hole pairs near the semiconductor surface. A delayed wide-band probe beam then measures changes in energy-resolved reflectance as a function of the delay time (see the figure). The changes in reflectance are caused by two direct effects on absorption: band filling caused by changes in carrier occupancy, and the Franz-Keldysh effect caused by electric fields. Both effects change the dielectric function of the absorber material according to the Kramers-Kronig relations.

Yang et al. measured three different p-type gallium indium phosphide ($GaInP_2$)–based samples: a bare sample, a platinum (Pt) Schottky-barrier sample, and a TiO_2-protected sample. They observed substantially different responses. The bare sample suffered from band filling–related dynamic absorption changes that decayed with the effective minority carrier lifetime of the sample, mainly because of surface trapping. The samples with junctions did not show the band filling–related signature within the time resolution of the experiment because charge separation occurred faster than 150 fs; that is,

Beams enter
Short wavelength pump beam and wide-band probe beam enter the cell

Reflectance
Delayed probe beam measures changes in absorption

Absorption
Pump beam creates electron-hole pairs
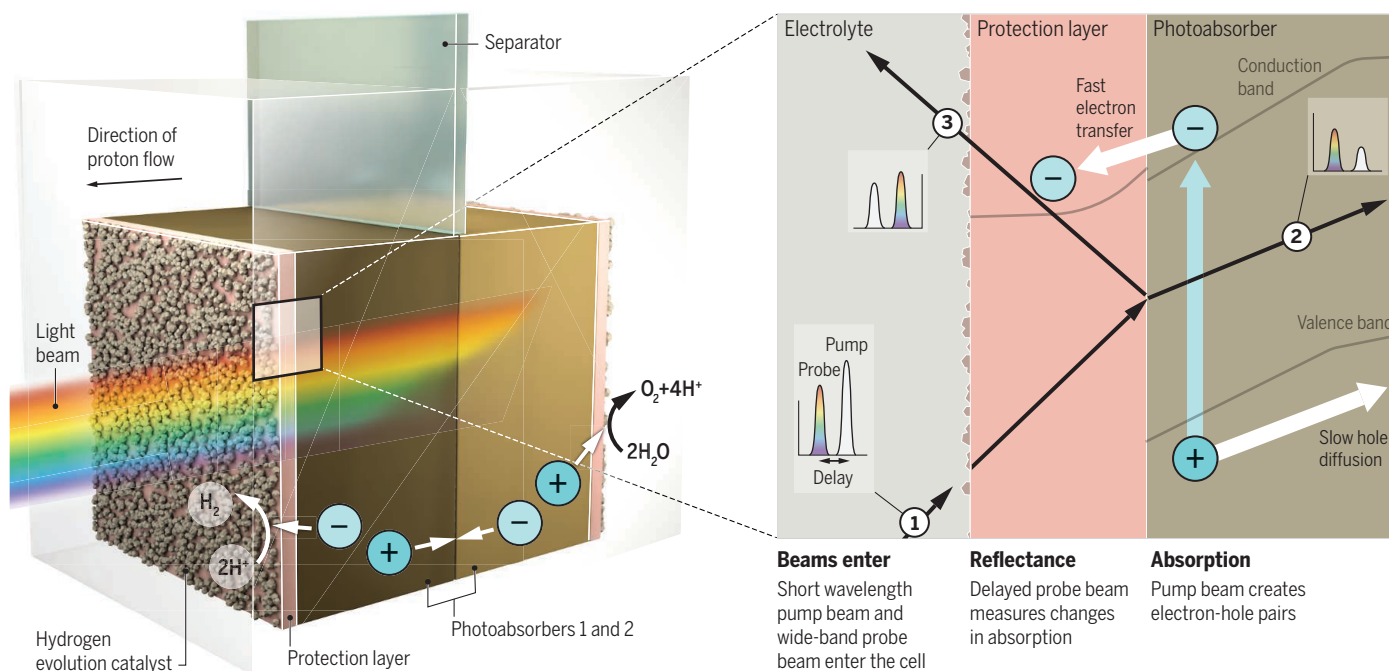
Interrogating solar fuel production. The schematic shows a tandem PEC cell with protected absorber layers; the combined photovoltage of the two absorbers drives the catalyzed water splitting. The magnification illustrates the pump-probe beams and shows that electron transfer is faster than hole diffusion in the charge separation process.

ILLUSTRATION: V. ALTOUNIAN/SCIENCE

the photogenerated electrons transferred extremely quickly to Pt or TiO_2. Instead, rather persistent, strong Franz-Keldysh–related features arising from the dipole electric field created by the carrier separation were seen. From these features, the static electric field, as well as the dynamics of the electric field near the junctions, can be deduced. The Franz-Keldysh features persisted an order of magnitude longer in the $GaInP_2/TiO_2$ sample (it hardly decayed in the time frame of the experiments) when compared to the $GaInP_2/Pt$ sample because of improved separation of the charge carriers.

Measurements of $GaInP_2/TiO_2$ samples with TiO_2 thicknesses from 0.5 to 35 nm also showed that the decay time for the Franz-Keldysh features increased with thicker layers, which improved charge separation by impeding recombination. Yang *et al.* measured the structures in the "dry" state without electrolyte, but in the relevant wavelength range, the electrolyte is highly transparent. Thus, measurements might be performed in operando (for example., at varying bias light intensity) under mild conditions, where gas bubble formation is negligible. Such measurements are important because for some protection layers, the charge transport mechanism is not even entirely clear—for example, TiO_2 protection of photoanodes (5, 6).

The detailed insight into the charge separation at the absorber-protection layer interface offered by the method of Yang *et al.* promises to speed up the development of well-behaved protected photoelectrodes, which then have to be coupled to good catalysts. The H_2 evolution catalyst could be the effective, earth-abundant MoS_2 catalyst (7), or even Pt, which was recently shown to be scalable to the terawatt power production level (8). Known O_2 evolution catalysts, however, are much less effective, and despite substantial efforts, very little progress has been made (9) because of scaling relations that limit all known inorganic catalysts (10). Nevertheless, some of the best metal-oxide catalyst materials for O_2 evolution (9), such as iron-treated NiO, are both efficient catalysts and protection layers (11). ∎

REFERENCES

1. A. Fujishima, K. Honda, *Nature* **238**, 37 (1972).
2. O. Khaselev, J. A. Turner, *Science* **280**, 425 (1998).
3. Y. Yang et al., *Science* **350**, 1061 (2015).
4. B. Seger et al., *Energy Environ. Sci.* **7**, 2397 (2014).
5. S. Hu et al., *Science* **344**, 1005 (2014).
6. B. Mei et al., *J. Phys. Chem. C* **119**, 15019 (2015).
7. T. F. Jaramillo et al., *Science* **317**, 100 (2007).
8. E. Kemppainen et al., *Energy Environ. Sci.* **8**, 2991 (2015).
9. C. C. L. McCrory et al., *J. Am. Chem. Soc.* **135**, 16977 (2013).
10. A. Vojvodic, J. K. Nørskov, *Natl. Sci. Rev.* **2**, 140 (2015).
11. B. Mei et al., *J. Phys. Chem. Lett.* **5**, 3456 (2014).

10.1126/science.aad6060
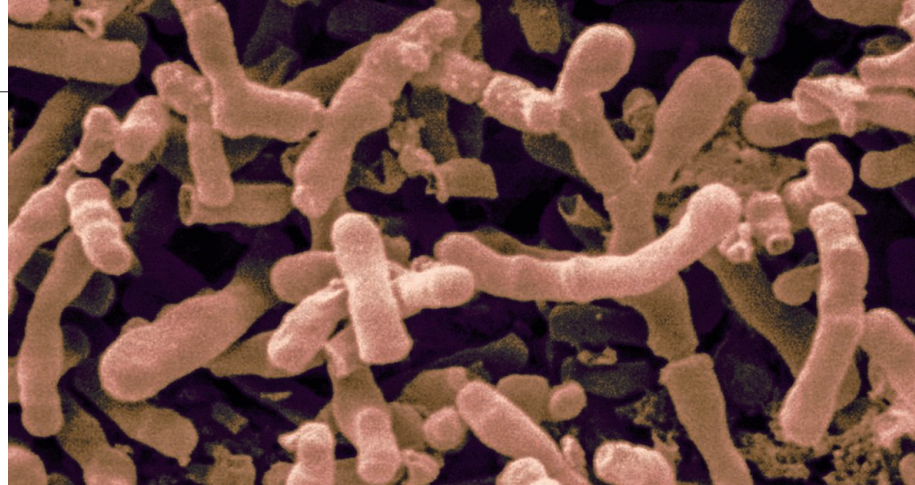
Gut microbe. *Bifidobacterium* is found in the intestines of most mammals, including humans.

IMMUNOTHERAPY

Could microbial therapy boost cancer immunotherapy?

Intestinal microbes affect immunotherapy responses in mouse models of cancer

By **Alexandra Snyder,[1] Eric Pamer,[2] Jedd Wolchok[3]**

Immunotherapies known as checkpoint blockades are rapidly changing standard treatment and outcomes for patients with advanced malignancies, as they lead to long-term disease control in a subset of patients (1). On pages 1084 and 1079 of this issue, Sivan *et al.* (2) and Vétizou *et al.* (3), respectively, illustrate an important role for the gut microbiome in modulating the efficacy of this treatment.
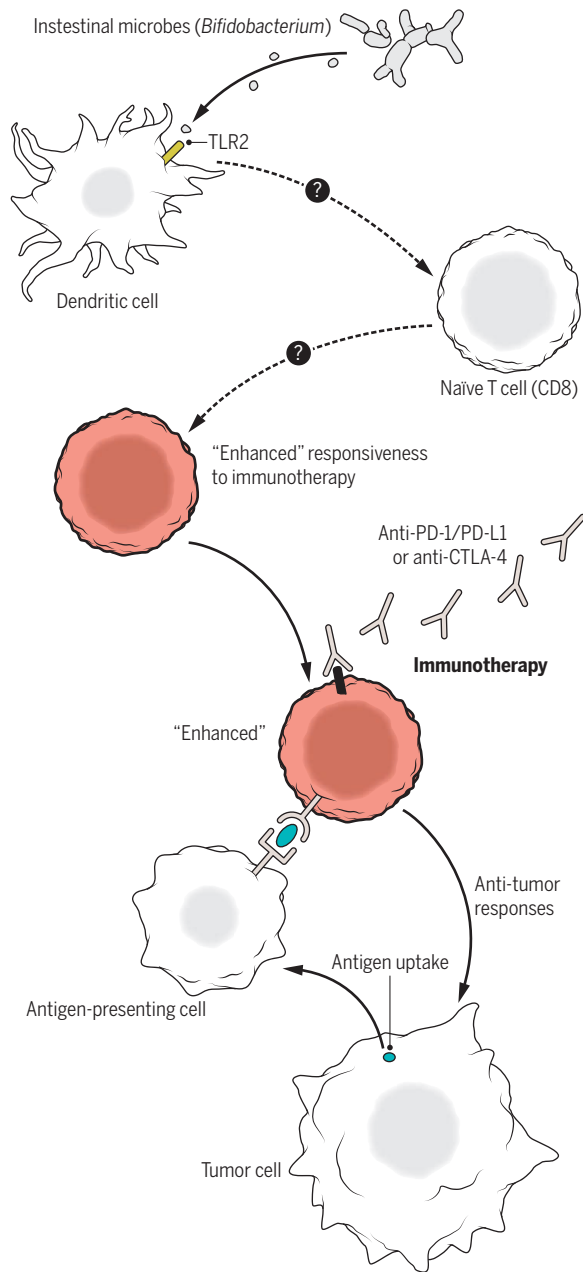
The checkpoint blockade immune therapies currently approved by the U.S. Food and Drug Administration for treating advanced melanoma and lung cancers target suppressive receptors on the surface of T cells. Anergic and/or exhausted T cells are released from inhibition by agents that target cytotoxic T lymphocyte–associated protein 4 (CTLA-4) or programmed cell death 1 (PD-1) so that T cells may then recognize and attack tumor cells. However, checkpoint blockade agents show variable efficacy within and across disease types. Concurrent with the development of these agents, tumor-intrinsic (4), tumor micro-

[1]Department of Medicine, Memorial Sloan Kettering Cancer Center, and Weill Cornell Medical College, New York, NY, USA. [2]Infectious Diseases Service, Department of Medicine, Memorial Sloan Kettering Cancer Center and Lucille Castori Center for Microbes, Inflammation, and Cancer, New York, NY, USA. [3]Swim across America–Ludwig Collaborative Research Laboratory, Department of Medicine, Memorial Sloan Kettering Cancer Center and Weill Cornell Medical College, Immunology Program, Ludwig Center for Cancer Immunotherapy, New York, NY, USA. E-mail: wolchokj@mskcc.org

environmental (5, 6), and circulating (7) factors have been associated with benefit or resistance to therapy; considered alone, each is an imperfect predictor.

Gut microbiota play a role in immune system development (8) and can affect the occurrence of autoimmunity (9, 10). In cancer, a diverse gut microbiome independently predicts for better outcomes after allogeneic stem cell transplant (11). The antibody against CTLA-4 (anti–CTLA-4) agent ipilimumab is thought to alter gastrointestinal immunity (12).

Sivan *et al.* and Vétizou *et al.* provide strong evidence for the role of stool microbiota (i.e., intestinal microbes) in response and resistance to immunotherapy. Sivan *et al.* illustrate the importance of *Bifidobacterium* to antitumor immunity and anti–PD-L1 antibody against (PD-1 ligand) efficacy in a mouse model of melanoma. The authors demonstrate that mice raised in two different facilities [Jackson Laboratory (JAX) and Taconic Farms (TAC)] that are known to harbor distinct microbiota exhibit differential tumor growth that disappears upon cohousing of the animals. Furthermore, when fecal material from JAX mice, whose tumors grow more slowly, was transferred into the intestine of TAC mice, the latter exhibited delayed tumor growth and enhanced CD8+ T cell infiltration of the tumor. Anti–PD-L1 therapy was more effective in JAX mice, and the combination of JAX fecal transfer to TAC mice undergoing anti–PD-L1 therapy was more effective than either intervention alone. Si-

Gut bacteria affect immunotherapy effectiveness. Intestinal microbes, such as *Bifidobacterium*, may cause dendritic cells to enhance T cell responsiveness to checkpoint blockade therapies, although the mechanism of this effect remains to be elucidated.

van *et al.* also show that *Bifidobacterium* confers nearly the same effect as JAX stool transfer and that bacteria must be alive for the treatment to be effective. Investigation of the underlying mechanisms of this effect reveals that *Bifidobacterium* alters dendritic cell activity, which in turn leads to improved tumor-specific CD8+ T cell function.

Vétizou *et al.* show that anti–CTLA-4 therapy is efficacious in mice housed in specific pathogen-free but not completely germ-free facilities. Several experiments underscore the importance of gut microbiota to response to therapy. Treatment with broad-spectrum antibiotics (ampicillin, colistin, and streptomycin) or imipenem dampened anti–CTLA-4 efficacy. When antibiotic-treated or germ-free–housed mice were fed *Bacteroides* isolates, the anticancer effect of anti–CTLA-4 was restored. Furthermore, adoptive transfer of *Bacteroides fragilis*–specific T helper 1 cells or injection of dendritic cells loaded with *B. fragilis*–purified polysaccharide into ACS-treated mice increased (although did not entirely restore) anti–CTLA-4 efficacy. Treatment with the antibiotic vancomycin also caused an increase of *Bacteroidales* and was associated with enhanced CTLA-4 efficacy. Transfer of feces from melanoma patients who harbored *Bacteroidales* species into the intestine of mice enhanced CTLA-4 efficacy in mice, which showed that boosting checkpoint blockade therapy can be mediated by bacteria colonizing the human gut.

In addition to illustrating the importance of microbiota to drug efficacy, Vétizou *et al.* also argue that, conversely, treatment with anti–CTLA-4 therapy alters the microbiome, specifically causing a decrease in *Bacteroidales* and *Burkholderiales* and an increase in *Clostridiales*. By demonstrating that Bacteroides also colonizes the small intestine, Vétizou *et al.* begin to address but leave unanswered many interesting questions regarding the localization of different bacterial taxa along the length of the gut and their relative impact on checkpoint blockade–driven immune responses.

Although they investigate different checkpoint blockade agents, Sivan *et al.* and Vétizou *et al.* come to a conceptually important conclusion: The composition of intestinal microbiota affects checkpoint blockade efficacy and can be manipulated to improve responses. Although both studies use the same mouse model of melanoma, the difference in specific bacteria that they identify (*Bacteroidales* or *Bifidobacterium*), may be due to the checkpoint blockade agents and specific experimental conditions used. However, this difference hints at the challenges that may arise in applying the findings to human patients. Human patients live in distinct environments with contrasting dietary habits and who consequently will exhibit substantial interpatient variability in microbiota relative to mouse models. Furthermore, the frequent exposure of cancer patients to antibiotics and antibiotic-resistant bacteria also shapes their intestinal microbiome (*13*). The dose, frequency, timing relative to immunotherapy, and content of administered bacteria or microbiome-altering antibiotics remain to be determined. These data will need to be integrated into the already-complex multidimensional model of tumor, tumor microenvironment, and host factors involved in therapeutic efficacy. Consequently, studying this question in humans will be challenging.

However, there is reason for optimism: The efficacy of fecal transplant for treating *Clostridium difficile* infection in humans has been conclusively shown (*14*), so it is now possible that this treatment strategy could be applied to immunotherapy-treated patients. Administration of specific bacterial species or combinations of bacteria that would enhance responses to therapy would be preferable but will require extensive development and testing. For now, additional studies on patient populations are warranted. Stool samples can be collected and the microbiota analyzed, and thus prospective collection from all members of phase 2 or 3 clinical study is feasible. The findings of Sivan *et al.* and Vétizou *et al.* show that collection of fecal samples should be considered going forward in immunotherapy studies to characterize and ultimately manipulate this factor to favor response in immunotherapy-treated cancer patients. ∎

REFERENCES

1. M. A. Postow *et al.*, *N. Engl. J. Med.* **372**, 2006 (2015).
2. A. Sivan *et al.*, *Science* **350**, 1094 (2015).
3. M. Vétizou *et al.*, *Science* **350**, 1079 (2015).
4. A. Snyder *et al.*, *N. Engl. J. Med.* **371**, 2189 (2014).
5. S. Spranger *et al.*, *Sci. Transl. Med.* **5**, 200ra116 (2013).
6. P. C. Tumeh *Nature* **515**, 568 (2014).
7. G. Y. Ku, J. Yuan *et al.*, *Cancer* **116**, 1767 (2010).
8. S. K. Mazmanian, C. H. Liu, A. O. Tzianabos, D. L. Kasper, *Cell* **122**, 107 (2005).
9. H. J. Wu *et al.*, *Immunity* **32**, 815 (2010).
10. R. Horai *et al.*, *Immunity* **43**, 343 (2015).
11. Y. Taur *et al.*, *Blood* **124**, 1174 (2014).
12. D. Berman *et al.*, *Cancer Immunity.* **10**, 11 (2010).
13. Y. Taur, E. G. Pamer, *Curr. Opin. Infect. Dis.* **26**, 332 (2013).
14. E. van Nood, *N. Engl. J. Med.* **368**, 407 (2013).

10.1126/science.aad7706

APPLIED OPTICS

Optical meta-atoms: Going nonlinear

Metamaterials are poised to transform nonlinear optics

By **Natalia M. Litchinitser** *and* **Jingbo Sun**

Nonlinear optics investigates the light-matter interactions in media, in which the dielectric polarization of the medium responds nonlinearly to the electric and/or magnetic field of the light. Materials with the potential for a large, fast, and broadband nonlinear response have been explored for decades; if realized, these would revolutionize nonlinear optics, leading to low-power, compact, and ultrafast applications. However, the materials now available are limited, either by relatively low nonlinear susceptibilities for ultrafast nonlinear processes or by slow response times attributable to photorefractive effect and thermal nonlinear phenomena. Moreover, growing demand for integration of multiple optoelectronic functionalities on a chip calls for nonlinear materials that are compatible with standard fabrication approaches, such as complementary metal-oxide semiconductor technology. Metamaterials have been predicted to enable a plethora of novel light-matter interactions, including magnetic nonlinear response, backward phase-matching, and the nonlinear mirror (*1*–*3*). Linear optical properties such as dielectric permittivity, magnetic permeability, and refractive index can be designed to be positive, negative, or even zero by properly tailoring various properties of meta-atoms (the unit cells of metamaterials). Engineering nonlinear properties of metamaterials beyond those available in nature may be feasible by judiciously designing their quantum, geometric, and topological properties (*4*).

A question of fundamental and practical importance that arises with these metamaterials is whether there is a limit to the nonlinear response characterized by macroscopic nonlinear susceptibilities or microscopic hyperpolarizabilities. Indeed, it has been shown that there are fundamental limits to the off-resonant, electronic,

nonlinear optical response (*5*). However, the largest hyperpolarizabilities of the best molecules fall short of the fundamental limit by a factor of $10^{3/2}$ (see the figure, panel A). Understanding of this discrepancy is still a subject of extensive research.

In parallel with the development of fundamental quantum models, a semi-empirical relationship (Miller's rule) has been proposed to estimate the nonlinear response from its linear counterpart (*6*). Because metamaterials gain their unique linear properties from their structural design, it is not obvious whether any generalized rules relating linear and nonlinear properties can be established. The initial studies of the second-order susceptibility of a plasmonic metasurface (see the figure, panel B) suggest that although Miller's rule may not provide accurate esti-

mates, the nonlinear scattering theory approach does agree well with experimental results (*7*).

To date, appreciable efforts have been devoted to enhancing the nonlinear optical response using various plasmonic metamaterials (*8*). However, a majority of these studies exploited local field enhancements rather than the actual design of the nonlinear optical response of meta-atoms. The local field enhancement inside each meta-atom leads to increased effective nonlinearity, but this comes at the expense of increased loss or decreased coherence length (*9*). Consequently, this approach might be useful for applications such as sensing, but not for ultrafast, low-power switching or wavelength conversion.

One of the first designs relying on the unique capabilities of metamaterials to enable effective nonlinear properties surpassing those of its ingredients was demonstrated by structuring three centrosymmetric dielectrics to realize nanolaminates exhibiting second-order nonlinearity not seen in any of the three components. Thin layers of A = Al_2O_3, B = TiO_2, and C = HfO_2 were arranged into a noncentrosymmetric ABC-stack such that the individual surface nonlinearities
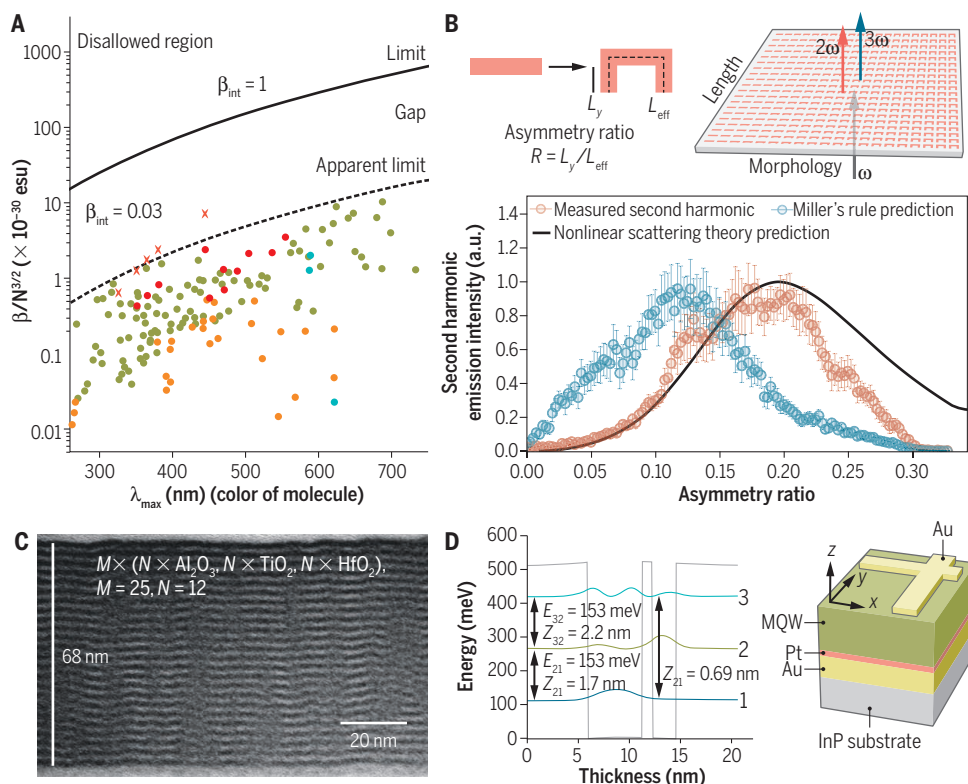
Nonlinearity at the limit. (**A**) Normalized values from the literature (points), maximum limit (solid line), and apparent limit (dashed line). (**B**) Top: Schematic of the metamaterial array with unit cell gradually changing from a symmetric bar to an asymmetric U-shape. Bottom: Second-harmonic intensity as a function of the asymmetry ratio. (**C**) The ABC-sample cross section. (**D**) Left: Conduction band diagram of one period of an $In_{0.53}Ga_{0.47}As/Al_{0.48}In_{0.52}As$ coupled quantum well structure designed for giant nonlinear response for second-harmonic generation. Right: Schematic of the unit cell of a semiconductor metasurface for second-harmonic generation.

Department of Electrical Engineering, State University of New York, Buffalo, NY 14260, USA. E-mail: natashal@buffalo.edu

GRAPHIC: P. HUEY/*SCIENCE*

originating at the boundary of neighboring materials do not sum to zero (see the figure, panel C) (10). Following this approach, artificial unidimensional crystals with the main component of their nonlinear susceptibility tensor of about 5 pm/V, comparable to well-established materials, were reported (11). These dielectric nonlinear metamaterials were grown by atomic layer deposition, making them compatible with standard fabrication technologies.

A different approach to realizing strongly nonlinear metamaterials is based on quantum engineering of electronic intersubband transitions in electron-doped multi–quantum-well semiconductor heterostructures (12). By controlling the widths of wells and barriers in these structures, the transition energy and dipole moments between electron subbands can be tailored in order to maximize the quantum mechanical expression for a particular nonlinear process. By combining quantum electronic engineering of intersubband nonlinearities with electromagnetic engineering of plasmonic nanoresonators, an ultrathin, planarized, highly nonlinear, optical 400-nm-thick metasurface with nonlinear susceptibility greater than 5×10^4 pm/V for second-harmonic generation at 8 μm has been demonstrated (see the figure, panel D).

Meta-atoms are poised to empower new strategies for optimizing the nonlinear response. Breakthroughs in the field of fast and strongly nonlinear materials are likely to be achieved by combining the current advances in both classical and quantum theory of artificial nanostructures, pattern optimization, and understanding how topology and geometry affect the nonlinearities. If realized, such materials would revolutionize nonlinear optics, leading to low-power, compact, and ultrafast applications of nonlinear effects in optical communications, optical computing, image processing, and quantum optical devices. ∎

REFERENCES AND NOTES

1. W. Cai, V. M. Shalaev, *Optical Metamaterials: Fundamentals and Applications* (Springer, 2010).
2. M. Lapine *et al.*, *Rev. Mod. Phys.* **86**, 1093 (2014).
3. I. V. Shadrivov *et al.*, *Nonlinear, Tunable and Active Metamaterials* (Springer, 2015).
4. R. Lytel *et al.*, *Phys. Rev. A* **87**, 043824 (2013).
5. M. G. Kuzyk, *Phys. Rev. Lett.* **85**, 1218 (2000).
6. R. C. Miller, *Appl. Phys. Lett.* **5**, 17 (1964).
7. K. O'Brien *et al.*, *Nat. Mater.* **14**, 379 (2015).
8. M. Kauranen, A. V. Zayats, *Nat. Photonics* **6**, 737 (2012).
9. J. B. Khurgin, G. Sun, *Opt. Express* **21**, 27460 (2013).
10. L. Alloatti *et al.*, *Appl. Phys. Lett.* **107**, 121903 (2015).
11. S. Clemmen *et al.*, http://arxiv.org/abs/1508.05275 (2015).
12. J. Lee *et al.*, *Nature* **511**, 65 (2014).

ENERGY AND ENVIRONMENT

Understanding China's non–fossil energy targets

Methodology standardization will improve comparability

By **Joanna I. Lewis**[1,2]*, **David G. Fridley**[2], **Lynn K. Price**[2], **Hongyou Lu**[2,3], **John P. Romankiewicz**[3]

POLICY

More than 130 countries have targets for increasing their share of renewable or nonfossil energy (1). These shares and targets are often reported without clear articulation of which energy accounting method was used to convert nonfossil electricity into units that allow comparison with other energy sources (2–4). Three commonly used conversion methods are well documented by organizations dealing in energy statistics, but often, the method is not clearly stated when countries translate national targets into international pledges or when organizations track and compare targets across nations. China—the world's largest energy producer, energy consumer, and emitter of energy-related carbon dioxide (CO_2)—uses a distinct fourth method that is unique, not well documented in the literature, and not

[1]Science, Technology, and International Affairs Program, Edmund A. Walsh School of Foreign Service, Georgetown University, Washington, DC, USA. [2]China Energy Group, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [3]Energy and Resources Group, University of California, Berkeley, CA, USA.
*Corresponding author. E-mail: Joanna.Lewis@georgetown.edu

transparent in policy documents. A single, standardized, and transparent methodology for any targets that are pledged as part of an international agreement is essential.

More than 120 national pledges in the form of Intended Nationally Determined Contributions (INDCs) have been submitted in advance of the 21st Conference of the Parties (COP21) to the United Nations Framework Convention on Climate Change (UNFCCC) beginning 30 November. Many of the pledges include targets for increasing the share of renewable or nonfossil electricity in the national energy mix. With this "bottom-up" system of pledges likely to be status quo for the near future, the absence of common methodological guidelines could lead to confusion and inaccuracies as requirements to track and report progress expand.

CONVERSION METHODS. "Primary energy" refers to energy in natural resources, fossil and nonfossil, before conversion into other forms, such as electricity. For nonfossil sources, it is not useful to calculate the primary solar or kinetic energy, for example, before conversion to electricity; thus, these sources are expressed in terms of electricity generated ("primary electricity"). Three methods are used internationally to convert primary electricity into standardized units

Less reliance on coal. Shizuishan City installed photovoltaic panels in 2014 to augment the power supply and to reduce consumption of coal.

of primary energy such as joules or British thermal units (Btus) (table S1).

(i) direct equivalent: electricity is considered the primary energy form in all cases, with 1 kWh of noncombustible electricity or heat equal to 3.6 MJ of primary energy [used by the Intergovernmental Panel on Climate Change (IPCC)] (2); (ii) substitution: calculates efficiency for all electricity production as if it had been generated by a fossil fuel power plant with an average electricity conversion factor [used by the U.S. Energy Information Administration (EIA) of the U.S. Department of Energy (DOE), BP, and the World Energy Council] (5–7); (iii) physical energy content: uses physical energy content of the primary energy source used for electricity production [used by the Organization for Economic Cooperation and Development (OECD), the International Energy Agency (IEA) and Eurostat] (8). For nuclear and geothermal, heat is considered the primary energy form; for other primary electricity production (hydroelectric, solar photovoltaics, or wave or tide), electricity is considered the primary form of energy.

A fourth method, power plant coal consumption (PPCC), in which conversion to standard units is based on the average heat rate of coal-fired power plants in that year, is used only in China. As coal has long been the dominant source of energy in China, it is not surprising that coal power plants are used as the baseline for conversion. However, it is not clear how China calculates the average coal consumption figure, as the derived heat rate matches neither the generation heat rate nor the supply heat rate as published in the *China Electric Power Yearbook* (9). Because this methodology is not transparently reported, it is difficult to compare China's

targets, and progress toward these targets, with those of other countries.

CHINA'S TARGETS. In recent years, China has set a number of domestic energy and climate targets, many of which are reflected in international pledges. In late 2009, China announced a goal to reduce carbon intensity [CO_2 per unit of Gross Domestic Product (GDP)] 40 to 45% below 2005 levels by 2020 and to increase the share of nonfossil energy to 15% in 2020. These goals were included in China's UNFCCC Copenhagen Accord pledge in early 2010 (10). In November 2014, as part of the U.S.-China Joint Announcement on Climate Change, China stated its intention to achieve peak CO_2 emissions around 2030, making best efforts to peak early, and to increase the share of nonfossil fuels in primary energy consumption to around 20% by 2030 (11). These targets, along with a pledge to lower carbon intensity by 60 to 65% from 2005 levels, were included in China's INDC in June 2015 (12).

In order to compare China's PPCC method–derived values to others used internationally, we reproduced China's methodology using published China energy balances expressed both in physical units [tonnes (metric tons), m³, terawatt-hours (TWh)] and in standard energy terms [metric tons of coal equivalent (tce)]. [For full details of the analysis, see the supplementary materials (SM).] By using the same energy and electricity data from China's 2010 National Energy Balance Table (13) but applying different methods of calculating the standard energy equivalent of primary electricity, the share of nonfossil electricity in China varies from a low of 3.4% (direct equivalent) to a high of 9.2% (substitution

method). Using China's PPCC method results in an 8.4% share, whereas using the physical energy content method results in a 4.2% share.

China's INDC goal of a 20% share of nonfossil energy in total energy by 2030 was calculated using the PPCC method. If the direct-equivalent method had been used, the share would have been different. Using the forecast electricity generation composition of the "continued-effort" scenario of a recent modeling forecast (14), if one assumes an average China coal power plant heat rate of 0.3098 kgce/kWh in 2030, the share of nonfossil energy in total energy in 2030, when using the direct-equivalent method, is 9% compared with 20% when using the PPCC method (details in SM).

Although the PPCC method that China uses calculates a higher share of primary electricity in total energy than the direct-equivalent method, its use increases the difficulty of reaching China's stated energy intensity (energy per unit of GDP) reduction goal. Since the 2030 target of reducing primary energy intensity by 60 to 65% compared with the 2005 level is calculated as total energy consumption (in tce) divided by GDP in 2030, use of the PPCC method increases the amount of primary electricity in the numerator energy total by a multiple defined by the ratio of the PPCC method coefficient to the direct-equivalent method coefficient (see SM). In 2010, this ratio was 2.6, and China's total primary energy supply calculated by using the PPCC method was 5% larger than if calculated using the direct-equivalent method. As the proportion of primary electricity increases over time, this gap will widen and increase the difficulty of reducing overall energy intensity.

TRANSPARENCY IN INTERNATIONAL NEGOTIATIONS. Because of the prevalence of countries adopting non–fossil energy targets, consistency in the reporting of methodologies for calculating energy-related mitigation targets will be increasingly important. The issue of comparability is complex and can include not just technical comparisons, but also procedural and political ones. Recognizing these complications, and the diversity of nationally determined rationales for setting a target in a specific format, translating these targets by using a standardized international convention when pledged or committed internationally would serve as an important first step.

China has been consistent in its use of the

> "[A]...more transparent, and comparable, metric for... nonfossil electricity generation is...electricity units."

PPCC method for many years, but since it is not publicly documented and not simple to derive, China's non–fossil energy targets are not easily comparable with those of other countries. China does now report the direct-equivalent number in its national energy balance but has not used it for formulating energy and climate targets. Since China's national non–fossil energy target was pledged within the context of the UNFCCC, it would be easier to understand if China followed UN and IPCC convention and reported on the basis of the direct-equivalent method (2). The same is true for any other countries submitting non–fossil share pledges to the UNFCCC to allow for comparability across all of the other UNFCCC parties.

An alternative, more-transparent, and comparable metric for understanding a country's share of nonfossil electricity generation is to report directly in electricity units. This avoids the need to convert from electricity to energy and, instead, clarifies exactly how much new nonfossil electricity will be generated to meet the target. For example, China's share of nonfossil electricity generation in total generation was 21% (842 TWh nonfossil of a total generation of 3937 TWh) in 2010. The continued-effort scenario referenced above results in a 39% share (4071 TWh nonfossil of a total generation of 10,490 TWh) of nonfossil electricity generation in 2030. So this near-doubling in the share of nonfossil generation between 2010 and 2030 represents an absolute increase of 3229 TWh between 2010 and 2030. Achieving such a level of nonfossil generation would require ~900 GW of new nonfossil power capacity to be installed in China between 2015 and 2030

(15), the actual amount dependent on the capacity factors of the mix of generation technologies. For context, the entire fossil and non–fossil electricity–generating capacity of the United States in 2015 is 1009 GW (16).

If reaching an agreement on guidelines for use of a standard methodology in the reporting of INDCs is not possible, then promulgating requirements for increased transparency to facilitate comparison and analysis would greatly improve understanding of different targets. Transparency has long been on the agenda in the UNFCCC negotiations, but it remains a complex and often politically sensitive topic. The issue raised in this paper is a tangible illustration of the need for methodological clarity, one aspect of the broader transparency discussion that is more technical than political. Addressing methodological transparency in the UNFCCC negotiations would have a positive impact on our ability to understand national pledges made at COP21 and beyond. ∎

REFERENCES AND NOTES

1. REN21, *Renewables Global Status Report 2015* (REN21 Secretariat, Paris, 2015).
2. W. P. Moomaw *et al.*, in *IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation*, O. Edenhofer *et al.*, Eds. (IPCC, New York, 2011), Annex II, Methodology.
3. H. D. Lightfoot, *Energy* **32**, 1478 (2007).
4. E. Martinot, C. Dienst, L. Weiliang, C. Qimin, *Annu. Rev. Environ. Resour.* **32**, 205 (2007).
5. EIA, DOE, in *Annual Energy Outlook 2011* (EIA, DOE, Washington, DC, 2012), table A6.
6. BP, *Statistical Review of World Energy* (BP, London, 2014).
7. World Energy Council, *World Energy Scenarios: Composing Energy Futures to 2050* (WEC, London, 2013).
8. OECD, IEA, Eurostat, *Energy Statistics Manual* (OECD, IEA, Eurostat, Paris and Luxembourg, 2004).
9. Editorial Board of the China Electric Power Yearbook, *China Electric Power Yearbook 2011* (Beijing: China Electric Power Press, 2011).
10. W. Su, "Letter including autonomous domestic mitigation actions," in *Copenhagen Accord* (UNFCCC Secretariat, Bonn, 2010), Appendix II.
11. The White House, "U.S.-China Joint Announcement on Climate Change, 11 November 2014" (White House, Washington, DC, 2014).
12. Department of Climate Change, National Development and Reform Commission of China, *Enhanced Actions on Climate Change: China's Intended Nationally Determined Contributions* (National Development and Reform Commission of China, Beijing, 2015).
13. National Bureau of Statistics, *China Energy Statistical Yearbook 2011* (China Statistics Press, Beijing, 2011).
14. X. Zhang, V. J. Karplus, T. Qi, D. Zhang, J. He, "Carbon emissions in China: How far can new efforts bend the curve?" (Report no. 267, MIT Joint Program on the Science and Policy of Global Change, Tsinghua-MIT China Energy and Climate Project, Cambridge, MA, 2014).
15. S. Fu, J. Zou, *Comments on China's INDC* [China National Center for Climate Change Strategy and International Cooperation, Beijing, 2015].
16. EIA, U.S. Department of Energy, "Annual energy outlook" (DOE, Washington, DC, 2015).

Learning from Africa's herbivores

Herbivore diversity plays a key role in grassland ecosystems

By **Jacquelyn L. Gill**

E arth's animals are downsizing. Since the end of the last ice age about 12,000 years ago, the largest animals on the planet have been hit disproportionately hard by what may have been the beginnings of the sixth mass extinction (1). We are only just beginning to appreciate the ecological impacts of this "trophic downgrading" (2): Both modern and paleoecological analyses are providing growing evidence that the extinction of Earth's largest animals has cascading ecological impacts across the globe (3). On page 1056 of this issue, Hempson *et al.* (4) provide a new tool for elucidating the ecological role of large herbivores at continental scales.

Finding the signal in the ecological noise is tricky when trying to understand large-scale patterns in Earth's biota. However, species often share characteristic traits, behaviors, or environments that allow generalization. Thus, grasslands in the Great Plains of North America and the Serengeti of Africa have functional similarities, even though they are made of different species. Such "syndromes" (a collection of features that occur together, as in medicine) have been long recognized for plants but have only recently been classified for global fire regimes, or pyromes (5). Surprisingly, knowing a region's biome or climate does not necessarily mean you can predict its fire regime. This finding was only made possible by taking a big-picture perspective of global pyromes.

A syndrome-based framework can be a useful abstraction, allowing biologists to model ecological and evolutionary relationships at large scales. This is especially important when species-level data are lacking or computationally impractical to model. Hempson *et al.* now define a new syndrome: the herbivome. Like fire, large herbivores are a key player in the distribution and function

School of Biology and Ecology, Climate Change Institute, University of Maine, Orono, ME 04469, USA. E-mail: jacquelyn. gill@maine.edu

of plant species, but, thus far, a generalizable conceptual framework to understand this has been lacking. By collapsing the 92 extant large animal species of sub-Saharan Africa into five distinct syndromes based on diet characteristics, behavior, and body size, the authors gain insights into large herbivore ecology. The approach also provides a tool for testing herbivore effects with a framework that can be scaled to many systems, past and present.

Herbivomes are constructed based on animal diets, behavior, and traits, as well as the relationship between animal biomass and environmental characteristics such as soil properties and precipitation levels. For example, the high variety and abundance of larger species (VALS) herbivome is defined by large species such as elephants and giraffes, which have unique relationships with the savannas they inhabit. Hempson *et al.* used the herbivomes to reconstruct the nature and intensity of herbivory pressure in Sub-Saharan Africa 1000 years ago, before widespread human agriculture in the region. This approach provides insights into African herbivore ecology that are difficult to obtain from field studies alone.

Elephants stand out in the analysis due to their large body size and disproportionate contribution to African mammalian biomass. This hints at the utility of the herbivome framework in quantifying the effects of prehistoric extinctions, because the largest members of Earth's ecosystems often play unique roles that cannot be replicated by the medium and small-bodied survivors (*6*). Also sobering is the realization that livestock, for which global data are much more readily available, are not good analogs for their wild counterparts. For example, millennia of artificial selection separate North American cow (*Bos*) and bison (*Bison*) in terms not only of their traits but also their ecological impacts. Through selective grazing and wallowing, bison create a mosaic of disturbance that supports a greater diversity of plants and animals compared with cattle-grazed regions. Thus, although bison and cattle can interbreed, they are not ecological equivalents (*7*). This difference—a difference of herbivomes— has implications for grassland restoration and management (*7*): Large, native herbivores can be a critical tool in conservation.

Hempsen *et al.*'s study joins a growing body of literature on large herbivore ecology, motivated by the recognition that the loss of Earth's fauna has widespread consequences (*8*). African herbivores survived the first wave of extinctions after the ice age, but their situation today is precarious; in central Africa, the regional elephant population has declined 64% in a decade (*9*). As we seek to protect Earth's endangered biota in the Anthropocene, biologists are increasingly looking to the Pleistocene (*10*) for insights into restoring lost ecosystem function through rewilding, or to predict the ecological consequences of species losses (*2*).

The grasslands of the Americas once supported even greater megafaunal biodiversity than the Serengeti does today (see the figure) (*11*), but we know little about the long-term impacts of those losses on modern biota. The elegance of the herbivome concept is its generalizability; it can be widely applied to other systems, past and present. A cross-continental analysis of herbivomes—one populated by the ghosts of herbivores past, the other the last refuge of Pleistocene megafauna—may provide important insights into the ecological causes and consequences of extinctions. It's too late for the mammoths and woolly rhinos, but not their Serengeti cousins. ∎

Anthropocene Serengeti
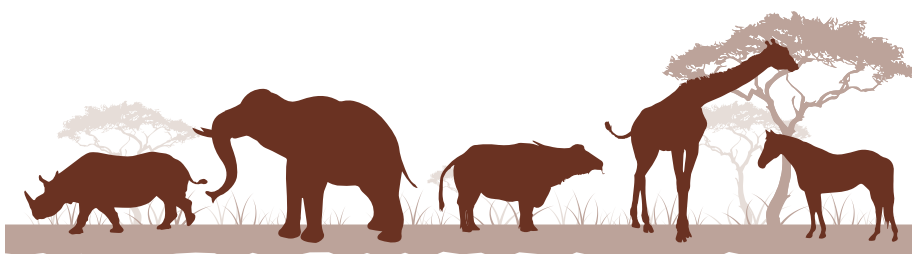
Pleistocene Great Plains

Low-diversity rangeland

Benefits of herbivore diversity. (**Top**) The Serengeti in Africa is one of Earth's most diverse communities of herbivores, but these species are under threat from poaching, disease, land-use change, and climate change. Without keystone species such as elephants and rhinos, the ecosystem would undergo dramatic shifts. (**Middle**) For millions of years, the animal diversity of the North American Great Plains rivaled that of the Serengeti. However, by 10,000 years ago, only bison and a few smaller herbivores (such as deer and pronghorn) remained. (**Bottom**) In much of the world, productive grasslands have been converted to rangelands for grazing cattle. The impacts of domestic herbivores on biodiversity are very different from those of wild herbivores and are often detrimental.

ILLUSTRATION: C. BICKEL/*SCIENCE*

REFERENCES
1. A. D. Barnosky *et al.*, *Nature* **471**, 51 (2011).
2. J. A. Estes *et al.*, *Science* **333**, 301 (2011).
3. E. S. Bakker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1502545112 (2015).
4. G. P. Hempson, S. Archibald, W. J. Bond, *Science* **350**, 1056 (2015).
5. S. Archibald, C. E. R. Lehmann, J. L. Gómez-Dans, R. A. Bradstock, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6442 (2013).
6. R. H. Peters, *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, UK, 1986).
7. A. K. Knapp *et al.*, *BioScience* **49**, 39 (1999).
8. R. Dirzo *et al.*, *Science* **345**, 401 (2014).
9. G. Wittemyer *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 13117 (2014).
10. J.-C. Svenning *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 10.1073/pnas.1502556112 (2015).
11. S. Faurby, J.-C. Svenning, *Divers. Distrib.* **21**, 1155 (2015).

EVOLUTION

How Victoria's fishes were knocked from their perch

Evolutionary innovations are not always beneficial

By **Geerat Vermeij**

Intuition can be a powerful force in science, but more often proves to be an unreliable guide to reality. This point is beautifully brought home in the report by McGee *et al.* (*1*) on page 1077 of this issue. The study concerns evolutionary innovations that enable members of an evolving lineage to exploit resources in ways inaccessible to their ancestors. Most biologists view such innovations as opening the evolutionary door to species proliferation and the longer persistence of lineages with the novel traits. This expectation is often realized. But, as McGee *et al.* show in their study of cichlid fishes in East Africa's Lake Victoria, this is not always the case.

Like several marine fishes, most cichlids break down food such as hard-shelled prey with a mill of pharyngeal bones in the throat. Compared to fishes without this capacity, cichlids have a much smaller mouth gape, limiting the size of food items that can be swallowed (see the photo). Although the innovation of pharyngeal mastication frees up the oral teeth to take on novel functions (*2, 3*), it constrains the evolutionary option of feeding on larger prey fishes.

This limitation proved disastrous for fish-eating cichlids in Lake Victoria when the voracious Nile perch (*Lates niloticus*), a species without pharyngeal mastication, was introduced into the lake in the 1950s. Intuition would suggest that the extinction of many of the lake's cichlids resulted from predation by the Nile perch. Instead, McGee *et al.* show that it was competition for fish prey that doomed the cichlids. The study also lends indirect support to the hypothesis (*4*) that the extraordinary proliferation of cichlid species in Lake Victoria had less to do with the pharyngeal innovation than with rampant sexual selection associated with the imperative to find and choose mates.

Besides unlocking Lake Victoria's secret, the results illustrate the point (*5*) that the short-term functional advantages of a novel trait do not always translate into a lineage's long-term survival or diversification. Adaptation by natural selection contributes to a good fit between an organism and its current environment (*6, 7*). What it cannot do is anticipate the future. An evolutionary innovation can promote a lineage's persistence and its potential to proliferate only if it confers adaptability, opening up new ways of life in unprecedented circumstances. It will be interesting to ascertain how often functional innovations that benefit individuals raise or lower a lineage's susceptibility to extinction.

The functional trade-off between mouth gape and throat mastication is the Achilles heel of the cichlid innovation, because it prevents lineages from attaining the status of apex predator, the pinnacle of competitiveness in animals. This finding prompts the question of how widespread such competition-limiting innovations are. Although the energetics of the pharyngeal innovation remain unknown, the cichlid trade-off implies that the allocation of resources to competing functions is constrained by a fixed metabolic rate in evolving cichlid lineages. Other innovations—such as endothermy (warm-bloodedness) in birds and mammals, high leaf-vein density in flowering plants, and complex social structure in insects and primates—permit a higher metabolic rate, enabling some species to rise to the top of the competitive heap (*8*). Such evolutionary breakthroughs can be expected in large, resource-rich settings such as rainforests, savannas, coral reefs, and waters where plankton is abundant.

In contrast, innovations arising under a fixed energy constraint are likely to characterize lineages in low-productivity settings, such as deserts, the deep sea, and isolated island-like settings. The latter include Lake Victoria before the Nile perch arrived. Even in productive, highly diverse environments, such innovations can be successful for organisms in subordinate roles, such as small herbivores.

A comparative approach will be needed to investigate these matters further. This should involve not only an understanding of how innovations come about through genetic and developmental changes, but also documenting historical patterns and probing aspects of novel resource allocation in relation to competition.

REFERENCES

1. M. D. McGee *et al.*, *Science* **350**, 1077 (2015).
2. K. F. Liem, *Syst. Zool.* **22**, 425 (1973).
3. C. D. Hulsey, F. J. García de León, R. Rodiles-Hernández, *Evolution* **60**, 2096 (2006).
4. F. Galis, J. A. J. Metz, *Trends Ecol. Evol.* **13**, 1 (1998).
5. L. Van Valen, *Evolution* **29**, 87 (1975).
6. W. Sterrer, *Evol. Cogn.* **1**, 101 (1992).
7. G. J. Vermeij, *The Evolutionary World: How Adaptation Explains Everything from Seashells to Civlization* (Thomas Dunne/St. Martin's, New York, 2010).
8. G. J. Vermeij, *Annu. Rev. Earth Planet. Sci.* **41**, 1 (2013).

Pundamilia nyererei, a cichlid from Lake Victoria.

Dept. of Earth and Planetary Sciences, University of California at Davis, Davis, CA 95616, USA. E-mail: gjvermeij@ucdavis.edu

Alan Hall (1952–2015)

The loss of a leading cell and cancer biologist

By **Catherine Nobes,**[1] **Alison Lloyd,**[2] **Mark Marsh**[2]

Alan Hall, a remarkable cell and cancer biologist, died suddenly on 3 May in New York City. He was an outstanding researcher, teacher, and colleague. In the 1980s, in his early career, Alan was one of a small group of molecular biologists who first revealed how genetic changes could cause cancer. Alan and his colleagues undertook pioneering work that revealed the mechanisms through which the Rho family of small guanosine triphosphatases (GTPases) regulate the cytoskeleton and thus how cells control their shapes and movement. He became one of the world's leading cell biologists and was a committed mentor to generations of young scientists.

Alan was from Barnsley, UK. He studied chemistry at Oxford University and earned a doctorate in biochemistry at Harvard University before taking up molecular biology as a postdoctoral fellow at the University of Edinburgh and the University of Zurich. In 1981, Alan joined Robin Weiss's group at the Institute of Cancer (ICR) in London. There he had the good fortune to meet Christopher Marshall, who became a life-long friend and collaborator. (Sadly, Chris died 3 months after Alan.) Together, Alan and Chris identified the oncogene *N-Ras*, a discovery that set them both on the trail of understanding the biology of Ras and the closely related Rho GTPases. Whereas some were disappointed at the discovery of another *Ras* gene, Alan's view was that finding a third family member indicated that Ras proteins must be worth studying and the closely related Rho proteins were likely to be just as important. This led Alan, together with Anne Ridley and Hugh Patterson, to discover the key roles of Rho, Rac, and Cdc42 as receptor-coupled molecular switches that control actin assembly and cell motility.

Over the years, Alan's reputation for original and important research enabled him to recruit outstanding postdoctoral fellows and graduate students (a scientific family tree can be found at www.ucl.ac.uk/lmcb/alan-hall-scientific-family-tree) who extended the work on Rho, Rac, and Cdc42 and identified many of the effectors through which these GTPases control a plethora of cell functions, including cell motility, wound healing, and cell polarity.

Alan's scientific achievements were widely recognized and acknowledged. He was a Fellow of the Royal Society and the Academy of Medical Sciences, a member of the European Molecular Biology Organization, and the winner of numerous prizes, including the Feldberg Foundation Prize, the Novartis Medal, the Gairdner Foundation International Award, and the Louis-Jeantet Prize for Medicine. He was also the Editor-in-Chief of the *Journal of Cell Biology* when he died.

Recently, a meeting was held at the Royal Society in London to celebrate Alan's work and scientific achievements. The event was attended by many of his colleagues and collaborators, and those who spoke not only described the science that Alan pioneered, inspired, and contributed to, but also highlighted the traits that made him such a successful and popular scientist. Alan was quietly confident; if he had an ego, you would not have known it. He was always positive, and this rubbed off on colleagues. Many recalled Alan's scientific rigor, integrity, and generos-

ity with his time, as well as his enthusiasm for science, which was for all the right reasons—he simply wanted to know the answer. He could spot something interesting, ask the key questions, and develop testable hypotheses. His experimental approaches were clean, insightful, and uncomplicated, and his sharing of reliable reagents led to many collaborations worldwide, which probably drove the rapid development of the Rho field in the early days. His special relationship with Chris Marshall was talked about frequently; their characters could not have been more different, but the combination worked perfectly for driving forward the Ras and Rho signaling fields over many years.

In 1993, Alan left the ICR and moved to the newly formed Laboratory for Molecular Cell Biology (LMCB) at University College London (UCL). The LMCB, a joint venture between King's College London and UCL and funded by the Medical Research Council (MRC), was the first UK research institute to focus on molecular cell biology. Together with the first director, Colin Hopkins, Martin Raff, and a group of young principal investigators, Alan created LMCB's open, interactive, and productive research environment. In particular, the new institute provided an opportunity to change graduate student training through an innovative 4-year graduate program that eventually became the prototype for many similar programs in the United Kingdom.

In 2000, Alan became the second director of the LMCB and oversaw the establishment of the MRC Cell Biology Unit at the core of the institute, ensuring ongoing MRC support and its continuing success. Alan was an exceptional director, leading by example through the excellence of his science, strong moral code, and mentorship skills. Although Alan left London in 2006 to take up the Chair of Cell Biology at Memorial Sloan Kettering Cancer Center in New York, he continued to mentor many at the LMCB and the United Kingdom more widely.

Alan was a proud Yorkshireman and would let you know it, even though he had lost his northern accent. He had a wonderful sense of humor and loved an opportunity to celebrate—he often had good reason to enjoy his own successes and those of his colleagues. He is sorely missed, but his legacy is assured and his groundbreaking accomplishments stand as a testament to his scientific career.

[1]*School of Biochemistry, University of Bristol, University Walk, Bristol BS8 1TD, UK.* [2]*MRC-Laboratory for Molecular Cell Biology, University College London, Gower Street, London WC1E 6BT, UK. E-mail: m.marsh@ucl.ac.uk.*

RETROSPECTIVE

Christopher Marshall (1949–2015)

A cell biologist's meticulous work drove the discovery of new cancer treatments

By **Richard Marais**

Christopher Marshall was a titan of cancer research. On 8 August he died of the very disease that he strived to understand. This was only 3 months after the unexpected death of his longtime friend and colleague, Alan Hall.

Chris was born in Coventry, UK, and studied natural sciences at the University of Cambridge. He went on to get his doctorate in cell biology at the University of Oxford, followed by postdoctoral work at the Imperial Cancer Research Fund in London and the Dana-Farber Cancer Institute in Boston. In the early 1980s, Chris Marshall and Alan Hall cloned *N-RAS*, the third member of the *RAS* family of human oncogenes. Over the next 35 years, Chris would contribute an enormous amount to our knowledge of the biology of this protein and the signaling pathways that it controls. Robin Weiss, the then newly appointed chief executive of the Institute of Cancer Research, London, recruited Chris in 1980 and within a few months also recruited Alan, establishing one of the most productive collaborative partnerships of our time. Chris and Alan joined the hunt for human oncogenes, and the cloning of *N-RAS* set them both on a course of studying cell signaling for the rest of their remarkable careers.

The realization that there are three human *RAS* oncogenes focused intense interest on these proteins, and within a few years an entire field had emerged to study them. Progress was rapid and Chris and Alan established themselves as major contributors to this new field. It was confirmed that cloned mutant genes could transform cells, and the endogenous genes were shown to be mutated in human cancer, confirming that the RAS proteins were oncogenes and likely to be drivers of cancer.

Chris was driven to understand how oncogenes cause cancer and rigorously dissected how RAS proteins transmit signals to the nucleus in the normal and cancerous state. In

the late 1980s and early 1990s, Chris and his newly formed "Oncogene Team" played a key role in elucidating the biochemical pathways that modify the C terminus of the individual RAS isoforms to take them to the plasma membrane. When it became clear that the signals that activated RAS also activated a newly discovered enzyme called extracellular signal–regulated kinase (ERK), Chris—with characteristic enthusiasm and drive—and his collaborators demonstrated that RAS and a
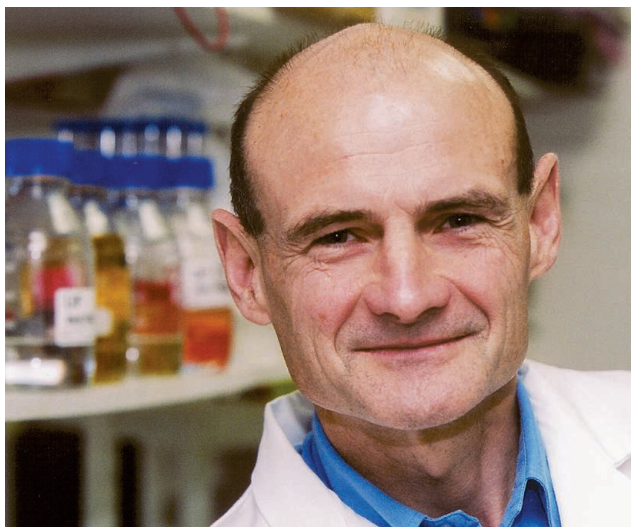
protein kinase called CRAF were necessary and sufficient to activate ERK. They mapped sites on MEK, also a protein kinase, that were phosphorylated by CRAF. This meticulous and pioneering work from Chris, together with biochemical and genetic data from groups working on worms and flies, elucidated the RAS signaling pathway that drives cell responses, showing that this cascade of events could mediate both proliferation and differentiation of cells, cellular decisions that were apparently mutually exclusive.

At the turn of the millennium, Chris contributed to the discovery of the BRAF oncogene. His work implicated a RAS-BRAF pathway in angiogenesis. Moreover, he established not only that cancer cells invade tissues by two distinct modes—amoeboid and elongated migration—but also that cells can switch between these modes in response to the tumor microenvironment and drug treatments. For a decade, Chris played a major role in elucidating the biochemical pathways underlying these distinct forms of migration

and how the cells switch between them.

The RAS signaling pathway presented many potential targets for cancer drug development, and Chris believed in translating basic research into therapeutics. He was appointed the director of research at the Institute of Cancer Research in London for several years. Chris received many awards throughout his career. He was a Fellow of the Royal Society, a member of the European Molecular Biology Organization, and a founding member of the Academy of Medical Sciences. Among his numerous awards were the Novartis Medal, the Buchanan Medal of the Royal Society, and the Biochemical Society Centenary Award, which he received this year.

Those who were privileged to know and work with Chris remember an enormous intellect who could, with apparent ease, see the new discovery or critical link among a great jumble of data. He taught his students and postdocs how to think scientifically and how to execute experiments with clarity. He would scold us when he saw room for improvement, sometimes even when we didn't deserve it, but when we were fragile he picked us up and tended to our wounds. Chris was supportive and protective and made sure to equip members of his lab with the tools needed to succeed. He encouraged us to work hard and to commit to discovering the answer, but most important of all, he made it fun and reminded us not to lose sight of the important things—our life and families outside of the laboratory. I had never seen Chris more excited than when he was about to welcome a new grandchild into his family.

Friends and collaborators are too many to mention, but Chris touched many of us in our professional lives and in our lives beyond science. He was a fantastic scientist, an excellent colleague, a very generous person, and a much-loved friend.

The loss of Chris Marshall and Alan Hall within a few months of each other has robbed the field of two of its giants. We will miss them, and we will miss their scientific rigor and insights.

CRUK Manchester Institute, University of Manchester, Manchester M20 4BX, UK, .E-mail: richard.marais@cruk.manchester.ac.uk

PHOTO: THE INSTITUTE OF CANCER RESEARCH/LONDON

BOOKS *et al.*

EXHIBITION

Humanity 2.0

A new exhibition explores how scientific advances are changing what it means to be human

By **Giovanni Frazzetto**

Complex and unpredictable, technological change invariably shakes our condition. Every new tool or discovery belittles past ambitions and sets renewed destinations. But how are we to keep up with the speed, direction, and magnitude of our own development?

HUMAN+, a new, ambitious, and sharply curated exhibition at the Centre for Contemporary Culture in Barcelona presented in collaboration with the Science Gallery Dublin, delves into the contemporary meaning of being human. It does so with a curiosity for the course our future might take in light of continuing advances in genetic and biomedical research technologies, as well as robotics, engineering, and artificial intelligence.

Hosting more than 50 works created by artists and scientists, the collection explores the strategies that we might use to transcend bodily and mental limits, our place in nature, and our social interactions, as well as redefinitions of birth and death. Although its title is suggestive of an optimistic upgrading for our species, the show is equally wary of taking the benefits of empowerment for granted. Bounc-

The reviewer is the author of Joy, Guilt, Anger, Love *(Penguin, 2014) and is a visiting research fellow at Trinity College, Dublin, Ireland. E-mail: gio@giovannifrazzetto.com*

ing between the concrete and the utopian, HUMAN+ invites visitors to contemplate their own future and measure the promise and peril of technological emancipation.

In Agatha Haines's stirring installation *Transfigurations*, a row of breathing baby dolls display surgical modifications that are intended to expand their fitness for a variety of biomedical, social, or environmental purposes. For instance, one child has been endowed with extra folds of skin on its scalp to facilitate the dissipation of heat, an asset in times of global warming.

What would it feel like if we could see the world through someone else's eyes? Based on theories of embodied cognition, virtual reality, and mirror neurons, *The Machine to Be Another*, created by the Be Another Lab artistic collective, is exhilarating and one of the most interactive pieces in the show. With the help of immersive goggles, pairs of visitors experience each other's view and gradually learn to anticipate and synchro-

In *Transfigurations*, Agatha Haines imagines a future in which infants are surgically endowed with unusual but advantageous physical features.

nize their movements and intentions. At once pragmatic and visionary, the experiment is deft at enabling empathy and has the potential to address themes such as kinship, gender, and ethnic identity, as well as conflict resolution at the individual and collective level.

By stimulating reward areas in the brain and inspiring proximity, direct eye-to-eye gaze is crucial in modulating social interactions. Named *Area V5*, in reference to the portion of the human cortex that is involved in the perception of motion, an installation by Louis-Philippe Demers probes our interaction with machines. Occupying an entire wall, a mesmerizing collection of robotic eyes roll frenetically as visitors pass by. With each shift in gaze comes the chance of tapping into the "Uncanny Valley," the feeling of discomfort that is elicited by strikingly anthropomorphic semblances.

What if there were ways to bring sophistication to the process of dying? In an enticing twist to the much debated and largely unresolved dilemma of euthanasia, Julijonas Urbonas's slick and morbid *Euthanasia Coaster* is a hypothetical machine that, in the words of the artist, would make people reach death "humanely—with elegance and euphoria." Resembling an amusement park roller coaster, the apparatus consists of an almost 2000-foot-tall tower followed by seven loops of decreasing diameter. Arousing feelings of elation and excitement, the descent from its top would bring the train to a speed of 220 mph, and the successive inversions would impose a deadly 10-g force, causing passengers to experience severe loss of oxygen, tunnel vision, and, ultimately, death.

At the entrance to the exhibition, a 17th-century painting by the Flemish artist Jacob Peter Gowy depicting the myth of Icarus acts as an iconic warning against scientific hubris. Icarus, you may recall, escaped imprisonment thanks to a pair of wings made out of feathers and wax. The wings were built by his father, Daedalus, who cautioned him not to fly too high, lest the sun melt the wax. Euphoric from the thrill of flight, Icarus ignores the admonishment, soars high, and then falls to his death.

Although humans will always push the limits of discovery, each advance warrants collective reflection about its societal, ethical, and cultural effects. Challenging in both its aesthetics and its content, HUMAN+ succeeds as a tinder for an engaging, open, and stimulating dialogue in that direction.

BIG DATA

Intimate details

A long-forgotten social science archive offers a lesson in responsible data management

By **Laura Stark**

T he U.S. Library of Congress holds a long-forgotten collection of intimate data, the fate of which Rebecca Lemov describes as "a parable for our time." Lemov's humane, hilarious, and smart new book, *Database of Dreams*, recovers the story of the enormous effort and expense that went into creating this analog database during the mid-20th century—only to have it fade from personal and institutional memory within decades. The book shows that, although some things are forgotten because they are unimportant, others lose importance because they are forgotten.

The database was the brainchild of Bert Kaplan, a Harvard Ph.D. and rising star in 1950s social science. His idea was to create an infrastructure to pool, preserve, and share "endangered data": the raw, primary documents that individual social scientists had collected for their own research projects and were apt to throw away in the days when filing cabinets were state-of-the-art data storage systems. Big funders bankrolled the project, and at its peak, upwards of 100 research institutions worldwide had bought the data set.

When the project started in the 1950s, it was literally a database full of dreams, accounts of the nighttime figments inside people's heads. Kaplan was among a cadre of anthropologists and psychologists who ventured out from universities to collect "fleeting thoughts, random asides, irreverent inquiries and sad memories, life stories and dreams" from people in the flux of modernization. Their aim was both patronizing and ambitious: to collect ephemeral data from communities that were likewise fading from modern life—living relics of vanishing worlds.

The database would eventually expand to include life histories and the results of Rorschach and other projective tests from members of a hodgepodge of communities around the world, including small-town Midwesterners, resettled American Indians, and natives of the Philippine Islands, to

The reviewer is at the Center for Medicine, Health, and Society, Vanderbilt University, Nashville, TN 37235, USA. E-mail: laura.stark@vanderbilt.edu

name a few. It can be hard to appreciate the authority once accorded to data that today seem the mere quirk and color of history. Yet, projective tests were taken so seriously in the 1940s and 1950s that scientists used them to access the psychic depths of Nazi war criminals.

The archive was also a "database of dreams" in a second sense. For social scientists, it represented the aspiration to catalog the human condition. Such was the enthusiasm for the project that in only 7 years, 60 researchers freely donated their data to be included in the archive.

Within decades, the database faded from use and memory, in part because of Kaplan's unlucky choice of platform: the Microcard. But there was more than platform at play, and to her credit, Lemov avoids the simplistic view that the commercial failure of the Microcard explains everything.

The great virtue, and fatal flaw, of the database was that it stored raw, uninterpreted intimate material that users could then enroll in the service of their own claims. In 1964, a *Newsweek* journalist did just this, imposing his own, politically provocative interpretation on data Kaplan had collected. The debacle enraged the communities

> *"The book shows that, although some things are forgotten because they are unimportant, others lose importance because they are forgotten."*
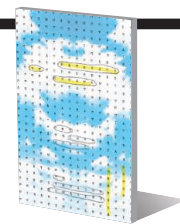
Kaplan studied, turned researchers against him, and distressed Kaplan himself. This event alone did not bring the demise of the database (although it certainly altered Kaplan's career). Yet, it was emblematic of the pervasive worry that users might fail to exercise restraint and good judgment with the database.

Today, many scholars are concerned about big data and the consequences of "biocapital," the new financial assets made possible through bioinformatics. Lemov, however, is interested in the responsibilities attached to intimate data in a world in which the long-term effects of personal revelations are impossible to predict.

The fundamental question that Lemov pursues is how the database of dreams can come to seem like a thing of the past. After all, the database still exists in the here and now: in the Library of Congress, in research libraries across the world, and even on Lemov's hard drive. It is the question of how to respond to the paradox of intimate data and is, Lemov writes, "a symptom of our own future buried in the present."

10.1126/science.aad2423

POPULAR SCIENCE
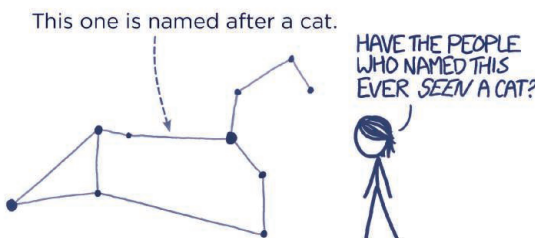
In *Thing Explainer*, xkcd creator Randall Munroe sets out to demystify a wide range of complex systems and natural phenomena using only the thousand most common words in the English language. The book features simple blueprints annotated with short descriptions that explain everything from the layout of the "shared space house" (international space station) to the causes of "great circle storms" (hurricanes) to the science behind the "food-heating radio box" (microwave). Munroe's signature humor and firm grasp on the underlying science and engineering make the book a delightful and informative read.

10.1126/science.aad7700

LETTERS

Edited by **Jennifer Sills**

Brazilian aquatic biodiversity in peril

THE ONGOING FISCAL and political crisis in Brazil has already caused deep cuts to science and education ("Fiscal crisis has Brazilian scientists scrambling," H. Escobar, In Depth, 28 August, p. 909), and now the environment is in the crosshairs. On 2 October, the Federal Government announced the closing of eight executive cabinets, including the Ministry of Fisheries, which was absorbed by the Ministry of Agriculture. Shortly afterward, on 9 October, a joint act of the Ministries of Agriculture and the Environment suspended the seasonal fishing closures that protect the spawning fish of several commercially and ecologically important species (*1*).

Without spawning closures, fishing is now allowed during the annual spawning migration of all freshwater species from six major hydrographic basins, including the Amazon, as well as during the reproductive season of several marine fishes and invertebrates. This decision was made in an effort to cut costs; an unknown percentage of approximately one million fishers were illegally receiving unemployment benefits during the closure season (*2*). Regional fisheries management councils, whose immediate future is now uncertain given the current political turmoil, will review the rules for fisheries closures, and the Ministry of Agriculture will reissue benefits for fishers by 9 February 2016.

Brazilian aquatic biodiversity will suffer from more than the suspension of spawning season closures. The fast-growing industrial lobby and misguided government actions have recently overthrown the Brazilian Red List of threatened species (*3*). Hundreds of hydroelectric dams are under construction or planned in several major river basins, severely threatening migratory fish species and the South American biota as a whole (*4*). Because of a series of controversial Acts, boats up to 20 tons are now considered artisanal vessels, and specific labor benefits for true traditional fishers have been cut, increasing their historical marginalization (*5*). On 15 October, high-level government officials and industrial fleet syndicate representatives were arrested for involvement in a bribery scheme to grant industrial fishing permits to companies that were fishing illegally, resulting in an estimated environmental loss of about US$365 million (*6*). Because fishers will no longer receive unemployment benefits, a growth on fishing efforts to maintain their income is predicted, hastening the demise of stocks that are already at the brink of collapse. The effect on the environment will be difficult to track, given the complete absence of fisheries statistics in Brazil (*3*).

Brazil is the first signatory of the Convention on Biological Diversity and allegedly endorses both the Food and Agriculture Organization's Code of Conduct for Responsible Fisheries and the Voluntary Guidelines for Securing Sustainable Small-Scale Fisheries in the Context of Food Security and Poverty Eradication. Instead of loosening the legal fisheries management framework to ameliorate a never-ending fiscal crisis, Brazil needs to improve fisheries and environmental management effectiveness. Baseline measures include the reestablishment of spawning closures, fisheries statistics, independent stock assessments, on-board scientific observer programs,

---

LIFE IN SCIENCE

Waste not, want not: Recycled science art

The trash can in my well-equipped Yale laboratory was overflowing with used gloves, tubes, and plastic tip boxes. As I neatly folded my own gloves and placed them in my drawer, I thought about how my experiences in my home of Jordan gave me a different perspective than that of many of my lab mates. Jordan's growing population, undergoing rapid modernization, generates an overwhelming volume of solid waste, and municipalities cannot afford modern solid waste collection, recycling, or successful landfill management. Solid waste mismanagement leads to public health risks. Meanwhile, purchasing lab consumables and chemicals is complicated by terrible bureaucracy in procurement and management of grants (*1*).

In Jordan, we would fight over a box of gloves.

I dreamed about taking all the waste home to Jordan to reuse, but I couldn't possibly fit it all in my luggage. Still, I couldn't bear to throw my materials away. One day, while staring at a pile of crinkled gloves, an idea occurred to me. Each glove looked like a cell. Together, they could form a picture of a stem cell colony.

Inspired to create a piece of recycled science art, I started a campaign to collect used gloves, tip boxes, and tubes (those that were not a biohazard, of course) all across the lab. The other members of my lab started collecting as well. The night before I left Yale to return to Jordan, my daughter and I spent all night creating our masterpiece, depicting a stem cell colony of mouse embryonic fibroblasts. We presented it to my lab group as a gift before leaving. Everyone loved it! One lab member even added a purple glove to represent a mycoplasma contaminating the colony.
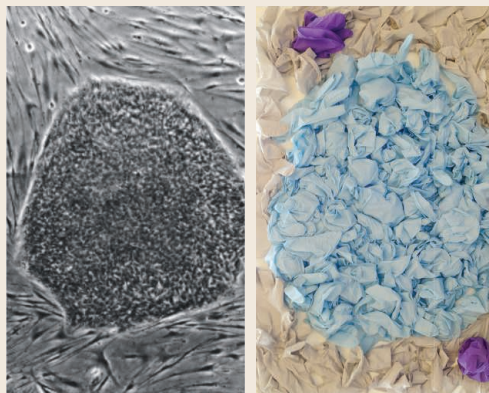
Once I returned to my country, I found a way to continue my artistic recycling efforts. When I teach cell biology to undergraduate students, I require them to create a piece of art from recycled material that represents a scientific concept. Perhaps we can help alleviate the growing solid waste problems in Jordan if we all embrace the idea that one scientist's trash can be another scientist's treasure.

Rana Dajani

Department of Biology and Biotechnology, Hashemite University, Zarqa, Jordan. E-mail: rdajani@hu.edu.jo

REFERENCE

1. MuslimScience, *Nature* **477**, 7 (2011).

and the effective operation of subnational fisheries management councils.

Hudson T. Pinheiro,[1,2]* Fabio Di Dario,[3] Leopoldo C. Gerhardinger,[4] Marcelo R. S. de Melo,[5] Rodrigo L. de Moura,[6] Roberto E. Reis,[7] Fábio Vieira,[8] Jansen Zuanon,[9] Luiz A. Rocha[1]

[1]Institute of Biodiversity Science and Sustainability, California Academy of Sciences, San Francisco, CA 94118, USA. [2]Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA 95060, USA. [3]NUPEM, Universidade Federal do Rio de Janeiro, 27910-970, Macaé, RJ, Brazil/ SAIAB, Grahamstown, 6140, South Africa. [4]University of the Region of Joinville, 89240-000, São Francisco do Sul, SC, Brazil. [5]Departamento de Oceanografia Biológica, Instituto Oceanográfico, Universidade de São Paulo, 05508-120, São Paulo, SP, Brazil. [6]Instituto de Biologia and SAGE/COPPE, Universidade Federal do Rio de Janeiro, 21944-970, Rio de Janeiro, RJ, Brazil. [7]PUCRS, Faculdade de Biociências, Laboratory of Vertebrate Systematics, 90619-900, Porto Alegre, RS, Brazil. [8]Centro de Transposição de Peixes/Coleção de Peixes, Universidade Federal de Minas Gerais, 31270-901 Belo Horizonte, MG, Brazil. [9]Instituto Nacional de Pesquisas da Amazônia, Manaus, Amazônas, Brazil.

*Corresponding author. E-mail: htpinheiro@gmail.com

REFERENCES

1. Brazil, Ministério da Agricultura, Pecuária e Abastecimento e Ministério do Meio Ambiente, Portaria N 192 de 5 de outubro de 2015, Diário Oficial da União– Seção 1, 194 (9 October 2015).
2. Brazil, Ministério da Agricultura, Pecuária e Abastecimento, "Portaria suspende seguro defeso por até 120 dias" (www.agricultura.gov.br/comunicacao/noticias/2015/10/portaria-suspende-seguro-defeso-por-ate-120-dias) [in Portuguese].
3. F. Di Dario *et al.*, *Science* **347**, 1079 (2015).
4. M. Benchimol, C. A. Peres, *PLOS ONE* **10**, e0129818 (31 March 2015).
5. Brazil, Presidência da República, Decreto nº 8.467 de 31 de março de 2015; 194º da Independência e 127º da República (2015).
6. G. Mascarenhas, Folha de São Paulo (15 October 2015); http://www1.folha.uol.com.br/poder/2015/10/1694234-operacao-da-pf-mira-venda-ilegal-de-documentos-emitidos-por-ministerio.shtml [in Portuguese].

Nurturing the microbiome field

THE UNIFIED MICROBIOME Initiative Consortium's desire to organize research efforts across disciplinary and geopolitical boundaries is much needed ("A unified initiative to harness Earth's microbiomes," A. P. Alivisatos *et al.*, Policy Forum, 30 October, p. 507; published online 28 October 2015). Previous "Big Science" efforts have had a top-down focus: Large amounts of money are invested in the ideas of a small number of individuals. For the microbiome project, I would propose a bottom-up approach, in which a broad vision is outlined and then as many researchers as possible address those concepts in their biological systems with their own talents and research networks.

Regardless of the structure of the initiative, this field belongs to its researchers, and we must nurture it to ensure continued funding and public interest. We must communicate our efforts to do robust hypothesis-driven research as well as the importance of descriptive studies that contribute to our still-incomplete census. The Consortium and the rest of us must lead by example in terms of transparency, robust methods, pre- and post-publication peer review, and acknowledgment of negative results. As much as we are enthusiastic about the role of the microbiome, we must also realize that it is not responsible for everything. It is exciting to see the nucleus of a grand vision for the field, but we must make the fundamentals of the scientific method central to our mission.

Patrick Schloss

Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI 48109, USA. E-mail: pschloss@umich.edu

J. Marshall Shepherd discusses the challenges of communicating complex climate topics with the public.

Fifty years after U.S. climate warning, scientists confront communication barriers

At the AAAS symposium, researchers searched for new avenues of public engagement to address the gap in climate change beliefs

By **Gavin Stern**

When it comes to climate change, a stark contrast persists between what the scientific community and the American public believe. Though 97% of climate scientists agree that human-caused climate change is occurring, according to the 2014 AAAS What We Know report, only about half of U.S. adults surveyed in an AAAS Pew report released the same year said that climate change is "mostly due to human activity."

Climate scientists have tried to bridge this knowledge gap by better informing the public. But challenging long-held beliefs with more scientific explanations doesn't work, said Katharine Hayhoe, director of the Texas Tech University Climate Science Center, at the "Climate Science, 50 Years Later" symposium in Washington, D.C.

"We've begun to realize that more facts are not going to fix the problem," she said. "Social science has shown that arguing over something as politically polarized as climate change only entrenches people's positions."

The 29 October event, at the Carnegie Institution for Science, commemorated a 1965 report to President Lyndon B. Johnson

Americans hold a variety of beliefs about climate change, said Katharine Hayhoe.

from the President's Council of Advisors on Science and Technology (PCAST). The PCAST report—the first climate warning to an American president—cautioned that the accumulation of atmospheric carbon dioxide from the burning of fossil fuels would "almost certainly cause significant changes" to the environment.

"This is a good time for us to look forward and to look backward. There were many in 1965 who thought that the sky was like the ocean—a metaphor for vastness that could not be changed by humans. We now know that we can," said Rush Holt, CEO of AAAS and executive publisher of the *Science* family of journals. "Today, the vast majority of climate scientists agree that climate change is under way, it's real, it's caused by us, and it's costly in lives and dollars. And it can be addressed."

AAAS and Carnegie Science organized the symposium with support from the American Meteorological Society and The Linden Trust for Conservation. It was a continuation of the AAAS What We Know climate change communications project, launched last year, and it also celebrated this year's launch of the Alan I. Leshner Leadership Institute for Public Engagement with Science.

John P. Holdren, assistant to the president for science and technology and director of the White House Office of Science and Technology Policy, said that the scientific community should have worked earlier to counter public misconceptions about climate change. Confronting this issue presents economic opportunities—not just costs, he said.

"1990 is the point where we really knew enough scientifically to justify the kinds of actions that we're only now talking about today 25 years later. I think the contrarians were incredibly effective in sowing doubt about the validity of the scientific conclusions," said Holdren, a past AAAS president.

Today, the signs of climate change are apparent. Global temperatures are approaching a 1 degree Celsius increase from

Climate Science Milestones Leading To 1965 PCAST Report

1824: French mathematician Joseph Fourier reasons that Earth would be colder in the absence of an atmosphere, describing the foundations of the greenhouse effect.

1856: Eunice Foote's unpublished research on the absorption of radiant energy by carbon dioxide and other gases in the atmosphere is read at the 10th AAAS Annual Meeting in Albany, New York.

1861: Irish physicist John Tyndall finds that gases such as water vapor, carbon dioxide, and methane trap heat efficiently, whereas oxygen and nitrogen gases do not.

1896: Svante Arrhenius, a Swedish physicist, chemist, and 1903 Nobel Prize winner, estimates that a doubling of atmospheric carbon dioxide would raise Earth's temperature by 5 to 6 degrees Celsius, averaged across all latitudes.

1938: English engineer Guy Stewart Callendar determines that a half-century of fuel combustion has added 150,000 million tons of carbon dioxide to the atmosphere.

1957: Former AAAS president Roger Revelle and Austrian geochemist Hans Suess show that absorption of atmospheric carbon dioxide by the ocean is slower than previously believed.

1958: American chemist Charles David Keeling begins recording concentrations of atmospheric carbon dioxide at Mauna Loa Observatory in Hawaii, confirming an increasing trend in atmospheric CO_2. Observations at Mauna Loa continue today.

1965: The President's Council of Advisors on Science and Technology (PCAST) report on environmental pollution to President Lyndon B. Johnson cautions that the accumulation of atmospheric carbon dioxide from the burning of fossil fuels would "almost certainly cause significant changes" to the environment.

(Sources: The Office of Science and Technology Policy, American Institute of Physics, Scripps Institution of Oceanography)

Science authors Katherine Mills of the Gulf of Maine Research Institute (GMRI), Janet Nye of Stony Brook University, and Andrew Pershing of GMRI discuss climate change impacts on cod stocks at the symposium.

pre-industrial levels, according a 9 November warning from the British Met Office. Extreme weather events are more common, and species are fleeing their habitats, symposium speakers said.

Chris Field, founding director of Carnegie Science's Department of Global Ecology and a professor of interdisciplinary environmental studies at Stanford University, called climate change "the defining issue of our time" but also one with great promise.

"We know now that warming caused by CO_2 is essentially permanent…it's not a question of finding a way to reallocate between countries. Carbon emissions everywhere eventually have to go to zero," Field said. "That's incredibly enabling for the future, rather than constricting. It's not about a race to be the last country to build a coal-fired power plant, but to be the first to deploy a sustainable, nonemitting 21st-century energy system."

Yet, there remains a gap in climate literacy—even among those people who accept the fact of climate change. The public views scientific concepts through its own personal lens, speakers said. Most people don't understand what peer-reviewed literature is or what terms like "bias" and "uncertainty" mean for scientists, said J. Marshall Shepherd, director of the University of Georgia atmospheric sciences program and host of Weather Geeks on the Weather Channel.

For example, people may see studies detailing both an increase in drought and more intense rain events as contradictory, said Shepherd, who joined Field at a Capitol Hill briefing organized by AAAS and the office of Sen. Edward Markey (D-Mass.) after the symposium.

"People will bring the climate change discussion to their level of understanding. If they don't understand anything besides averages, maximums, and minimums, they'll put it into that simple context when the science has much more complexity to it," Shepherd said. "We have to be outside of the ivory tower. We have to engage in the media and forums to move the meter. Because if we don't, people skilled in messaging will."

Communicating about climate change was also the focus of a 22 October AAAS Colloquium Series presentation by Susan Joy Hassol, director of Climate Communications. Like Hayhoe and Shepherd, Hassol said that effective communication requires knowing the audience, and "connecting on values" as a first step toward building trust.

Hassol cited surveys that show Americans favor funding clean-energy research, even though acceptance of human-caused climate change still falls along partisan lines. The issue could be better framed by talking about innovation and ingenuity, she said, rather than the need to regulate emissions and reduce energy use.

"It's an opportunity for us to break the partisan gridlock," she said, "and focus on solutions instead of the problem."

Most Americans are at least open to talking about climate change, according to the "Six Americas" study by researchers at Yale and George Mason universities. The 2014 edition places Americans into six climate change perception categories: alarmed (13%), concerned (31%), cautious (23%), disengaged (7%), doubtful (13%), and dismissive (13%).

The vocal minority in the "dismissive" category have found receptive outlets for their beliefs, particularly on social media. But climate change communicators shouldn't focus on that group, Hayhoe said.

Instead, she suggested finding common ground with those who are cautious, disengaged, or doubtful. Scientists can connect with them on issues that they can identify with: parents caring for their children's future, effects on activities like hunting, and principles of religious faith.

"I believe that just about every human being living on this planet has all the values they need to care about climate change. We just need to connect the dots," Hayhoe said. ■

AAAS Members Elected as Fellows

In October 2015, the AAAS Council elected 347 members as Fellows of AAAS. These individuals will be recognized for their contributions to science and technology at the Fellows Forum to be held on 13 February 2016 during the AAAS Annual Meeting in Washington, D.C. Presented by section affiliation, they are:

Section on Agriculture, Food, and Renewable Resources

Arthur G. Appel, Auburn Univ.

Kathryn J. Boor, Cornell Univ.

Judith K. Brown, Univ. of Arizona

Ronald David Green, Univ. of Nebraska-Lincoln

Michael Andrew Grusak, USDA-ARS/Baylor College of Medicine

Jan W. Hopmans, Univ. of California, Davis

Ken Lee, Ohio State Univ.

Andrew (Sandy) M. Liebhold, U.S. Forest Service Northern Research Station

Thomas L. Marsh, Washington State Univ.

Philip Gordon Pardey, Univ. of Minnesota

Randall S. Prather, Univ. of Missouri-Columbia

James M. Reecy, Iowa State Univ.

Mark R. Riley, Univ. of Nebraska-Lincoln

R. Michael Roberts, Univ. of Missouri

G. Philip Robertson, Michigan State Univ.

Daniel P. Schachtman, Univ. of Nebraska-Lincoln

Ravi Prakash Singh, International Maize and Wheat Improvement Center (Mexico)

Zhiyong Wang, Carnegie Institution for Science

Frank F. White, Univ. of Florida

Roger Philip Wise, USDA-ARS/Iowa State Univ.

Section on Anthropology

Timothy G. Bromage, New York Univ.

Bruce M. Latimer, Case Western Reserve Univ.

Thomas W. McDade, Northwestern Univ.

Kathleen A. O'Connor, Univ. of Washington

Michael D. Rose, Rutgers New Jersey Medical School/Duke Univ.

Richard R. Wilk, Indiana Univ.

Section on Astronomy

Sarbani Basu, Yale Univ.

Edmund Bertschinger, Massachusetts Institute of Technology

Matthew Colless, Australian National Univ.

Lucy-Ann A. McFadden, NASA Goddard Space Flight Center

Timothy A. McKay, Univ. of Michigan

Section on Atmospheric and Hydrospheric Sciences

Claudia Benitez-Nelson, Univ. of South Carolina

David H. Bromwich, Ohio State Univ.

Qiang Fu, Univ. of Washington

William H. Hooke, American Meteorological Society

Michael D. King, Univ. of Colorado Boulder

Brian K. Lamb, Washington State Univ.

Zhanqing Li, Univ. of Maryland

Jean Lynch-Stieglitz, Georgia Institute of Technology

Sergey A. Nizkorodov, Univ. of California, Irvine

Ronald S. Oremland, U.S. Geological Survey

Marc B. Parlange, Univ. of British Columbia (Canada)

Cora Einterz Randall, Univ. of Colorado

Jose M. Rodriguez, NASA Goddard Space Flight Center

Dennis Tirpak, World Resources Institute

Compton J. Tucker, NASA Goddard Space Flight Center

Section on Biological Sciences

Cheryl H. Arrowsmith, Univ. of Toronto (Canada)

Irina Artsimovitch, Ohio State Univ.

Jacques Balthazart, Univ. of Liege (Belgium)

Tamar Barkay, Rutgers, The State Univ. of New Jersey

Michelle Barton, Univ. of Texas MD Anderson Cancer Center

Steven R. Beissinger, Univ. of California, Berkeley

Michael A. Bell, Stony Brook Univ.

Keith Louis Bildstein, Hawk Mountain Sanctuary

Douglas L. Black, Univ. of California, Los Angeles

Julie A. Brill, The Hospital for Sick Children (Canada)

Emily A. Buchholtz, Wellesley College

Zachary F. Burton, Michigan State Univ.

Andrea Califano, Columbia Univ.

Yury Chernoff, Georgia Institute of Technology

Karlene A. Cimprich, Stanford Univ. School of Medicine

Jerry David Cohen, Univ. of Minnesota

David O. Conover, Stony Brook Univ.

Nancy J. Cox, Vanderbilt Univ. Medical Center

Gerald R. Crabtree, Stanford Univ. School of Medicine

Kendra L. Daly, Univ. of South Florida

Hans G. Dam, Univ. of Connecticut

Frédéric J. de Sauvage, Genentech

John DiGiovanni, Univ. of Texas at Austin

Beverly M. Emerson, Salk Institute for Biological Studies

Christoph J. Fahrni, Georgia Institute of Technology

Melville Brockett Fenton, Univ. of Western Ontario (Canada)

Gerald R. Fink, Whitehead Institute/ Massachusetts Institute of Technology

Maureen Anne Gannon, Vanderbilt Univ.

Scott Lyell Gardner, Univ. of Nebraska-Lincoln

Pamela K. Geyer, Univ. of Iowa

Helen Haskell Hobbs, Univ. of Texas Southwestern Medical Center

Leonard (Jim) S. Jefferson, Pennsylvania State Univ.

Hailing Jin, Univ. of California, Riverside

Paul Stephen Keim, Northern Arizona Univ.

Hannah L. Klein, New York Univ. School of Medicine

Alberto R. Kornblihtt, Univ. of Buenos Aires/ National Council of Research and Technology (Argentina)

Damian J. Krysan, Univ. of Rochester

Dennis Edward Kyle, Univ. of South Florida

David L. Lentz, Univ. of Cincinnati

Arthur Mallay Lesk, Pennsylvania State Univ.

Maria C. Linder, California State Univ., Fullerton

Erik A. Lundquist, Univ. of Kansas

Pierre P. Massion, Vanderbilt Univ. School of Medicine

Michael J. Matunis, Johns Hopkins Univ.

William W. Metcalf, Univ. of Illinois at Urbana-Champaign

Armin Philipp Moczek, Indiana Univ.

Cynthia Casson Morton, Brigham and Women's Hospital

Robert K. Moyzis, Univ. of California, Irvine

Steven A. Murawski, Univ. of South Florida

Gerd P. Pfeifer, Van Andel Research Institute

Ellen K. Pikitch, Stony Brook Univ.

David W. Piston, Washington Univ. in St. Louis

Carolyn M. Price, Univ. of Cincinnati College of Medicine

Jan A. Randall, San Francisco State Univ.

Joseph C. Reese, Pennsylvania State Univ.

Louise A. Rollins-Smith, Vanderbilt Univ. Medical Center

David G. Schatz, Yale Univ.

Brian Silliman, Duke Univ.

Temple F. Smith, Boston Univ.

C. Neal Stewart Jr., Univ. of Tennessee

Beth A. Sullivan, Duke Univ. Medical Center

LuZhe Sun, Univ. of Texas Health Science Center at San Antonio

Song Tan, Pennsylvania State Univ.

Carolyn M. Teschke, Univ. of Connecticut

Joe M. Tohme, International Center for Tropical Agriculture (Colombia)

Kenneth W. Turteltaub, Lawrence Livermore National Laboratory

Ratna K. Vadlamudi, Univ. of Texas Health Science Center at San Antonio

Ilya Vakser, Univ. of Kansas

William Edward Walden, Univ. of Illinois at Chicago College of Medicine

Wenyi Wei, Harvard Medical School

Beverly Wendland, *Johns Hopkins Univ., Krieger School of Arts and Sciences*

Dong Xu, *Univ. of Missouri*

Philip L. Yeagle, *Univ. of Connecticut*

William E. Zamer, *National Science Foundation*

Janos Zempleni, *Univ. of Nebraska-Lincoln*

Ning Zheng, *Univ. of Washington*

Section on Chemistry

David B. Berkowitz, *Univ. of Nebraska-Lincoln*

William F. Carroll Jr., *Occidental Chemical Corporation (Retired)/Indiana Univ.*

Arthur J. Carty, *Univ. of Waterloo (Canada)*

Xi Chen, *Univ. of California, Davis*

Luis A. Colón, *Univ. at Buffalo, SUNY*

Gregg B. Fields, *Florida Atlantic Univ.*

Leonard W. Fine, *Science Foundation Arizona*

Edith M. Flanigen, *UOP LLC, a Honeywell Company*

Cassandra L. Fraser, *Univ. of Virginia*

Ingrid Fritsch, *Univ. of Arkansas*

Peter F. Green, *Univ. of Michigan*

Masa-aki Haga, *Chuo Univ. (Japan)*

Patrick Harran, *Univ. of California, Los Angeles*

Craig Jon Hawker, *Univ. of California, Santa Barbara*

Christopher P. Jaroniec, *Ohio State Univ.*

Michael Lawrence Klein, *Temple Univ.*

Lon B. Knight Jr., *Furman Univ.*

Shu Kobayashi, *Univ. of Tokyo (Japan)*

Amnon Kohen, *Univ. of Iowa*

Bern Kohler, *Montana State Univ.*

Jay A. LaVerne, *Univ. of Notre Dame*

Dennis L. Lichtenberger, *Univ. of Arizona*

Timothy E. Long, *Virginia Tech*

Walter Loveland, *Oregon State Univ.*

Michael J. Maroney, *Univ. of Massachusetts Amherst*

E.W. "Bert" Meijer, *Eindhoven Univ. of Technology (Netherlands)*

Joseph S. Merola, *Virginia Tech*

David E. Morris, *Los Alamos National Laboratory*

Kate H. Murashige, *Morrison & Foerster LLP*

Ralph G. Nuzzo, *Univ. of Illinois at Urbana-Champaign*

M. Parans Paranthaman, *Oak Ridge National Laboratory/Univ. of Tennessee, Knoxville*

Robin N. Perutz, *Univ. of York (UK)*

Reuben J. Peters, *Iowa State Univ.*

Eli Pollak, *Weizmann Institute of Science (Israel)*

Andrzej Rajca, *Univ. of Nebraska-Lincoln*

Tariq M. Rana, *Univ. of California, San Diego School of Medicine*

Louis Terminello, *Pacific Northwest National Laboratory*

Rao M. Uppu, *Southern Univ. and A&M College*

R. Bruce Weisman, *Rice Univ.*

Paul H. Wine, *Georgia Institute of Technology*

Jay R. Winkler, *California Institute of Technology*

Jackie Y. Ying, *Institute of Bioengineering and Nanotechnology (Singapore)*

Timothy S. Zwier, *Purdue Univ.*

Section on Dentistry and Oral Health Sciences

William Giannobile, *Univ. of Michigan*

Francis L. Macrina, *Virginia Commonwealth Univ.*

Thomas E. Van Dyke, *Forsyth Institute*

Section on Education

Carol Bender, *Univ. of Arizona*

Lin Chambers, *NASA*

Edward E. Geary, *Western Washington Univ.*

Theodore Hodapp, *American Physical Society*

Watson M. Laetsch, *Univ. of California, Berkeley (Retired)*

Rochelle D. Schwartz-Bloom, *Duke Univ. Medical Center*

Samuel M. Taylor, *Center for Science and Public Life*

Section on Engineering

James Hiram Aylor, *Univ. of Virginia*

David F. Bahr, *Purdue Univ.*

Ian Baker, *Dartmouth College*

Karl K. Berggren, *Massachusetts Institute of Technology*

Venkat R. Bhethanabotla, *Univ. of South Florida*

R. Byron Bird, *Univ. of Wisconsin-Madison*

Marc Cahay, *Univ. of Cincinnati*

Shu Chien, *Univ. of California, San Diego*

Lalit Chordia, *Thar Energy, LLC*

Ted Allen Conway, *Florida Institute of Technology*

Dennis E. Discher, *Univ. of Pennsylvania*

Jon Dobson, *Univ. of Florida*

Dominique M. Durand, *Case Western Reserve Univ.*

Greg Evans, *Univ. of Toronto (Canada)*

Yuguang "Michael" Fang, *Univ. of Florida*

Donald P. Gaver III, *Tulane Univ.*

K. Jane Grande-Allen, *Rice Univ.*

Vincent G. Harris, *Northeastern Univ.*

Kevin J. Hemker, *Johns Hopkins Univ.*

Marwan K. Khraisheh, *Qatar Foundation*

Timothy L. Killeen, *Univ. of Illinois*

Shankar Mahalingam, *Univ. of Alabama in Huntsville*

William H. Mischo, *Univ. of Illinois at Urbana-Champaign*

Sushanta K. Mitra, *York Univ. (Canada)*

Babatunde Ogunnaike, *Univ. of Delaware*

Martin Richardson, *Univ. of Central Florida*

Ian M. Robertson, *Univ. of Wisconsin-Madison*

Shelly E. Sakiyama-Elbert, *Washington Univ. in St. Louis*

Mark. T. Swihart, *Univ. at Buffalo, SUNY*

Michael W. Vannier, *Univ. of Chicago Medical Center*

Haiyan Wang, *Texas A&M Univ.*

Hong Yang, *Univ. of Illinois at Urbana-Champaign*

Tian C. Zhang, *Univ. of Nebraska-Lincoln*

Yuwen Zhang, *Univ. of Missouri*

Andrew L. Zydney, *Pennsylvania State Univ.*

Section on General Interest in Science and Engineering

Barry Aprison, *Univ. of Chicago*

Carla Carlson, *Univ. of Minnesota*

Monica M. Metzler, *Illinois Science Council*

David Evans Shaw, *Black Point Group*

Section on Geology and Geography

Marie-Pierre Aubry, *Rutgers, The State Univ. of New Jersey*

Annalisa Berta, *San Diego State Univ.*

David P. Dethier, *Williams College*

Jacqueline Eaby Dixon, *Univ. of South Florida*

Timothy H. Dixon, *Univ. of South Florida*

Frank R. Ettensohn, *Univ. of Kentucky*

Janet Franklin, *Arizona State Univ.*

Andrew J. Friedland, *Dartmouth College*

Steve Frolking, *Univ. of New Hampshire*

Nancy L. Jackson, *New Jersey Institute of Technology*

R. James Kirkpatrick, *Michigan State Univ.*

Paul L. Koch, *Univ. of California, Santa Cruz*

M. Susan Lozier, *Duke Univ.*

Michael Mann, *Pennsylvania State Univ.*

Douglas Burton Richardson, *Association of American Geographers*

J. Donald Rimstidt, *Virginia Tech*

David Stahle, *Univ. of Arkansas*

Section on History and Philosophy of Science

Anne Fagot-Largeault, *College de France*

Sandra D. Mitchell, *Univ. of Pittsburgh*

Section on Industrial Science and Technology

Hiroshi Nagano, *National Graduate Institute for Policy Studies (Japan)/Japan Science and Technology Agency*

Shigeo (Ted) Oyama, *Virginia Tech/Univ. of Tokyo (Japan)*

Philip Shapira, *Univ. of Manchester (UK)/ Georgia Institute of Technology*

Thomas Zacharia, *Oak Ridge National Laboratory*

Section on Information, Computing, and Communication

Legand L. Burge III, *Howard Univ.*

James W. Demmel, *Univ. of California, Berkeley*

Bruce Randall Donald, *Duke Univ.*

Kenneth M. Ford, *Institute for Human and Machine Cognition*

Bruce Hendrickson, *Sandia National Laboratories*

Anna W. Topol, *IBM Research*

Erik Winfree, *California Institute of Technology*

Victor W. Zue, *Massachusetts Institute of Technology*

Section on Linguistics and Language Sciences

Victor Golla, *Humboldt State Univ.*

John J. McCarthy, *Univ. of Massachusetts Amherst*

Section on Mathematics

Daniel L. Goroff, *Alfred P. Sloan Foundation*

Peter Kuchment, *Texas A&M Univ.*

Reinhard C. Laubenbacher, *Univ. of Connecticut Health Center/Jackson Laboratory for Genomic Medicine*

Howard A. Levine, *Iowa State Univ.*

Section on Medical Sciences

Joey V. Barnett, *Vanderbilt Univ. School of Medicine*

Peter A. Barry, *Univ. of California, Davis*

Gordon R. Bernard, *Vanderbilt Univ. School of Medicine*

Robert David Burk, *Albert Einstein College of Medicine*

John M. Carethers, *Univ. of Michigan Health System*

Gen-Sheng Feng, *Univ. of California, San Diego*

Michael A. Frohman, *Stony Brook Univ.*

Stephen J. Galli, *Stanford Univ. School of Medicine*

Gabriel Hortobagyi, *Univ. of Texas MD Anderson Cancer Center*

Robert E. Hurst, *Univ. of Oklahoma Health Sciences Center*

Christopher D. Kontos, *Duke Univ. Medical Center*

Mitchell Kronenberg, *La Jolla Institute for Allergy and Immunology*

Calvin J. Kuo, *Stanford Univ. School of Medicine*

E. Douglas Lewandowski, *Univ. of Illinois at Chicago/Sanford Burnham Prebys Medical Discovery Institute*

Asrar B. Malik, *Univ. of Illinois at Chicago*

Robert J. Matusik, *Vanderbilt Univ. Medical Center*

Lopa Mishra, *George Washington Univ./Univ. of Texas MD Anderson Cancer Center*

Beverly S. Mitchell, *Stanford Univ. School of Medicine*

Vincent Monnier, *Case Western Reserve Univ.*

Hugh M. O'Brodovich, *Stanford Univ. School of Medicine*

Paul A. Offit, *Children's Hospital of Philadelphia*

Mark E. Peeples, *Nationwide Children's Hospital/Ohio State Univ.*

Mark R. Philips, *New York Univ. School of Medicine*

David Joseph Pintel, *Univ. of Missouri-Columbia*

Vito Quaranta, *Vanderbilt Univ. School of Medicine*

Peter S. Rabinovitch, *Univ. of Washington*

William N. Rom, *New York Univ. Langone Medical Center*

Jeffrey M. Rosen, *Baylor College of Medicine*

Victor L. Schuster, *Albert Einstein College of Medicine*

Alan L. Scott, *Johns Hopkins Univ.*

Melinda Wharton, *Centers for Disease Control and Prevention*

Mark Yeager, *Univ. of Virginia School of Medicine*

Qing Yi, *Cleveland Clinic*

Section on Neuroscience

Edwin George (Ted) Abel III, *Univ. of Pennsylvania*

Robert E. Burke, *Columbia Univ. Medical Center*

John R. Huguenard, *Stanford Univ. School of Medicine*

Anumantha G. Kanthasamy, *Iowa State Univ.*

Edward H. Koo, *Univ. of California, San Diego/National Univ. of Singapore*

Shinn-Zong Lin, *China Medical Univ.*

Paul S. Mischel, *Univ. of California, San Diego/Ludwig Institute for Cancer Research*

Louis J. Ptáček, *Univ. of California, San Francisco*

Thomas A. Rando, *Stanford Univ. School of Medicine*

Anna Wang Roe, *Vanderbilt Univ./Zhejiang Univ. (China)*

William W. Seeley, *Univ. of California, San Francisco*

Michael N. Shadlen, *Columbia Univ.*

Pamela Jean Shaw, *Univ. of Sheffield (UK)*

Steven A. Siegelbaum, *Columbia Univ. Medical Center*

Ethan R. Signer, *CHDI Foundation/CHDI Management*

Edward L. Stuenkel, *Univ. of Michigan*

Kent E. Vrana, *Pennsylvania State Univ. College of Medicine*

Section on Pharmaceutical Sciences

James T. Dalton, *Univ. of Michigan*

Courtney V. Fletcher, *Univ. of Nebraska Medical Center*

Stephen B. Howell, *Univ. of California, San Diego*

Patricia D. Kroboth, *Univ. of Pittsburgh*

Cynthia Moreton Kuhn, *Duke Univ. School of Medicine*

Richard R. Neubig, *Michigan State Univ.*

Alan Paau, *Guangda Cooperation International Technology Center (China)*

Doodipala Samba Reddy, *Texas A&M Univ.*

Section on Physics

Mark A. Eriksson, *Univ. of Wisconsin-Madison*

Raymond Fonck, *Univ. of Wisconsin-Madison*

Martin Greven, *Univ. of Minnesota*

Andreas J. Heinrich, *IBM Almaden Research Center*

A.T. Charlie Johnson Jr., *Univ. of Pennsylvania*

Steven M. Kahn, *Stanford Univ.*

Duncan McBride, *National Science Foundation (Retired)*

Adilson E. Motter, *Northwestern Univ.*

Sekazi Kauze Mtingwa, *Massachusetts Institute of Technology (Retired)/Triangle Science, Education & Economic Development, LLC*

Beate Schmittmann, *Iowa State Univ.*

Gary Shiu, *Univ. of Wisconsin-Madison*

Michelle Yvonne Simmons, *Univ. of New South Wales (Australia)*

Mark Trodden, *Univ. of Pennsylvania*

Section on Psychology

Rebecca D. Burwell, *Brown Univ.*

Laurence Baker Leonard, *Purdue Univ.*

David G. Myers, *Hope College*

Scott Plous, *Wesleyan Univ.*

J. Edward Russo, *Cornell Univ.*

Paula R. Skedsvold, *Federation of Associations in Behavioral and Brain Sciences*

Kiran K. Soma, *Univ. of British Columbia (Canada)*

Jeffrey M. Zacks, *Washington Univ. in St. Louis*

Section on Social, Economic, and Political Sciences

Chloe E. Bird, *RAND*

Thomas A. DiPrete, *Columbia Univ.*

Jerald Hage, *Univ. of Maryland, College Park*

Marie Currie Thursby, *Georgia Institute of Technology*

Section on Societal Impacts of Science and Engineering

Cathleen A. Campbell, *CRDF Global*

David Goldston, *Natural Resources Defense Council*

Owen D. Jones, *Vanderbilt Univ.*

Kathleen M. Rest, *Union of Concerned Scientists*

Section on Statistics

Michael Paul Cohen, *American Institutes for Research*

Bruce A. Craig, *Purdue Univ.*

Patricia A. Jacobs, *Naval Postgraduate School*

Alan F. Karr, *RTI International*

Stephen Portnoy, *Univ. of Illinois at Urbana-Champaign*

James Matthew Robins, *Harvard Univ.*

Daniel O. Stram, *Keck School of Medicine of Univ. of Southern California*

Chih-Ling Tsai, *Univ. of California, Davis*

Alyson G. Wilson, *North Carolina State Univ.*

AAAS Kavli Science Journalism Award winners named

This year's winners included stories on the health impacts of urban violence, local signs of global climate change, and West Africa's Ebola epidemic

*By **Earl Lane***

Stories on the stressful impact of urban violence on children, the shared aptitudes of humans and songbirds for vocal learning, and the impact of climate change on the forests of Minnesota and beyond, are among the winners of the 2015 AAAS Kavli Science Journalism Awards.

For the first time in the 70-year history of the program, entries were accepted from journalists around the globe in all award categories. The expansion was made possible by a generous doubling of the program endowment by The Kavli Foundation, which established the endowment in 2009. The new funds also permitted two awards in each of the eight categories for the first time—a Gold Award ($5,000) and a Silver Award ($3,500).

Just under 40% of the winners were international entries, comparable to the percentage of international entries received.

"The new global era for the AAAS Kavli awards is off to a great start," said Rush Holt, chief executive officer of AAAS and executive publisher of the *Science* family of journals. "The breadth of the winning work and the diversity of outlets in which it appeared demonstrate the vitality of science journalism at a time when public understanding of science is more important than ever. I expect the awards will prove to mean as much for international science writers as they have over the years for science writers in the United States."

The winners of the 2015 AAAS Kavli Science Journalism Awards are:

Large Newspaper (Circulation of 150,000 or more) Gold Award: Andrea K. McDaniels, *The Baltimore Sun*, for "Collateral Damage" series: "City's Violence Can Take Hidden Toll, " 14 December 2014; "Some Wounded Wind Up at Home," 18 December 2014; and "Hidden Sorrows," 21 December 2014.

Large Newspaper Silver Award: Nathaniel Herzberg, *Le Monde*, for "Stéthoscope, Il n'a Plus le Monopole du Coeur" (The stethoscope no longer holds a monopoly over our hearts), 26 November 2014; "Anguilles, Sept mille lieues sous les mers" (Eels, 7000 leagues under the sea), 17 December 2014; and "La Souris, Reine Contestée des Labos" (The mouse, challenged queen of the lab), 18 February 2015.

Small Newspaper (Circulation less than 150,000) Gold Award: Matthew Miller, *Lansing State Journal*, for "Battle of the Ash Borer," 27 July 2014.

Small Newspaper Silver Award: Helga Rietz, *Neue Zürcher Zeitung* (Switzerland), for "Arien für die Wissenschaft" (Arias for Science), 24 December 2014.

Magazine Gold Award: Alexandra Witze, *Nature and Science News*, for "The Quake Hunters" (in *Nature*), 9 July 2015; "The Pluto Siblings" (in *Nature*), 26 February 2015; "Let the River Run" (in *Science News*), 10 January 2015.

Magazine Silver Award: Amanda Gefter, *Nautilus*, for "The Man Who Tried to Redeem the World with Logic," March/April 2015.

Television Spot News/Feature Reporting (20 minutes or less) Gold Award: Katie Campbell, KCTS 9 (Seattle), for "Is Alaska Safe for Sea Stars?" 8 October 2014.

Television Spot News/Feature Reporting Silver Award: Miles O'Brien, PBS NewsHour, for "Will a Robotic Arm Ever Have the Full Functionality of a Human Limb?" 12 February 2015; and "Can Modern Prosthetics Actually Help Reclaim the Sense of Touch?" 13 February 2015.

Television In-Depth Reporting (more than 20 minutes) Gold Award: Jonathan Renouf and Alex Freeman, BBC, for "Climate Change by Numbers," 2 March 2015.

Television In-Depth Reporting Silver Award: Lone Frank and Pernille Rose Grønkjær, Danish Broadcasting Corporation, for "Genetic Me," 26 November 2014.

Radio Gold Award: Rami Tzabar and Angela Saini, BBC Radio 4 and BBC World Service, for "What the Songbird Said," 11 May 2015.

Radio Silver Award: Dan Kraker and Elizabeth Dunbar, Minnesota Public Radio, for "Climate Change in Minnesota: More Heat. More Big Storms," 2 February 2015; "A Forest Dilemma: What Will Grow in a Changing Climate?" 3 February 2015; and "As State Warms, a Few Spots Keep Their Cool," 3 February 2015.

Online Gold Award: Mark Harris, Backchannel, for "How a Lone Hacker Shredded the Myth of Crowdsourcing," 9 February 2015.

Online Silver Award: Kevin Sack, Sheri Fink, Pam Belluck, and Adam Nossiter, with Daniel Berehulak, Dan Edge (for Frontline), and *The New York Times* graphics team, *The New York Times*, for "How Ebola Roared Back," 29 December 2014.

Children's Science News Gold Award: Stephen Ornes, Science News for Students (online site), for "Where Will Lightning Strike?" 16 September 2014.

Children's Science News Silver Award: Joan Cartan-Hansen, Idaho Public Television, for "Science Trek: Bats—White Nose Syndrome," 16 September 2014.

The awards, administered by AAAS since their inception in 1945, go to professional journalists for distinguished reporting for a general audience. Independent panels of science journalists pick the winners, who will receive their awards at the 2016 AAAS Annual Meeting in Washington, DC, in February. More information on the winning entries is available at: www.aaas.org/sja2015. ∎

For the first time this year, the AAAS Kavli Awards included international entries in all categories and two prizes in each category.

PHOTO: AAAS

Protein folding starts in the ribosome

Holtkamp et al., *p. 1104*
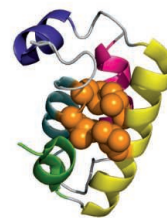
IN *SCIENCE* JOURNALS

Edited by **Caroline Ash**

EVOLUTION

Evolutionary routes to blue

Animal coloration is vital for mate selection, camouflage, warning, and temperature regulation. The evolutionary pressure to get color right leads to tremendous variation among species. The spectacular iridescent blue of some invertebrates is created by microscopic surface structures that interfere with incident light. Hsiung *et al.* found that nearly identical blue coloration evolved at least eight times in tarantulas via a variety of surface nanostructures. Because these spiders' vision is poor, the blue color is unlikely to be for sexual display. Natural selection, rather than sexual selection, seems to be operating on the spiders via some as yet unknown pressure. — RLL

Sci. Adv. 10.1126.sciadv.1500709 (2015).

Many variants of surface nanostructure cause blue coloration

CANCER IMMUNOTHERAPY

Gut microbes affect immunotherapy

The unleashing of antitumor T cell responses has ushered in a new era of cancer treatment. Although these therapies can cause dramatic tumor regressions in some patients, many patients inexplicably see no benefit. Mice have been used in two studies to investigate what might be happening. Specific members of the gut microbiota influence the efficacy of this type of immunotherapy (see the Perspective by Snyder *et al.*). Vétizou *et al.* found that optimal responses to anticytotoxic T lymphocyte antigen blockade required specific *Bacteroides* spp. Similarly, Sivan *et al.* discovered that *Bifidobacterium* spp. enhanced the efficacy of antiprogrammed cell death ligand 1 therapy. — KLM

Science, this issue, pp. 1079 and 1089; see also p. 1031

ECOLOGY

How herbivores affect ecosystems

Abiotic forces, such as fire and water, have powerful effects on ecosystem structure and function. Animals that eat plants also have strong effects in natural systems, but their impacts are harder to assess. Hempson *et al.* measured how herbivores affect vegetation across Africa (see the Perspective by Gill).

Four distinct herbivory regimes emerge from the analysis, characterized by forest antelopes, arid-region gazelles, high-diversity savanna fauna, and bulk feeders (such as elephants), which have had equivalent impact to those of fire and water on shaping ecosystems. — SNV

Science, this issue p. 1056; see also p. 1036

GEOLOGY

Salted away no longer?

Rock salt deposits are thought to be impermeable to fluid flow and so are candidates for nuclear waste repositories. Ghanbarzadeh *et al.* found that some salt deposits in the Gulf of Mexico are infiltrated by oil and other hydrocarbons. If these salt domes are not completely isolated from the surrounding environment, they will not be suitable for deep geological waste storage sites. — BG

Science, this issue p. 1069

GENOME EDITING

Virally cleansing the pig genome

Transplants from pigs could be a solution to a shortage of human organs for transplantation. Unfortunately, porcine endogenous retroviruses (PERVs) are rife in pigs and can be transmitted to humans, risking disease. L. Yang *et al.* integrated CRISPR-Cas into the pig cell genome, where continuous induction of the Cas9 editing enzyme

resulted in the mutation of every single PERV reverse transcriptase gene. This prevented replication of all copies of PERV, viral infection, and transmission to human cells. — GR

Science, this issue p. 1101

MALARIA
How malaria parasites infect the liver

Early in infection, malaria parasites establish themselves within hepatocytes in the liver. Inside these cells, the parasites occupy a so-called parasitophorous vacuole. Kaushansky *et al.* show that malaria parasites prefer to create vacuoles within hepatocytes that express the EphA2 receptor. Hepatocytes with low levels of this receptor were less conducive to malaria infection. — SMH

Science, this issue p. 1089

ECONOMICS
Predicting unmeasurable wealth

In developing countries, collecting data on basic economic quantities, such as wealth and income, is costly, time-consuming, and unreliable. Taking advantage of the ubiquity of mobile phones in Rwanda, Blumenstock *et al.* mapped mobile phone metadata inputs to individual phone subscriber wealth. They applied the model to predict wealth throughout Rwanda and show that the predictions matched well with those from detailed boots-on-the-ground surveys of the population. — GJC

Science, this issue p. 1073

NANOMATERIALS
Brighter molybdenum layers

The confined layers of molybdenum disulphide (MoS_2) exhibit photoluminescence that is attractive for optoelectronic applications. In practice, efficiencies are low, presumably because defects trap excitons before they can recombine and radiate light. Amani *et al.* show

that treatment of monolayer MoS_2 with a nonoxidizing organic superacid, bis(trifluoromethane) sulfonimide, increased luminescence efficiency in excess of 95%. The enhancement mechanism may be related to the shielding of defects, such as sulfur vacancies. — PDS

Science, this issue p. 1065

PHOTOPHYSICS
Charge separation viewed in reflection

When light strikes a semiconductor, excited electrons travel across the interface. Y. Yang *et al.* applied ultrafast reflection spectroscopy to probe this process in a gallium indium phosphide system used for hydrogen generation from water (see the Perspective by Hansen *et al.*). Platinum and titanium dioxide (TiO_2) coatings enhanced charge separation of the excited electrons from the positive holes they left behind. TiO_2, however, was more effective at suppressing the reverse process of unproductive recombination. — JSY

Science, this issue p. 1061; see also p. 1030

DIABETES IMMUNOTHERAPY
Tweaking T regulatory affairs

In patients with type 1 diabetes (T1D), immune cells destroy the insulin-producing beta cells of the pancreas. Consequent prolonged exposure to high blood sugar can damage organs and lead to heart disease and kidney failure. Regulatory T cells (T_{regs}) are known to be defective in autoimmune diseases, such as type 1 diabetes. Bluestone *et al.* report a phase 1 trial of adoptive T_{reg} immunotherapy to repair or replace these cells in type 1 diabetics. The cultured T_{regs} were long-lived after transfer and retained a broad T_{reg} phenotype. Moreover, the trial showed that transfer therapy was safe, endorsing efficacy testing in further trials. — ACC

Sci. Transl. Med. **7**, 315ra189 (2015).

Species with more neurons, such as this Gelada baboon, may need less sleep

NEUROSCIENCE
More neurons mean less need for sleep

Sleep is seemingly universal among animals. Daily sleep time varies considerably between mammalian species and also during mammalian development, yet we still don't know what drives this variation. Herculano-Houzel hypothesized that sleep-inducing metabolites produced during waking hours accumulate more slowly in brains that have a smaller density of neurons underneath a unit surface area that gets washed by cerebrospinal fluid during waking. In 24 mammalian species and several postnatal stages in the developing rat, there was indeed a correlation between the ratio of neuronal density to brain surface area and daily sleep duration. The evolutionary addition of neurons may have decreased the need for sleep, allowing a species to feed for longer, and thus facilitated further increases in neuronal numbers. — PRS

Proc. R. Soc. London Ser. B **282**, 1816 (2015).

WORKFORCE DIVERSITY
Diversity through ADVANCEment

The NSF ADVANCE program aims to increase the advancement of women in academic careers, usually through the implementation of work/life support policies. Tower and Dilks developed a policy rating scheme to measure the level of parental leave, tenure clock extension, availability of child

PHOTO: MICHAEL NICHOLS/NATIONAL GEOGRAPHIC CREATIVE

ALSO IN *SCIENCE* JOURNALS

Edited by **Caroline Ash**

BIOIMAGING
Imaging with molecular vibrations

The vibrational spectra of biomolecules could in principle image cells and tissue without added markers. Practically, several technical problems need to be overcome to achieve sufficient imaging depths, resolution, and data acquisition speed. Cheng and Xie review emerging bioimaging methods for use in the lab and the clinic. — PDS

Science, this issue p. 1054

NEUROSCIENCE
A census of neocortical neurons

Despite the importance of the brain's neocortex, we still do not completely understand the diversity and functional connections of its cell types. Jiang *et al.* recorded, labeled, and classified over 1200 interneurons and more than 400 pyramidal neurons in the mature mouse visual cortex. Fifteen major classes of interneurons fell into three types: some connect to all neurons, some connect to other interneurons, and some form synapses with pyramidal neurons. — PRS

Science, this issue p. 1055

PROTEIN FOLDING
Proteins shape up in the ribosome

Proteins consist of linear chains of amino acids. These chains must fold into complex three-dimensional shapes to become functional. Holtkamp *et al.* "watched" how a small helical protein folds as it is being synthesized by the ribosome. The lengthening polypeptide passes out through the ribosome exit tunnel where folding starts. The initially compact structure quickly rearranges into a native three-dimensional structure as the polypeptide emerges from the tunnel. — GR

Science, this issue p. 1104

EVOLUTIONARY BIOLOGY
The downside of innovation

Evolutionary innovation allows a species to invade a new niche or environment. Generally, the emergence of adaptive traits is thought to lead to diversification. Support for this process in nature, however, is mixed. McGee *et al.* show that the evolution of secondary jaws in fish may be an example of how innovation can reduce diversification (see the Perspective by Vermeij). Fish with secondary jaws are less able to rapidly ingest fish prey, which puts them at a competitive disadvantage to regularly jawed fish. Such competition could lead to reduced variation and the extinction of lineages with the trait. — SNV

Science, this issue p. 1077; see also p. 1038

GENOMICS
Zeroing in on essential human genes

More powerful genetic techniques are helping to define the list of genes required for the life of a human cell. Two papers used the CRISPR genome editing system and a gene trap method in haploid human cells to screen for essential genes (see the Perspective by Boone and Andrews). Wang *et al.*'s analysis of multiple cell lines indicates that it may be possible to find tumor-specific dependencies on particular genes. Blomen *et al.* investigate the phenomenon in which nonessential genes are required for fitness in the absence of another gene. Hence, complexity rather than robustness is the human strategy. — LBR

Science, this issue p. 1096; see also p. 1028

NEUROSCIENCE
Antidepressants suppress DNA methylation

Some depressed patients show increased DNA methylation and decreased expression of a gene encoding BDNF, a secreted factor important for synaptic plasticity. In cells obtained from depressed patients before treatment, Rein *et al.* found that an antidepressant, paroxetine, increased expression of *BDNF* via inhibition of the DNA methyltransferase DNMT1. Paroxetine has similar effects on human blood cells; hence a simple blood test may aid in personalizing treatment for depression. — LKF

Sci. Signal. **8**, ra119 (2015).

care, and related support programs at 124 ADVANCE institutions. Results showed that ADVANCE universities were highly progressive regarding new parent support (80% offer benefits beyond FMLA protections to birth mothers) and basic child care (only 8% had no direct benefits). Additionally, a 1-year extension of the tenure clock was available to birth mothers at 44% of the institutions. As promising as these results seem, the authors caution that simply having work/life policies on the books is not enough to ensure their use. — MM

J. Divers. High. Educ. **8,** 157 (2015).

PLANT SCIENCE
Cell size matters to meristems

In the meristem that generates flowers for the plant *Arabidopsis thaliana*, the cells are all generally the same size. Serrano-Mislata *et al.* asked what happens when that regularity is perturbed. Some irregularity occurs naturally, as cell divisions were often a bit unequal, producing daughters of different sizes. Experimental manipulation of cell cycle progression introduced other irregularities. Persistent disruption of cell-size controls correlated with irregular or absent definition of the floral organs generated by the meristem. Regardless of how the irregularities were generated, the meristem corrected size irregularities and brought daughter cells into the same regularized size. The authors hypothesize that without a controlled unit size, the signaling that establishes developmental fates in the meristem becomes scrambled and vague. — PJH

Curr. Biol. **25,** 1 (2015).

PALEOCLIMATE
A shifting wet girdle around the tropics

The Intertropical Convergence Zone (ITCZ) causes heavy rains
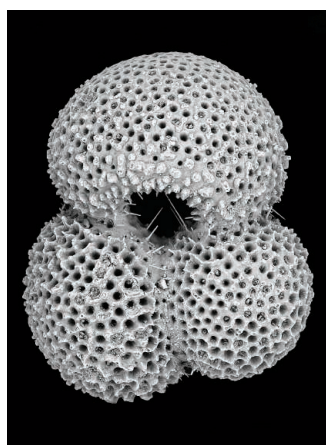
A fossil *Globigerinoides ruber*, used to reconstruct past precipitation changes

to fall in a seasonally migrating band around the globe near the equator. Because it delivers so much precipitation to so many regions, it is a vital component of climate that affects many ecosystems and human populations, so any possible changes in its position could have major implications for them. Liu *et al.* show that, for the past 280,000 years at least, the average position of the ITCZ in the western Pacific has been controlled by a combination of solar obliquity and precession. The dependence that they see on the thermal

Changes in the neuronal network may consolidate memory, as seen in the *Rosy Tritonia*

state of the atmosphere may provide insights into possible effects on the ITCZ from anthropogenic global warming. — HJS

Nat. Commun. 10.1038/ ncomms10018 (2015).

NEUROSCIENCE
Committing to memory

Neuronal networks in the brain can expand during memory formation, but where do the additional neurons come from? Hill *et al.* visually tracked this process in the marine mollusk *Tritonia diomedea*, as the animal became more sensitized to a stimulus that evokes its escape swim response. Inactive neurons appeared to be "pre-positioned" for rapid recruitment into an existing network as the animal displayed a learned response. As the memory faded, some new recruits remained committed to the network, while some original neurons departed. Continual change in the cellular constituents of a network may be a mechanism of memory formation. — LDC

Curr. Biol. **25,** 1 (2015).

MITOCHONDRIA
Mitochondrial quality control

Mitochondria contain proteins encoded by the nucleus and by their own mitochondrial genome. Sometimes protein synthesis within the mitochondria goes wrong, which can damage both mitochondrial and cellular fitness. Richter *et al.* interfered with intramitochondrial protein synthesis by means of an antibiotic and found that misfolded proteins accumulated at the inner mitochondrial membrane. This compromised the mitochondrial membrane potential, kicking into gear a mechanism to put the brakes on mitochondrial protein synthesis. The reduction in mitochondrial protein synthesis protected the mitochondria from further damage. This quality-control mechanism would be expected to protect mitochondria in the event of short-term perturbations of mitochondrial translation. — SMH

J. Cell Biol. **211,** 373 (2015).

CARBENE CHEMISTRY
Crystal structure of a rhodium carbene

It often is challenging in chemistry to characterize highly reactive compounds. The trouble is that modifications that render them sufficiently stable to study detract from the very properties underlying the reactivity of interest. Werlé *et al.* now report success in the low-temperature preparation and isolation of crystals of a dirhodium carbene complex. They elucidated its structure by x-ray diffraction of the compound in the solid state, as well as by using several spectroscopic techniques in solution. Moreover, they confirmed that it indeed manifested the reactivity for which this class of compounds is prized: transfer of the divalent carbon to an olefin to form a cyclopropyl ring. They also observed facile transfer of the carbene from rhodium to gold. — JSY

Angew. Chem. Int. Ed. 10.1002/ anie.201506902 (2015).

REVIEW SUMMARY

BIOIMAGING

Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine

Ji-Xin Cheng* and X. Sunney Xie*

BACKGROUND: Biomolecules can serve as natural labels for microscopy by measuring their molecular vibration spectra in living cells and tissues. However, the transition from spectroscopy of molecules in cuvettes to spectroscopic imaging of living systems requires more than putting spectrometers on microscopes. A series of technical challenges must be addressed, such as delivery of light beams for sample excitation and scattering of signals, which limit the probe depth of spectroscopic imaging. There are also trade-offs in how large a spectral window can be measured at a pixel (the amount of chemical information) in a given amount of time (the recording speed).

ADVANCES: Several technical advances made by different groups have pushed the boundary of the vibrational spectroscopic imaging field in terms of spectral acquisition speed, detection sensitivity, spatial resolution, and penetration depth. Specifically, coherent Raman scattering microscopy has emerged as a high-speed vibrational imaging platform. Single-frequency coherent anti-Stokes Raman scattering (CARS) and stimulated Raman scattering (SRS) microscopes have reached a video-rate imaging speed. Multiplex CARS by broadband excitation has reached a speed of 3.5 ms per pixel; multiplex SRS covering a window of 200 wave numbers has reached a speed of 32 μs per pixel.

The marriage of high-speed coherent Raman microscopy and Raman-sensitive tags of large cross sections has enabled real-time imaging of small molecules at micromolar concentration. Nanoscale vibrational imaging has been demonstrated by integration of atomic force microscopy and vibrational spectroscopy. Vibrational imaging of deep tissue has been enabled by acoustic detection of overtone transitions or optical detection of diffuse photons. Biological applications of CARS and SRS microscopy have generated new insights into myelin biology, lipid droplet biology, intracellular drug delivery, and single-cell metabolism. Meanwhile, clinical applications of vibrational spectroscopic imaging are enabling molecule-based diagnosis of cancer and heart disease without the need for any exogenous contrast agent. Examples include intravascular vibrational photoacoustic imaging of lipid-laden plaques and spatially offset Raman spectroscopic detection of cancer margins.

OUTLOOK: There remain two central challenges facing the field. One is to increase the detection sensitivity of vibrational microscopy to micromolar or even nanomolar levels, so that low-concentration biomolecules in a living system can be mapped. The other is to increase the vibrational imaging depth to tens of centimeters for noninvasive molecule-based medical diagnosis.

With continuous developments, high-resolution, high-speed vibrational microscopy will cultivate unexpected discoveries in cell biology. The findings may lead to the development of new therapies for currently incurable diseases. Meanwhile, with further improvement of penetration depth and progressive reduction of instrument size, vibrational spectroscopic imaging devices are expected to become fundamental clinical tools for disease diagnosis and therapy effectiveness evaluation. ■
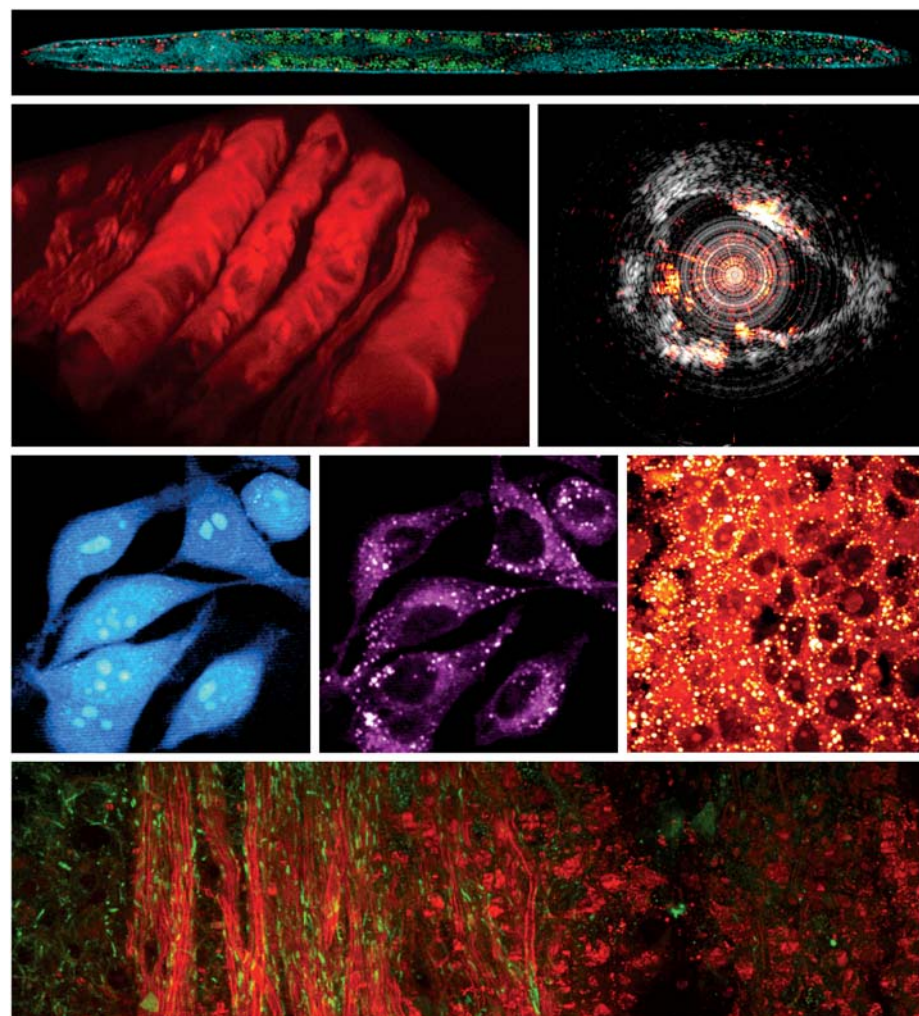
Molecular fingerprints for biology and medicine through imaging. Recent efforts focused on pushing the fundamental limits of vibrational spectroscopic imaging in terms of spectral acquisition speed, detection sensitivity, spatial resolution, and penetration depth. The resulting platforms are enabling transformative applications in functional analysis of single living cells and noninvasive diagnosis of human diseases with biomarker sensitivity.

BIOIMAGING

Vibrational spectroscopic imaging of living systems: An emerging platform for biology and medicine

Ji-Xin Cheng[1]* and X. Sunney Xie[2]*

Vibrational spectroscopy has been extensively applied to the study of molecules in gas phase, in condensed phase, and at interfaces. The transition from spectroscopy to spectroscopic imaging of living systems, which allows the spectrum of biomolecules to act as natural contrast, is opening new opportunities to reveal cellular machinery and to enable molecule-based diagnosis. Such a transition, however, involves more than a simple combination of spectrometry and microscopy. We review recent efforts that have pushed the boundary of the vibrational spectroscopic imaging field in terms of spectral acquisition speed, detection sensitivity, spatial resolution, and imaging depth. We further highlight recent applications in functional analysis of single cells and in label-free detection of diseases.

A century ago, stains developed by Golgi allowed Ramón y Cajal to picture neuronal cells in the central nervous system. More recently, the discovery of green fluorescent proteins facilitated the imaging of protein dynamics in living cells and animals. The development of superresolution fluorescence microscopy has further enabled close examination of cellular structures at the nanometer scale. Despite such advances, the labeling approach has various limitations: (i) Labels may perturb the function of a biological molecule or structure; (ii) the labeling approach offers limited capacity of discovery because it is only applicable to known species; (iii) delivery of labels to a target could be difficult, especially under in vivo conditions; and (iv) the potential toxicity often prevents the use of labels in human subjects.

The emerging field of in vivo vibrational spectroscopic imaging delivers a way to circumvent such barriers. The concept of vibrational spectroscopic imaging can be explained by the children's book *Where's Waldo*. In this book, by simple pattern recognition (his distinctive cap), children locate Waldo in an image of crowds without a need for labels. Similarly, molecules can be recognized by their distinctive signature, or "fingerprint," produced by quantized vibrations of chemical bonds. Fingerprint vibrational spectra of molecules in a single cell can be recorded by infrared (IR) or Raman spectroscopy. Nonetheless, a cell is not a bag of molecules, but rather a spatially and temporally organized dynamic system. If a finger-print spectrum could be recorded at every pixel with sufficient speed and a multivariate analysis performed on the resulting data set, a given "Waldo" molecule could be located and monitored in real time in a living system, thereby gaining insight into its function.

The transition from spectroscopy to spectroscopic imaging of living systems, however, cannot be accomplished simply by combining a spectrometer and a microscope. A series of technical difficulties need to be addressed: Can we increase the speed of spectral acquisition to the microsecond scale in order to capture the dynamics in a living cell? Can we achieve enough sensitivity to detect target molecules in a tissue environment (whose propensity for scattering light degrades images) without harming the cells? Can we have superior resolution to enable imaging of biological structures at the nanoscale? Finally, can we detect a spectrum from a centimeter-deep tissue for in vivo diagnosis? Advances toward overcoming these technical difficulties are discussed below.

Molecular fingerprint vibration spectra can be recorded through measurements of linear IR absorption or inelastic Raman scattering. Fourier-transform IR microscopy has enabled chemical imaging of histological slices [reviewed in (*1*)]. Live-cell imaging applications of IR spectroscopy are, however, hampered by the strong water absorption of IR light and also by the low spatial resolution given the long wavelength (a few micrometers) of IR light. Raman spectroscopy, which uses shorter-wavelength visible light for excitation, has been extensively used for analysis of cells and tissues [reviewed in (*2*)]. Feld, Puppels, Popp, and others have demonstrated broad biomedical applications of Raman spectroscopy. A Raman microscope, first reported in the 1970s, is now commercially available and has reached a speed of milliseconds per pixel. Nonetheless, because of the extremely small cross section of spontaneous Raman scattering, the bioimaging speed of a state-of-the-art Raman microscope is limited to tens of minutes per frame (*3*), insufficient for capturing the dynamics (e.g., movement of organelles) inside living cells.

Coherent Raman scattering (CRS) microscopy (*4*) is an emerging technique that overcomes the frame rate limitation of spontaneous Raman microscopy. In most CRS imaging experiments, two excitation fields are used, denoted as pump (ω_p) and Stokes (ω_s). When the beating frequency ($\omega_p - \omega_s$) between the pump and Stokes fields is resonant with a Raman-active molecular vibration, four major CRS processes occur simultaneously, namely coherent anti-Stokes Raman scattering (CARS) at a new frequency of ($\omega_p - \omega_s$) + ω_p, coherent Stokes Raman scattering at a new frequency of $\omega_s - (\omega_p - \omega_s)$, stimulated Raman gain (SRG) at ω_s, and stimulated Raman loss (SRL) at ω_p. CARS is a parametric phenomenon in which the input and output photons exchange energy while the quantum state of the molecule is left unchanged. As a vibrationally enhanced four-wave mixing process, the CARS signal is generated at a new frequency apart from input beams and is accompanied by a nonresonant signal caused by electronic contributions to the four-wave mixing process. SRG and SRL belong to the process of stimulated Raman scattering (SRS), in which the Stokes beam experiences a gain in intensity and the pump beam experiences an intensity loss. SRS is a dissipative process in which energy corresponding to the beating frequency ($\omega_p - \omega_s$) of input photons is transferred to the molecule for vibrational excitation. The SRS signal appears at the same wavelengths as the excitation fields and is commonly extracted through a lock-in amplifier. Figure 1 shows the schematics of spontaneous Raman scattering, CARS, and SRS.

CARS as a third-order nonlinear process was documented in 1965 by Terhune and Maker at the Ford Motor Company (*5*). Coherent anti-Stokes Raman spectroscopy (*6*) has been widely used for monitoring chemical reaction dynamics and remains a powerful tool for combustion analysis. The CARS microscope was reported in 1982 by Duncan *et al.* using a two-dimensional (2D) imaging device (*7*). In 1999, Zumbusch *et al.* demonstrated 3D CARS imaging of live cells by raster-scanning two collinearly overlapped and tightly focused femtosecond pulse excitation beams (*8*). This work triggered the development of modern CARS microscopy. By spectrally focusing the excitation energy on a single Raman band by means of picosecond pulses, Cheng *et al.* demonstrated CARS imaging of cells in the fingerprint vibration region (*9*) and further developed a high-speed laser-scanning CARS microscope for visualization of cell mitosis and apoptosis (*10*). Epi-detected (i.e., backward-detected) CARS was developed to minimize the detection of nonresonant signals (*11*) and was later used for imaging live animals at video rate (*12*). A Green's function model was developed to interpret the radiation pattern of CARS from a 3D object and thereby explain the contrast mechanism (*13*). At the same time, multiplex CARS

[1]Weldon School of Biomedical Engineering and Department of Chemistry, Purdue University, West Lafayette, IN 47907, USA. [2]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA.
*Corresponding author. E-mail: jcheng@purdue.edu (J.-X.C.); xie@chemistry.harvard.edu (X.S.X.)
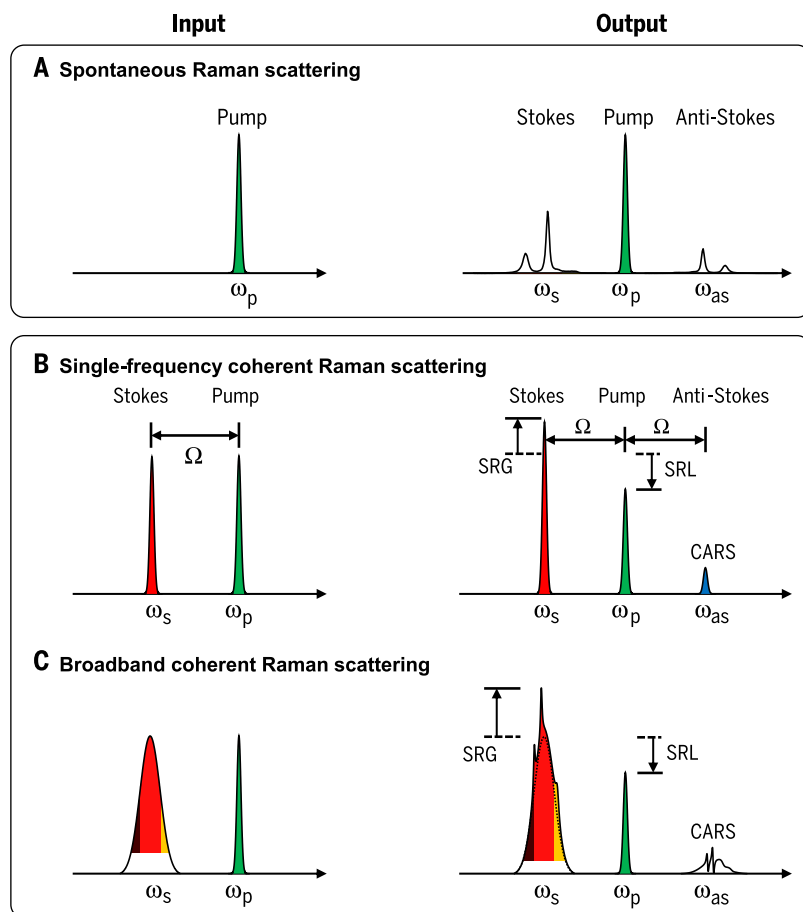
Fig. 1. Spontaneous versus coherent Raman scattering process. (**A**) A narrowband pump laser generates spontaneous Raman scattering. (**B**) Two narrowband lasers generate single-frequency coherent Raman scattering. (**C**) A narrowband laser and a broadband laser simultaneously excite multiple Raman transitions. SRG, stimulated Raman gain; SRL, stimulated Raman loss; CARS, coherent anti-Stokes Raman scattering. ω_p, ω_s, and ω_{as} denote the frequencies of the pump, Stokes beam, and anti-Stokes beam, respectively; Ω is the fundamental.

microscopy was developed for spectroscopic imaging (*14*, *15*). The strong signal contributed by C–H stretching vibrations has led to the application of CARS microscopy in lipid biology [reviewed in (*16*)].

The SRS phenomenon was reported in 1962 (*17*). Owyoung and co-workers developed SRS spectroscopy with continuous-wave lasers in the 1970s (*18*). Ploetz *et al*. demonstrated SRS imaging of polymer beads with a broadband, low–repetition rate laser in 2007 (*19*). In 2008, the Xie group reported high-speed, high-sensitivity single-frequency SRS imaging by megahertz modulation of a picosecond laser and modulation transfer to the other laser (*20*), and this work was quickly followed by reports from the Volkmer group, the Ozeki group, and others (*21–25*). The SRS signal can be extracted at the optical modulation frequency by either a lock-in amplifier (*20–25*) or a tuned amplifier (*26*). SRS is free of nonresonant background and provides spectral profiles that are nearly identical to spontaneous Raman spectroscopy. Moreover, the SRS intensity is linearly dependent on molecular concentration. These fea-

tures render SRS microscopy a robust and quantitative method for chemical imaging [reviewed in (*27*, *28*)].

The vibrational signal can also be detected indirectly with acoustic signatures generated by optical absorption. The photoacoustic effect, first documented by Bell in 1880 (*29*), offers an elegant route toward deep-tissue molecular imaging. When photons are absorbed by a molecule, part of the absorbed energy is converted into heat. The local heating results in a thermal expansion that creates a transient pressure wave detectable by an ultrasonic transducer. The time of flight of the photoacoustic signal carries the information about the location of the absorbers, and thus the distribution of the molecular absorbers can be mapped through image reconstruction. Photoacoustic tomography and microscopy have found wide biomedical applications based on the electronic absorption of hemoglobin [reviewed in (*30*, *31*)].

In an effort to overcome the limited imaging depth of CRS microscopy, which is on the scale of 100 μm, two photoacoustic vibrational imaging

platforms—one based on overtone absorption (*32*) and the other on SRS (*33*)—have been demonstrated. In the SRS-based platform, the energy difference of the pump and Stokes beam, $\hbar(\omega_p - \omega_s)$ (where $\hbar$ is Planck's constant divided by 2π), is absorbed by the interacted molecules (*34*) to generate acoustic waves (*33*, *35*). The SRS-based platform relies mostly on ballistic photons, which limits its potential for deep-tissue photoacoustic imaging. In overtone-based photoacoustic imaging (*36*), pulsed near-IR light induces overtone vibrational absorption. Through the heat dissipation process, such vibrational absorption is converted into a detectable transient acoustic wave. Unlike the strong mid-IR absorption by water in living organisms, the overtone vibrational bands are located in the near-IR region, where spectral windows of minimal water absorption exist for deep-tissue, bond-selective imaging (*32*). Moreover, unlike the SRS process, both diffuse and ballistic photons contribute to overtone absorption equally. This platform has led to the development of a high-speed intravascular vibrational photoacoustic (IVPA) catheter for label-free imaging of lipid-laden plaques (*37*) and vibration-based photoacoustic tomography (*38*) for label-free detection of breast cancer margin (*39*) and nerves (*40*). Besides the photoacoustic detection scheme, vibrational imaging of deep tissues has been pursued via spatially offset Raman spectroscopy (*41*), Raman tomography (*42*), and near-IR diffuse optical tomography (*43*).

Below, we discuss strategies contributed by different research groups that have pushed the boundary of the vibrational spectroscopic imaging field in terms of acquisition speed, detection sensitivity, spatial resolution, and imaging depth. We also highlight important applications enabled by these technical advances.

Pushing the limit of imaging speed

The imaging speed in spontaneous Raman, CARS, and SRS microscopy is dictated by the signal-to-noise ratio (SNR) at certain pixel dwell times. In all of these microscopic techniques, three major sources of noise are typically present: photon shot noise, laser intensity excess noise, and detector Johnson noise. Generally, Raman spectroscopy has a very low level of signal photons and the Johnson noise is the SNR-limiting factor. Thus, it takes milliseconds to seconds to record a Raman spectrum, and minutes to acquire a Raman image. CARS microscopy gains speed through coherent addition of the radiation fields, which renders the signal large and highly directional (*13*). The laser beams in CARS microscopy are usually not modulated. Thus, the laser intensity excess noise is transferred to the signal and becomes the SNR-limiting factor. In highly concentrated conditions, such as imaging lipid bodies by C–H stretch vibrations, a pixel dwell time of <1 μs is achieved. Experimentally, video-rate CARS imaging of skin lipids in living mice has been reported (*12*).

SRS microscopy gains speed through signal amplification by the presence of a local oscillator at the detection wavelength. SRS imaging usually involves intensity modulation of one laser and extraction of modulated signals from the other.

The shot noise in SRS microscopy is much greater than that in CARS. Meanwhile, because of the modulation and demodulation at MHz frequency, contribution from the excess laser noise is substantially reduced. Thus, the shot noise limit can be reached at a pixel dwell time of 100 ns, as demonstrated by video-rate single-frequency SRS imaging (*44*).

The above discussion considers CARS or SRS resonance with a single vibrational frequency. Without providing spectral information, this configuration is only applicable to mapping known species in a specimen. Hyperspectral CARS and SRS microscopy has been developed by sweeping the laser wavelength. A spectral resolution better than 10 cm^{-1} has been reached by femtosecond pulse-shaping technology (*45*, *46*). With these methods, it takes seconds to minutes to record an entire stack of images for reconstruction. Such speed is not sufficient to capture dynamics in a living system without spectral distortion.

Many research groups are developing multiplex CARS and SRS microscopy, in which a spectrum is instantaneously recorded at each pixel. First reported in 2002, multiplex CARS microscopy and microspectroscopy use a broadband and a narrowband beam to produce a CARS spectrum recorded by a spectrometer (*14*, *15*). Later, it was realized that the nonresonant background can be used as a local oscillator to enhance the resonant signal in multiplex CARS imaging (*47*). Two methods, maximum entropy (*48*) and Kramers-Kronig transformation (*49*), were adopted to extract the Raman spectrum from the CARS signal. Alternatively, with the addition of a third beam, interferometric CARS was developed to produce Raman spectral data at each pixel (*50*). Broadband CARS covering the entire Raman window was developed by using supercontinuum laser sources (*51*). Using this method, spectral acquisition on the order of a few milliseconds per pixel was recently achieved by Cicerone and co-workers (*51*).

In a further advance toward multiplex SRS detection, Fu *et al.* demonstrated multicolor imaging by modulation multiplexing and parallel detection of three spectral channels (*52*). Marx *et al.* compared various multichannel detectors that can be potentially used for SRS microscopy (*53*). Rock *et al.* reported the recovery of an SRS spectrum from a complementary metal-oxide semiconductor array with 20-ms integration time (*54*). Seto *et al.* developed a multiplex SRS microscope through multichannel lock-in detection, with a moderate detection sensitivity of 10^{-4} modulation (*55*). With a 32-channel resonant amplifier array, Liao *et al.* demonstrated lock-in free multiplex SRS imaging, within a spectral window defined by the femtosecond pulse, with pixel dwell time of 32 μs and detection sensitivity of 10^{-6} modulation (*56*). By spatial frequency multiplexing of a femtosecond pulse, Liao *et al.* further demonstrated microsecond-scale vibrational spectroscopic imaging by single photodiode detection of diffuse photons (*57*).

Along with efforts to improve instrumentation, various multivariate analysis methods were adopted to decompose the spectroscopic image into chemical maps of major components. For samples of known compositions, least-squares fitting provides a solid and rapid means to generate concentration maps (*58*). For samples with limited prior knowledge, principal components analysis (*59*), *k*-means clustering (*60*), independent components analysis (*45*), or spectral phasor analysis (*61*) can identify the major components in a CARS or SRS image. Concentration maps of these major components can be produced by further multivariate curve resolution analysis (*62*). Similar methods were developed by the Cicerone group (*63*) and the Langbein group (*64*).

At present, taking a CARS spectrum of the full window (3000 cm^{-1}) requires a few milliseconds per pixel. SRS spectra covering a window of hundreds of wave numbers can be recorded in microseconds, whereas single-frequency CARS or SRS signals can be acquired in nanoseconds.

Pushing the limit of detection sensitivity

With picosecond pulse excitation, single-frequency CARS microscopy has reached the detection sensitivity of ~1 million C-H bonds in a lipid membrane (*65*), which corresponds to millimolar molecular concentration in a femtoliter excitation volume. For weaker Raman bands, the CARS signal is often buried in a large nonresonant background contributed by the medium. Various advanced detection schemes have been developed to suppress the nonresonant background, including polarization-sensitive detection, time-resolved detection, interferometric CARS, and frequency modulation, but the complex configurations of these schemes have limited their application [reviewed in (*66*)]. Epi-detected CARS (*11*), which avoids the forward-propagating nonresonant signal contributed by the solvent, allows quantitative imaging of domains in a single supported lipid layer (*67*). Recently, the Potma group demonstrated surface-enhanced CARS detection of a single molecule located in a gold dumbbell by time-resolved measurement (*68*).

Under the shot noise–limited condition, the detection sensitivities of SRS and CARS are comparable (*21*). By excitation with a time-lens laser source synchronized with a femtosecond pulse, a hyperspectral SRS imaging sensitivity of 5×10^{-6} modulation depth for C-H vibrations in dimethyl sulfoxide (DMSO) was reached at a pixel dwell time of 2 μs, which corresponds to 19 mM DMSO in water (*69*). In hyperspectral SRS imaging, denoising (*70*) and subsequent multivariate data analysis allowed separation of spectroscopic signal from the unwanted cross-phase modulation, which improved the SNR in final chemical maps. For single-frequency SRS microscopy, advanced-phase or spectral modulation schemes were developed to remove the non-Raman background [reviewed in (*28*)].

Raman tags with large scattering cross sections offer an effective approach to improve molecular selectivity and boost detection sensitivity. Relative to fluorescent labels, Raman tags have smaller volumes and thus provide a promising way of imaging small biomolecules without perturbing their intracellular functions. Raman tags use vibrational signatures of the carbon-deuterium (C-D) bond, the cyano bond (C≡N), or the alkyne bond (C≡C), which are spectrally isolated from the endogenous Raman bands. The alkyne tag has allowed direct Raman visualization of DNA synthesis and mobile small molecules in living cells (*71*, *72*), at a speed of 50 min per frame of 127 × 127 pixels. CRS microscopy was deployed for faster imaging of C-D–labeled fatty acids (*23*, *73*), amino acids (*74*), and drugs (*75*) in living cells. More recently, SRS imaging of alkyne-tagged molecules has been reported (*76*, *77*), with detection sensitivity at the level of 200 μm.

Pushing the limit of spatial resolution

As nonlinear optical processes, CARS and SRS offer submicrometer lateral spatial resolution and an axial sectioning capability of ~1 μm. In order to study subcellular structures, partition of chemicals within an organelle, and cell functional activity, it is desirable to have higher imaging resolution. The use of two laser fields for signal generation opens opportunities to exceed the optical diffraction limit in CARS and SRS microscopy. Superresolution CARS imaging that achieved 130-nm lateral resolution for imaging nanostructured materials has been reported (*78*). By using ground-state population depletion, superior resolution was demonstrated through saturated transient absorption microscopy (*79*) and has been proposed for SRS microscopy (*80*). In this modality (*79*), a donut-shaped saturation beam was used to deplete the ground-state population, which restricted the pump-probe volume to the very center of the focus. Experimental demonstration of superresolution SRS imaging is expected in the near future.

Near-field scanning optical microscopy has achieved imaging resolution well below 100 nm and has been deployed for nanoscale vibrational imaging. By combining atomic force microscopy and IR absorption spectroscopy, Keilmann and co-workers used a near-field probe tip (rather than laser beam focusing) to obtain surface-enhanced contrast in vibrational absorption of chemical compounds (*81*). Here, a silicon cantilevered probing tip was metallized with a 100-nm layer of gold to achieve a diffraction-limited imaging resolution and enhanced IR absorption around the tip's apex. Application of this nanoscale spectroscopic absorption signal to life science has not yet been reported.

Pushing the limit of imaging depth

As the number of ballistic photons rapidly decreases with depth in a turbid tissue, current approaches of deep-tissue vibrational imaging exploit the interaction between diffuse photons and molecules to extract vibrational spectroscopic information. These methods are based on the detection of (i) overtone vibrational signal via near-IR diffuse optical tomography, (ii) vibrationally scattered photons through spatially offset Raman spectroscopy or Raman tomography, or (iii) the ultrasound signal induced by overtone vibrational absorption.

In near-IR diffuse optical tomography, photons interact with molecular overtone and
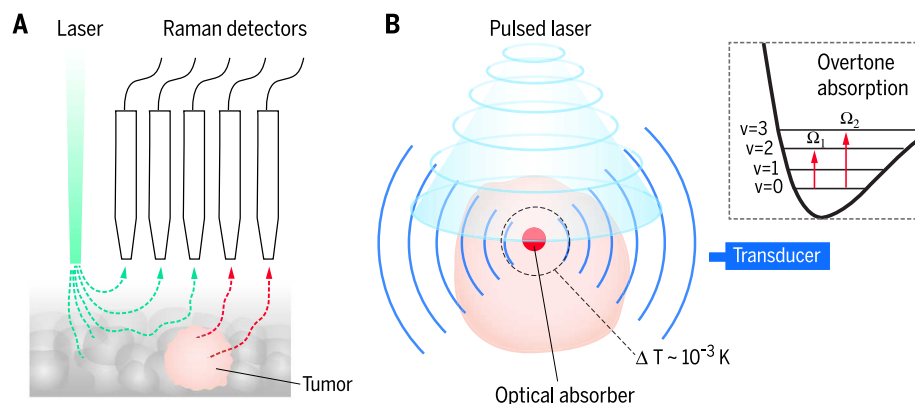
Fig. 2. Modalities for vibrational imaging of deep tissue. (**A**) Spatially offset Raman spectroscopy entails a spatial offset between the incident laser and the Raman signal probe, which allows detection of objects such as a tumor deep under the tissue surface. The green and red dotted lines indicate incident and Raman photons, respectively. (**B**) Vibration-based photoacoustic tomography is a modality in which a pulsed laser induces molecular overtone transitions inside a tissue. Subsequent relaxation of vibrational energy into heat generates acoustic waves detectable by an ultrasound transducer. Ω_1 and Ω_2 are the first and second overtone transition frequencies.

combinational vibrational transitions, which are allowed by the anharmonicity of chemical bond vibration (*82*). The transition frequency of an overtone band is described by $\Omega = \Omega_0 n - \chi\Omega_0(n + n^2)$, where Ω_0 is the frequency of a fundamental vibration, χ is the anharmonicity, and $n = 2, 3, \ldots$ represents the first overtone, second overtone, and so on. By modeling photon diffusion in turbid tissue, molecule-specific information can be obtained through approximately extracting absorption from the elastic scattering of photons (*83*). Spatial distribution of the molecule can be tomographically reconstructed by solving the "inverse problem" in a diffusion equation.

Unlike SRS, spontaneous Raman scattering is a linear inelastic scattering process and is less dependent on tight focusing for signal generation. Thus, diffuse photons that travel deep into the tissue could contribute substantially to a Raman spectrum. An intuitive way of extracting molecular information from different depths is to use ultrafast temporal gating (*84*, *85*), which measures the time-of-flight Raman signal generated from diffused photons. On the basis of these studies, Matousek and co-workers demonstrated spatially offset Raman spectroscopy as an effective and relatively simple method to obtain molecular information from deep tissue (*41*) (Fig. 2A). In this method, when collecting a Raman signal from a surface position that is laterally offset from the position of excitation, a larger portion of backscattered Raman photons from a deeper layer is detected rather than those from an upper layer (*41*). With modeling of light propagation and multivariate analysis of the Raman spectrum, molecule-specific information at a tissue depth of a few millimeters can be extracted (*86*–*89*). Raman tomography (*90*, *91*), developed by Morris and Pogue, works in a manner similar to diffuse optical imaging, where the propagation of the inelastically scattered photons is described by diffusion equations. By detecting the

diffusely propagated Raman-scattered photons in a 360° plane, the spatial distribution of molecules at depths greater than 1 cm can be reconstructed (*92*, *93*).

Excited by near-IR pulsed lasers, the molecular vibrational energy is quickly relaxed into heat and converted into mechanical energy via tissue expansion and retraction. The generated pressure waves can then be detected with an acoustic transducer (Fig. 2B). This overtone-based photoacoustic imaging method has been used for mapping objects that are rich in C-H bonds (*32*, *36*, *94*). The second and first overtones of C-H stretch vibration—located at ~1210 nm and ~1730 nm, respectively—were chosen by taking advantage of their relatively large absorption coefficients (*36*, *95*–*97*). As the wavelength increases, water absorption starts to become a factor and attenuates the light that can arrive to deep tissue. Thus, selection of the spectral window is critical to maximize the contrast of C-H bonds over O-H bonds while maintaining a certain imaging depth (*98*). Providing a stronger photoacoustic signal, the first overtone of C-H bonds is favorable in the applications where millimeter-scale penetration depth is needed. When dealing with applications requiring a penetration depth greater than a few millimeters, the second overtone becomes a better choice (*96*–*98*).

Shedding light on cellular machinery

CRS microscopy provides submicrometer resolution and can resolve intracellular organelles, which is especially important for single-cell studies. Small biomolecules including amino acids, nucleic acids, lipids, and carbohydrates play essential roles in cell proliferation and function. Because of their small sizes, their activity inside cells would be considerably changed by the bulky fluorescent labels. Applications of CARS and SRS microscopy have contributed to new understanding of lipid droplet biology, myelin biology, cell metabolism, cell cycle,

drug delivery, and tissue organization, as summarized below. Although we focus on mammalian cells and tissues here, we note that SRS microscopy has also been applied to botanical science (*99*, *100*).

In a label-free manner, CARS and SRS microscopy have been widely used to study two important organelles—lipid droplets and the myelin sheath—that generate strong signals from C-H stretching vibrations. In 2004, Nan *et al.* reported CARS imaging of the process of adipogenesis in live 3T3-L1 cells (*101*). In 2007, Hellerer *et al.* reported CARS imaging of fat storage in live *Caenorhabditis elegans* (*102*). These studies were followed by multiple studies on lipid droplet biology, including trafficking (*103*), lipolysis (*104*), growth (Fig. 3A) (*105*), and dietary fat absorption through the small intestine (*16*). Without interference with non-resonant background, SRS microscopy provides a new way to study lipid metabolism in model organisms. Wang *et al.* coupled SRS microscopy and RNA interference screening to identify genetic regulators of fat storage in *C. elegans* (*106*). Dou *et al.* found active transport of lipid droplets in early fruit fly embryos by femtosecond pulse–stimulated Raman-loss imaging (*107*). Using hyperspectral SRS imaging of fingerprint Raman bands, Wang *et al.* resolved fat droplets, cholesterol-rich lysosomes, and lipid oxidation in wild-type and mutant worms (*108*). Multiplex SRS microscopy allowed separation of fat droplets from protein-rich organelles, using their spectroscopic signatures of the CH_2 bending mode (*56*). By integrating single-frequency CARS microscopy with Raman microspectroscopy for compositional analysis (*109*) (Fig. 3B), Yue *et al.* reported that lipid droplets in human patient prostate cancer cells are rich in cholesteryl ester (*110*). This finding presents a new way to treat aggressive cancers by abrogating the activity of cholesterol esterification.

The myelin sheath, a multilamellar membrane surrounding the axon, is responsible for rapid propagation of action potentials in a saltatory manner. Myelin sheath degradation leads to serious disorders such as multiple sclerosis. The myelin sheath is abundant in lipids and produces a strong CARS signal that allows label-free imaging of myelin in its natural tissue environment (*111*, *112*). The 3D sectioning capability of CARS allowed visualization of detailed myelin structure, such as the node of Ranvier (*113*) (Fig. 3C). Different groups have used CARS microscopy to characterize myelin degradation and regeneration under different conditions of diseases and/or white matter injury (*111*, *113*, *114*). CARS has been further used to longitudinally monitor demyelination and remyelination in an injured spinal cord in vivo (*113*). More recently, Hu *et al.* demonstrated in vivo SRS imaging of single myelin sheaths in early-stage tadpoles (*115*). These studies collectively demonstrate a CRS-based platform for mechanistic study of myelin diseases.

Both CARS and SRS microscopy have enabled dynamic imaging of metabolic conversions in single living cells. By hyperspectral SRS imaging of fingerprint bands and multivariate analysis, Wang *et al.* monitored multiple metabolites in

small intestine cells of live *C. elegans* (*108*) (Fig. 4A). Liao *et al.* observed the conversion of retinol into retinoic acids in cancer cells (*56*). Isotope substitution has been used to selectively monitor the fate of specific metabolites. The C-D bond produces Raman peaks around 2100 cm^{-1}, spectrally isolated from endogenous Raman modes. SRS imaging of deuterated fatty acids administered to live cells revealed that oleic acids facilitate the conversion of palmitic acid into neutral lipid stored in lipid droplets (*23*). Min and co-workers used deuterated amino acids as substrates for mapping protein synthesis by SRS microscopy (*74*) (Fig. 4B). They further extended this work to monitor the process of protein degradation using ^{13}C-labeled phenylalanine (*116*). By using deuterated glucose, Li *et al.* demonstrated direct visualization of de novo lipogenesis in pancreatic cancer cells, which occurs at a much lower rate in immortalized normal pancreatic epithelial cells (*117*). A similar strategy has been used to monitor choline metabolism at the single-cell level (*118*). The alkyne-based Raman tags open a new avenue to monitor DNA and RNA synthesis and drug distribution in living cells (*76*, *77*). By synthesis of phenol-diyne cholesterol, Lee *et al.* reported SRS imaging of cholesterol storage in *C. elegans* (Fig. 4C) and lysosomal storage of cholesterol in a cellular model of Niemann-Pick type C disease, with a detection sensitivity of 31 μM phenol-diyne cholesterol (*119*).

High-speed CRS microscopy has enabled chemical imaging of cell mitosis and apoptosis. In 2002, CARS microscopy was used to detect cell apoptosis according to signals from the plasma membrane and nucleus (*10*). Such capability has been amplified by selective imaging of lipids and proteins based on signals from CH$_2$ and CH$_3$ vibrations (*120*). Recently, SRS microscopy allowed real-time monitoring of the cell cycle based on selective imaging of DNA, using its distinctive signature in the high–wave number C-H vibration region (*121*) (Fig. 4D).

CRS microscopy is becoming a powerful tool for mapping spatial and temporal dynamics of drug molecules for which fluorescent labeling would alter their physical properties. A number of groups applied CARS and SRS to map the distribution and follow the release of active pharmaceutical ingredients (Fig. 4E) (*122*–*125*). Guy and co-workers monitored molecular diffusion into the human nail (*126*). Fu *et al.* applied hyperspectral SRS microscopy to visualize drug distribution inside cells using fingerprint Raman bands (*127*).

CRS microscopy has further enabled chemical imaging of tissues with 3D sectioning capability. Three major advances have been made: (i) CARS-based multimodal nonlinear optical microscopy has been developed to visualize multiple species in a tissue environment (*128*); (ii) histology-mimicking two-color contrast has been achieved by combined imaging of CH$_2$ vibration for lipids and CH$_3$ vibration for proteins (*129*); and (iii) hyperspectral CARS and SRS microscopy has allowed chemical imaging of multiple molecules using spectrally overlapped fingerprint Raman

bands (*49*, *64*, *108*). These studies laid the groundwork for the development of mobile imaging devices that would be useful for in vivo diagnosis applications in clinical settings.

Moving into the clinic

High-speed, high-resolution vibrational spectroscopic imaging is expected to affect the current paradigm of medical practice by enabling molecule-based diagnosis without the need for any contrast agent. Generally, CRS microscopy provides submicrometer spatial resolution, together with an imaging depth and a field of view both on the order of 100 μm. On the basis of these characteristics, CARS and SRS microscopy are considered to be suitable for in vivo imaging of skin or exposed tissue.

Toward this direction, a few groups have reported preliminary results in developing CARS and SRS microendoscopy (*130*–*132*). For clinical applications, the integration of CARS with multiphoton fluorescence and second-harmonic generation modalities allowed microscopic examination of lipids, endogenous fluorophores, and collagen fibers in human skin (*133*). Koenig *et al.* demonstrated clinical CARS imaging to provide label-free in vivo skin biopsy (*134*, *135*) (Fig. 5A). Potma and co-workers applied SRS microscopy to identify squamous cell carcinoma in human skin (*136*) (Fig. 5B). Ji *et al.* applied two-color SRS microscopy to map CH$_2$ and CH$_3$ vibrations at Raman shifts of 2845 and 2930 cm^{-1} (*137*). Using a linear recombination method, they extracted maps of lipids and proteins that reflected the proper distribution of cellular bodies and extracellular structures. This two-color SRS imaging method was also applied in vivo in mice to reveal brain tumor margins that were not detectable under standard operative conditions (*137*) (Fig. 5C). We note that fiber-optic Raman spectroscopy has been used for intraoperative brain tumor detection during surgery on a human patient (*138*). Relative to single-point Raman spectroscopy, SRS microscopy provides visual information that resembles hematoxylin and eosin–based histology (*110*, *129*).

Various deep-tissue vibrational imaging methods are opening new ways for minimally invasive diagnosis of human diseases. By photoacoustic detection of C-H bonds' overtone transition, several groups are developing IVPA catheters for visualization of lipid deposition inside the blood vessel wall, with an imaging depth of a few millimeters (*32*, *95*, *139*). Multivariate analysis of the overtone spectrum has been applied to extract chemically specific information. Differentiation of CH$_2$-rich fat and CH$_3$-rich collagen in atherosclerotic arteries was achieved by vibrational photoacoustic imaging through the distinctive spectral profiles of CH$_2$ and CH$_3$ groups (*36*, *96*, *97*). Clinical translation of IVPA imaging has been hampered by the lack of suitable high–repetition rate lasers for excitation of the first overtone of C-H bonds at 1730 nm or the second overtone around 1200 nm. By shifting the wavelength from a 1064-nm YAG laser via the SRS process, a Raman laser offers a solution

Fig. 3. Biological applications of coherent anti-Stokes Raman scattering (CARS) microscopy. (**A**) Real-time CARS imaging reveals neutral lipid depletion in 3T3-L1 adipocytes. Left: 0 min; right: 2.5 hours. Scale bars, 10 μm. (**B**) Integrating coherent Raman imaging with spontaneous Raman spectroscopy. Left: CARS image of lipid bodies in a living 3T3-L1 cell. Right: Raman spectrum of a single lipid droplet (marked with a cross in the CARS image) reveals its chemical composition. Scale bar, 20 μm. (**C**) Left and center: Ex vivo CARS imaging reveals chemically induced myelin sheath swelling in a spinal cord tissue. Right: In vivo CARS image of a node of Ranvier inside the white matter of a rat spinal cord. Scale bars, 10 μm.
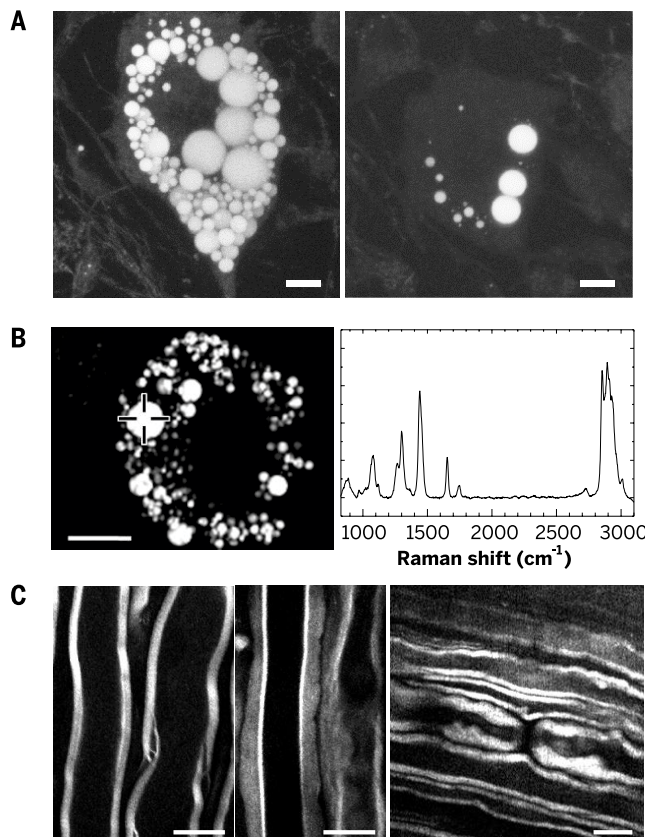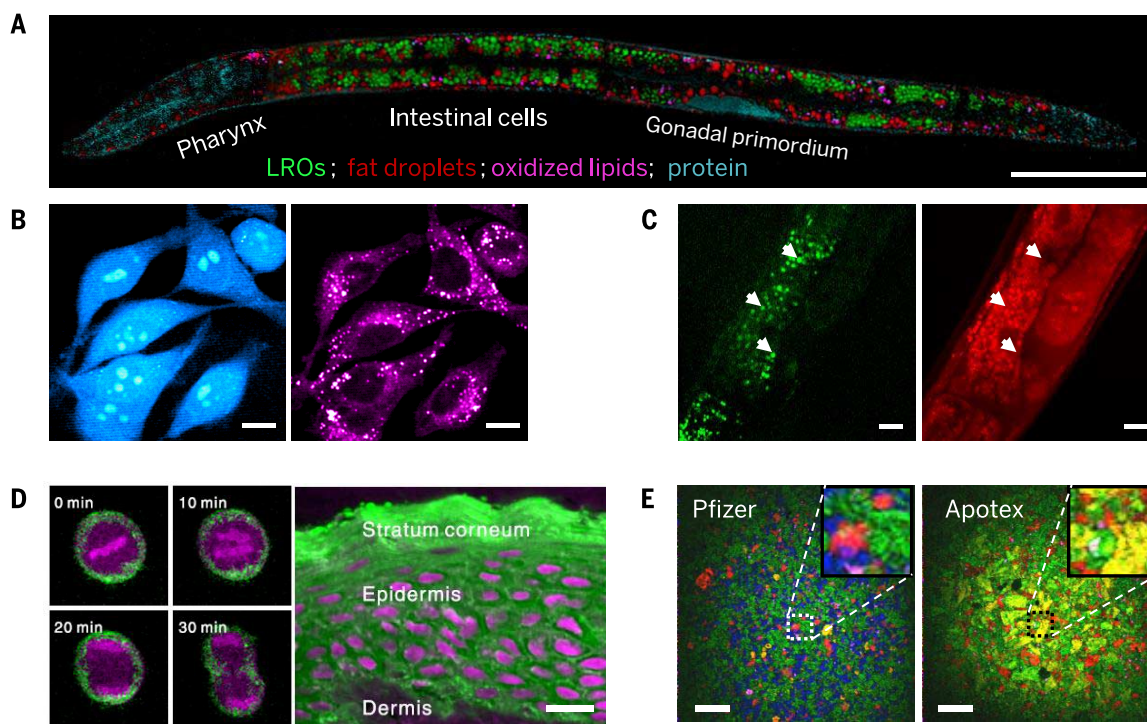
Fig. 4. Biological applications of stimulated Raman scattering (SRS) microscopy. (**A**) Chemical maps of intracellular compartments in *C. elegans* generated by hyperspectral SRS imaging and multivariate curve resolution analysis. The green, red, magenta, and cyan colors represent lysosome-related organelles (LROs), neutral fat droplets, oxidized lipids, and proteins, respectively, in the body of a *daf-2* mutant. Scale bar, 50 μm. (**B**) SRS imaging of live HeLa cells using signals from C-D vibration at 2133 cm^{-1} (left) and C-H vibration at 2845 cm^{-1} (right). The C-D signal shows newly synthesized proteins by metabolic incorporation of a deuterium-labeled set of amino acids. Scale bars, 10 μm. (**C**) SRS imaging of phenyl-diyne cholesterol in live *C. elegans*. Left: SRS image at C≡C vibrational mode at 2247 cm^{-1} reveals a distinct cholesterol store in the worm. Right: SRS image at 2885 cm^{-1} reveals C-H–rich fat stores. Scale bars, 10 μm. (**D**) Left: Time-lapse SRS images of DNA (magenta) and lipids (green) in a HeLa cell undergoing cell division. Right: SRS images of DNA (magenta) and lipids (green) in a normal human skin tissue section. Scale bar, 20 μm. (**E**) Large-area SRS images of drug tablets from Pfizer (left) and Apotex (right). Green, blue, red, yellow/orange, and magenta colors represent microcrystalline cellulose, dibasic calcium phosphate anhydrous, amlodipine besylate (drug), sodium starch glycolate, and magnesium stearate, respectively. Scale bars, 200 μm.

for vibrational photoacoustic imaging of lipids (*140*). By constructing a 2-kHz Raman laser at 1197 nm, Wang *et al.* demonstrated high-speed IVPA imaging of lipid-laden plaque in an intact artery at a speed of 1 frame per second (Fig. 5D) (*37*). Development of kilohertz lasers at 1730 nm is important for in vivo IVPA imaging.

In the tomography domain, Tromberg and co-workers deployed molecular vibration–based diffuse optical imaging for quantitation of water and lipid content within turbid tissue (*141, 142*). In the optical window typically from 600 to 1000 nm, lipid has a peak at ~930 nm, which is assigned to the third overtone of the CH$_2$ stretching mode, and water has a peak at ~980 nm, which is assigned to the combination of symmetric and asymmetric stretching modes of the O–H bond (*82*). The quantification of water and lipid according to their absorption signature has been used to determine the malignancy of breast tissue in vivo, where an imaging depth of a few centimeters has been achieved (*143*). With high chemical specificity, both spatially offset Raman spectroscopy and Raman tomography have shown promise as tools for breast cancer margin assessment and bone characterization (*91, 93, 144–146*). Coupling overtone absorption of lipids with elec-

tronic absorption contrast from hemoglobin, photoacoustic tomography was applied to breast tumor margin assessment (Fig. 5E) and peripheral nerve visualization (*39, 40*).

Outlook

Through integration with advances in other fields, new technical breakthroughs are expected that will further increase detection sensitivity, acquisition speed, imaging depth, and spatial resolution. The current detectable concentration of CARS or SRS microscopy is above 1.0 mM for endogenous molecules. Such sensitivity prohibits wide biological applications of coherent Raman microscopy, because most biomolecules inside cells are at a nanomolar to micromolar level. Because CARS and SRS microscopy have reached the shot noise limit, a possible way of improving the detection sensitivity could be through non-optical detection. For example, through force measurement of SRS-induced chemical bond vibration, single-molecule sensitivity has been reported (*147*). Extension of this method to biological imaging remains to be demonstrated. If single-molecule detection sensitivity is reached for vibrational imaging of living cells, then we can, in principle, reach nanoscale spatial resolution as

in fluorescence microscopy. Along the direction of high-speed imaging, with parallel modulation schemes, frequency-multiplexed SRS could potentially reach a speed of 1000 frames per second to enable monitoring of very fast processes. Along the direction of deeper imaging, integrating vibrational spectroscopy with techniques that enable active focusing of light through a turbid tissue, such as time reversal of ultrasound-coded light (*148, 149*), might allow deep-tissue vibrational imaging for in vivo diagnosis.

For single-cell functional analysis, we expect high-speed vibrational spectroscopic imaging to be applied in studies of cell activities. Examples include monitoring the propagation of action potential using intrinsic vibrational signal from the neuronal membrane, following the transport and release of neurotransmitters, and measuring changes of protein conformation inside smooth muscle cells under mechanical stress.

Finally, we believe that cooperation among physicists, engineers, entrepreneurs, and clinicians is important for clinical translation of in vivo vibrational imaging technologies. We expect advances in the reduction of technology cost and instrument size. Progress in ultrafast fiber laser technology, although not covered in this review,
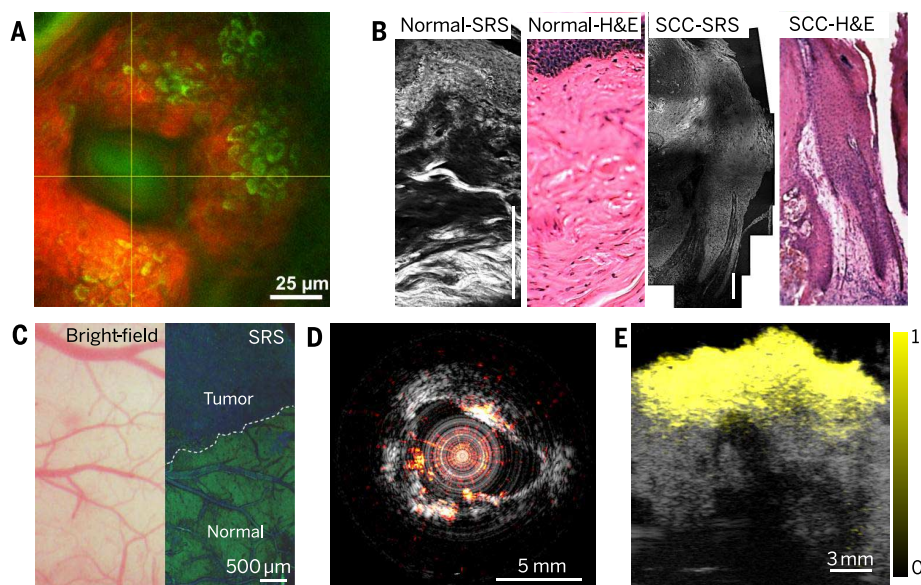
Fig. 5. Clinical applications of vibrational spectroscopic imaging. (**A**) In vivo cellular imaging shows structures of an excretory duct of the sweat gland from healthy human skin. Green, two-photon excitation fluorescence; red, CARS. (**B**) From left to right: Stimulated Raman scattering (SRS) image of healthy human skin; hematoxylin and eosin (H&E)–stained specimen of an adjacent healthy skin section; SRS image of superficial squamous cell carcinoma (SCC); H&E-stained specimen of an adjacent SCC section. Scale bars, 100 μm. (**C**) Bright-field (left) and SRS (right) images of a margin between a brain tumor and normal brain tissue in a human glioblastoma multiforme xenografted mouse. The contrast in the SRS image provides label-free detection of the tumor margin. (**D**) Intravascular photoacoustic (fire color) and ultrasound (gray) imaging of an atherosclerotic artery. The photoacoustic signal shows the distribution of lipids deposited inside the arterial wall. (**E**) Photoacoustic (yellow) and ultrasound (gray) tomographic imaging of a human breast tumor tissue section. The photoacoustic signal arising from C-H overtone absorption demonstrates the lipid-rich tumor margin.

promises to enable mobile CARS and SRS imaging systems. Invenio Imaging Inc. has produced a fiber laser–based commercial system for SRS imaging (*150*). Likewise, nanosecond lasers with kilohertz repetition rates have shown promise for high-speed intravascular imaging of lipid-laden plaques (*37*). These efforts are expected to convert vibrational spectroscopic imaging platforms into fundamental clinical tools for in vivo detection of residual disease or longitudinal evaluation of therapy effectiveness.

REFERENCES AND NOTES

1. R. Bhargava, Infrared spectroscopic imaging: The next generation. *Appl. Spectrosc.* **66**, 1091–1120 (2012). doi: 10.1366/12-06801; pmid: 23031693
2. Z. Movasaghi, S. Rehman, I. U. Rehman, Raman spectroscopy of biological tissues. *Appl. Spectrosc. Rev.* **42**, 493–541 (2007). doi: 10.1080/05704920701551530
3. A. F. Palonpon et al., Raman and SERS microscopy for molecular imaging of live cells. *Nat. Protoc.* **8**, 677–692 (2013). doi: 10.1038/nprot.2013.030; pmid: 23471112
4. J.-X. Cheng, X. S. Xie, *Coherent Raman Scattering Microscopy* (CRC Press, Boca Raton, FL, 2013).
5. R. W. Terhune, P. D. Maker, C. M. Savage, Measurements of nonlinear light scattering. *Phys. Rev. Lett.* **14**, 681–684 (1965). doi: 10.1103/PhysRevLett.14.681
6. R. Begley, A. Harvey, R. L. Byer, Coherent anti-Stokes Raman spectroscopy. *Appl. Phys. Lett.* **25**, 387 (1974). doi: 10.1063/1.1655519
7. M. D. Duncan, J. Reintjes, T. J. Manuccia, Scanning coherent anti-Stokes Raman microscope. *Opt. Lett.* **7**, 350–352 (1982). doi: 10.1364/OL.7.000350; pmid: 19714017
8. A. Zumbusch, G. R. Holtom, X. S. Xie, Three-dimensional vibrational imaging by coherent anti-Stokes Raman

scattering. *Phys. Rev. Lett.* **82**, 4142–4145 (1999). doi: 10.1103/PhysRevLett.82.4142
9. J.-X. Cheng, L. D. Book, X. S. Xie, Polarization coherent anti-Stokes Raman scattering microscopy. *Opt. Lett.* **26**, 1341–1343 (2001). doi: 10.1364/OL.26.001341; pmid: 18049602
10. J.-X. Cheng, Y. K. Jia, G. Zheng, X. S. Xie, Laser-scanning coherent anti-Stokes Raman scattering microscopy and applications to cell biology. *Biophys. J.* **83**, 502–509 (2002). doi: 10.1016/S0006-3495(02)75186-2; pmid: 12080137
11. A. Volkmer, J.-X. Cheng, X. S. Xie, Vibrational imaging with high sensitivity via epidetected coherent anti-Stokes Raman scattering microscopy. *Phys. Rev. Lett.* **87**, 023901 (2001). doi: 10.1103/PhysRevLett.87.023901
12. C. L. Evans et al., Chemical imaging of tissue in vivo with video-rate coherent anti-Stokes Raman scattering microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 16807–16812 (2005). doi: 10.1073/pnas.0508282102; pmid: 16263923
13. J.-X. Cheng, X. S. Xie, Theoretical and experimental characterization of coherent anti-Stokes Raman scattering microscopy. *J. Opt. Soc. Am. B* **19**, 1363 (2002). doi: 10.1364/JOSAB.19.001363
14. J.-X. Cheng, A. Volkmer, L. D. Book, X. S. Xie, Multiplex coherent anti-Stokes Raman scattering microspectroscopy and study of lipid vesicles. *J. Phys. Chem. B* **106**, 8493–8498 (2002). doi: 10.1021/jp025771z
15. M. Müller, J. M. Schins, Imaging the thermodynamic state of lipid membranes with multiplex CARS microscopy. *J. Phys. Chem. B* **106**, 3715–3723 (2002). doi: 10.1021/jp014012y
16. T. T. Le, S. Yue, J.-X. Cheng, Shedding new light on lipid biology with coherent anti-Stokes Raman scattering microscopy. *J. Lipid Res.* **51**, 3091–3102 (2010). doi: 10.1194/jlr.R008730; pmid: 20713649
17. E. J. Woodbury, W. K. Ng, Ruby operation in the near IR. *Proc. Inst. Radio Eng.* **50**, 2367 (1962).
18. A. Owyoung, E. D. Jones, Stimulated Raman spectroscopy using low-power cw lasers. *Opt. Lett.* **1**, 152–154 (1977). doi: 10.1364/OL.1.000152; pmid: 19680362

19. E. Ploetz, S. Laimgruber, S. Berner, W. Zinth, P. Gilch, Femtosecond stimulated Raman microscopy. *Appl. Phys. B* **87**, 389–393 (2007). doi: 10.1007/s00340-007-2630-x
20. C. W. Freudiger et al., Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science* **322**, 1857–1861 (2008). doi: 10.1126/science.1165758; pmid: 19095943
21. Y. Ozeki, F. Dake, S. Kajiyama, K. Fukui, K. Itoh, Analysis and experimental assessment of the sensitivity of stimulated Raman scattering microscopy. *Opt. Express* **17**, 3651–3658 (2009). doi: 10.1364/OE.17.003651; pmid: 19259205
22. P. Nandakumar, A. Kovalev, A. Volkmer, Vibrational imaging based on stimulated Raman scattering microscopy. *New J. Phys.* **11**, 033026 (2009). doi: 10.1088/1367-2630/11/3/033026
23. D. Zhang, M. N. Slipchenko, J.-X. Cheng, Highly sensitive vibrational imaging by femtosecond pulse stimulated Raman loss. *J. Phys. Chem. Lett.* **2**, 1248–1253 (2011). doi: 10.1021/jz200516n; pmid: 21731798
24. E. R. Andresen, P. Berto, H. Rigneault, Stimulated Raman scattering microscopy by spectral focusing and fiber-generated soliton as Stokes pulse. *Opt. Lett.* **36**, 2387–2389 (2011). doi: 10.1364/OL.36.002387; pmid: 21725420
25. H. T. Beier, G. D. Noojin, B. A. Rockwell, Stimulated Raman scattering using a single femtosecond oscillator with flexibility for imaging and spectral applications. *Opt. Express* **19**, 18885–18892 (2011). doi: 10.1364/OE.19.018885; pmid: 21996830
26. M. N. Slipchenko, R. A. Oglesbee, D. Zhang, W. Wu, J.-X. Cheng, Heterodyne detected nonlinear optical imaging in a lock-in free manner. *J. Biophotonics* **5**, 801–807 (2012). doi: 10.1002/jbio.201200005; pmid: 22389310
27. W. Min, C. W. Freudiger, S. Lu, X. S. Xie, Coherent nonlinear optical imaging: Beyond fluorescence microscopy. *Annu. Rev. Phys. Chem.* **62**, 507–530 (2011). doi: 10.1146/annurev.physchem.012809.103512; pmid: 21453061
28. D. Zhang, P. Wang, M. N. Slipchenko, J.-X. Cheng, Fast vibrational imaging of single cells and tissues by stimulated Raman scattering microscopy. *Acc. Chem. Res.* **47**, 2282–2290 (2014). doi: 10.1021/ar400331q; pmid: 24871269
29. A. G. Bell, On the production and reproduction of sound by light. *Am. J. Sci.* **20**, 305–324 (1880). doi: 10.2475/ajs.s3-20.118.305
30. V. Ntziachristos, Going deeper than microscopy: The optical imaging frontier in biology. *Nat. Methods* **7**, 603–614 (2010). doi: 10.1038/nmeth.1483; pmid: 20676081
31. L. V. Wang, S. Hu, Photoacoustic tomography: In vivo imaging from organelles to organs. *Science* **335**, 1458–1462 (2012). doi: 10.1126/science.1216210; pmid: 22442475
32. H.-W. Wang et al., Label-free bond-selective imaging by listening to vibrationally excited molecules. *Phys. Rev. Lett.* **106**, 238106 (2011). doi: 10.1103/PhysRevLett.106.238106; pmid: 21770549
33. V. V. Yakovlev et al., Stimulated Raman photoacoustic imaging. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 20335–20339 (2010). doi: 10.1073/pnas.1012432107; pmid: 21059930
34. H. Wang, Y. Fu, J.-X. Cheng, Experimental observation and theoretical analysis of Raman resonance-enhanced photodamage in coherent anti-Stokes Raman scattering microscopy. *J. Opt. Soc. Am. B* **24**, 544 (2007). doi: 10.1364/JOSAB.24.000544
35. V. V. Yakovlev, G. D. Noojin, M. L. Denton, B. A. Rockwell, R. J. Thomas, Monitoring stimulated Raman scattering with photoacoustic detection. *Opt. Lett.* **36**, 1233–1235 (2011). doi: 10.1364/OL.36.001233; pmid: 21479040
36. P. Wang, J. R. Rajian, J.-X. Cheng, Spectroscopic imaging of deep tissue through photoacoustic detection of molecular vibration. *J. Phys. Chem. Lett.* **4**, 2177–2185 (2013). doi: 10.1021/jz400559a; pmid: 24073304
37. P. Wang et al., High-speed intravascular photoacoustic imaging of lipid-laden atherosclerotic plaque enabled by a 2-kHz barium nitrite raman laser. *Sci. Rep.* **4**, 6889 (2014). doi: 10.1038/srep06889; pmid: 25366991
38. J. R. Rajian, R. Li, P. Wang, J.-X. Cheng, Vibrational photoacoustic tomography: Chemical imaging beyond the ballistic regime. *J. Phys. Chem. Lett.* **4**, 3211–3215 (2013). doi: 10.1021/jz401638e; pmid: 24224071
39. R. Li et al., Assessing breast tumor margin by multispectral photoacoustic tomography. *Biomed. Opt. Express* **6**, 1273–1281 (2015). doi: 10.1364/BOE.6.001273; pmid: 25909011

40. R. Li, E. Phillips, P. Wang, C. J. Goergen, J.-X. Cheng, Label-free in vivo imaging of peripheral nerve by multispectral photoacoustic tomography. *J. Biophoton.* 10.1002/jbio.201500004 (2015). doi: 10.1364/BOE.6.001273; pmid: 25909011

41. P. Matousek et al., Subsurface probing in diffusely scattering media using spatially offset Raman spectroscopy. *Appl. Spectrosc.* **59**, 393–400 (2005). doi: 10.1366/0003702053641450; pmid: 15901323

42. M. V. Schulmerich et al., Noninvasive Raman tomographic imaging of canine bone tissue. *J. Biomed. Opt.* **13**, 020506 (2008). doi: 10.1117/1.2904940; pmid: 18465948

43. A. H. Hielscher et al., Near-infrared diffuse optical tomography. *Dis. Markers* **18**, 313–337 (2002). doi: 10.1155/2002/164252; pmid: 14646043

44. B. G. Saar et al., Video-rate molecular imaging in vivo with stimulated Raman scattering. *Science* **330**, 1368–1370 (2010). doi: 10.1126/science.1197236; pmid: 21127249

45. Y. Ozeki et al., High-speed molecular spectral imaging of tissue with stimulated Raman scattering. *Nat. Photonics* **6**, 845–851 (2012). doi: 10.1038/nphoton.2012.263

46. B. Liu et al., Vibrational fingerprint mapping reveals spatial distribution of functional groups of lignin in plant cell wall. *Anal. Chem.* **87**, 9436–9442 (2015). doi: 10.1021/acs.analchem.5b02434; pmid: 26291845

47. M. Müller, A. Zumbusch, Coherent anti-Stokes Raman scattering microscopy. *ChemPhysChem* **8**, 2156–2170 (2007). doi: 10.1002/cphc.200700202; pmid: 17768730

48. E. M. Vartiainen, H. A. Rinia, M. Müller, M. Bonn, Direct extraction of Raman line-shapes from congested CARS spectra. *Opt. Express* **14**, 3622–3630 (2006). doi: 10.1364/OE.14.003622; pmid: 19516509

49. M. T. Cicerone, K. A. Aamer, Y. J. Lee, E. Vartiainen, Maximum entropy and time-domain Kramers–Kronig phase retrieval approaches are functionally equivalent for CARS microspectroscopy. *J. Raman Spectrosc.* **43**, 637–643 (2012). doi: 10.1002/jrs.3169

50. D. L. Marks, S. A. Boppart, Nonlinear interferometric vibrational imaging. *Phys. Rev. Lett.* **92**, 123905 (2004). doi: 10.1103/PhysRevLett.92.123905; pmid: 15089675

51. C. H. Camp Jr. et al., High-speed coherent Raman fingerprint imaging of biological tissues. *Nat. Photonics* **8**, 627–634 (2014). doi: 10.1038/nphoton.2014.145; pmid: 25621002

52. D. Fu et al., Quantitative chemical imaging with multiplex stimulated Raman scattering microscopy. *J. Am. Chem. Soc.* **134**, 3623–3626 (2012). doi: 10.1021/ja210081h; pmid: 22316340

53. B. Marx, L. Czerwinski, R. Light, M. Somekh, P. Gilch, Multichannel detectors for femtosecond stimulated Raman microscopy–ideal and real ones. *J. Raman Spectrosc.* **45**, 521–527 (2014). doi: 10.1002/jrs.4528

54. W. Rock, M. Bonn, S. H. Parekh, Near shot-noise limited hyperspectral stimulated Raman spectroscopy using low energy lasers and a fast CMOS array. *Opt. Express* **21**, 15113–15120 (2013). doi: 10.1364/OE.21.015113; pmid: 23842298

55. K. Seto, Y. Okuda, E. Tokunaga, T. Kobayashi, Development of a multiplex stimulated Raman microscope for spectral imaging through multi-channel lock-in detection. *Rev. Sci. Instrum.* **84**, 083705 (2013). doi: 10.1063/1.4818670; pmid: 24007071

56. C.-S. Liao et al., Microsecond scale vibrational spectroscopic imaging by multiplex stimulated Raman scattering microscopy. *Light Sci. Appl.* **4**, e265 (2015). doi: 10.1038/lsa.2015.38; pmid: 26167336

57. C.-S. Liao et al., Spectrometer-free vibrational imaging by retrieving stimulated Raman signal from highly scattered photons. *Sci. Advances* **1**, e1500738 (2015). doi: 10.1126/sciadv.1500738

58. D. Fu, G. Holtom, C. Freudiger, X. Zhang, X. S. Xie, Hyperspectral imaging with stimulated Raman scattering by chirped femtosecond lasers. *J. Phys. Chem. B* **117**, 4634–4640 (2013). doi: 10.1021/jp308938t; pmid: 23256635

59. C.-Y. Lin et al., Picosecond spectral coherent anti-Stokes Raman scattering imaging with principal component analysis of meibomian glands. *J. Biomed. Opt.* **16**, 021104 (2011). doi: 10.1117/1.3533716; pmid: 21361667

60. C. Krafft et al., A comparative Raman and CARS imaging study of colon tissue. *J. Biophotonics* **2**, 303–312 (2009). doi: 10.1002/jbio.200810063; pmid: 19434617

61. D. Fu, X. S. Xie, Reliable cell segmentation based on spectral phasor analysis of hyperspectral stimulated Raman scattering imaging data. *Anal. Chem.* **86**, 4115–4119 (2014). doi: 10.1021/ac500014b; pmid: 24684208

62. D. Zhang et al., Quantitative vibrational imaging by hyperspectral stimulated Raman scattering microscopy and multivariate curve resolution analysis. *Anal. Chem.* **85**, 98–106 (2013). doi: 10.1021/ac3019119; pmid: 23198914

63. Y. J. Lee, D. Moon, K. B. Migler, M. T. Cicerone, Quantitative image analysis of broadband CARS hyperspectral images of polymer blends. *Anal. Chem.* **83**, 2733–2739 (2011). doi: 10.1021/ac103351q; pmid: 21395296

64. F. Masia, A. Glen, P. Stephens, P. Borri, W. Langbein, Quantitative chemical imaging and unsupervised analysis using hyperspectral coherent anti-Stokes Raman scattering microscopy. *Anal. Chem.* **85**, 10820–10828 (2013). doi: 10.1021/ac402303g; pmid: 24099603

65. E. O. Potma, X. Xie, Detection of single lipid bilayers with coherent anti-Stokes Raman scattering (CARS) microscopy. *J. Raman Spectrosc.* **34**, 642–650 (2003). doi: 10.1002/jrs.1045

66. J.-X. Cheng, Coherent anti-Stokes Raman scattering microscopy. *Appl. Spectrosc.* **61**, 197–208 (2007). doi: 10.1366/000370207781746044; pmid: 17910784

67. L. Li, H. Wang, J.-X. Cheng, Quantitative coherent anti-Stokes Raman scattering imaging of lipid distribution in coexisting domains. *Biophys. J.* **89**, 3480–3490 (2005). doi: 10.1529/biophysj.105.065607; pmid: 16126824

68. S. Yampolsky et al., Seeing a single molecule vibrate through time-resolved coherent anti-Stokes Raman scattering. *Nat. Photonics* **8**, 650–656 (2014). doi: 10.1038/nphoton.2014.143

69. K. Wang et al., Time-lens based hyperspectral stimulated Raman scattering imaging and quantitative spectral analysis. *J. Biophotonics* **6**, 815–820 (2013). pmid: 23840041

70. C.-S. Liao, J. H. Choi, D. Zhang, S. H. Chan, J.-X. Cheng, Denoising stimulated Raman spectroscopic images by total variation minimization. *J. Phys. Chem. C* **119**, 19397–19403 (2015). doi: 10.1021/acs.jpcc.5b06980

71. H. Yamakoshi et al., Imaging of EdU, an alkyne-tagged cell proliferation probe, by Raman microscopy. *J. Am. Chem. Soc.* **133**, 6102–6105 (2011). doi: 10.1021/ja108404p; pmid: 21443184

72. H. Yamakoshi et al., Alkyne-tag Raman imaging for visualization of mobile small molecules in live cells. *J. Am. Chem. Soc.* **134**, 20681–20689 (2012). doi: 10.1021/ja308529n; pmid: 23198907

73. X. S. Xie, J. Yu, W. Y. Yang, Living cells as test tubes. *Science* **312**, 228–230 (2006). doi: 10.1126/science.1127566; pmid: 16614211

74. L. Wei, Y. Yu, Y. Shen, M. C. Wang, W. Min, Vibrational imaging of newly synthesized proteins in live cells by stimulated Raman scattering microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11226–11231 (2013). doi: 10.1073/pnas.1303768110; pmid: 23798434

75. G. Bergner et al., Quantitative detection of C-deuterated drugs by CARS microscopy and Raman microspectroscopy. *Analyst* **136**, 3686–3693 (2011). doi: 10.1039/c0an00956c; pmid: 21785774

76. L. Wei et al., Live-cell imaging of alkyne-tagged small biomolecules by stimulated Raman scattering. *Nat. Methods* **11**, 410–412 (2014). doi: 10.1038/nmeth.2878; pmid: 24584195

77. S. Hong et al., Live-cell stimulated Raman scattering imaging of alkyne-tagged biomolecules. *Angew. Chem. Int. Ed.* **53**, 5827–5831 (2014). doi: 10.1002/anie.201400328; pmid: 24753329

78. H. Kim, G. W. Bryant, S. J. Stranick, Superresolution four-wave mixing microscopy. *Opt. Express* **20**, 6042–6051 (2012). doi: 10.1364/OE.20.006042; pmid: 22418482

79. P. Wang et al., Far-field imaging of non-fluorescent species with subdiffraction resolution. *Nat. Photonics* **7**, 449–453 (2013). doi: 10.1038/nphoton.2013.97; pmid: 24436725

80. L. Gong, H. Wang, Breaking the diffraction limit by saturation in stimulated-Raman-scattering microscopy: A theoretical study. *Phys. Rev. A* **90**, 013818 (2014). doi: 10.1103/PhysRevA.90.013818

81. B. Knoll, F. Keilmann, Near-field probing of vibrational absorption for chemical microscopy. *Nature* **399**, 134–137 (1999). doi: 10.1038/20154

82. H. W. Siesler, Y. Ozaki, S. Kawata, H. M. Heise, *Near-Infrared Spectroscopy: Principles, Instruments, Applications*. (Wiley-VCH, 2002).

83. T. Durduran, R. Choe, W. B. Baker, A. G. Yodh, Diffuse optics for tissue monitoring and tomography. *Rep. Prog. Phys.* **73**, 076701 (2010). doi: 10.1088/0034-4885/73/7/076701; pmid: 26120204

84. N. Everall, T. Hahn, P. Matousek, A. W. Parker, M. Towrie, Picosecond time-resolved Raman spectroscopy of solids: Capabilities and limitations for fluorescence rejection and the influence of diffuse reflectance. *Appl. Spectrosc.* **55**, 1701–1708 (2001). doi: 10.1366/0003702011954053

85. N. Everall, T. Hahn, P. Matousek, A. W. Parker, M. Towrie, Photon migration in Raman spectroscopy. *Appl. Spectrosc.* **58**, 591–597 (2004). doi: 10.1366/000370204774103426; pmid: 15165336

86. M. D. Keller, S. K. Majumder, A. Mahadevan-Jansen, Spatially offset Raman spectroscopy of layered soft tissues. *Opt. Lett.* **34**, 926–928 (2009). doi: 10.1364/OL.34.000926; pmid: 19340173

87. I. E. Iping Petterson, P. Dvořák, J. B. Buijs, C. Gooijer, F. Ariese, Time-resolved spatially offset Raman spectroscopy for depth analysis of diffusely scattering layers. *Analyst* **135**, 3255–3259 (2010). doi: 10.1039/c0an00611d; pmid: 20941438

88. K. A. Esmonde-White, F. W. Esmonde-White, M. D. Morris, B. J. Roessler, Fiber-optic Raman spectroscopy of joint tissues. *Analyst* **136**, 1675–1685 (2011). doi: 10.1039/c0an00824a; pmid: 21359366

89. K. Buckley, J. G. Kerns, A. W. Parker, A. E. Goodship, P. Matousek, Decomposition of in vivo spatially offset Raman spectroscopy data using multivariate analysis techniques. *J. Raman Spectrosc.* **45**, 188–192 (2014). doi: 10.1002/jrs.4434

90. M. V. Schulmerich, W. F. Finney, R. A. Fredricks, M. D. Morris, Subsurface Raman spectroscopy and mapping using a globally illuminated non-confocal fiber-optic array probe in the presence of Raman photon migration. *Appl. Spectrosc.* **60**, 109–114 (2006). doi: 10.1366/000370206776023340; pmid: 16542561

91. S. Srinivasan et al., Image-guided Raman spectroscopic recovery of canine cortical bone contrast in situ. *Opt. Express* **16**, 12190–12200 (2008). doi: 10.1364/OE.16.012190; pmid: 18679495

92. S. Sil, S. Umapathy, Raman spectroscopy explores molecular structural signatures of hidden materials in depth: Universal multiple angle Raman spectroscopy. *Sci. Rep.* **4**, 5308 (2014). doi: 10.1038/srep05308; pmid: 24930768

93. J.-L. H. Demers, F. W. Esmonde-White, K. A. Esmonde-White, M. D. Morris, B. W. Pogue, Next-generation Raman tomography instrument for non-invasive in vivo bone imaging. *Biomed. Opt. Express* **6**, 793–806 (2015). doi: 10.1364/BOE.6.000793; pmid: 25798304

94. T. J. Allen, A. Hall, A. P. Dhillon, J. S. Owen, P. C. Beard, Spectroscopic photoacoustic imaging of lipid-rich plaques in the human aorta in the 740 to 1400 nm wavelength range. *J. Biomed. Opt.* **17**, 061209 (2012). doi: 10.1117/1.JBO.17.6.061209; pmid: 22734739

95. K. Jansen, A. F. W. van der Steen, H. M. M. van Beusekom, J. W. Oosterhuis, G. van Soest, Intravascular photoacoustic imaging of human coronary atherosclerosis. *Opt. Lett.* **36**, 597–599 (2011). doi: 10.1364/OL.36.000597; pmid: 21368919

96. P. Wang, P. Wang, H.-W. Wang, J.-X. Cheng, Mapping lipid and collagen by multispectral photoacoustic imaging of chemical bond vibration. *J. Biomed. Opt.* **17**, 96010–96011 (2012). pmid: 23085911

97. K. Jansen, M. Wu, A. F. van der Steen, G. van Soest, Photoacoustic imaging of human coronary atherosclerosis in two spectral bands. *Photoacoustics* **2**, 12–20 (2014). doi: 10.1016/j.pacs.2013.11.003; pmid: 25302152

98. P. Wang, H.-W. Wang, M. Sturek, J.-X. Cheng, Bond-selective imaging of deep tissue through the optical window between 1600 and 1850 nm. *J. Biophotonics* **5**, 25–32 (2012). doi: 10.1002/jbio.201100102; pmid: 22125288

99. B. G. Saar et al., Label-free, real-time monitoring of biomass processing with stimulated Raman scattering microscopy. *Angew. Chem. Int. Ed.* **49**, 5476–5479 (2010). doi: 10.1002/anie.201000900; pmid: 20589818

100. J. C. Mansfield et al., Label-free chemically specific imaging in planta with stimulated Raman scattering microscopy. *Anal. Chem.* **85**, 5055–5063 (2013). doi: 10.1021/ac400266a; pmid: 23581493

101. X. Nan, J.-X. Cheng, X. S. Xie, Vibrational imaging of lipid droplets in live fibroblast cells with coherent anti-Stokes Raman scattering microscopy. *J. Lipid Res.* **44**, 2202–2208 (2003). doi: 10.1194/jlr.D300022-JLR200; pmid: 12923234

102. T. Hellerer et al., Monitoring of lipid storage in Caenorhabditis elegans using coherent anti-Stokes Raman scattering (CARS) microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 14658–14663 (2007). doi: 10.1073/pnas.0703594104; pmid: 17804796

103. R. K. Lyn, D. C. Kennedy, A. Stolow, A. Ridsdale, J. P. Pezacki, Dynamics of lipid droplets induced by the hepatitis C virus

core protein. *Biochem. Biophys. Res. Commun.* **399**, 518–524 (2010). doi: 10.1016/j.bbrc.2010.07.101; pmid: 20678475

104. J. P. Day, G. Rago, K. F. Domke, K. P. Velikov, M. Bonn, Label-free imaging of lipophilic bioactive molecules during lipid digestion by multiplex coherent anti-Stokes Raman scattering microspectroscopy. *J. Am. Chem. Soc.* **132**, 8433–8439 (2010). doi: 10.1021/ja102069d; pmid: 20507119

105. M. Paar *et al.*, Remodeling of lipid droplets during lipolysis and growth in adipocytes. *J. Biol. Chem.* **287**, 11164–11173 (2012). doi: 10.1074/jbc.M111.316794; pmid: 22311986

106. M. C. Wang, W. Min, C. W. Freudiger, G. Ruvkun, X. S. Xie, RNAi screening for fat regulatory genes with SRS microscopy. *Nat. Methods* **8**, 135–138 (2011). doi: 10.1038/nmeth.1556; pmid: 21240281

107. W. Dou, D. Zhang, Y. Jung, J.-X. Cheng, D. M. Umulis, Label-free imaging of lipid-droplet intracellular motion in early *Drosophila* embryos using femtosecond-stimulated Raman loss microscopy. *Biophys. J.* **102**, 1666–1675 (2012). doi: 10.1016/j.bpj.2012.01.057; pmid: 22500767

108. P. Wang *et al.*, Imaging lipid metabolism in live *Caenorhabditis elegans* using fingerprint vibrations. *Angew. Chem. Int. Ed.* **53**, 11787–11792 (2014). doi: 10.1002/anie.201406029; pmid: 25195517

109. M. N. Slipchenko, T. T. Le, H. Chen, J.-X. Cheng, High-speed vibrational imaging and spectral analysis of lipid bodies by compound Raman microscopy. *J. Phys. Chem. B* **113**, 7681–7686 (2009). doi: 10.1021/jp902231y; pmid: 19422201

110. S. Yue *et al.*, Cholesteryl ester accumulation induced by PTEN loss and PI3K/AKT activation underlies human prostate cancer aggressiveness. *Cell Metab.* **19**, 393–406 (2014). doi: 10.1016/j.cmet.2014.01.019; pmid: 24606897

111. H. Wang, Y. Fu, P. Zickmund, R. Shi, J.-X. Cheng, Coherent anti-stokes Raman scattering imaging of axonal myelin in live spinal tissues. *Biophys. J.* **89**, 581–591 (2005). doi: 10.1529/biophysj.105.061911; pmid: 15834003

112. Y. Fu, H. Wang, T. B. Huff, R. Shi, J. X. Cheng, Coherent anti-Stokes Raman scattering imaging of myelin degradation reveals a calcium-dependent pathway in lyso-PtdCho-induced demyelination. *J. Neurosci. Res.* **85**, 2870–2881 (2007). doi: 10.1002/jnr.21403; pmid: 17551984

113. Y. Shi *et al.*, Longitudinal in vivo coherent anti-Stokes Raman scattering imaging of demyelination and remyelination in injured spinal cord. *J. Biomed. Opt.* **16**, 106012 (2011). doi: 10.1117/1.3641988; pmid: 22029359

114. E. Bélanger *et al.*, In vivo evaluation of demyelination and remyelination in a nerve crush injury model. *Biomed. Opt. Express* **2**, 2698–2708 (2011). doi: 10.1364/BOE.2.002698; pmid: 22091449

115. C.-R. Hu, D. Zhang, M. N. Slipchenko, J.-X. Cheng, B. Hu, Label-free real-time imaging of myelination in the *Xenopus laevis* tadpole by in vivo stimulated Raman scattering microscopy. *J. Biomed. Opt.* **19**, 086005 (2014). doi: 10.1117/1.JBO.19.8.086005; pmid: 25104411

116. Y. Shen, F. Xu, L. Wei, F. Hu, W. Min, Live-cell quantitative imaging of proteome degradation by stimulated Raman scattering. *Angew. Chem. Int. Ed.* **53**, 5596–5599 (2014). doi: 10.1002/anie.201310725; pmid: 24737659

117. J. Li, J.-X. Cheng, Direct visualization of de novo lipogenesis in single living cells. *Sci. Rep.* **4**, 6807 (2014). doi: 10.1038/srep06807; pmid: 25351207

118. F. Hu, L. Wei, C. Zheng, Y. Shen, W. Min, Live-cell vibrational imaging of choline metabolites by stimulated Raman scattering coupled with isotope-based metabolic labeling. *Analyst* **139**, 2312–2317 (2014). doi: 10.1039/c3an02281a; pmid: 24555181

119. H. J. Lee *et al.*, Assessing cholesterol storage in live cells and *C. elegans* by stimulated Raman scattering imaging of phenyl-Diyne cholesterol. *Sci. Rep.* **5**, 7930 (2015). doi: 10.1038/srep07930; pmid: 25608867

120. A. Pliss, A. N. Kuzmin, A. V. Kachynski, P. N. Prasad, Biophotonic probing of macromolecular transformations during apoptosis. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 12771–12776 (2010). doi: 10.1073/pnas.1006374107; pmid: 20615987

121. F.-K. Lu *et al.*, Label-free DNA imaging in vivo with stimulated Raman scattering microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11624–11629 (2015). doi: 10.1073/pnas.1515121112; pmid: 26324899

122. E. Kang *et al.*, In situ visualization of paclitaxel distribution and release by coherent anti-Stokes Raman scattering microscopy. *Anal. Chem.* **78**, 8036–8043 (2006). doi: 10.1021/ac061218s; pmid: 17134137

123. M. N. Slipchenko *et al.*, Vibrational imaging of tablets by epi-detected stimulated Raman scattering microscopy. *Analyst* **135**, 2613–2619 (2010). doi: 10.1039/c0an00252f; pmid: 20625604

124. C. M. Hartshorn *et al.*, Multicomponent chemical imaging of pharmaceutical solid dosage forms with broadband CARS microscopy. *Anal. Chem.* **85**, 8102–8111 (2013). doi: 10.1021/ac400671p; pmid: 23855585

125. A. L. Fussell, F. Grasmeijer, H. W. Frijlink, A. H. de Boer, H. L. Offerhaus, CARS microscopy as a tool for studying the distribution of micronised drugs in adhesive mixtures for inhalation. *J. Raman Spectrosc.* **45**, 495–500 (2014). doi: 10.1002/jrs.4515

126. W. S. Chiu *et al.*, Molecular diffusion in the human nail measured by stimulated Raman scattering microscopy. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7725–7730 (2015). doi: 10.1073/pnas.1503791112; pmid: 26056283

127. D. Fu *et al.*, Imaging the intracellular distribution of tyrosine kinase inhibitors in living cells with quantitative hyperspectral stimulated Raman scattering. *Nat. Chem.* **6**, 614–622 (2014). doi: 10.1038/nchem.1961; pmid: 24950332

128. S. Yue, M. N. Slipchenko, J.-X. Cheng, Multimodal nonlinear optical microscopy. *Laser Photon. Rev.* **5**, 496–512 (2011). doi: 10.1002/lpor.201000027; pmid: 24353747

129. C. W. Freudiger *et al.*, Multicolored stain-free histopathology with coherent Raman imaging. *Lab. Invest.* **92**, 1492–1502 (2012). doi: 10.1038/labinvest.2012.109; pmid: 22906986

130. F. Légaré, C. L. Evans, F. Ganikhanov, X. S. Xie, Towards CARS endoscopy. *Opt. Express* **14**, 4427–4432 (2006). doi: 10.1364/OE.14.004427; pmid: 19516594

131. B. G. Saar, R. S. Johnston, C. W. Freudiger, X. S. Xie, E. J. Seibel, Coherent Raman scanning fiber endoscopy. *Opt. Lett.* **36**, 2396–2398 (2011). doi: 10.1364/OL.36.002396; pmid: 21725423

132. Z. Liu *et al.*, Development of fibre bundle probe for coherent anti-Stokes Raman scattering microendoscopy. *Electron. Lett.* **49**, 522 (2013). doi: 10.1049/el.2013.0220

133. S. Heuke *et al.*, Multimodal mapping of human skin. *Br. J. Dermatol.* **169**, 794–803 (2013). pmid: 23927013

134. K. König *et al.*, Optical skin biopsies by clinical CARS and multiphoton fluorescence/SHG tomography. *Laser Phys. Lett.* **8**, 465–468 (2011). doi: 10.1002/lapl.201110014

135. M. Weinigel *et al.*, In vivo histology: Optical biopsies with chemical contrast using clinical multiphoton/coherent anti-Stokes Raman scattering tomography. *Laser Phys. Lett.* **11**, 055601 (2014). doi: 10.1088/1612-2011/11/5/055601

136. R. Mittal *et al.*, Evaluation of stimulated Raman scattering microscopy for identifying squamous cell carcinoma in human skin. *Lasers Surg. Med.* **45**, 496–502 (2013). pmid: 23996592

137. M. Ji *et al.*, Rapid, label-free detection of brain tumors with stimulated Raman scattering microscopy. *Sci. Transl. Med.* **5**, 201ra119 (2013). doi: 10.1126/scitranslmed.3005954; pmid: 24005159

138. M. Jermyn *et al.*, Intraoperative brain cancer detection with Raman spectroscopy in humans. *Sci. Transl. Med.* **7**, 274ra19 (2015). doi: 10.1126/scitranslmed.aaa2384; pmid: 25673764

139. B. Wang *et al.*, Intravascular photoacoustic imaging of lipid in atherosclerotic plaques in the presence of luminal blood. *Opt. Lett.* **37**, 1244–1246 (2012). doi: 10.1364/OL.37.001244; pmid: 22466209

140. R. Li, M. N. Slipchenko, P. Wang, J.-X. Cheng, Compact high power barium nitrite crystal-based Raman laser at 1197 nm for photoacoustic imaging of fat. *J. Biomed. Opt.* **18**, 040502 (2013). doi: 10.1117/1.JBO.18.4.040502; pmid: 23536057

141. A. Cerussi *et al.*, In vivo absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy. *J. Biomed. Opt.* **11**, 044005 (2006). doi: 10.1117/1.2337546; pmid: 16965162

142. S. Kukreti, A. Cerussi, B. Tromberg, E. Gratton, Intrinsic near-infrared spectroscopic markers of breast tumors. *Dis. Markers* **25**, 281–290 (2008). doi: 10.1155/2008/651308; pmid: 19208946

143. A. E. Cerussi *et al.*, Diffuse optical spectroscopic imaging correlates with final pathological response in breast cancer neoadjuvant chemotherapy. *Phil. Trans. R. Soc. A* **369**, 4512–4530 (2011). doi: 10.1098/rsta.2011.0279; pmid: 22006904

144. N. Stone, R. Baker, K. Rogers, A. W. Parker, P. Matousek, Subsurface probing of calcifications with spatially offset Raman spectroscopy (SORS): Future possibilities for the diagnosis of breast cancer. *Analyst* **132**, 899–905 (2007). doi: 10.1039/b705029a; pmid: 17710265

145. M. V. Schulmerich, K. A. Dooley, T. M. Vanasse, S. A. Goldstein, M. D. Morris, Subsurface and transcutaneous Raman spectroscopy and mapping using concentric illumination rings and collection with a circular fiber-optic array. *Appl. Spectrosc.* **61**, 671–678 (2007). doi: 10.1366/000370207781393307; pmid: 17697459

146. P. Matousek, N. Stone, Recent advances in the development of Raman spectroscopy for deep non-invasive medical diagnosis. *J. Biophotonics* **6**, 7–19 (2013). doi: 10.1002/jbio.201200141; pmid: 23129567

147. I. Rajapaksa, H. Kumar Wickramasinghe, Raman spectroscopy and microscopy based on mechanical force detection. *Appl. Phys. Lett.* **99**, 161103–1611033 (2011). doi: 10.1063/1.3652760; pmid: 22087048

148. X. Xu, H. Liu, L. V. Wang, Time-reversed ultrasonically encoded optical focusing into scattering media. *Nat. Photonics* **5**, 154–157 (2011). doi: 10.1038/nphoton.2010.306; pmid: 21532925

149. Y. M. Wang, B. Judkewitz, C. A. Dimarzio, C. Yang, Deep-tissue focal fluorescence imaging with digitally time-reversed ultrasound-encoded light. *Nat. Commun.* **3**, 928 (2012). doi: 10.1038/ncomms1925; pmid: 22735456

150. C. W. Freudiger *et al.*, Stimulated Raman scattering microscopy with a robust fibre laser source. *Nat. Photonics* **8**, 153–159 (2014). doi: 10.1038/nphoton.2013.360; pmid: 25313312

ACKNOWLEDGMENTS

RESEARCH ARTICLE SUMMARY

NEUROSCIENCE

Principles of connectivity among morphologically defined cell types in adult neocortex

Xiaolong Jiang,* Shan Shen, Cathryn R. Cadwell, Philipp Berens, Fabian Sinz, Alexander S. Ecker, Saumil Patel, Andreas S. Tolias*

INTRODUCTION: The intricate microcircuitry of the cerebral cortex is thought to be a critical substrate from which arise the impressive capabilities of the mammalian brain. Until now, our knowledge of the stereotypical connectivity in neocortical microcircuits has been pieced together from individual studies of the connectivity between small numbers of neuronal cell types. Here, we provide unbiased, large-scale profiling of neuronal cell types and connections to reveal the essential building blocks of the cortex and the principles governing their assembly into cortical circuits. Using advanced techniques for tissue slicing, multiple simultaneous whole-cell recording, and morphological reconstruction, we are able to provide a comprehensive view of the connectivity between diverse types of neurons, particularly among types of γ-aminobutyric acid–releasing (GABAergic) interneurons, in the adult animal.

RATIONALE: We took advantage of a method for preparing high-quality slices of adult tissue and combined this technique with octuple simultaneous, whole-cell recordings followed by an improved staining method that allowed detailed recovery of axonal and dendritic arbor morphology. These data allowed us to perform a census of morphologically and electrophysiologically defined neuronal types (primarily GABAergic interneurons) in neocortical layers 1, 2/3, and 5 (L1, L23, and L5, respectively) and to observe their connectivity patterns in adult animals.

RESULTS: Our large-scale, comprehensive profiling of neocortical neurons differentiated 15 major types of interneurons, in addition to two lamina-defined types of pyramidal neurons (L23 and L5). Cortical interneurons comprise two types in L1 (eNGC and SBC-like), seven in L23 (L23MC, L23NGC, BTC, BPC, DBC, L23BC, and ChC), and six in L5 (L5MC, L5NGC, L5BC, SC, HEC, and DC) (see the figure). Each type has stereotypical electrophysiological properties and morphological features and can be differentiated from all

others by cell type–specific axonal geometry and axonal projection patterns. Importantly, each type of neuron has its own characteristic input-output connectivity profile, connecting with other constituent neuronal types with varying degrees of specificity in postsynaptic targets, laminar location, and synaptic characteristics. Despite specific connection patterns for each cell type, we found that a small number of simple connectivity motifs are repeated across layers and cell types defining a canonical cortical microcircuit.

CONCLUSION: Our comprehensive profiling of neuronal cell types and connections in adult neocortex provides the most complete wiring diagram of neocortical microcircuits to date. Compared with current genetic labels for cell class, which paint the cortex in broad strokes, our analysis of morphological and electrophysiological properties revealed new cell classes and allowed us to derive a small number of simple connectivity rules that were repeated across layers and cell types. This detailed blueprint of cortical wiring should aid efforts to identify specific circuit abnormalities in animal models of brain disease and may eventually provide a path toward the development of comprehensive circuit-based, cell type–specific interventions. ∎

Connectivity among morphologically defined cell types in adult neocortex. (**A**) Simultaneous octuple whole-cell recording to study connectivity followed by morphological reconstruction. (**B**) Synaptic connectivity between morphologically distinct types of neurons, including pyramidal (P) neurons. (**C**) Connectivity from neurogliaform cells (NGCs) to other cell types. This connectivity is believed to be nonsynaptic and mediated by volume transmission. Martinotti cell, MC; basket cell, BC; single-bouquet cell–like cell, SBC-like; bitufted cell, BTC; bipolar cell, BPC, double-bouquet cell, DBC; chandelier cell, ChC; shrub cell, SC; horizontally elongated cell, HEC; deep-projecting cell, DC.
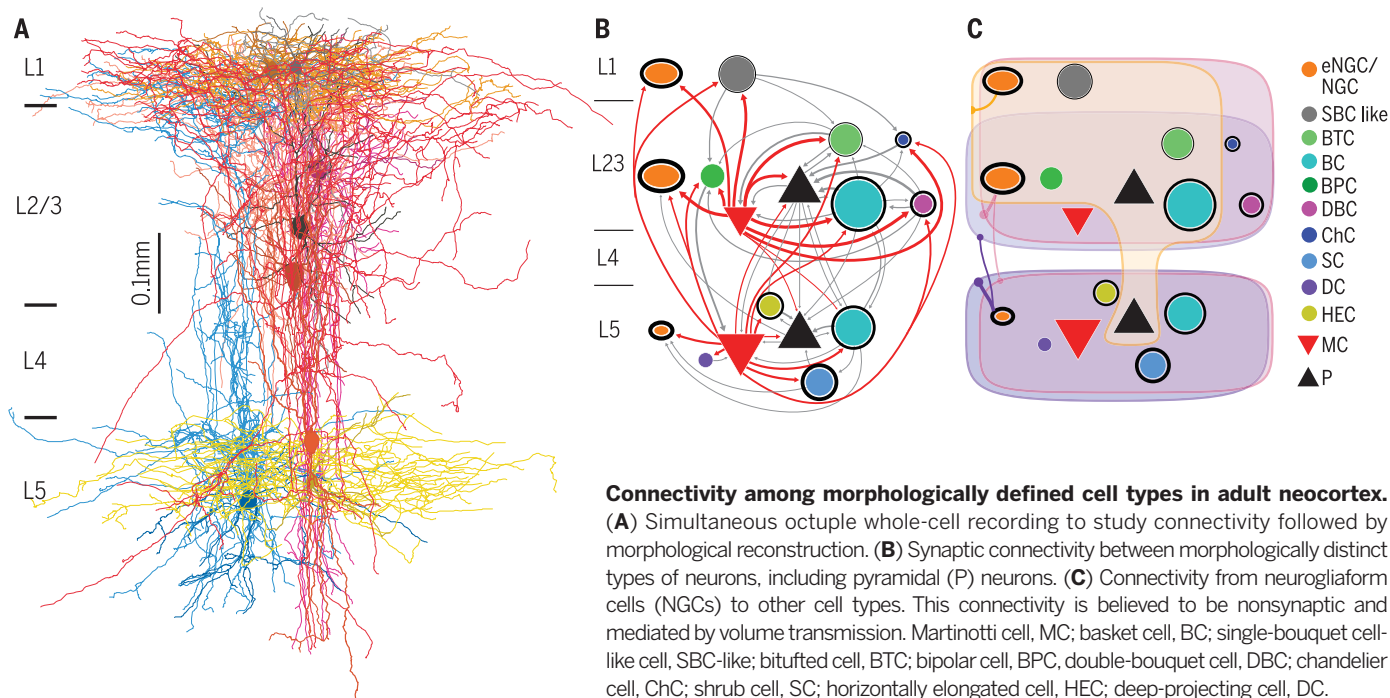
NEUROSCIENCE

Principles of connectivity among morphologically defined cell types in adult neocortex

Xiaolong Jiang,[1]* Shan Shen,[1] Cathryn R. Cadwell,[1] Philipp Berens,[1,2,3,4] Fabian Sinz,[1] Alexander S. Ecker,[1,2,4,5] Saumil Patel,[1] Andreas S. Tolias[1,2]*

Since the work of Ramón y Cajal in the late 19th and early 20th centuries, neuroscientists have speculated that a complete understanding of neuronal cell types and their connections is key to explaining complex brain functions. However, a complete census of the constituent cell types and their wiring diagram in mature neocortex remains elusive. By combining octuple whole-cell recordings with an optimized avidin-biotin-peroxidase staining technique, we carried out a morphological and electrophysiological census of neuronal types in layers 1, 2/3, and 5 of mature neocortex and mapped the connectivity between more than 11,000 pairs of identified neurons. We categorized 15 types of interneurons, and each exhibited a characteristic pattern of connectivity with other interneuron types and pyramidal cells. The essential connectivity structure of the neocortical microcircuit could be captured by only a few connectivity motifs.

Despite its importance, we are still far from completely understanding the extensive diversity of cell types in the neocortex and how they are connected into functional circuits. Neocortical neurons fall into two broad classes. Excitatory glutamatergic neurons form the majority and exhibit relatively stereotypical properties, whereas inhibitory γ-aminobutyric acid–releasing (GABAergic) interneurons are highly diverse (1). GABAergic interneurons show many distinct morphological, electrophysiological, neurochemical, and synaptic wiring features (1–6). Due to the complex and multifaceted nature of interneurons, there is still no consensus on how many types of interneurons exist in the neocortex, and unequivocal identification of the features that distinguish one type of neuron from another is a matter of considerable interest and debate (2, 4). Maturation of GABAergic interneurons takes longer than for pyramidal cells, and their continuous development throughout adolescence into adulthood often further obscures our understanding (7, 8). Therefore, it is imperative to study the mature neocortex to gain a true understanding of interneuron cell types in the neocortex. However, the vast majority of in vitro slice electrophysiology experiments to date were carried out in juvenile animals due to the technical difficulties of preparing high-quality slices from adult tissue.

Morphologically defined interneuron types in neocortex

We took advantage of a recently developed adult animal slicing method (9) to study mature neocortical circuitry. We performed simultaneous octuple whole-cell recordings in acute slices prepared from the primary visual cortex (area V1) of adult mice (≥2 months old) and focused on neurons from layers 1, 2/3, and 5 (L1, L23, and L5, respectively) because these layers are particularly enriched with GABAergic interneurons (10). We recorded both GABAergic neurons (n = 1654) and pyramidal neurons (n = 547) to identify their connections, followed by post hoc morphological recovery of each neuron with an optimized avidin-biotin-peroxidase staining method (fig. S1 and supplementary materials). Using this technique, we recovered the morphology of all recorded pyramidal neurons (n = 547/547, 100%), and almost all recorded GABAergic interneurons (n = 1566/1654, ≥ 95%), especially their fine axonal arbors (fig. S1A).

The morphologies of interneurons were highly diverse, whereas the morphologies of pyramidal neurons in L23 and L5 were relatively uniform (for a discussion of the morphological diversity of pyramidal cells, see the supplementary text). Morphologically recovered interneurons were visually assessed or reconstructed if necessary under microscopy and then were classified following a widely used classification scheme based on their axonal geometry (thickness, tortuosity, bifurcation angle, branching order, and the shape of terminal branches) and their axonal projection pattern (1, 4, 5). We could differentiate recorded interneurons into 15 major types: two types in L1, seven in L23, and six in L5 (Fig. 1, A to C). To quantitatively support our classification, a subset of neurons from each type were fully reconstructed to carry out dendritic and axonal length density analysis (figs. S2 and S3) (n = 24 in L1, n = 96 in L23, and n = 73 in L5). We extracted simple features from the axon and dendrite density maps of the reconstructed neurons using principal component analysis (11) and trained cross-validated sparse logistic regression classifiers to distinguish between any two types of interneurons as labeled manually (see the supplementary materials). The classifiers could separate almost all cell type pairs with an accuracy of >90% and an average performance of ~97%, supporting the notion that our 15 classes represent distinct types (Fig. 1D). The classifiers mostly used features computed from the axon density map, giving support to the idea that interneurons can primarily be distinguished based on their axonal features (fig. S4).

Two types of layer 1 interneurons

The neurons recorded in L1 of mature V1 could be morphologically divided into two major types, which were generally similar to those described in juvenile animals (5, 12). One type of neuron was neurogliaform cells (NGCs, n = 84) with elongated axonal arborization (Figs. 1A and 3A and figs. S2A, S2D, S5A, S6B, and S7). We follow the previous nomenclature naming them elongated NGCs (eNGC). The vast majority of eNGCs (n = 77/84 neurons) fired late-spiking (LS) action potentials (APs) followed by a deep, wide afterhyperpolarization (AHP) (Fig. 2A, Type 1, and table S1). The eNGCs constituted about one-third of all L1 neurons (Fig. 1E). The remaining L1 neurons had axonal projection patterns (Figs. 1A and 3A and figs. S2A, S5A, S6B, and S7), dendritic features (fig. S3, A and D), and electrophysiological properties (Fig. 2A and tables S1 and S2) that were clearly distinct from eNGCs. Some of them (~40%) had an axon that was very similar to single-bouquet cells (SBCs) reported in juvenile rat somatosensory cortex (5, 12). However, many of them (~60%) had atypical axonal projection patterns compared to those previously described for SBCs, and their axon arborized mostly within L1, with only one or two side branches extending to deep layers (not deeper than L4). Despite their variable axonal projection patterns, non-neurogliaform L1 neurons shared similar dendritic and electrophysiological features (tables S1 and S2) and similar connectivity profiles (table S3) that correspond to SBCs in rat somatosensory cortex (5, 12). We thus refer to this group as SBC-like cells. A detailed description of the morphology, firing patterns, and intrinsic membrane properties of these two major types of L1 neurons can be found in the supplementary text.

Seven types of layers 2/3 interneurons

Interneurons from L23 of mature V1 could be grouped into seven major types based on axonal

[1]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA. [2]Bernstein Centre for Computational Neuroscience, Tübingen, Germany. [3]Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany. [4]Werner Reichardt Center for Integrative Neuroscience and Institute of Theoretical Physics, University of Tübingen, Tübingen, Germany. [5]Max Planck Institute for Biological Cybernetics, Tübingen, Germany.
*Corresponding author. E-mail: astolias@bcm.edu (A.S.T.); xiaolonj@bcm.edu (X.J.)
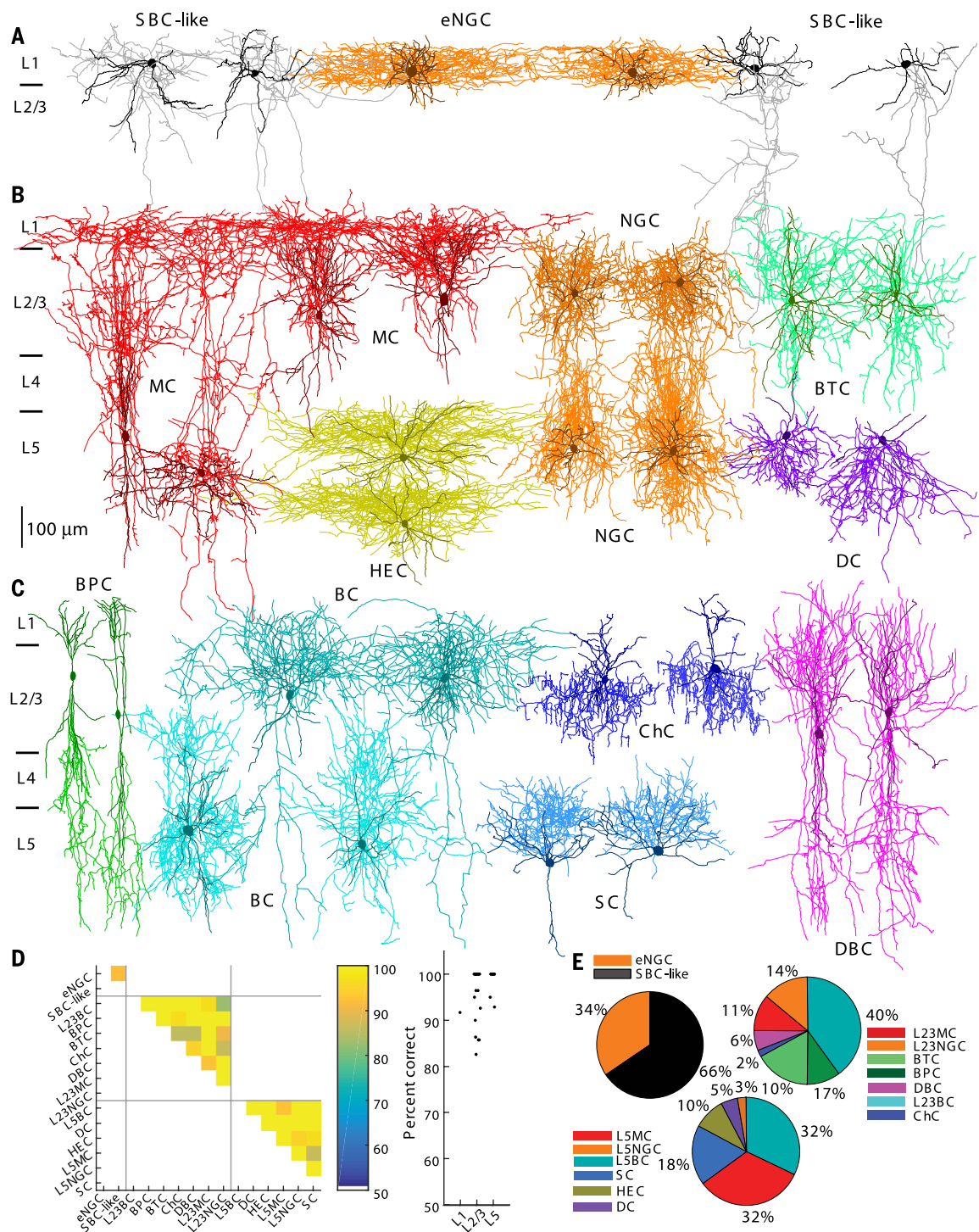
Fig. 1. Morphologically distinct GABAergic interneurons in L1, L23, and L5 of V1. (**A**) Two eNGCs (two middle; axon in dark orange and dendrite in brown) and four SBC-like cells. SBC-like cells on the left (two leftmost; axon in gray) have the axon arborizing mostly within layer 1, whereas SBC-like cells on the right (two rightmost; axon in gray) have the axon projecting mostly toward the deep layers. (**B**) Four MCs (both L23 and L5; axon in red and dendrite in dark red), four NGCs (L23 and L5; axon in orange and dendrite in brown), two HECs (axon in yellow and dendrite in dark yellow), two BTCs (axon in aquamarine and dendrite in green), and two DCs (axon in blue violet and dendrite in dark blue). (**C**) Two BPCs (axon in lime and dendrite in green), four BCs (both L23 and L5; axons in cyan and dendrite in dark cyan), two ChCs (axon in blue and dendrite in dark blue), two DBCs (axon in magenta and dendrite in purple), and two SCs (axon in dodge blue and dendrite in dark blue). (**D**) (Left) Cross-validated classification performance for each pair of cell types within a layer. (Right) Classification performance collapsed within each layer. (**E**) The proportion of each morphologically distinct type of interneurons.
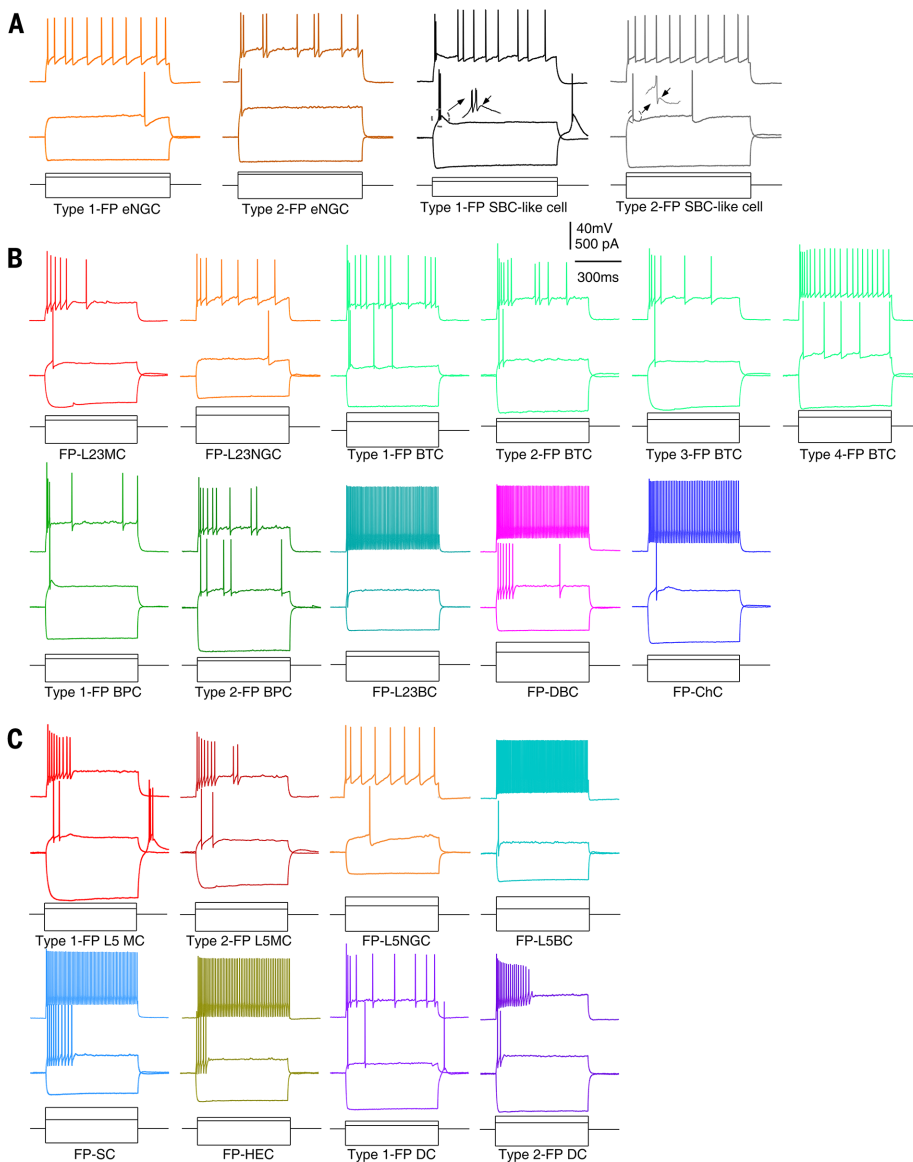
Fig. 2. Firing patterns (FPs) of morphologically distinct types of interneurons in V1. (**A**) Responses to hyperpolarizing, near-threshold, and suprathreshold current injections are shown for L1 eNGCs and SBC-like cells. The eNGCs had two types of firing patterns [LS, orange (leftmost); non-LS, dark orange], and SBC-like cells had two types of firing patterns (burst, black; no burst, gray). The eNGCs can be differentiated from SBC-like cells based on the absence of ADPs (see inset). (**B**) Responses to hyperpolarizing, near-threshold and suprathreshold current injections are shown for L23MCs, L23NGCs, BTCs, BPCs, L23BCs, DBCs, and ChCs. BTCs had four major types of firing patterns with subtle differences, and BPCs had two types of firing patterns that differ in the capability of burst. (**C**) Responses to hyperpolarizing, near-threshold, and suprathreshold current injections are shown for L5MCs, L5NGCs, L5BCs, SCs, HECs, and DCs. L5MCs and DCs had two types of firing patterns with subtle differences. The intrinsic membrane properties for each type of L1, L23, and L5 interneurons are presented in table S1.

morphology (Fig. 1, B and C): L23 Martinotti cells (L23MCs, *n* = 182), L23 neurogliaform cells (L23NGCs, *n* = 102), bitufted cells (BTCs, *n* = 118), bipolar cells (BPCs, *n* = 85), double bouquet cells (DBCs, *n* = 46), L23 basket cells (L23BCs, *n* = 322), and chandelier cells (ChCs, *n* = 18). All of these L23 neuronal types have been previously reported in juvenile rodents (*1*, *5*, *13*, *14*). (For a detailed description of their axonal and dendritic mor-

phology, see the supplementary text). For some of these types (for instance, BPC and BC), the morphology in mature neocortex did not completely match the morphological description for this cell type in the developing neocortex, which could be due to differences in either age or species (mouse versus rat; see the supplementary text). Each type of L23 interneuron had a characteristic axonal projection pattern (Fig. 1, B and C), which was

confirmed by axonal length density analysis and pairwise type classification (Fig. 1D and figs. S2, B and E, and S4). Neurons within a type tended to have the same dendritic arborization patterns (see the supplementary text), but these patterns were often not type-specific (figs. S3, B and E, and S4). The morphological types varied greatly in population size, with L23BCs being the largest population of L23 interneurons (40%) and ChC being the smallest (2%) (Fig. 1E).

Firing patterns and intrinsic membrane properties of L23 interneurons were mostly stereotypical within a cell type (see the supplementary text) but again were often not cell type–specific (Fig. 2B and table S1). For instance, L23BCs and ChCs were both fast-spiking (FS), consistent with previous observations. All DBCs in mature V1 were FS, which had not been previously reported (Fig. 2B and table S1). All BPCs exhibited an irregular spiking pattern with or without initial burst (Fig. 2B), and this irregular spiking pattern was also seen in a large proportion of BTCs (Fig. 2B). BTCs exhibited the most diverse firing patterns, and some of them had a firing pattern similar to MCs (Fig. 2B). Nevertheless, L23MCs and L23NGCs appeared to have firing patterns and intrinsic membrane properties that were mostly cell type–specific and could be used to differentiate them from most of the other neuronal cell types (except for a small number of BTCs) (see Fig. 2B).

Six types of layer 5 interneurons

Interneurons in L5 of mature V1 could be grouped into six major types based on their axonal morphology, only one of which had been reported and characterized before in the developing neocortex (Martinotti cells; L5MCs) (*15*, *16*). The remaining five types have not been previously described in L5, and we named them as follows: neurogliaform cells (L5NGCs), basket cells (L5BCs), shrub cells (SCs), horizontally elongated cells (HECs), and deep-projecting cells (DCs) (Fig. 1, B and C). Each of these types had a characteristic axonal projection pattern (Fig. 1, B to D, and fig. S4), which was confirmed by axonal length density analysis and pairwise type classification (Fig. 1D and figs. S2, C and F, and S4). Neurons within a type also tended to have the same dendritic arborization pattern and electrophysiological properties, but as for L23 interneurons, these properties were often not cell type–specific (figs. S3, C and F, and S4). L5 morphological types also varied greatly in population size, with L5BCs and L5MCs being the largest population of L5 interneurons (32%) and L5NGCs being the smallest (3%) (Fig. 1E). Because most of these L5 interneuron types have not been previously reported, we describe the morphology and electrophysiology of each type in detail below.

L5 Martinotti cell

L5MCs in mature neocortex (*n* = 174) were very similar to L5MCs previously reported in the developing neocortex (*15*, *16*). The majority of L5MCs had bitufted somatodendritic morphology with an elaborate dendritic tree that ascended
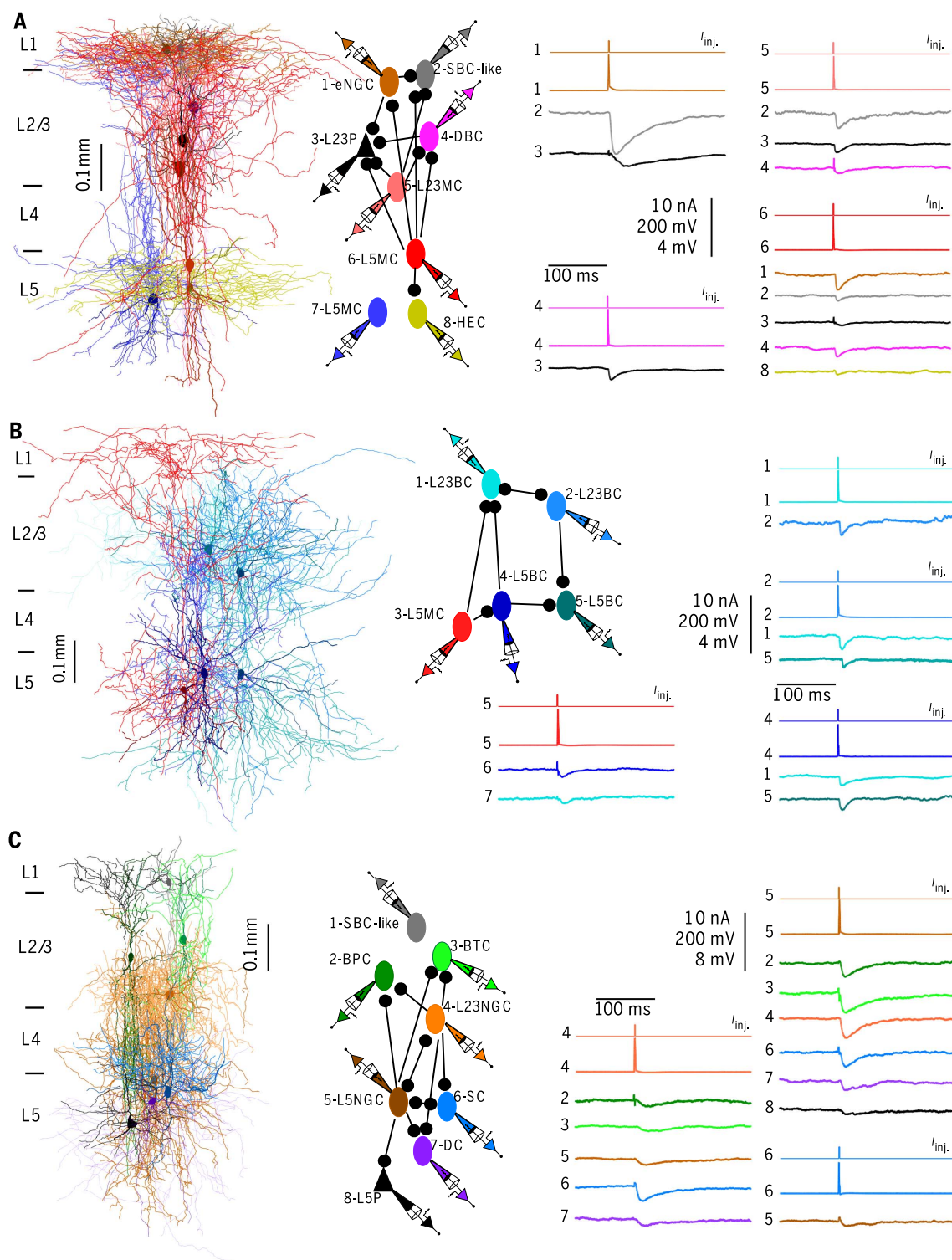
Fig. 3. Connections between morphologically distinct types of interneurons in V1. Right-most vertical scale bars from top to bottom show amplitudes of current injection ($I_{inj.}$ in nA), APs (mV), and uIPSPs (mV). (**A**) (Left) Connections between eight simultaneously recorded neurons, including one eNGC, one SBC-like cell, one DBC, one L23MC, one L23Pyr, one HEC, and two L5MCs. Each neuron was spatially separated in fig. S7A. (Middle) Connection diagram of the eight reconstructed neurons. (Right) APs elicited in presynaptic neurons and response traces of IPSPs evoked in postsynaptic neurons for each connection. (**B**) (Left) Connections between five simultaneously recorded neurons, including two L23BCs, two L5BCs, and one L5MC. Each neuron was spatially separated in fig. S7B. (Middle) Connection diagram of the five neurons. (Right) APs elicited in presynaptic neurons and response traces of IPSPs evoked in postsynaptic neurons for each connection. (**C**) (Left) Connections between eight simultaneously recorded neurons, including one SBC-like cell, one BPC, one BTC, one L23NGC, one L5Pyr, one L5NGC, one DC, and one SC. Each neuron was spatially separated in fig. S7C. (Middle) Connection diagram of the eight neurons. (Right) APs elicited in presynaptic neurons and response traces of IPSPs evoked in postsynaptic neurons for each connection.

Fig. 4. Connectivity matrix for adult mouse V1. (**A**) Color-coded matrix showing the probability of finding a connected pair of neurons between two specific types within and across layers. For total connections found/total connections tested for each type of connection and the mean amplitude of the connections, see fig. S14, table S6, and the supplementary text. (**B**) A simple model incorporating three connectivity rules (master, ISI, and PTI) can explain most of the observed data. Bar height corresponds to the negative likelihood of the model per cell pair in bits, *d* denotes the degrees of freedom, and error bars show the standard deviation over 50 bootstrapped data sets. Models increase in complexity from left to right; pictograms on top of the bars show the corresponding connectivity matrix. The first four models assume that the connection probability is uniform, uniform within a layer, and uniform within a layer and excitatory/inhibitory neuron, respectively. The next three models include the connectivity rules. The full model has an individual connection probability for each type of connection. The Hinton plots on the right depict the connection probabilities according to the full model and the model including all three connectivity principles.
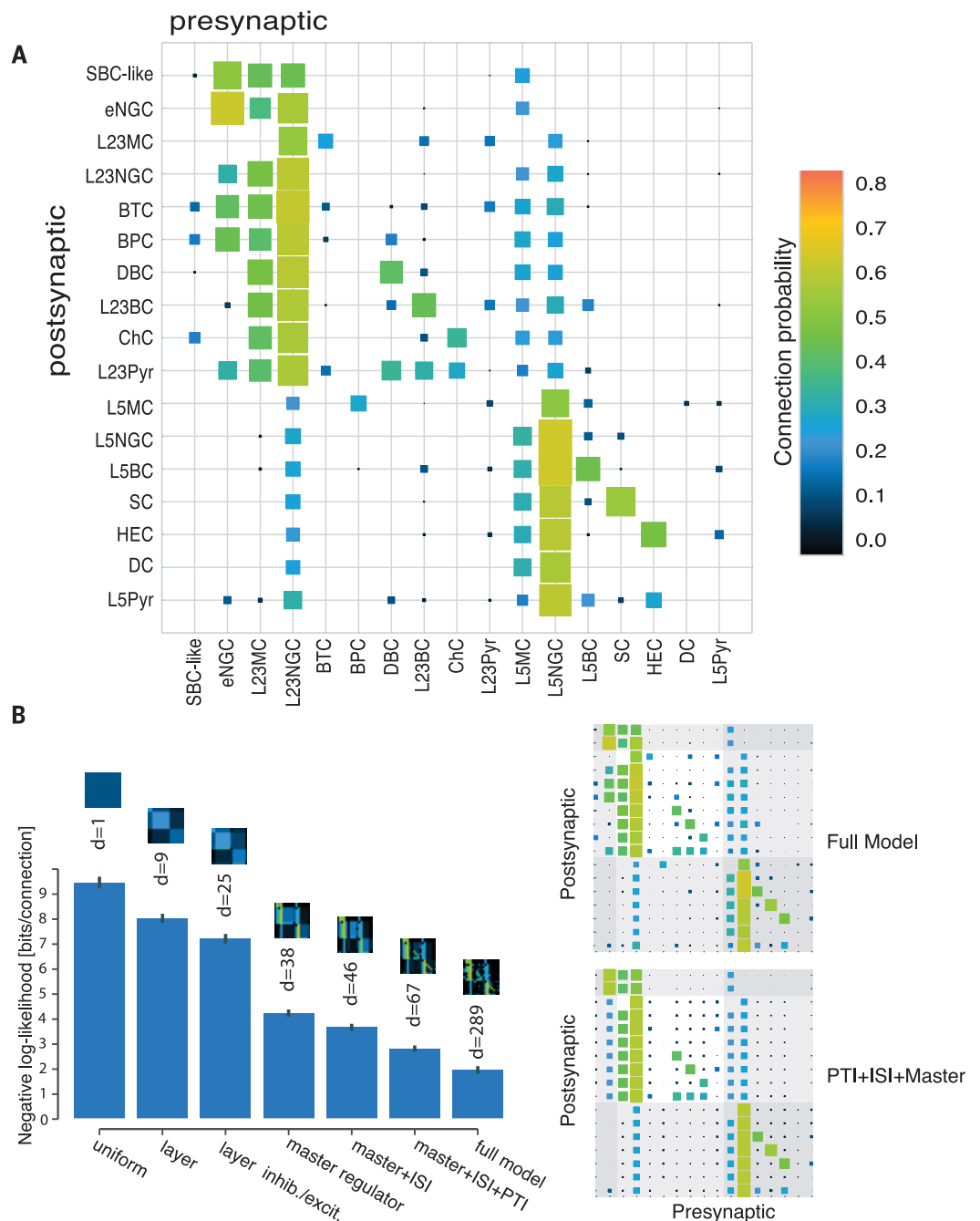
and descended toward L23 and L6, respectively (Figs. 1B and 3A; and figs. S3, C and F; S5, A and D; S7; S8B; fig. S9, D and E; and fig. S10A), and a small number of L5MCs from deeper L5 (L5B) had multipolar somatodendritic morphology (Figs. 1B and 3, A and B, and fig. S7). The axons of L5MC originated from the pia side of the somata or one of the primary dendrites, then ascended to form large axonal clusters in L1 and L4 (or, less often, in L1 and L5) (Figs. 1B and 3, A and B; and figs. S2, C and F; S5, A and D; S7; S8B, S9, D and E; and S10A). All L5MCs exhibited a firing pattern similar to their young counterparts (*13*, *16*, *17*) except that they could not sustain continuous firing in response to prolonged current injection (Fig.

2C and table S1). About one-fourth of L5MCs could generate a rebound burst after cessation of the hyperpolarization step (Fig. 2C, Type 1, 44/164), reminiscent of low-threshold-spiking interneurons in the developing neocortex (*16*, *17*). The remaining three-fourths of L5MCs did not exhibit a rebound burst (Type 2, 120/164). In addition, L5MC, similar to L23MCs, had a characteristic intrinsic membrane property: All L5MCs had a very large membrane time constant (table S1).

L5 neurogliaform cell

We called this L5 cell type neurogliaform cells (L5NGC, *n* = 17) because they had axonal and

dendritic geometry typical of NGCs in other layers. However, their axonal arborization was vertically elongated, different from NGCs in L1 and L23 (Figs. 1B and 3C and figs. S2, C and F, and S7). All L5NGCs had a late-spiking firing pattern, similar to the firing patterns of NGCs in L1 (Type 1) and L23 (Figs. 2C and table S1).

L5 basket cell

We called this L5 cell type basket cells (L5BC, *n* = 151) based on their similarity to L23BCs (see the supplementary text). They had a roughly vertically oriented, bitufted-like dendritic tree (Figs. 1C and 3B; and figs. S3, C and F; S5, B to D; S7; S9B, and S10, A and B) and a thick axon that
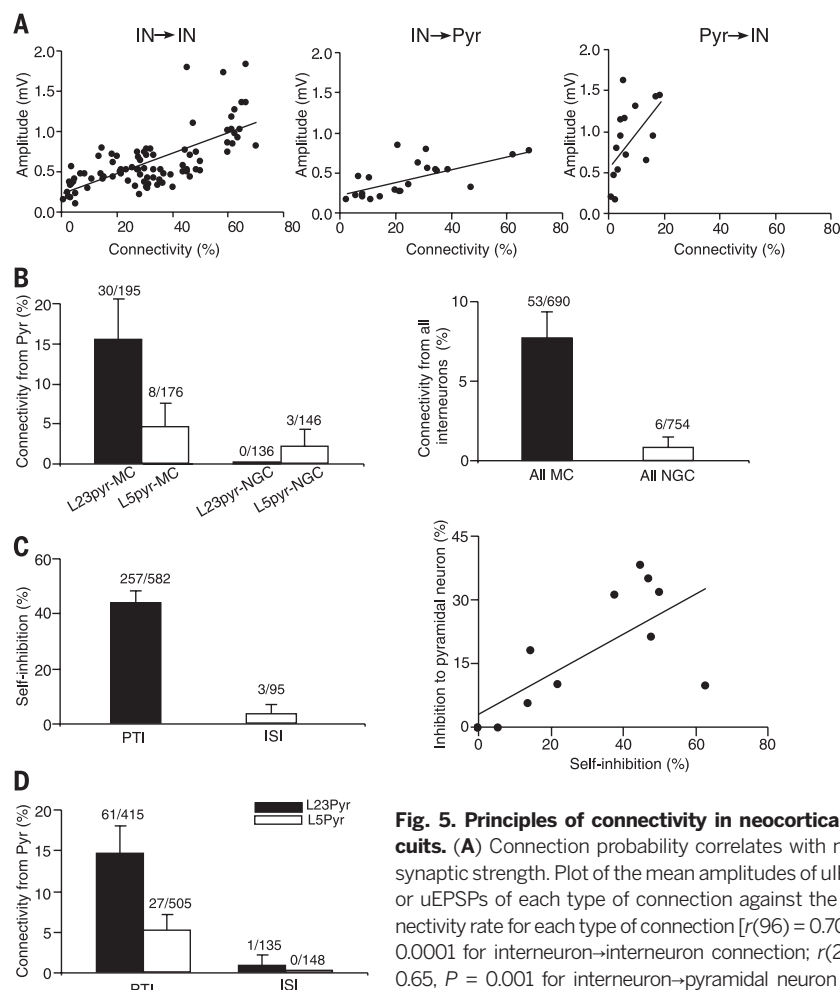
Fig. 5. Principles of connectivity in neocortical circuits. (**A**) Connection probability correlates with mean synaptic strength. Plot of the mean amplitudes of uIPSPs or uEPSPs of each type of connection against the connectivity rate for each type of connection [$r(96) = 0.70$, $P < 0.0001$ for interneuron→interneuron connection; $r(20) = 0.65$, $P = 0.001$ for interneuron→pyramidal neuron connection; $r(14) = 0.56$, $P = 0.03$ or pyramidal neuron→ interneuron connection]. (**B**) (Left) Connectivity rate from excitatory cells to MCs and NGCs. (Right) Connectivity rate from all (non-MC, non-NGC) interneurons to MCs and NGCs. (**C**) (Left) The self-connections of PTI occurs in 44.2% (257/582; pooling across all PTIs), whereas self-inhibition between ISIs occurs only in 3.2% (3/95) of tested connections (GLM with factor of class of interneuron (PTI vs. ISI); $P < 0.0001$). (Right) There is a strong positive correlation between inhibition of pyramidal neurons and self-inhibition [$r(11) = 0.72$; $p = 0.01$]. Note that origin contains three points. (**D**) Pyramidal cells provided input to PTIs in 14.7% (61/415; pooling across all PTIs, from L23Pyr) and 5.3% (27/505, from L5Pyr) of the tested connections, whereas pyramidal cells rarely provided input to ISIs (0.74%, 1/135; 0.0%, 0/148 for L23 and L5 pyramidal neurons, respectively; GLM with factors of class of interneuron (PTI vs. ISI) and layer of pyramidal neuron; effect of interneuron class: $P = 0.005$; layer: $P = 0.009$; interaction: $P = 0.76$).

typically originated from the apical side of the somata and ascended into L23 but never extended into L1 (Figs. 1C and 3B; and figs. S2, C and F; S5, B to D; S7; S9B; and S10, A and B). All L5BCs were FS (Fig. 2C and table S1).

Shrub cell

This type of interneuron in L5 was named based on the shrub-like axonal field that they exhibited (SC, $n = 77$). These cells had a characteristic asymmetric dendritic tree, with most of the dendritic branches localized above the somata (Figs. 1C and 3C; and figs. S3, C and F; S5, A and B; S7; and S10B). The axon typically originated from the apical side of the somata and ascended for a short distance before bifurcating into several main branches that then arborized into a short shrub-like axonal field above the somata (Figs. 1C

and 3C; and figs. S2, C and F; S5, A and B; S7; and S10B). All SCs were FS (Fig. 2C and table S1).

Horizontally elongated cell

This type was named based on the horizontally elongated axonal arbors (HEC, $n = 54$). HECs had a multipolar somatodendritic morphology (Fig. 3A and figs. S3, C and F and S5C) and a thick axon that originated from either the base or the apical side of the somata and arborized into a narrow, horizontally elongated axonal field (Figs. 1B and 3A; and figs. S2, C and F; S5C; S7; and S10B). The vast majority of HECs (53/54) had a FS firing pattern (Fig. 2C and table S1).

Deep-projecting cell

This cell type had a multipolar somatodendritic morphology (fig. S3, C and F), and their axons

primarily descended toward deeper cortical areas, even to layer 6 (Figs. 1B and 3C; and figs. S2, C and F, and S7). Therefore, we refer to them as deep-projecting cells (DCs; $n = 27$). Although all DCs discharged irregularly, their firing patterns could be differentiated into two types with subtle differences [Type 1, 15/22, with afterdepolarization (ADP); and Type 2, a deeper and faster AHP and no ADP, 7/22] (Fig. 2C and table S1).

Overlap of morphological types with Cre driver lines

In many experiments, we used Viaat-Cre/Ai9 mice ($n = 81$) to facilitate targeting of GABAergic interneurons (18). All unlabeled neurons recorded from L23 ($n = 120$) and L5 ($n = 105$) in these transgenic mice were morphologically and electrophysiologically confirmed as pyramidal neurons, and none were interneurons, suggesting that indeed the entire population of GABAergic interneurons in L23 and L5 was labeled in these mice. Almost all labeled neurons in L23 were morphologically confirmed as interneurons with distinct morphology (401/402) (table S4). The vast majority of labeled neurons in L5 were also confirmed as interneurons (269/289), although a small percentage (~7%, 20/289) were pyramidal neurons (table S4). In contrast to L23 and L5, a considerable proportion of L1 interneurons were unlabeled (~26%, $n = 25/95$), and, interestingly, all unlabeled interneurons were SBC-like cells.

Several Cre lines label subpopulations of interneurons. To determine the morphological cell types labeled by specific Cre driver lines, we recovered the morphology of parvalbumin-expressing (PV+), somatostatin-expressing (SOM+) and vasointestinal peptide-expressing (VIP+) interneurons using PV-Cre/Ai9 ($n = 18$ mice), SOM-Cre/Ai9 ($n = 30$ mice), and VIP-Cre/Ai9 ($n = 5$ mice) transgenic mice, respectively. The vast majority of recovered PV+ neurons in L23 were BCs ($n = 80/82$) (fig. S10B and table S4), and only two were ChCs. In L5, most recovered PV+ neurons were BCs as well (60%, 26/43) (fig. S10B and table S4), and the rest were either SCs (23%, $n = 10/43$) (fig. S10B) or HECs (19%, $n = 8/43$) (fig. S10B and table S4). Most recovered SOM+ neurons in both L23 and L5 were MCs (~70%, 133/190) (fig. S10A and table S4), consistent with the long-held contention that all MCs express somatostatin (13). However, some SOM+ neurons were fast spiking and had a morphology corresponding to BCs (21.5%, 41/190) (fig. S10A and table S4) (1, 19, 20), indicating that some BCs were labeled in SOM-transgenic mice. In addition, a few SOM+ neurons in L23 had a morphology corresponding to BTCs ($n = 10/100$, 10%) (fig. S10A and table S4), consistent with previous observations that some BTCs express SOM (1). The VIP+ neurons in L23 were either BTCs (55%, 22/40) or BPCs (45%, 18/40), consistent with previous reports (fig. S10B and table S4) (1, 21, 22). VIP+ neurons in L5 were very sparse, and we did not successfully recover any L5 VIP+ neurons.
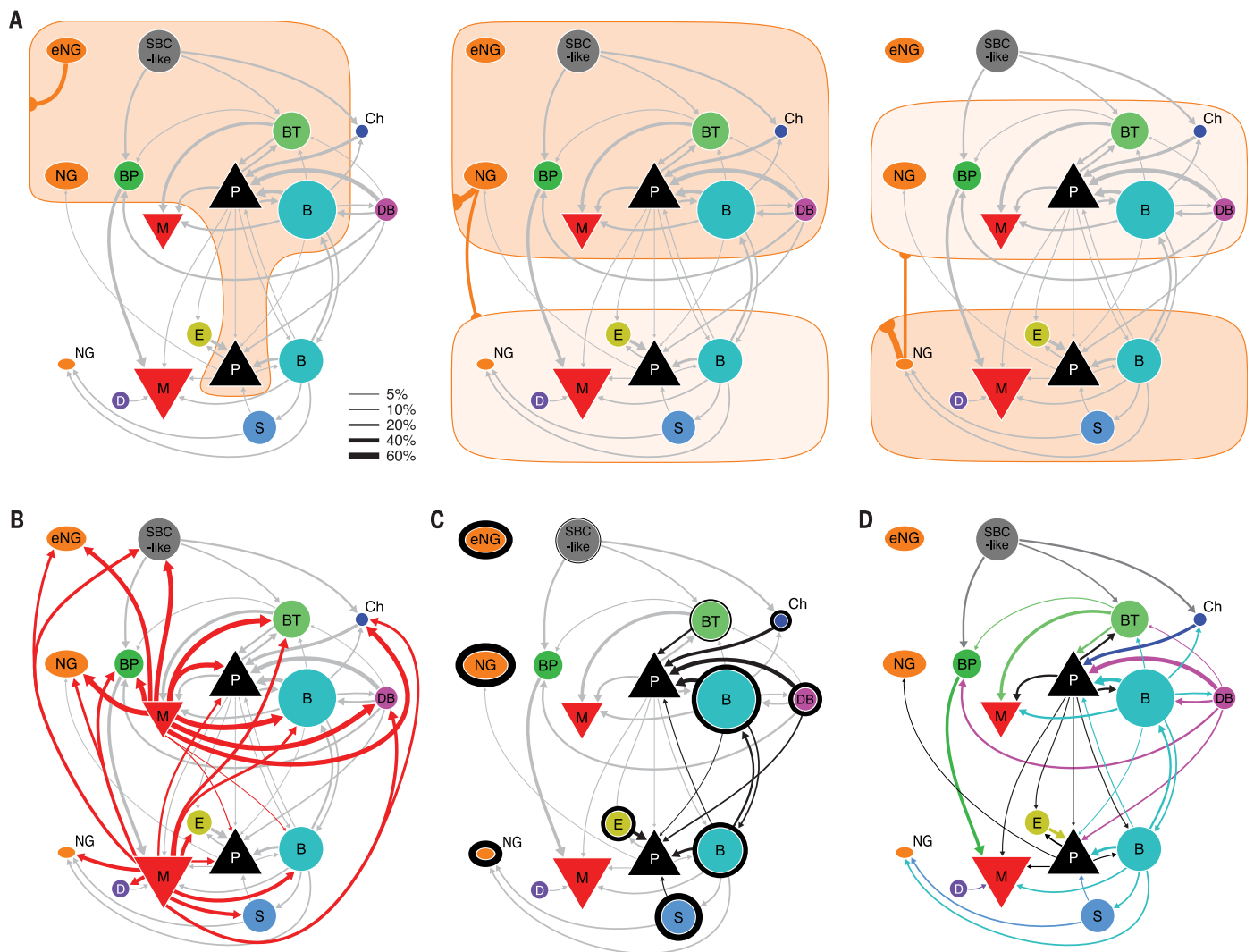
Fig. 6. Wiring diagram of V1 microcircuit. Different connectivity panels highlight different connectivity rules. Connectivity rates are indicated by line width (see legend). Connections with at least 5% connectivity rate are shown. Area of symbols indicate proportion of each cell type in the respective layer. Area of triangle depicting pyramidal cells does not represent proportion of pyramidal cells. (**A**) These diagrams illustrate volume transmission by eNGCs (left), L23NGCs (middle), and L5NGCs (right). The connections made by other neurons are shown in gray. (**B**) Connection of MCs to other cell types. (**C**) Self-inhibition is illustrated by the outline around each cell type (thickness illustrates the connectivity rate). Inhibition to pyramidal cells is highlighted in black. Connections from NGCs and MCs are omitted for clarity. (**D**) Connectivity between interneuron types and pyramidal cells is highlighted. Color of arrow is according to presynaptic cell type. Self-inhibition, volume transmission, and connections from MCs are omitted for clarity.

The overlap of morphological types with Cre driver lines suggested that specific genetic markers are expressed in some morphologically defined interneuron type (table S5). Some of the morphologically defined interneuron types could be recovered from two lines (for instance, BTC and BC), whereas some types were not found in any of the PV[+], SOM[+], or VIP[+] Cre driver lines (table S5). Specifically, all DBCs exhibited a FS firing pattern, suggesting that they might be PV[+]. However, we did not recover any DBCs from PV[+] interneurons. Similarly, the firing pattern of DCs suggested that they may express VIP (*23*, *24*), but we did not recover any DCs from VIP[+] interneurons. Finally, L1 interneurons and all NGCs across the layers could not be recovered from any of the three Cre driver lines (table S5), underscoring

the importance of identifying specific molecular markers for these cell types.

Principles of connectivity in neocortical microcircuit

Synaptic connections were identified by evoking unitary excitatory or inhibitory postsynaptic potentials [uE(I)PSPs] on postsynaptic neurons with brief depolarizing current pulses applied in presynaptic neurons (fig. S1B). We identified a total of 1680 connections (both inhibitory and excitatory) from 11,771 putative connections tested between 175 L1, 1139 L23, and 782 L5 morphologically identified neurons (Fig. 3). For inhibitory connections, the latencies of uIPSPs were typically less than 2 ms (1.84 ± 0.3 ms, $n = 1020$) unless the presynaptic neuron was a neuroglia-

form cell (latency: 4.74 ± 0.13 ms, $n = 275$) (fig. S11A) (*5*, *25*). The short latency of u(E)IPSPs and the response of the uIPSP latency to glutaminergic antagonists suggested that all identified connections were monosynaptic (fig. S11, B and C and supplementary materials).

Because interneurons generally have localized axonal arbors, their interactions with other neurons should depend on the distance between neurons. We thus computed the inhibitory connection probability as a function of intersoma Euclidean, horizontal (X) or vertical (Y) distance for all pairs (figs. S1A and S12A). Connection probability decreased sharply with distance ($P < 0.02$; Kruskal–Wallis test), leveling off after ~250 μm but not quite reaching zero, primarily due to vertical connections between

cells in different cortical layers (fig. S12, A and B). The connectivity dropped more quickly in the horizontal plane than in the vertical (fig. S12A), indicating that the inhibition from an interneuron was more far-reaching in the vertical than in the horizontal direction. The connectivity was higher between L23 neurons than between L5 neurons (fig. S12B). Because intersomatic distance is an important factor determining the connectivity, only within-layer pairs with a Euclidean distance < 250 μm and across-layer pairs with a horizontal distance < 150 μm were included in the subsequent connectivity analysis (figs. S12, A to C, and S13).

Next, we examined the connectivity at the level of the morphologically defined cell types (both interneurons and pyramidal cells) (Figs. 3 and 4A, fig. S14, and supplementary text). The connectivity of each type was not random, but highly predictable, and each morphologically distinct type of neuron had its own characteristic input-output connectivity profile, connecting with other constituent neuronal types with varying degrees of specificity in postsynaptic targets, layer location, and synaptic characteristics (Figs. 3 and 4A and fig. S14) (see the supplementary text for a detailed description). The probability of a connection between two specific types of neurons strongly correlated with their average synaptic connection strength (Fig. 5A), and this correlation applied to all types of connections, including interneuron→interneuron, interneuron→pyramidal neuron, and pyramidal neuron→interneuron connections (Fig. 5A and table S6).

Three major groups of interneurons with distinct output connection rules

Although each type of interneuron had a specific output connectivity profile, the 15 morphologically distinct interneuron types generally followed three output connection rules and thus could be divided into three major groups. The first group consisted of interneurons that appeared to project nonspecifically to almost all neuronal types within the home layer and some of them even to almost all neuronal types across several layers (for instance, L5MCs). Because of this connection pattern, we call this group "master regulators." It includes NGCs and MCs in different layers (Figs. 4A and 6, A and B; fig. S14; and supplementary text). The second group consisted of interneuron types that only projected to other interneurons and not to pyramidal cells, and we refer to them as interneuron-selective interneurons (ISIs), which included SBC-like cell, BPCs, and DCs (Figs. 4A and 6D, fig. S14, and supplementary text). The remaining interneuron types projected to local pyramidal neurons, and we refer to them as pyramidal-neuron-targeting interneurons (PTIs), which included L23BCs, L5BCs, DBCs, ChCs, HECs, SCs, and BTCs (Figs. 4A and 6C and fig. S14). Moreover, most PTIs (L23BCs, L5BCs, DBCs, ChCs, HECs, and SCs) also preferentially projected to interneurons of the same morphological type (Figs. 4A and 6C and fig. S14). In addition to projecting to pyramidal neurons and themselves, some types of PTIs projected to certain

specific types of interneuron with distinct connectivity (for instance, BTCs and BCs projected to MCs) (Figs. 4A and 6D and fig. S14). Neurons from each of these three classes (master, PTI, and ISI) could be found in each cortical layer. For instance, MCs in L23 and L5 provided a uniform inhibition to every neuronal type (except MCs) within L23 and L5, respectively, and the layers above them (Fig. 4A and 6B and fig. S14); NGCs in L1, L23, and L5 provided a uniform inhibition to every neuronal type within L1, L23, and L5, respectively, and in the nearby layers (Fig. 4A and 6B and fig. S14). The three ISIs are equally distributed between layers and serve a disinhibitory role. Several L23 PTIs appeared to have closely matched cell types in L5 (Fig. 4A and fig. S14), and the connectivity motifs of PTIs are recycled across L23 and L5.

Two master regulators with distinct input profiles

NGCs and MCs projected to almost every neuronal type with a similar high connection probability (Figs. 4A and 6, A and B, and fig. S14), indicating that NGCs and MCs act as master regulators by providing nonspecific inhibition to the local neocortical circuit. However, NGCs and MCs have very different input connectivity profiles and output mechanisms, suggesting that these two master regulators operate in fundamentally different ways.

Within their home layer, a single AP in NGCs elicited large, slow uIPSPs in almost all simultaneously recorded neurons, regardless of their type (Fig. 3C). This was observed in NGCs from all layers (L1, L23, and L5) (Fig. 3, A and C; and figs. S5A; S6, A and B; and S14). The connection probability from NGCs reached up to 90% if postsynaptic somata were located ≤100 μm apart. This exceptionally high connectivity, the lack of postsynaptic specificity, and the unusual, slow synaptic events evoked by NGCs support the notion that NGCs use volume transmission (25, 26) (Fig. 6A). Projection patterns from L23 and L5 NGCs to nearby layers are also consistent with this idea (Figs. 4A and 6A and fig. S14). Nevertheless, L1 eNGCs exhibited a certain degree of postsynaptic specificity in projecting to L23 neurons (they preferentially project to BPC, BTC, and NGCs in L23) (Figs. 4A and 6A and fig. S14), which seems to argue against volume transmission for this cell type (27). However, because the axon of eNGCs is restricted to L1, for any L23 neuron to interact with an eNGC, it must have a dendritic tree extending into L1. The preferential projection to BPC, BTC, and NGCs thus may simply reflect that more of their dendritic trees extend into L1, making them more likely to be exposed to a cloud of GABA released there than other types of L23 interneurons. Therefore, the projection pattern from L1 eNGCs to L23 may still be consistent with the mechanism of volume transmission (Fig. 6A).

Similar to NGCs, a single spike in MCs elicited uIPSPs in all non-MC neuronal types within home layers and the layers above them (Figs. 3, A and B, and 6B; and figs. S5, A and D; S6, A and C;

S8, A to C; and S9A). However, MCs differ from NGCs in several important ways. First, MCs use synaptic transmission rather than volume transmission, as evidenced by the fast kinetics of their uIPSPs (fig. S11) and their ability to selectively avoid inhibiting other MCs (Figs. 4A and 6B and fig. S14). Second, MCs received strong inputs from local pyramidal neurons with strongly facilitating synapses (15.4% and 4.5% from L23 and L5 pyramidal neurons, respectively, pooled across all MC types) (Figs. 5B and 6D and fig. S14), whereas NGCs received very little input from local pyramidal neurons, especially from L23 pyramidal neurons (0.0% and 2.1% from L23 and L5 pyramidal neurons, respectively; pooled across all NGC types) (Figs. 4A and 5B and supplementary text) [generalized linear model (GLM) with factors cell type and layer; effect of cell type: $P = 0.002$; layer: $P = 0.0006$; interaction: $P = 0.03$; for details of model fitting, see the supplementary materials]. In addition, MCs received very specific inhibition from local circuits: BCs projected to MCs within L23 and L5, BTCs projected to L23MCs, and BPCs projected almost exclusively to L5MCs (Figs. 4A and 6D and fig. S14), whereas NGCs received little specific inhibition from local circuits (aside from NGC and MC input, which is nonspecific) (Fig. 5B; $P < 0.0001$ compared with MC, Chi-square test). Therefore, MCs are locally controlled, in contrast to NGCs, which may be primarily controlled by long-range inputs. Although both NGCs and MCs serve as master regulators for the cortical microcircuit, the specific computational operations implemented by these two types may be fundamentally different.

Interneuron-selective interneurons are neither self-inhibitory nor locally recruited

ISIs included L1 SBC-like cells (0% connection to pyramidal neurons, 0/152), BPCs (0% connection to pyramidal neurons, 0/115), and DCs (0% connection to pyramidal neurons, 0/46) (Figs. 4A and 6D and fig. S14). Although all ISIs targeted only interneurons, the three ISIs had different preferences in postsynaptic targets: BPCs and DCs projected almost exclusively to L5MCs, whereas SBC-like cells targeted several interneuron types in L23 but avoided MCs (Figs. 4A and 6D and fig. S14). However, because the axon of DCs projects toward the deeper cortical areas, even into L6, they may additionally project to the interneurons in L6. All these types generally did not exhibit self-inhibition (Figs. 4A; 6, C and D; 5C; and fig. S14), which was seen in all PTI interneuron types. Interneuron-targeting interneurons also received very little input from local pyramidal neurons and specific inhibitory inputs, very different from other groups of interneurons (Fig. 5, C and D). ISIs thus may act primarily as disinhibitors of the local cortical microcircuit and may be primarily controlled by long-range inputs.

Pyramidal-neuron-targeting interneurons are self-inhibitory and locally recruited

PTIs (the seven types described above) connected to pyramidal neurons and to other interneurons

of the same morphological type (self-inhibition) (Figs. 4A and 6C and fig. S14). For most of them (except BTCs), self-inhibition was the strongest output connection. This connectivity motif occurred not only within a layer but also across layers (L23BCs and L5 BCs preferentially connect with BCs and pyramidal neurons in both L23 and L5) (Figs. 4A and 6C and fig. S14). The connection rule shared by this group of neurons is thus very different from that of master regulators and ISIs (Figs. 4A; 5C; and 6, C and D; and fig. S14). In addition, there was a strong positive correlation between connectivity to pyramidal neurons and self-inhibition probability (Fig. 5C) [$r(11)$ = 0.72; P = 0.01], indicating that a PTI's ability to inhibit local pyramidal neurons faces comparative competition from interneurons of the same morphological type.

Many PTI cell types not only inhibited local pyramidal neurons but also received strong excitatory inputs from these neurons via synapses that depressed upon repetitive stimulation (Figs. 4A and 6D and figs. S9F and S14). This reciprocal connectivity motif also occurred between interneurons and pyramidal neurons located in different cortical layers. For example, BCs in L23 projected to pyramidal neurons in L5, and they also received excitatory inputs from L5 pyramidal neurons; the same was true for L5BCs (Figs. 4A and 6D and fig. S14). This reciprocal connectivity motif between PTIs and pyramidal neurons is very different from ISIs. ISIs did not target local pyramidal neurons, and they also received very little inputs from local pyramidal neurons (Figs. 4A, 5D, and 6D and fig. S14). These observations tend to indicate that local pyramidal neurons tend to recruit interneuron types that inhibit pyramidal neurons. Nevertheless, local pyramidal neurons do not necessarily recruit all interneuron types that inhibit pyramidal neurons. There were several types of PTIs, such as DBC, that did not receive innervation from pyramidal neurons.

A simple model incorporating three connectivity rules (master, ISI, and PTI) captures most of the connectivity structure

To determine whether the three connectivity rules (master, ISI, and PTI) that we observed represent general connection rules with explanatory power, we devised a series of simple models of the connectivity matrix between the reported cell types (Fig. 4B and supplementary materials). We started from a uniform model that assumes that connectivity probabilities for each type of connections are equal (1 parameter). Then we gradually added details (layer identity, inhibitory or excitatory cell types, and the three principles) into the model, obtaining a series of models. We compared the likelihoods of these models with that of the full model, which has an individual connection probability parameter for each type of connections (289 parameters) (Fig. 4B). We wanted to see under what restrictions the model would still recapitulate the essential connectivity structure of the full model. When incorporating

the three connectivity rules (master, ISI, and PTI), the model came within one bit of the performance of the full model with only about a fourth of the degrees of freedom (d = 67) (Fig. 4B), indicating that the identified rules captured the essential connectivity structure among the types of neurons.

Comparison to connectivity of juvenile cell types

Connectivity patterns of three molecularly identified interneuronal types (PV[+], SOM[+], and VIP[+]) in the developing neocortex have been recently revealed by optogenetic studies (28–30). Although these studies are informative about the broad pattern of connections between molecularly identified interneuronal types within a layer in the developing neocortex, they are unable to capture the full functional diversity of interneurons, given that interneurons labeled in the same Cre line (PV-Cre, VIP-Cre, and SOM-Cre) in mature cortex had distinct morphologies (fig. S10, A and B, and tables S4 and S5), and the morphologically distinct interneurons from the same Cre lines had different connectivity profiles within and across layers (fig. S10C and tables S7 to S9). For instance, the L5 PV[+] neurons could be morphologically classified into L5BCs, SCs, and HECs (fig. S10B), and L5BCs, SCs, and HECs, regardless of the mouse lines that they were recovered from, are FS (Fig. 2C). In addition, the connectivity of L5BCs, SCs, and HECs from different mouse lines (Viatt/ai9 and wild-type) was very similar to the connectivity of the corresponding cell types in the PV[+] Cre line (fig. S10C and table S7). All these observations strongly suggested that all L5BCs, SCs, and HECs express PV, but these L5 PV-expressing neurons do not have the same connectivity profiles within and across the layers (Fig. 4A and fig. S14). In addition, it has been increasingly recognized that the SOM-Cre driver line (SOM-IRES-Cre) (31) labels a population of neurons that can be grouped into distinct types both functionally (19, 32) and morphologically (fig. S10A and tables S3 and S4) (20). In contrast to what has been reported in optogenetic studies [SOM[+] neurons avoid each other (28)], there is a high connectivity between SOM[+] MCs and SOM[+] non-MCs. Within the SOM[+] class, only MCs avoid connecting to each other (table S8). Moreover, several morphologically defined types could be labeled in several Cre lines (tables S4 and S5), but connectivity patterns of the same morphological types from different lines appeared to be very similar. For instance, BTCs could be labeled in both VIP-Cre and SOM-Cre lines, and we did not observe any morphological differences between VIP[+] BTC and SOM[+] BTC (fig. S10, A and B). Although differences in other neuronal features may exist between them, they appeared to have very similar connectivity profiles (they both connected to pyramidal neurons and MCs and received input from pyramidal neurons) (tables S8 and S9 and fig. S10C). These observations provide another line of evidence that the axonal morphology of an interneuron indeed determines its connection rule.

In addition to the above-mentioned discrepancies, we noticed that several connection patterns we observed in mature neocortex appeared to be different from that in the developing neocortex. First, NGCs received very little input from local excitatory neurons (33). However, in the developing somatosensory cortex, NGCs receive excitatory inputs from L23 (34, 35). Second, there is a fair connection probability from BCs to MCs in both L23 and L5 (Figs. 4 and 6C and fig. S14), indicating that PV[+] cells do project to SOM[+] cells in mature neocortex, in contrast to a study performed in the developing neocortex (28). Third, inhibitory projections to pyramidal neurons in mature neocortex appeared to be different from the developing neocortex. In this study, we did not identify any connection from BPCs to L23 or L5 pyramidal neurons or from BTCs to L5 pyramidal neurons (Figs. 4 and 6C and fig. S14), but these types of connections have been frequently identified in the developing neocortex (5). Finally, the connectivity among mature pyramidal neurons, particularly among L5 pyramidal neurons, was much lower than the connectivity among pyramidal neurons within the same range of intersoma distance in juvenile slices [figs. S13B (average, 91 ± 4μm) and S14 and supplementary text] (36, 37). Although other experimental differences might explain some of these discrepancies (e.g., methods, cortical area, species, and low connection probability), the most compelling and consistent difference across experiments is the age of the animals tested, suggesting that mature cortical circuits are not identical to developing circuits.

Conclusions

Establishing a complete census of cell types in the neocortex and their wiring diagrams poses a tremendous technical challenge but is key to mechanistically understanding complex cognitive functions, such as perception, memory, and decision-making. This study provides the most comprehensive wiring diagram of the adult neocortical microcircuit to date at the level of morphologically and electrophysiologically defined types of cells and reveals that neocortical microcircuits are built from a small number of simple connectivity motifs that are recycled across the layers to generate the essential cortical connectivity structure. The finding that the complex functional cortical architecture can be broken down into a small number of connectivity motifs may lead to breakthroughs in our understanding of cortical computation at the circuit level. Given that numerous neuropsychiatric and neurological diseases—such as autism spectrum disorders and epilepsy—may be associated with cell type–specific connectivity change in cortical microcircuits (18, 38, 39), such a detailed blueprint of cortical microcircuits could serve as an invaluable platform for screening and pinpointing specific circuit abnormalities in animal models of disease, thus providing a path to the development of comprehensive circuit-based, cell type–specific interventions that are otherwise not addressable with current treatments.

REFERENCES AND NOTES

1. H. Markram et al., Interneurons of the neocortical inhibitory system. Nat. Rev. Neurosci. 5, 793–807 (2004). doi: 10.1038/nrn1519; pmid: 15378039

2. J. DeFelipe et al., New insights into the classification and nomenclature of cortical GABAergic interneurons. Nat. Rev. Neurosci. 14, 202–216 (2013). pmid: 23385869

3. D. Dumitriu, R. Cossart, J. Huang, R. Yuste, Correlation between axonal morphologies and synaptic input kinetics of interneurons from mouse visual cortex. Cereb. Cortex 17, 81–91 (2007). doi: 10.1093/cercor/bhj126; pmid: 16467567

4. G. A. Ascoli et al., Petilla terminology: Nomenclature of features of GABAergic interneurons of the cerebral cortex. Nat. Rev. Neurosci. 9, 557–568 (2008). doi: 10.1038/nrn2402; pmid: 18568015

5. X. Jiang, G. Wang, A. J. Lee, R. L. Stornetta, J. J. Zhu, The organization of two new cortical interneuronal circuits. Nat. Neurosci. 16, 210–218 (2013). doi: 10.1038/nn.3305; pmid: 23313910

6. A. Kepecs, G. Fishell, Interneuron cell types are fit to function. Nature 505, 318–326 (2014). doi: 10.1038/nature12983; pmid: 24429630

7. C. Le Magueresse, H. Monyer, GABAergic interneurons shape the functional maturation of the cortex. Neuron 77, 388–405 (2013). doi: 10.1016/j.neuron.2013.01.011; pmid: 23395369

8. Z. J. Huang, G. Di Cristo, F. Ango, Development of GABA innervation in the cerebral and cerebellar cortices. Nat. Rev. Neurosci. 8, 673–686 (2007). doi: 10.1038/nrn2188; pmid: 17704810

9. J. Ting, T. Daigle, Q. Chen, G. Feng, in Patch-Clamp Methods and Protocols, M. Martina, S. Taverna, Eds. (Springer New York, 2014), vol. 1183, chap. 14, pp. 221-242.

10. H. S. Meyer et al., Inhibitory interneurons in a cortical column form hot zones of inhibition in layers 2 and 5A. Proc. Natl. Acad. Sci. U.S.A. 108, 16807–16812 (2011). doi: 10.1073/pnas.1113648108; pmid: 21949377

11. G. S. Jefferis et al., Comprehensive maps of Drosophila higher olfactory centers: Spatially segregated fruit and pheromone representation. Cell 128, 1187–1203 (2007). doi: 10.1016/j.cell.2007.01.040; pmid: 17382886

12. S. Hestrin, W. E. Armstrong, Morphology and physiology of cortical neurons in layer I. J. Neurosci. 16, 5290–5300 (1996). pmid: 8757242

13. Y. Wang et al., Anatomical, physiological and molecular properties of Martinotti cells in the somatosensory cortex of the juvenile rat. J. Physiol. 561, 65–90 (2004). doi: 10.1113/jphysiol.2004.073353; pmid: 15331670

14. A. Gupta, Y. Wang, H. Markram, Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex. Science 287, 273–278 (2000). doi: 10.1126/science.287.5451.273; pmid: 10634775

15. K. A. Buchanan et al., Target-specific expression of presynaptic NMDA receptors in neocortical microcircuits. Neuron 75, 451–466 (2012). doi: 10.1016/j.neuron.2012.06.017; pmid: 22884329

16. M. Le Bon-Jego, R. Yuste, Persistently active, pacemaker-like neurons in neocortex. Front. Neurosci. 1, 123–129 (2007). doi: 10.3389/neuro.01.1.1.009.2007; pmid: 18982123

17. Z. Xiang, D. A. Prince, Heterogeneous actions of serotonin on interneurons in rat visual cortex. J. Neurophysiol. 89, 1278–1287 (2003). doi: 10.1152/jn.00533.2002; pmid: 12626611

18. H.-T. Chao et al., Dysfunction in GABA signalling mediates autism-like stereotypies and Rett syndrome phenotypes. Nature 468, 263–269 (2010). doi: 10.1038/nature09582; pmid: 21068835

19. H. Hu, J. Z. Cavendish, A. Agmon, Not all that glitters is gold: Off-target recombination in the somatostatin-IRES-Cre mouse line labels a subset of fast-spiking interneurons. Front. Neural Circuits 7, 195 (2013). doi: 10.3389/fncir.2013.00195; pmid: 24339803

20. M. Nassar et al., Diversity and overlap of parvalbumin and somatostatin expressing interneurons in mouse presubiculum. Front. Neural Circuits 9, 20 (2015). doi: 10.3389/fncir.2015.00020; pmid: 26005406

21. J. T. Porter et al., Properties of bipolar VIPergic interneurons and their excitation by pyramidal neurons in the rat neocortex. Eur. J. Neurosci. 10, 3617–3628 (1998). doi: 10.1046/j.1460-9568.1998.00367.x; pmid: 9875341

22. B. Cauli et al., Molecular and physiological diversity of cortical nonpyramidal cells. J. Neurosci. 17, 3894–3906 (1997). pmid: 9133407

23. Y. Kawaguchi, Y. Kubota, Physiological and morphological identification of somatostatin- or vasoactive intestinal polypeptide-containing cells among GABAergic cell subtypes in rat frontal cortex. J. Neurosci. 16, 2701–2715 (1996). pmid: 8786446

24. S. Lee, J. Hjerling-Leffler, E. Zagha, G. Fishell, B. Rudy, The largest group of superficial neocortical GABAergic interneurons expresses ionotropic serotonin receptors. J. Neurosci. 30, 16796–16808 (2010). doi: 10.1523/JNEUROSCI.1869-10.2010; pmid: 21159951

25. G. Tamás, A. Lorincz, A. Simon, J. Szabadics, Identified sources and targets of slow inhibition in the neocortex. Science 299, 1902–1905 (2003). doi: 10.1126/science.1082053; pmid: 12649485

26. S. Oláh et al., Regulation of cortical microcircuits by unitary GABA-mediated volume transmission. Nature 461, 1278–1281 (2009). doi: 10.1038/nature08503; pmid: 19865171

27. R. Chittajallu, K. A. Pelkey, C. J. McBain, Neurogliaform cells dynamically regulate somatosensory integration via synapse-specific modulation. Nat. Neurosci. 16, 13–15 (2013). doi: 10.1038/nn.3284; pmid: 23222912

28. C. K. Pfeffer, M. Xue, M. He, Z. J. Huang, M. Scanziani, Inhibition of inhibition in visual cortex: The logic of connections between molecularly distinct interneurons. Nat. Neurosci. 16, 1068–1076 (2013). doi: 10.1038/nn.3446; pmid: 23817549

29. S. Lee, I. Kruglikov, Z. J. Huang, G. Fishell, B. Rudy, A disinhibitory circuit mediates motor integration in the somatosensory cortex. Nat. Neurosci. 16, 1662–1670 (2013). doi: 10.1038/nn.3544; pmid: 24097044

30. H. J. Pi et al., Cortical interneurons that specialize in disinhibitory control. Nature 503, 521–524 (2013). doi: 10.1038/nature12676; pmid: 24097352

31. H. Taniguchi et al., A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. Neuron 71, 995–1013 (2011). doi: 10.1016/j.neuron.2011.07.026; pmid: 21943598

32. J. Reimer et al., Pupil fluctuations track fast switching of cortical states during quiet wakefulness. Neuron 84, 355–362 (2014). doi: 10.1016/j.neuron.2014.09.033; pmid: 25374359

33. C. Wozny, S. R. Williams, Specificity of synaptic connectivity between layer 1 inhibitory interneurons and layer 2/3 pyramidal neurons in the rat neocortex. Cereb. Cortex 21, 1818–1826 (2011). doi: 10.1093/cercor/bhq257; pmid: 21220765

34. N. V. De Marco García, R. Priya, S. N. Tuncdemir, G. Fishell, T. Karayannis, Sensory inputs control the integration of neurogliaform interneurons into cortical circuits. Nat. Neurosci. 18, 393–401 (2015). doi: 10.1038/nn.3946; pmid: 25664912

35. X. Xu, E. M. Callaway, Laminar specificity of functional input to distinct types of inhibitory cortical neurons. J. Neurosci. 29, 70–85 (2009). doi: 10.1523/JNEUROSCI.4104-08.2009; pmid: 19129386

36. H. Markram, J. Lübke, M. Frotscher, A. Roth, B. Sakmann, Physiology and anatomy of synaptic connections between thick tufted pyramidal neurones in the developing rat neocortex. J. Physiol. 500, 409–440 (1997). doi: 10.1113/jphysiol.1997.sp022031; pmid: 9147328

37. R. Perin, T. K. Berger, H. Markram, A synaptic organizing principle for cortical neuronal groups. Proc. Natl. Acad. Sci. U.S.A. 108, 5419–5424 (2011). doi: 10.1073/pnas.1016051108; pmid: 21383177

38. D. R. Dajani, L. Q. Uddin, Local brain connectivity across development in autism spectrum disorder: A cross-sectional investigation. Autism Res. xxx, xxx (2015). pmid: 26058882

39. K. Menuz, R. A. Nicoll, Loss of inhibitory neuron AMPA receptors contributes to ataxia and epilepsy in stargazer mice. J. Neurosci. 28, 10599–10603 (2008). doi: 10.1523/JNEUROSCI.2732-08.2008; pmid: 18923036

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/350/6264/aac9462/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S10
References (40–64)

ECOLOGY

A continent-wide assessment of the form and intensity of large mammal herbivory in Africa

Gareth P. Hempson,[1,2]* Sally Archibald,[2] William J. Bond[1,3]

Megafaunal extinctions and a lack of suitable remote sensing technology impede our understanding of both the ecological legacy and current impacts of large mammal herbivores in the Earth system. To address this, we reconstructed the form and intensity of herbivory pressure across sub-Saharan Africa ~1000 years ago. Specifically, we modeled and mapped species-level biomass for 92 large mammal herbivores using census data, species distributions, and environmental covariates. Trait-based classifications of these species into herbivore functional types, and analyses of their biomass surfaces, reveal four ecologically distinct continental-scale herbivory regimes, characterized by internally similar forms and intensities of herbivory pressure. Associations between herbivory regimes, fire prevalence, soil nutrient status, and rainfall provide important insights into African ecology and pave the way for integrating herbivores into global-scale studies.

The past 15 years have witnessed a revolution in our understanding of how fire shapes ecosystems at global scales (1–4). However, the relevance of another major consumer of vegetation, the large mammalian herbivores, has been largely neglected. This is in part because remote sensing tools, such as the Moderate Resolution Imaging Spectroradiometer (MODIS), provide near-daily global data sets that allow fire extent, frequency, and intensity to be quantified at 250-m resolution (5, 6); equivalent tools are not available for directly quantifying herbivore community composition and biomass. An important application of the well-developed understanding of fire at macroscales has been its implementation in Dynamic Global Vegetation Models (DGVMs). These global models allow researchers to explore the implications of global climate change for fire regimes and the concomitant impacts on global vegetation patterns and Earth system feedbacks (7–10). Looking backward in time, an understanding of climate–vegetation–fire relationships also allows reconstruction of the environmental contexts that have shaped biotic evolutionary trajectories (11–13). Like fire, large herbivores can have a substantial impact on the structure and function of ecological communities (14–18), and past alterations in the composition and degree of mammalian herbivory (e.g., Pleistocene megafaunal extinctions) have affected global vegetation and Earth system

processes (19–22). There is thus an urgent need to develop tools that explore how the biotic interactions of herbivores and other consumers have shaped current and past ecological function, processes, and resilience across global ecosystems (23–25).

Early ecologists organized the diverse vegetation on Earth into biomes, or global vegetation units that internally share similar vegetation functional traits and respond in the same way to environmental drivers (26, 27). Plant functional types form the basis of these classifications by simplifying vast species-level diversity into groups based on common biological attributes (e.g., deciduous versus evergreen trees) (28–30). More recently, access to global-scale information about fire has enabled the delineation of global fire regimes, or "pyromes," parts of the globe grouped by their shared fire characteristics and responses to environmental change (31). Key to both of these classification systems is the assessment of syndromes of plant functional types or fire characteristics that, because of energetic or environmental constraints, tend to occur together. This categorization can be useful in assessing how the distributions of biomes and pyromes might change in response to environmental change, and how closely linked fire regimes and vegetation distributions are. Using the same approach to assess global syndromes of herbivory can inform the understanding of interactions between herbivores and other components of the Earth system.

Determining large-scale herbivory patterns

We reconstructed herbivore biomass surfaces across sub-Saharan Africa for 92 large mammal herbivore species, using statistical models fitted to census data from protected areas. Rainfall, soil

nutrient status, and vegetation patterns were fitted as covariates and used to predict species' biomass across their distribution range, which was then refined by fine-scale habitat preferences (figs. S1 and S2 and tables S1 and S2) (32). Our aim was to quantify herbivore biomass before human hunting with guns, and thus potential herbivory surfaces, ~1000 years ago. Data availability and model fit determined whether species densities (individuals/km²) were modeled (using analysis of covariance) at (i) vegetation group level (forest, caesalpinioid savanna, mixed savanna, grassland, or shrubland) or (ii) across their full distribution range. Alternately, median densities were applied at (iii) vegetation group or (iv) overall levels; otherwise, (v) an expert opinion estimate was applied. Elephant density estimates were based on 95th-quantile regression to account for reduced current-day densities, which are due to their large home ranges exceeding the extent of protected areas and considerable current and historical hunting pressure. Species densities were reduced to 10 or 0% of predicted values for low-suitability or unsuitable habitats and converted to biomass (kg/km²) using average adult body mass. This analysis provides the first large-scale mapped information on wild herbivore biomass [although the United Nations Food and Agriculture Organization (FAO) has produced spatial data on livestock densities for several continents] (Fig. 1). Previous similar studies have fitted species- (33) and community-level (34–36) regression equations relating biomass to rainfall and soil nutrient status [(37) analyzes large herbivore diversity patterns]. These studies restricted their focus to large mammal herbivores in the savanna biome and did not attempt to create spatial biomass surfaces.
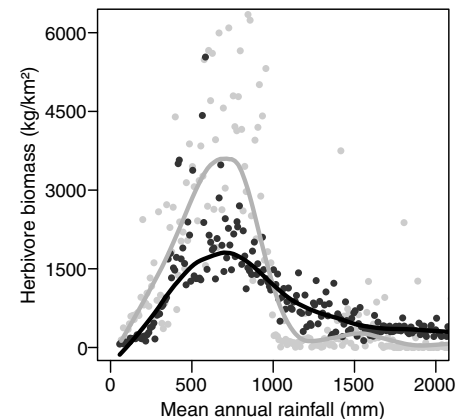
Fig. 1. Relationships between African herbivore biomass and rainfall. Reconstructed historical total wildlife biomass (black points and black line) and current FAO livestock biomass estimates (gray points and gray line) are shown in relation to rainfall. Elephants overwhelm the signal from the other species and are excluded from the figure. Points represent median values for 10-mm rainfall intervals and are shown with locally weighted scatterplot smoothing regression lines.

[1]Department of Biological Sciences, University of Cape Town, Private Bag X1, Rondebosch, 7701, South Africa. [2]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Wits, 2050, South Africa. [3]South African Environmental Observation Network, care of the Department of Biological Sciences, University of Cape Town, Private Bag X1, Rondebosch, 7701, South Africa.
*Corresponding author. E-mail: ghempson@gmail.com

We reduced species-level complexity by trait-based classifications of species into herbivore functional types (HFTs), analogous to grass, shrubs, and trees in a vegetation context. Hierarchical cluster analysis was used to partition species into HFTs based on five traits: body mass, diet, gut type, herd size, and water dependence (see Fig. 2 and tables S3 and S4 for detailed trait information). These traits are important for understanding consumption patterns in terms of the amount (body mass and gut type), type (diet and gut type), and spatial patterning (herd size and water dependence) of forage offtake. Furthermore, in a tropical and subtropical African context, these traits largely capture the forage resource and habitat characteristics of species (including thermal tolerance) that will shape their responses to global environmental changes. Our HFT classifications should thus prove useful for populating DGVMs with large mammal herbivores (38). HFT biomass surfaces were created by combining species-level biomass surfaces, enabling each 1° grid square (~12,000 km² at the equator) to be characterized by its total herbivore biomass (kg/km²) and that of each HFT. Grouping 1° squares with similar herbivore biomass characteristics using cluster analyses allows delimitation of different large mammal herbivory regimes, or "herbivomes," analogous to biome and pyrome classifications. These large-scale patterns in the relative abundance of HFTs and total herbivore biomass show how the form and intensity of consumptive offtake by large mammal herbivores vary across sub-Saharan Africa.

Biomass surfaces

Elephants dominate African herbivore biomass, often having biomasses equivalent to those of all other species combined (fig. S3). Excluding elephants, which obscure biomass patterns, reveals a unimodal relationship between potential African mammal herbivore biomass and mean annual rainfall (MAR), peaking at ~1700 kg/km² and ~700 mm MAR (Fig. 1). This finding contrasts with previous studies that report a constant linear increase in the log biomass–log MAR relationship from ~150 to 1200 mm (33–35), implying an accelerating increase in herbivore biomass gain with MAR. Our analyses include data from these earlier studies, and disparities are attributable to differences in spatial resolution, our incorporation of vegetation attributes and habitat preference, and the sensitivity of their analyses to biomass declines at high rainfalls when analyzing log-transformed data. The general functional form of our biomass-rainfall curve is in close agreement with present-day estimates of livestock biomass (www.fao.org). Peak densities of livestock are almost two times higher between 500 and 750 mm MAR [perhaps reflecting management interventions such as the provision of drinking water and supplemental feed and the suppression of disease and predation (39)], but modeled precolonial wildlife densities are higher above 1000 mm MAR. The inclusion of elephant biomass predictions further increases the disparity between wildlife and livestock biomass at high rainfall. An elephant's massive body size and mixed diet enable effective utilization of diverse and even low-quality forage typical of high-rainfall areas, and elephants thus exhibit a lack of response to soil nutrient status and vegetation composition (14). This is reflected in their former continent-wide distribution, with biomass patterns responding primarily to rainfall at our modeling resolution (36). Surface-water availability is a key determinant of the distribution of this wide-ranging but water-dependent species
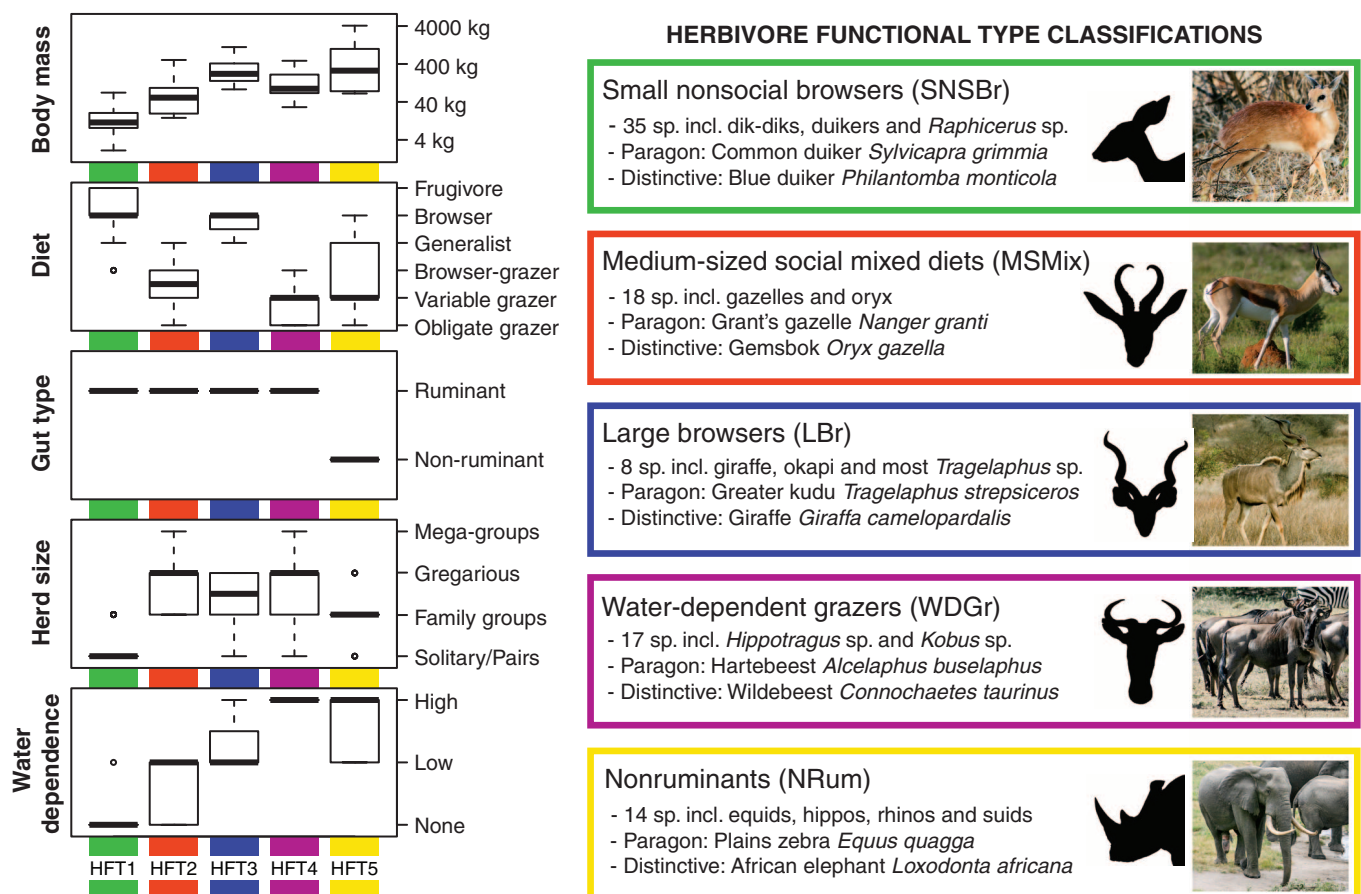
Fig. 2. HFT classifications. Vertical color bands with the trait boxplots (left) match the boxes bounding the HFT descriptions (right). Dietary classifications follow (54): obligate grazer (>90% monocots, not variable), variable grazer (60 to 90% monocots, variable), browser-grazer intermediate (30 to 70% dicots and monocots, <20% fruits), generalist (>20% of all food types), browser (>70% dicots), and frugivore (>70% fruits, few or no monocots). Paragon species are close to the centroid of the HFT grouping, and distinctive species are furthest from other HFT groupings.
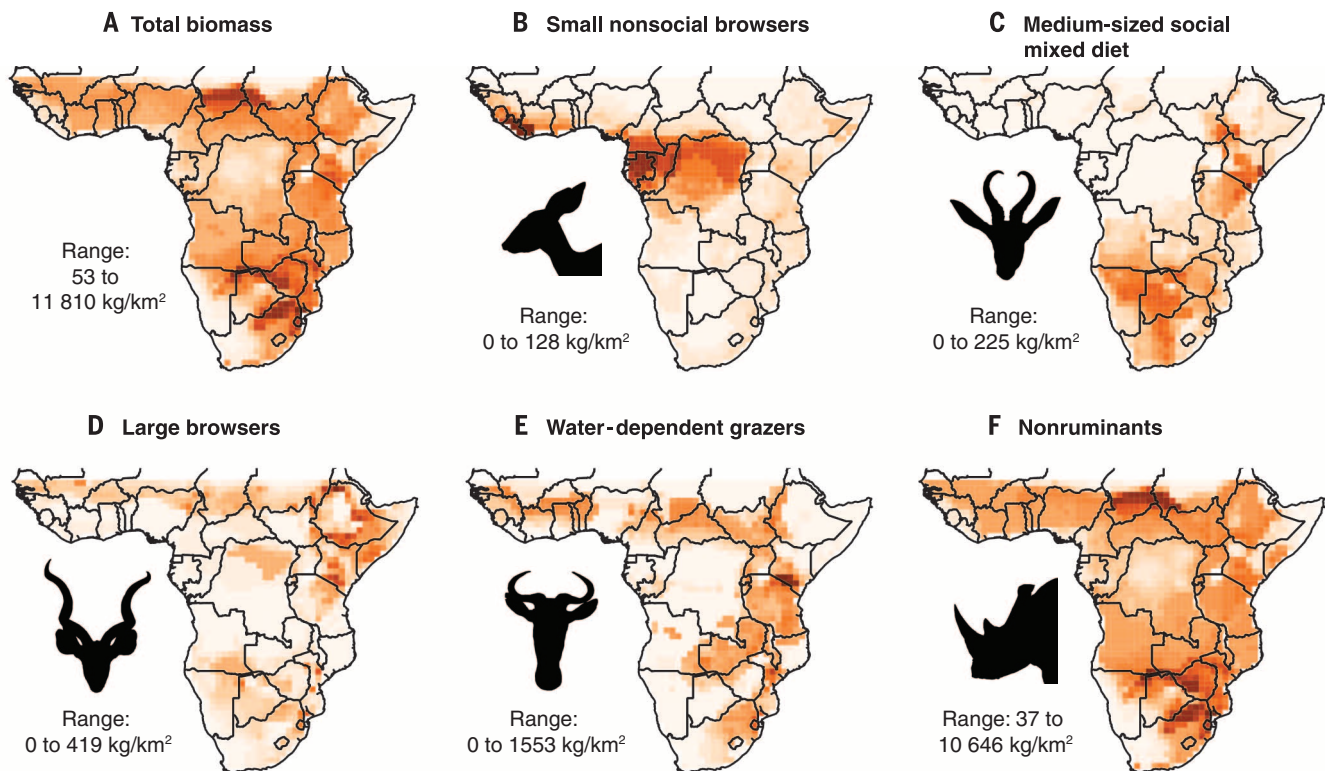
Fig. 3. Total and HFT biomass. (**A** to **F**) Species-level biomass estimates (kg/km^2) were combined to reconstruct total and HFT biomass surfaces ~1000 years ago. In each panel, the darkest colors correspond to the maximum biomass in the given range.

(40), but spatial data sets of the requisite detail are not yet available at the continental scale.

HFT classification

We identified five HFTs (Fig. 2 and figs. S4 and S5), with each of the five traits contributing significantly to their delimitation (table S5). Small nonsocial browser species form the largest grouping (35 species) and are selective foragers that can meet their water requirements from their high-quality food sources. This group includes species ranging in habitat from the rainforest (e.g., forest duikers) to the desert (e.g., dik-diks). The medium-sized social mixed-diet group (18 species) tends to consist of species from drier, seasonal environments (e.g., the springbok, Thomson's gazelle, and Grant's gazelle), many of which will switch from selective utilization of new grass or forb growth during the wet season to a browse-dominated diet in the dry season. Eight large-browser species group together (e.g., the greater kudu and giraffe) and are larger, more social, and more water-dependent than the small nonsocial browser group. Water-dependent grazing ruminant species form a cluster of 17 species, many of which are also highly social (e.g., the common wildebeest). Most floodplain-associated species are included in the water-dependent grazer group (e.g., Nile, red, and black lechwes; the puku; and the sitatunga). The final grouping consists of all 14 nonruminant species, including elephants. Aside from the importance of gut type in delimiting this group, these species also tend to be large and

water-dependent, and most incorporate both grass and browse plants in their diet. A full list of species traits and functional type classifications is provided in table S1.

HFT biomass

Combining species-level biomass surfaces by their HFT classification reveals distinct patterns in their relative biomass (Fig. 3). Small nonsocial browsers (Fig. 3B) occur at their highest densities in the forested regions of tropical Africa and display little high-density overlap with either medium-sized social mixed-diet (Fig. 3C) or water-dependent grazer species (Fig. 3E). Medium-sized social mixed-diet species are the only group with relatively high biomass in the western parts of southern Africa and are of perhaps unexpectedly low abundance in open habitats in West Africa. Large browsers (Fig. 3D) and water-dependent grazers species are widely spread in open grassy habitats across the continent, although large browsers are more prominent in the horn of Africa. These patterns (Fig. 3, B to E), however, are overwhelmed by the dominance of nonruminant species (Fig. 3F), which almost exclusively determine patterns in the total biomass of African large mammal herbivores (Fig. 3A).

Classification of large mammal herbivomes

We identified four African large mammal herbivomes (Fig. 4) from our cluster analysis of 1° grid squares, characterized by their total and HFT biomasses (i.e., the surfaces in Fig. 3; figs.

S6 and S7; and table S6). The clearest association between an herbivome and a single HFT is the classification of the forested tropical regions as the "forest duiker herbivome," which closely matches the area of high biomass of small nonsocial browsers (Figs. 3B and 5B). The "arid gazelle herbivome" comprises arid southwestern Africa, the horn of Africa, and the northern parts of the Sahel and has the lowest total herbivore biomass (Fig. 5A). Medium-sized social mixed-diet species are the only HFT to have a significant positive association with this herbivome, but large parts of East Africa with high total biomass and medium-sized social mixed-diet species biomass are not included in this herbivome (Fig. 5). The lack of perfect mapping between HFTs and herbivomes is expected. By analogy, grass and trees occur in forest, savanna, and desert biomes; it is the relative abundance of functional types that is important in recognizing the biome. The arid gazelle herbivome is well resolved at the 1° grid square resolution (Fig. 4B), with very little clustering uncertainty at the 95% level (Fig. 4C). The "bulk feeder herbivome" has the largest total area (Fig. 4A), but this decreases substantially when classification uncertainty is recognized (Fig. 4B). The core distribution of this herbivome comprises two east-west bands at about 10°N and 10°S and also includes the South African Highveld and Ethiopian Highlands. High relative abundances of nonruminant and water-dependent grazer species characterize the bulk feeder herbivome (Fig. 5),
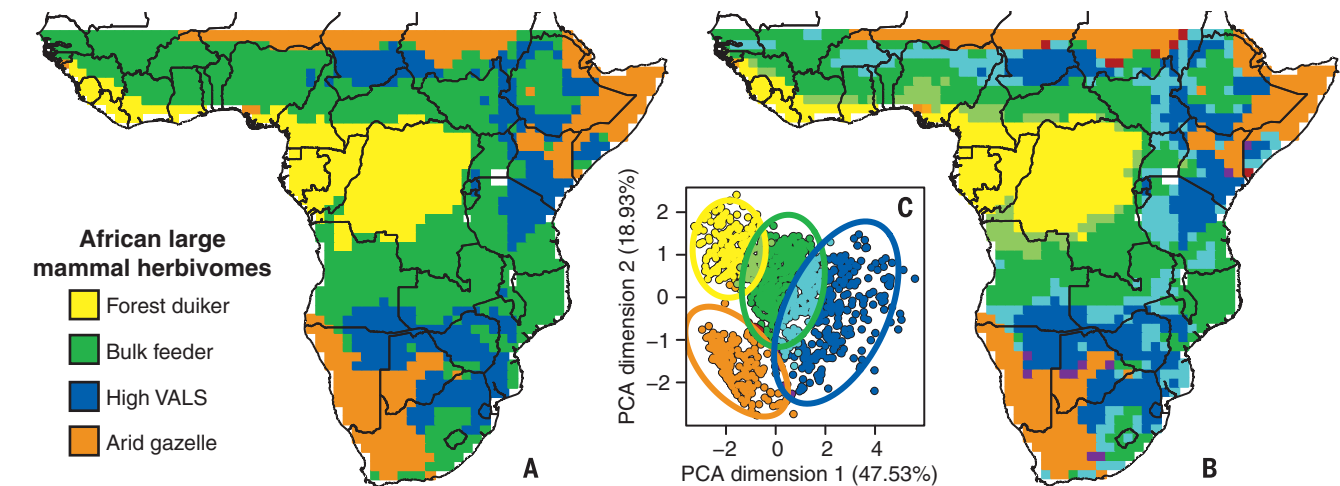
Fig. 4. African large mammal herbivomes. Herbivomes were classified by cluster analysis of total and HFT biomass estimates for each 1° grid square. Grid-square herbivome classifications are shown (**A**) without and (**B**) with classification uncertainty, as determined by (**C**) overlap of 95% confidence ellipses constructed around the mean herbivome position on dimensions 1 and 2 of the principal components analysis. Colors in (B) correspond with those in (C) (purple, arid gazelle–high VALS overlap; dark red, arid gazelle–bulk feeder overlap; orange, arid gazelle–forest duiker overlap; light blue, bulk feeder–high VALS overlap; light green, bulk feeder–forest duiker overlap).
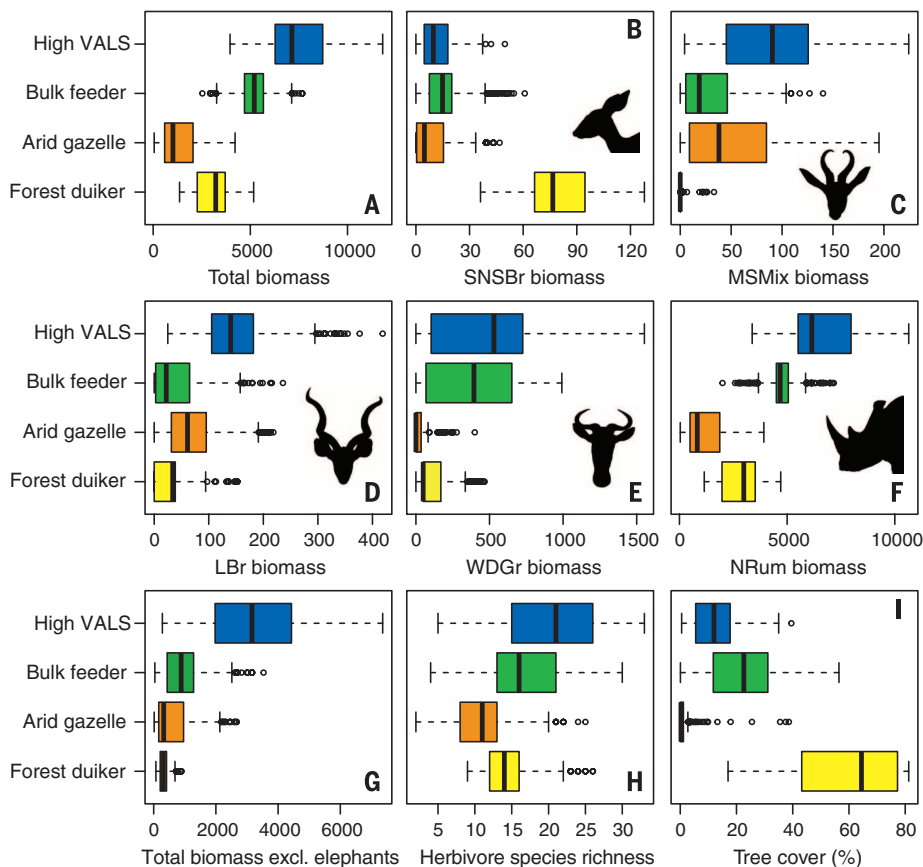
Fig. 5. Characteristics of large mammal herbivomes. (**A**) Total biomass estimates for all 92 species assessed. (**B** to **F**) HFT biomass estimates. (**G**) Total biomass with elephant biomass excluded. (**H**) Herbivore species richness per 1° grid square. (**I**) Percentage of tree cover from (*55*). All herbivore biomass estimates are in kg/km².

which has a high and elephant-dominated total biomass (Fig. 5, A versus G, and fig. S3). The final grouping is the "high VALS herbivome"

(high variety and abundance of larger species), so named because of its positive association with, and high biomass of, each of the four larger-sized

HFTs. This herbivome showcases the diversity of African savanna large mammals (Fig. 5H) and is home to many of East and southern Africa's world-renowned protected areas (e.g., the Masai Mara National Reserve and the Serengeti, Chobe, and Kruger National Parks). It is somewhat patchily distributed across the continent and largely absent from West Africa. The high VALS herbivome differs from the bulk feeder herbivome, which dominates West African savannas, by having a high proportion of large-browser and medium-sized social mixed-diet species (Fig. 5, C and D) and a lower proportion of elephants in the total biomass (Fig. 5, A versus G). Heterogeneity in soil and water distribution at scales smaller than the study resolution probably accounts for the large area of classification uncertainty between these two herbivomes (Fig. 4, B and C).

Environmental correlates

MAR distinguishes the four herbivomes (Fig. 6), with the forest duiker and arid gazelle herbivomes being the wettest and driest, respectively, and the bulk feeder herbivome being wetter than the high VALS herbivome. Mean annual temperature has little discriminatory power (Fig. 6A). Soil nutrient availability separates the bulk feeder and high VALS herbivomes (Fig. 6B), which correspond roughly to the respective distributions of nutrient-poor and nutrient-rich savannas in East and southern Africa (*41*). The bulk feeder herbivome is typified by low-nutrient soils, often associated with high rates of soil leaching in wetter regions. A combination of high rainfall and low nutrients produces a tall grass layer with high C:N stems to provide the structural support to successfully compete for light. Cold winters in the high-lying and wetter regions of South Africa and Ethiopia will also result in a high biomass of cured grass of low forage quality
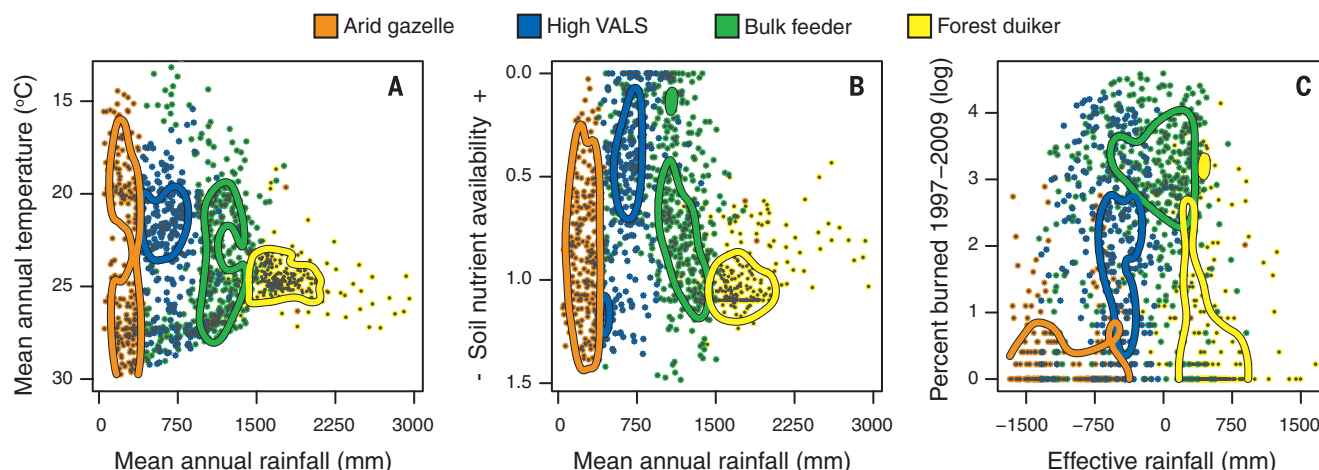
Fig. 6. Herbivomes in environmental space. (**A**) Whittaker plot, as traditionally used to distinguish vegetation biomes (*27*). (**B**) Replacement of mean annual temperature with the natural log of the soil nutrient availability index from the FAO Harmonized World Soils Database (smaller numbers reflect higher nutrient statuses). (**C**) An index of fire activity derived from the Global Fire Emissions Database, shown against effective rainfall (MAR minus potential evapotranspiration). Lines show the 90th quantile of density of unambiguously classified points for each herbivome.

in the winter dry season. Consequently, greater grass fuel loads make the bulk feeder herbivome more fire-prone than the high VALS herbivome (Fig. 6C), and the low-quality grazing necessitates larger body sizes with longer gut passage times to allow for digestion and nutrient extraction (*42*). At similar body sizes, nonruminant gut systems are also better suited to using high-abundance, low-quality resources than ruminant gut systems are, because their less efficient but higher-throughput digestive strategy optimizes nutrient intake rates when forage quantity is not limiting (*43*). The bulk feeder herbivome corresponds to the higher wildlife densities relative to livestock densities above 1000 mm MAR (Fig. 1), suggesting that the full set of species-trait combinations typical of this herbivome is not present in domesticated livestock. Tree cover (Fig. 5I) correlates with the rainfall of each herbivome. Water constraints (Fig. 6C) probably account for the higher tree cover in the bulk feeder herbivome than in the high VALS herbivome (*44*), yet both have the climate potential to produce closed woody canopies (*2, 27*). This suggests that consumer control determines vegetation structure and composition in these herbivomes (*45*), with the role of large mammal herbivory probably increasing as fire becomes less prevalent in the high VALS herbivome (Fig. 6C). The correlation of the high VALS herbivome with low-fire, high-nutrient regions supports long-held theories on the relative importance of fire versus herbivory as consumers in arid and mesic savannas in Africa (*41*).

Implications

Global maps of herbivory are still unobtainable, but quantitative data on large mammal herbivores in Africa will substantially bolster studies seeking to understand their role in driving biome distributions (*4*) and controlling vegetation structure (*16, 17, 46, 47*), as well as how they function as evolutionary and ecological drivers of community assembly (*1, 48*). Furthermore, such data

open the door to understanding the cascading effects of continental-scale trophic downgrading due to the loss of apex predators and other megafauna (*23*). This is of particular relevance, given the growing awareness of the crucial role of biotic interactions in shaping both current and paleoecosystem dynamics, and thus in informing forecasts of species, community, and ecosystem responses to climate change (*24, 25, 49*). Our findings also bring rangeland ecology under new scrutiny, because the replacement of wild herbivores by livestock, with the latter's reduced set of functional traits, might confront regional floras with a novel combination or biomass of HFTs and could expand the scope for fire as a management tool in, for example, the bulk feeder herbivome. Our HFT classifications offer a formal description of the intuitive differences among African herbivores and should considerably facilitate current efforts to include herbivores in DGVMs (*38*). Further exploration of the potential for herbivores to shape vegetation patterns at large scales, which will require estimates of herbivore biomass, should improve our ability to predict the consequences of re-wilding projects (*50, 51*) and the naturalization of novel herbivore groups for ecosystems (*52, 53*). The African continent provides the best example of a near-intact megafaunal assemblage that resembles conditions before the Anthropocene. Extending the herbivome concept beyond Africa's tropical to subtropical ecosystems is likely to reveal a more important role for temperature in delimiting new herbivomes associated with temperate and polar systems, with traits such as the ability to store large fat reserves being important to identifying additional HFTs. Recent research on the consequences of megafaunal extinctions has had to grapple with the relative importance of herbivory and fire (*20, 21*). Integrating the extent and selectivity of vegetation consumption by herbivores and fire at large scales will allow exploration of their respective and synergistic roles in

shaping past, current, and future global vegetation patterns and Earth system feedbacks.

REFERENCES AND NOTES

1. W. J. Bond, J. E. Keeley, *Trends Ecol. Evol.* **20**, 387–394 (2005).
2. W. J. Bond, F. I. Woodward, G. F. Midgley, *New Phytol.* **165**, 525–538 (2005).
3. B. P. Murphy *et al.*, *J. Biogeogr.* **40**, 1048–1058 (2013).
4. C. E. R. Lehmann *et al.*, *Science* **343**, 548–552 (2014).
5. E. Dwyer, S. Pinnock, J. Gregoire, J. M. C. Pereira, *Int. J. Remote Sens.* **21**, 1289–1302 (2000).
6. G. van der Werf *et al.*, *Atmos. Chem. Phys. Discuss.* **10**, 16153–16230 (2010).
7. W. J. Bond, G. F. Midgley, F. I. Woodward, M. T. Hoffman, R. M. Cowling, *S. Afr. J. Bot.* **69**, 79–91 (2003).
8. D. M. J. S. Bowman *et al.*, *Science* **324**, 481–484 (2009).
9. S. Scheiter, S. I. Higgins, *Glob. Change Biol.* **15**, 2224–2246 (2009).
10. M. Martin Calvo, I. C. Prentice, *New Phytol.* **208**, 987–994 (2015).
11. D. J. Beerling, C. P. Osborne, *Glob. Change Biol.* **12**, 2023–2031 (2006).
12. W. J. Bond, A. C. Scott, *New Phytol.* **188**, 1137–1150 (2010).
13. S. Scheiter *et al.*, *New Phytol.* **195**, 653–666 (2012).
14. N. Owen-Smith, *Megaherbivores* (Cambridge Univ. Press, Cambridge, 1988).
15. S. A. Zimov *et al.*, *Am. Nat.* **146**, 765–794 (1995).
16. G. P. Asner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **106**, 4947–4952 (2009).
17. M. Sankaran, D. J. Augustine, J. Ratnam, *J. Ecol.* **101**, 1389–1399 (2013).
18. F. Keesing, T. P. Young, *Bioscience* **64**, 487–495 (2014).
19. D. A. Burney, T. F. Flannery, *Trends Ecol. Evol.* **20**, 395–401 (2005).
20. J. L. Gill, J. W. Williams, S. T. Jackson, K. B. Lininger, G. S. Robinson, *Science* **326**, 1100–1103 (2009).
21. S. Rule *et al.*, *Science* **335**, 1483–1486 (2012).
22. J. L. Gill, *New Phytol.* **201**, 1163–1169 (2014).
23. J. A. Estes *et al.*, *Science* **333**, 301–306 (2011).
24. J. L. Blois, P. L. Zarnetske, M. C. Fitzpatrick, S. Finnegan, *Science* **341**, 499–504 (2013).
25. M. S. Wisz *et al.*, *Biol. Rev. Camb. Philos. Soc.* **88**, 15–30 (2013).
26. L. R. Holdridge, *Science* **105**, 367–368 (1947).
27. R. Whittaker, *Communities and Ecosystems* (MacMillan, New York, ed. 2, 1975).
28. S. Lavorel, S. McIntyre, J. Landsberg, T. D. A. Forbes, *Trends Ecol. Evol.* **12**, 474–478 (1997).
29. S. Díaz, M. Cabido, *Trends Ecol. Evol.* **16**, 646–655 (2001).
30. F. I. Woodward, M. R. Lomas, C. K. Kelly, *Philos. Trans. R. Soc. London Ser. B* **359**, 1465–1476 (2004).

31. S. Archibald, C. E. R. Lehmann, J. L. Gómez-Dans,
    R. A. Bradstock, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6442–6447
    (2013).
32. Materials and methods are available as supplementary
    materials on *Science* Online.
33. R. East, *Afr. J. Ecol.* **22**, 245–270 (1984).
34. M. J. Coe, D. H. Cumming, J. Phillipson, *Oecologia* **22**, 341–354
    (1976).
35. H. Fritz, P. Duncan, *Proc. Biol. Sci.* **256**, 77–82 (1994).
36. H. Fritz, P. Duncan, I. J. Gordon, A. W. Illius, *Oecologia* **131**,
    620–625 (2002).
37. H. Olff, M. E. Ritchie, H. H. T. Prins, *Nature* **415**, 901–904
    (2002).
38. A. Pachzelt, A. Rammig, S. Higgins, T. Hickler, *Ecol. Modell.*
    263, 92–102 (2013).
39. M. Oesterheld, O. E. Sala, S. J. McNaughton, *Nature* **356**,
    234–236 (1992).
40. S. Chamaillé-Jammes, M. Valeix, H. Fritz, *J. Appl. Ecol.* **44**,
    625–633 (2007).
41. R. J. Scholes, B. H. Walker, *An African Savanna: Synthesis
    of the Nylsvley Study* (Cambridge Univ. Press, Cambridge,
    1993).
42. M. W. Demment, P. J. Van Soest, *Am. Nat.* **125**, 641–672
    (1985).
43. A. W. Illius, I. J. Gordon, *Oecologia* **89**, 428–434
    (1992).
44. M. Sankaran *et al.*, *Nature* **438**, 846–849 (2005).
45. W. J. Bond, *J. Veg. Sci.* **16**, 261–266 (2005).
46. G. Bucini, N. P. Hanan, *Glob. Ecol. Biogeogr.* **16**, 593–605
    (2007).
47. M. Sankaran, J. Ratnam, N. P. Hanan, *Glob. Ecol. Biogeogr.* **17**,
    236–245 (2008).
48. T. Charles-Dominique, A. C. Staver, G. F. Midgley, W. J. Bond,
    S. Afr. J. Bot. (2015).
49. P. L. Zarnetske, D. K. Skelly, M. C. Urban, *Science* **336**,
    1516–1518 (2012).
50. J. Donlan, *Nature* **436**, 913–914 (2005).
51. S. A. Zimov, *Science* **308**, 796–798 (2005).
52. S. Legge, M. S. Kennedy, R. Lloyd, S. A. Murphy, A. Fisher, *Aust.
    Ecol.* **36**, 791–799 (2011).
53. J. C. Z. Woinarski, A. A. Burbidge, P. L. Harrison, *Proc. Natl.
    Acad. Sci. U.S.A.* **112**, 4531–4540 (2015).
54. M. Gagnon, A. E. Chew, *J. Mammal.* **81**, 490–511 (2000).
55. M. C. Hansen *et al.*, *Earth Interact.* **7**, 1–15 (2003).

REPORTS

PHOTOPHYSICS

Semiconductor interfacial carrier dynamics via photoinduced electric fields

Ye Yang,[1]*† Jing Gu,[1]† James L. Young,[1,2] Elisa M. Miller,[1] John A. Turner,[1]
Nathan R. Neale,[1] Matthew C. Beard[1]*

Solar photoconversion in semiconductors is driven by charge separation at the interface of the
semiconductor and contacting layers. Here we demonstrate that time-resolved photoinduced
reflectance from a semiconductor captures interfacial carrier dynamics. We applied this
transient photoreflectance method to study charge transfer at p-type gallium-indium
phosphide (p-GaInP$_2$) interfaces critically important to solar-driven water splitting. We
monitored the formation and decay of transient electric fields that form upon photoexcitation
within bare p-GaInP$_2$, p-GaInP$_2$/platinum (Pt), and p-GaInP$_2$/amorphous titania (TiO$_2$)
interfaces. The data show that a field at both the p-GaInP$_2$/Pt and p-GaInP$_2$/TiO$_2$ interfaces
drives charge separation. Additionally, the charge recombination rate at the p-GaInP$_2$/TiO$_2$
interface is greatly reduced owing to its p-n nature, compared with the Schottky nature of the
p-GaInP$_2$/Pt interface.

emiconductor photoelectrodes used in pho-
toelectrochemical (PEC) cells directly con-
vert sunlight into stored chemical potential
(*1–5*). Junctions that form between a semi-
conductor and a contact layer are the key to
charge separation that drives photoconversion
processes. Equilibration of chemical potential at
such junctions creates an internal electric field
(referred to as the built-in field) and establishes a
region where mobile charges are driven away

known as the depletion region. Absorption of
light within the depletion region results in charge
separation by the built-in field. As a result, photo-
generated electrons (holes) transfer across the
interface to participate in a reduction (oxidation)
reaction and holes (electrons) are transported to
the counter electrode via the external circuit for the
oxidation (reduction) reaction. However, the photo-
carriers can also recombine across the same inter-
face, and such recombination reduces the energy
conversion efficiency. Thus, the carrier dynamics—
charge separation and recombination across
junctions—represent a key determining factor in
the PEC performance. Time-resolved spectroscopies
are normally employed to study ultrafast carrier
dynamics in semiconductors (*6–8*) and semicon-

ductor nanostructures (*9–12*) by probing the ki-
netics of the spectral features of initial and/or
final states. Time-resolved surface-sensitive spec-
troscopy also has been exploited to study the
semiconductor surface dynamics (*13*). However,
isolating spectral signatures and/or the carrier dy-
namics that are specific to junctions—often strong-
ly affecting subsurface depths of tens of nanometers
such as in photoelectrodes—is challenging.

Here, we introduce transient photoreflectance
(TPR) spectroscopy to probe the dynamics of the
transient electric field (ΔF) caused by charge sep-
aration and recombination at a junction of inte-
rest. In the TPR method, the change in reflectance
(ΔR) of a broadband probe from a specific inter-
face is monitored as a function of pump-probe
delay (Fig. 1). We applied TPR to study the p-type
gallium-indium-phosphide (p-GaInP$_2$) photoelec-
trode system, a well-known photocathode for light-
driven hydrogen evolution (*3*, *14*). We compare
TPR spectral and dynamical signatures of bare
p-GaInP$_2$ with p-GaInP$_2$/platinum (Pt) and p-
GaInP$_2$/amorphous titania (TiO$_2$) photoelec-
trodes. We demonstrate that TPR can extract the
dynamics of ΔF upon photoexcitation, as well as
the magnitude of the built-in field (F) in these
junctions.

TPR spectroscopy is a pump-probe technique
(Fig. 1) and in our experiment, the temporal re-
sponse was ~150 fs. The pump photon energy of
3.1 eV corresponded to a pump penetration depth
of 29 nm according to the absorption coefficient
(fig. S1). The probe was a white-light continuum
with photon energy $\hbar\omega$ between 1.55 and 2.40 eV,
corresponding to effective detection depths be-
tween 18 to 10 nm, as approximated by $\lambda/4\pi n$
(where λ is wavelength and n is refractive index).
The photocarrier density was calculated from the
absorption coefficient and pump photon flux. The
pump and probe beams were overlapped at the
electrode surface, and the reflected probe directed
to a spectrometer for reflectance spectrum (R)
detection. The photoelectrode consisted of a ~1-μm-
thick top layer of p-GaInP$_2$ epitaxially grown on a
GaAs substrate. The absorption coefficient and
refractive index of p-GaInP$_2$ were determined

[1]National Renewable Energy Laboratory, Chemistry and
Nanoscience Center, Golden, CO, 80401, USA. [2]Material
Science and Engineering Program, University of Colorado,
Boulder, CO, 80309, USA.
*Corresponding author. E-mail: ye.yang@nrel.gov (Y.Y); matt.beard@
nrel.gov (M.C.B) †These authors contributed equally to this work.

from ellipsometry characterization (fig. S1). The doping concentration was measured to be 2.7 × 10^{17} cm^{-3} from Mott-Schottky analysis (fig. S2). The platinum and amorphous TiO$_2$ layers were deposited by atomic layer deposition.

The normalized change in reflectance ($\Delta R/R$) at different probe photon energies ($\hbar\omega$) was recorded as a function of pump-probe delay. Reflectance modulations could arise from band filling by photogenerated free-carriers (15) and/or transient field changes arising from electro-optic effects (16, 17). Pseudocolor images of $\Delta R/R$ as a function of probe photon energy and pump-probe

delay are displayed for the three samples in Fig. 2, A, C, and E, and their representative spectra at the specified pump-probe delays are shown in Fig. 2, B, D, and F.

For the bare p-GaInP$_2$ surfaces (Fig. 2, A and B), the data exhibit a negative and a positive band below and above the bandgap (~1.8 eV), respectively, for delays less than ~10 ps (Fig. 2B, blue circles); such spectral features can be attributed to band filling (15). Because the effective mass of electrons is much lighter than that of holes (18), the spectra are dominated by filling of conduction band states (photogenerated electrons) (19).
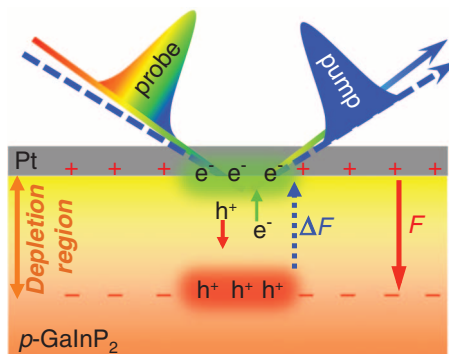
Thus, in the low-excitation range, $\Delta R/R$ is proportional to the surface carrier density, and surface-specific carrier dynamics can be extracted (20, 21). As carriers undergo surface trapping or diffuse away from the surface, the TPR spectra evolve over ~100 ps to a derivative-like shape near the bandgap (Fig. 2B, red triangles). For bare semiconductor surfaces with flat bands, little to no field is present before photoexcitation, but a surface field forms upon photoinduced surface trapping of minority carriers (22). Because of the low surface trapping density, ΔF can be considered to be in the low-field regime, and thus, $\Delta R/R$ should resemble the third-derivative of the refractive index with respect to photon energy, $\frac{\Delta R}{R}(\hbar\omega) \propto \frac{\partial^3 n}{\partial(\hbar\omega)^3}$, according to the electro-reflectance effect (16, 23). Thus, in the absence of free carriers that occupy band-edge states, $\Delta R/R$ results from the transient surface electric field (24). We simulated typical TPR spectra at short and long delay times (2 ps and 1 ns; Fig. 2B, dashed traces) based on the band-filling model and low-field electro-reflectance model, respectively [details in supplementary materials (SM) section 3.2 and 3.3]. We find good agreement between the simulation and the experimental data.

The situation is quite different in the presence of a charge-separating junction. In the p-GaInP$_2$/Pt

Fig. 1. Schematic illustration of TPR spectroscopy applied to p-GaInP$_2$. A broadband probe pulse spanning the semiconductor bandgap is reflected from an interface of interest. An above-bandgap monochromatic pump pulse modulates the reflectance, either via band filling due to the presence of free carriers or via surface field due to charge separation across the interface. For this experiment, the pump pulse frequency was tuned to ensure an optical penetration depth shallower than that of the depletion region so that all of the photoinduced carriers were separated.

Fig. 2. TPR spectra. Pseudocolor image and spectral snapshot of TPR spectra for (**A** and **B**) p-GaInP$_2$, (**C** and **D**) p-GaInP$_2$/Pt, and (**E** and **F**) p-GaInP$_2$/TiO$_2$. Intensities of red and blue in pseudocolor images represent the magnitude of the reduced and increased reflectance, respectively. The blue and red spectra in (B), (D), and (F) are snapshots from the image at 2 ps and 1 ns delays, indicated by the dashed blue and red lines in (A), (C), and (E). For the p-GaInP$_2$ sample, the spectra evolve over time owing to diffusion and surface trapping effects. The oscillations at 2.3 eV [red dash-line boxes in (D) and (F)] are assigned to the transition from the valence band edge to the upper conduction band (fig. S3). The black-dash traces are simulations discussed in the text and detailed in the supplementary material (SM section 3).

case, a built-in electric field (F) is present in the dark owing to the Schottky junction and drives ultrafast photogenerated charge separation, which produces a ΔF and thus ΔR. For p-GaInP$_2$/Pt (Fig. 2, C and D), the carrier-induced spectral features are not observed at early delay times (as they are for p-GaInP$_2$ for delays <100 ps). The absence of features associated with band-filling suggests that the electron injection into Pt is too fast to be resolved (<150 fs). Instead, a set of oscillatory features emerges immediately after excitation and remains for the whole delay time (~5 ns), with little spectral evolution occurring other than a gradual decrease of the amplitude.

Under low-fluence conditions, F is the dominant field and ΔF can be considered as a small perturbation. Because of the large built-in field, R should exhibit periodic oscillations above the semiconductor bandgap, known as Franz-Keldysh oscillations (FKO) (24), and $\Delta R/R$ will represent a small perturbation to those oscillations. According to Franz-Keldysh theory, the periodicity is determined by F, whereas the amplitude is proportional to ΔF (25).

$$\frac{\Delta R}{R}(\hbar\omega) \propto \Delta F \cdot A(F) \qquad (1)$$

where $A(F)$ contains Airy functions, exhibiting the oscillatory behavior. Thus, in TPR spectroscopy, the dynamics of ΔF can be probed by following

the FKO kinetics and F can be determined (SM section 3.4). Similar FKOs have been reported in photoreflectance and electroreflectance studies of p-GaInP$_2$ junctions (26, 27). The TPR spectrum near the p-GaInP$_2$ bandgap can be simulated by Franz-Keldysh theory (Fig. 2D, dashed trace). In our simulations, F is determined to be 158 ± 4 kV cm^{-1}, corresponding to a depletion region width (w) of ~42 nm ($w = F\varepsilon/\rho$, where ε and ρ are the dielectric constant and doping charge density). There is an additional oscillatory feature at 2.3 eV (Fig. 2, D and F, red box, not resolved in the spectra of bare p-GaInP$_2$) that corresponds to the electro-optic effect of the optical transition connecting the upper conduction band and the bottom valence band (fig. S3) (23), which is not included in the simulation. The periodicity of the oscillation remains constant for the various delays, indicative of the pertubative nature of ΔF.

The pseudo color image of TPR spectra for p-GaInP$_2$/TiO$_2$ (TiO$_2$ thickness is ~35 nm) is shown in Fig. 2E. Similar to p-GaInP$_2$/Pt, the data display FKOs, which are the signature of the electric field modulation. At early delay times, the band-filling–induced features are also not observed, implying that the electron transfer rate is faster than our temporal resolution. The TPR spectra were simulated with Franz-Keldysh theory (Fig. 2F, black-dash trace), and F was determined to be 139 ± 2 kV cm^{-1} (corresponding to a depletion re-

gion of ~37 nm). Owing to the observation of the field in p-GaInP$_2$/TiO$_2$, we conclude that the junction between p-GaInP$_2$ and TiO$_2$ can best be described as a p-n heterojunction.

To correlate the relationship between TPR magnitude, ΔF, and carrier density (N), we carried out TPR measurements for different excitation intensities. For p-GaInP$_2$, the magnitude of $\Delta R/R$ at short delay (2 ps) is proportional to N (Fig. 3A, blue circles and dashed fit), consistent with theoretical prediction (see SM section 5) for the band-filling effect. In contrast, the low-field TPR signal at a long delay time (5 ns) is nearly constant with increasing N (Fig. 3A, red squares). For p-GaInP$_2$/Pt and p-GaInP$_2$/TiO$_2$, the FKO amplitude is plotted as function of carrier density (Fig. 3B) on a logarithm scale. In the low–carrier density region ($N < 5 \times 10^{16}$ cm^{-3}), the FKO amplitude is proportional to log(N). In keeping with the photovoltaic effect, ΔF also should be proportional to log(N). Thus, the FKO amplitude linearly depends on ΔF, which can also be theoretically deduced from the Franz-Keldysh theory (Eq. 1). The carrier-density–dependent measurements further reveal the nature of the different spectral features. The FKO amplitude begins to saturate for both the Pt and TiO$_2$ samples when $N > 5 \times 10^{16}$ cm^{-3}. In that case, ΔF becomes significant and the perturbation approximation is not valid. The FKO amplitude is higher for the p-GaInP$_2$/TiO$_2$ at a given photon

Fig. 3. Carrier density dependence and TPR kinetics. (**A**) Band-filling–induced and low-field–modulated TPR signal as a function of carrier density for p-GaInP$_2$. (**B**) FKO amplitude as a function of carrier density plotted on a logarithmic scale. For each pump intensity, the data points represent the peak-to-peak amplitude between the first positive and negative peaks in the spectra at 50 ps delay time. The black dashed lines show a linear dependence on log(N). (**C**) Kinetics of the band-filling–induced TPR in p-GaInP$_2$ (recorded at 2.2 eV), which represent the dynamics of the surface carrier density. The black dashed trace is a diffusion model discussed in the text. (**D**) Kinetics of FKO in p-GaInP$_2$/Pt and p-GaInP$_2$/TiO$_2$ (recorded at 1.82 eV), which represent the dynamics of the transient field. The black dashed lines represent a model dis-

cussed in the text. All of the kinetic traces are scaled by normalizing the maximum to 1. In (C) and (D), the data are plotted on a linear scale for delays less than 10 ps and on a logarithmic scale from 10 ps to 4 ns (indicated by the vertical line).
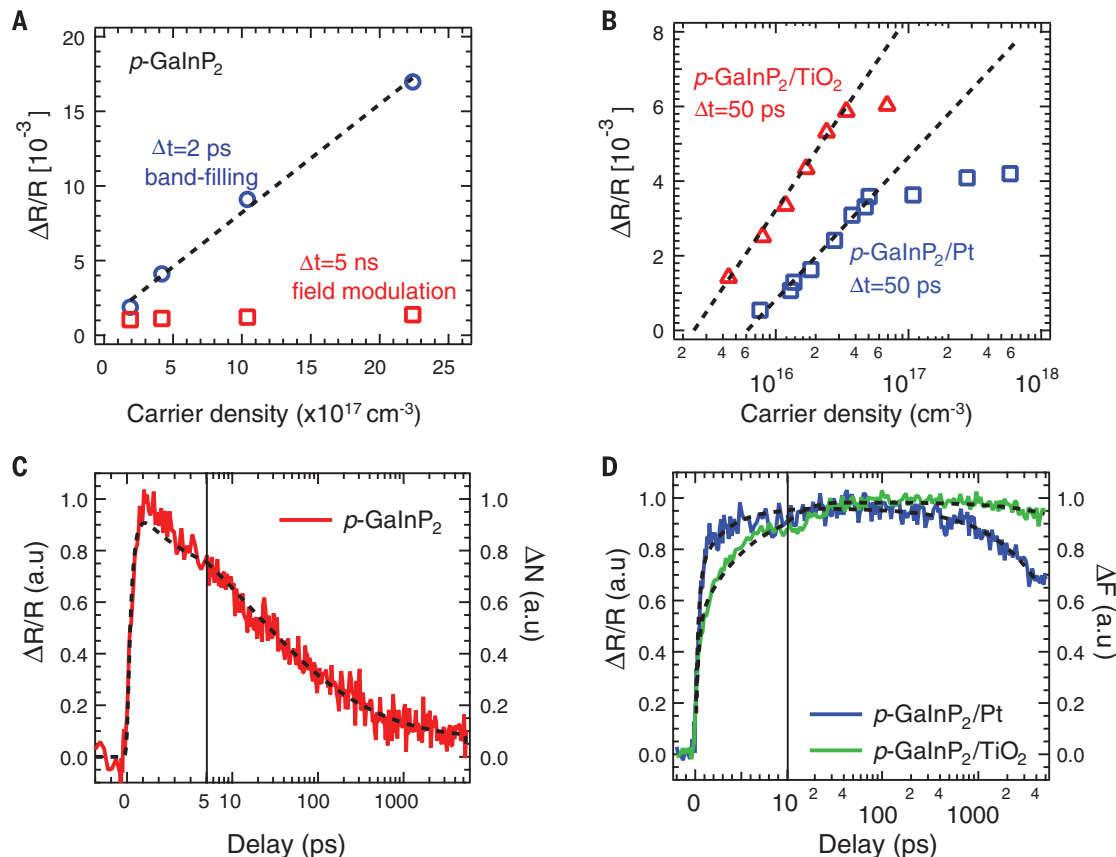
flux because of the antireflective nature of the TiO$_2$ layer.

Because the detection depth is less than 20 nm, the kinetics of $\Delta R/R$ caused by band-filling (p-GaInP$_2$) represent the dynamics of photogenerated electrons in the detection region (surface carriers). The kinetics (Fig. 3C) were recorded at 2.2 eV to avoid overlapping with the trapping-induced derivative-like features (Fig. 2B, red trace). The bulk lifetime in p-GaInP$_2$ at low carrier density is ~3 ns (28); therefore, surface trapping and carrier diffusion into the bulk are responsible for the short surface lifetime. Furthermore, the surface carrier dynamics are independent of excitation density (fig. S6), suggesting that the fast decay kinetics are not affected by bulk recombination (21). We fit a diffusion model to the kinetic trace with the ambipolar diffusion coefficient and the surface recombination velocity (SRV) as best-fit parameters with values of 0.5 cm^2 s^{-1} and 500 cm s^{-1}, respectively. Because the electron diffusion coefficient is much higher than the determined ambipolar diffusion coefficient (28), carrier diffusion is limited by holes, not electrons. The low SRV also implies a low density of surface trapping states. The rise of the signal corresponds to carrier cooling as the carrier distribution narrows in k-space, and the cooling time is determined to be 0.30 ± 0.01 ps.

Because of the linear relationship between FKO amplitude and ΔF, the dynamics of ΔF for p-GaInP$_2$/

Pt and p-GaInP$_2$/TiO$_2$ can be directly probed by following the FKO kinetics (Fig. 3D). The increase in ΔF results from the interfacial charge separation, and the kinetics of both samples show biphasic behavior. Our modeling indicates a common fast component (0.3 ps) for both kinetic traces, which we attribute to carrier cooling. The time constant of the slow component for p-GaInP$_2$/TiO$_2$ (5.7 ± 0.3 ps) is longer than that for p-GaInP$_2$/Pt (2.8 ± 0.3 ps). The formation of ΔF results from electron and hole separation, and the electron transfer time is shorter than the time resolution (~150 fs); therefore, the slow component is attributed to holes that drift from the surface to the bulk. The electron and hole drift time can be estimated from their respective mobilities, distance over which they travel, and F. The average distance for electrons was approximated as the pump penetration depth (29 nm); the average distance for holes was determined by the depletion region width (~43 nm), and F is determined by the FKO simulations discussed above. The mobilities in p-GaInP$_2$ are on the order of 10^3 cm^2 V^{-1} s^{-1} for electrons and 10 cm^2 V^{-1} s^{-1} for holes, and the corresponding drift times are ~20 fs and ~3 ps for electrons and holes, respectively, which is consistent with our measurement. Thus, the slower growth of ΔF in p-GaInP$_2$/TiO$_2$ arises from the smaller built-in electric field. The decay of ΔF reflects the interfacial charge recombination. We determined transient field decay time constants

of 104 ± 25 and 11.3 ± 0.5 ns for p-GaInP$_2$/TiO$_2$ and p-GaInP$_2$/Pt, respectively. According to the relationship between N and $\Delta R/R$ (Fig. 3B), the interfacial charge recombination time can also be extracted from the field dynamics (fig. S7). We summarize our findings for p-GaInP$_2$, p-GaInP$_2$/Pt, and p-GaInP$_2$/TiO$_2$, by schematic illustrations in Fig. 4.

Our results uncover key factors that enhance the photon conversion process in amorphous TiO$_2$-coated electrodes. For the bare electrode, surface trapping leads to localized electrons at lower energy states and diffusion transports electrons away from the surface, both of which are harmful to efficient photoconversion. Thus, appropriate junctions at the electrode interface (formed with Pt and TiO$_2$) produce surface fields that can drive electron-hole separation and promote high charge separation yields. However, the nature of the junction (Schottky or p-n) has a profound influence on the carrier dynamics following separation. Compared with Pt, the amorphous TiO$_2$ greatly retards the interfacial charge recombination without appreciably sacrificing the charge separation rate. The barrier for electron-hole recombination at the Pt interface is determined by the Schottky barrier height, which from our data we estimate as ~0.3 eV ($\phi = F \cdot w/2$), derived solely from p-GaInP$_2$ valence band bending (Fig. 4B). In contrast, both electrons and holes are driven away from the interface in p-GaInP$_2$/TiO$_2$ (Fig. 4C) by the built-in fields that exist on either side of the junction. We estimate that the charges become separated by ~72 nm (total depletion width in GaInP$_2$ and TiO$_2$), establishing a substantial kinetic barrier to recombination. In addition, x-ray photoelectron spectroscopy (XPS) and ultraviolet photoelectron spectroscopy (UPS) characterization also show that the thermodynamic barrier to electron (hole) injection into p-GaInP$_2$ (TiO$_2$) for bulk recombination is unlikely, owing to the large band offset (1.88 eV in valence band and 0.64 eV in conduction band, Fig. 4C), and thus the recombination should occur at the p-GaInP$_2$/TiO$_2$ interface. In the latter scenario, both an electron in TiO$_2$ and a hole in p-GaInP$_2$ must overcome barriers of >0.3 eV (measured by XPS and UPS; see SM section 8 for details) and 0.26 eV, respectively, and find one another likely assisted by an interfacial recombination center. Thus, the interfacial recombination probability is further reduced by the energy barrier and low density of interfacial defect sites. Therefore, we conclude that the observed slower decay of ΔF in p-GaInP$_2$/TiO$_2$ versus p-GaInP$_2$/Pt can be attributed to both kinetic and thermodynamic barriers that result from the p-n nature of this junction.

We investigated the effect of the amorphous TiO$_2$ layer thickness on the interfacial carrier dynamics to explore the limiting thickness required to observe the beneficial effects of the p-n junction. For TiO$_2$ thicknesses from 0.5 to 5 nm, the TPR spectral shape and magnitude are similar to the low-field modulation features found in p-GaInP$_2$ at long delays (fig. S9A). The modulation field arises from charge separation via electron transfer from p-GaInP$_2$ to TiO$_2$. As the TiO$_2$ thickness increases from 10 to 35 nm, the TPR spectra of
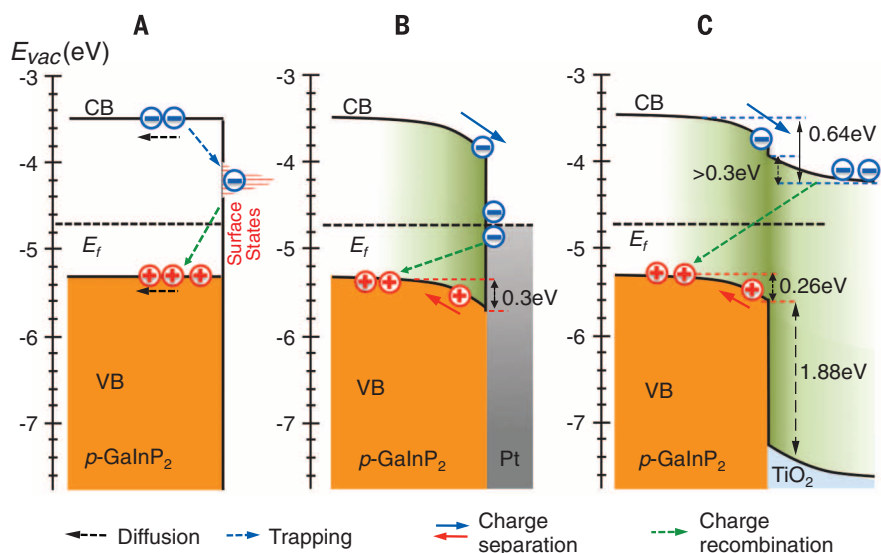
Fig. 4. Energy band and charge flow diagram. Band bending and carrier dynamics at the surface or interface for (**A**) p-GaInP$_2$, (**B**) p-GaInP$_2$/Pt, and (**C**) p-GaInP$_2$/TiO$_2$. The energy band positions are determined by XPS and UPS characterization and the optical bandgaps. The energy scale bar (left) is referenced to the p-GaInP$_2$ vacuum level. (A) For p-GaInP$_2$, the photocarriers are initially generated near the surface and are primarily depopulated by diffusion to the bulk and surface trapping. The trapped electron and the remaining hole form a weak transient electric field at the surface, which modulates the reflectance at longer delay times. In contrast, (B) the Schottky (p-GaInP$_2$/Pt) and (C) p-n (p-GaInP$_2$/TiO$_2$) junctions establish a depletion region width larger than the pump penetration depth. These built-in fields accelerate electrons and holes toward Pt/TiO$_2$ and the bulk, respectively. The electron transfers from the depletion region to the interfacial layer within the time resolution (~150 fs) of the experiment, whereas the hole requires several picoseconds to drift out of the depletion region. The charge separation screens the built-in field and forms the FKOs in TPR spectra. Though the charge separation process is similar, the recovery of ΔF arising from charge recombination at the Schottky (p-GaInP$_2$/Pt) junction is faster (by about one order of magnitude) than that at the p-n (p-GaInP$_2$/TiO$_2$) junction.

p-GaInP$_2$/TiO$_2$ exhibit oscillations similar to that for p-GaInP$_2$/Pt, meaning that the surface field increases substantially when the TiO$_2$ thickness increases from 0.5 to 35 nm. The formation and decay time constant of ΔF for these samples are extracted from the corresponding TPR kinetics (fig. S9B). Best-fit parameters are tabulated in table S1. Thicker TiO$_2$ layers exhibit slightly faster field formation but slower decay, which is likely due to the larger built-in field that drives carriers apart and separates them at a greater distance, both of which lead to slower recombination. We find that the kinetics are effectively unperturbed once a sufficient amorphous TiO$_2$ thickness has been reached, suggesting that thicker layers would not drastically influence the photoconversion performance from a charge dynamics perspective. A thick TiO$_2$ layer may still be necessary for other reasons (such as elimination of pinholes) that affect stabilization against photocorrosion, as has been found for 140-nm-thick amorphous TiO$_2$ layers on Si, GaAs, and GaP photoanodes (2).

Our results uncover key beneficial roles of amorphous TiO$_2$ in the energy-conversion process that have come under intense investigation after several recent reports of TiO$_2$-stablized photoelectrodes (2, 29, 30). The TPR technique developed here furthermore introduces a general method to understand charge transfer at semiconductor junctions.

REFERENCES AND NOTES

1. A. Fujishima, K. Honda, *Nature* **238**, 37–38 (1972).
2. S. Hu et al., *Science* **344**, 1005–1009 (2014).
3. O. Khaselev, J. A. Turner, *Science* **280**, 425–427 (1998).
4. M. Grätzel, *Nature* **414**, 338–344 (2001).
5. N. S. Lewis, D. G. Nocera, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15729–15735 (2006).
6. M. C. Beard, G. M. Turner, C. A. Schmuttenmaer, *J. Phys. Chem. B* **106**, 7146–7159 (2002).
7. R. Ulbricht, E. Hendry, J. Shan, T. F. Heinz, M. Bonn, *Rev. Mod. Phys.* **83**, 543–586 (2011).
8. J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures* (Springer, Berlin and Heidelberg, ed. 2, 1999).
9. Y. Yang, T. Lian, *Coord. Chem. Rev.* **263-264**, 229–238 (2014).
10. K. Tvrdy, P. A. Frantsuzov, P. V. Kamat, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 29–34 (2011).
11. K. E. Knowles, M. D. Peterson, M. R. McPhail, E. A. Weiss, *J. Phys. Chem. C* **117**, 10229–10243 (2013).
12. V. I. Klimov, *Annu. Rev. Phys. Chem.* **58**, 635–673 (2007).
13. W. A. Tisdale et al., *Science* **328**, 1543–1547 (2010).
14. Z. Chen et al., *J. Mater. Res.* **25**, 3–16 (2010).
15. B. R. Bennett, R. A. Soref, J. A. del Alamo, *IEEE J. Quantum Electron.* **26**, 113–122 (1990).
16. D. E. Aspnes, *Phys. Rev. Lett.* **28**, 168–171 (1972).
17. S. Marcinkevičius et al., *AIP Adv* **2**, 042148 (2012).
18. P. Emanuelsson et al., *Appl. Phys. Lett.* **64**, 2849–2851 (1994).
19. K. L. Vodopyanov, H. Graener, C. C. Phillips, T. J. Tate, *Phys. Rev. B* **46**, 13194–13200 (1992).
20. A. J. Sabbah, D. M. Riffe, *J. Appl. Phys.* **88**, 6954–6956 (2000).
21. Y. Yang et al., *Nat. Commun.* **6**, 7961 (2015).
22. H. Shen, M. Dutta, *J. Appl. Phys.* **78**, 2151–2176 (1995).
23. T. Kippenberg et al., *Phys. Rev. B* **60**, 4446–4449 (1999).
24. F. H. Pollak, *Surf. Interface Anal.* **31**, 938–953 (2001).
25. H. Shen, F. H. Pollak, *Phys. Rev. B* **42**, 7097–7102 (1990).
26. T. Kita et al., *J. Appl. Phys.* **94**, 6487–6490 (2003).
27. Y. S. Huang et al., *Appl. Phys. Lett.* **73**, 214–216 (1998).
28. F. J. Schultes et al., *Appl. Phys. Lett.* **103**, 242106 (2013).
29. B. Seger et al., *J. Am. Chem. Soc.* **135**, 1057–1064 (2013).
30. B. Seger et al., *RSC Adv.* **3**, 25902–25907 (2013).

NANOMATERIALS

Near-unity photoluminescence quantum yield in MoS$_2$

Matin Amani,[1,2]* Der-Hsien Lien,[1,2,3,4]* Daisuke Kiriya,[1,2]* Jun Xiao,[5,2]
Angelica Azcatl,[6] Jiyoung Noh,[6] Surabhi R. Madhvapathy,[1,2] Rafik Addou,[6]
Santosh KC,[6] Madan Dubey,[7] Kyeongjae Cho,[6] Robert M. Wallace,[6] Si-Chen Lee,[4]
Jr-Hau He,[3] Joel W. Ager III,[2] Xiang Zhang,[5,2,8] Eli Yablonovitch,[1,2] Ali Javey,[1,2]†

Two-dimensional (2D) transition metal dichalcogenides have emerged as a promising material system for optoelectronic applications, but their primary figure of merit, the room-temperature photoluminescence quantum yield (QY), is extremely low. The prototypical 2D material molybdenum disulfide (MoS$_2$) is reported to have a maximum QY of 0.6%, which indicates a considerable defect density. Here we report on an air-stable, solution-based chemical treatment by an organic superacid, which uniformly enhances the photoluminescence and minority carrier lifetime of MoS$_2$ monolayers by more than two orders of magnitude. The treatment eliminates defect-mediated nonradiative recombination, thus resulting in a final QY of more than 95%, with a longest-observed lifetime of 10.8 ± 0.6 nanoseconds. Our ability to obtain optoelectronic monolayers with near-perfect properties opens the door for the development of highly efficient light-emitting diodes, lasers, and solar cells based on 2D materials.

Monolayer transition metal dichalcogenides (TMDCs) have properties that make them highly suitable for optoelectronics (1, 2), including the ability to form van der Waals heterostructures without the need for lattice matching (3, 4), circular dichroism arising from the direct band gap occurring at the K and K′ points of the Brillouin zone (5), and widely tunable band structure through the application of external forces such as electric field and strain (6). Unlike III-V semiconductors, the optical properties of TMDCs are dominated by excitons with strong binding energies (on the order of 300 meV) (7–9) and large radii (~1.6 nm) (10). However, TMDCs have exhibited poor luminescence quantum yield (QY)—that is, the number of photons the material radiates is much lower than the number of generated electron-hole pairs. QY values ranging from 0.01 to 6% have been reported, indicating a high density of defect states and mediocre electronic quality (11–13). The origin of the low quantum yield observed in these materials is attributed to defect-mediated nonradiative recombination and biexcitonic recombination at higher excitation powers (11, 13).

Two-dimensional (2D) monolayers are amenable to surface passivation by chemical treatments. We studied a wide range of chemical treatments and describe here an air-stable, solution-based process using an organic superacid that removes the contribution of defect-mediated nonradiative recombination acting on electronically active defect sites by uniformly passivating them, repairing them, or both. With the use of this process, the photoluminescence (PL) in MoS monolayers

[1]Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720, USA. [2]Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA. [3]Computer, Electrical and Mathematical Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia. [4]Department of Electrical Engineering, Institute of Electronics Engineering, National Taiwan University, Taipei 10617, Taiwan, Republic of China. [5]National Science Foundation Nanoscale Science and Engineering Center, University of California, Berkeley, Berkeley, CA 94720, USA. [6]Department of Materials Science and Engineering, University of Texas, Dallas, Richardson, TX 75080, USA. [7]Sensors and Electron Devices Directorate, U.S. Army Research Laboratory, Adelphi, MD 20723, USA. [8]Department of Physics, King Abdulaziz University, Jeddah 21589, Saudi Arabia.
*These authors contributed equally to this work. †Corresponding author. E-mail: ajavey@eecs.berkeley.edu

Fig. 1. Enhancement of PL by chemical treatment. (**A**) PL spectrum for both the as-exfoliated and TFSI-treated MoS_2 monolayers measured at an incident power of 1×10^{-2} W cm^{-2}. The inset shows normalized spectra. (**B** and **C**) PL images of a MoS_2 monolayer before (B) and after treatment (C). Insets show optical micrographs.
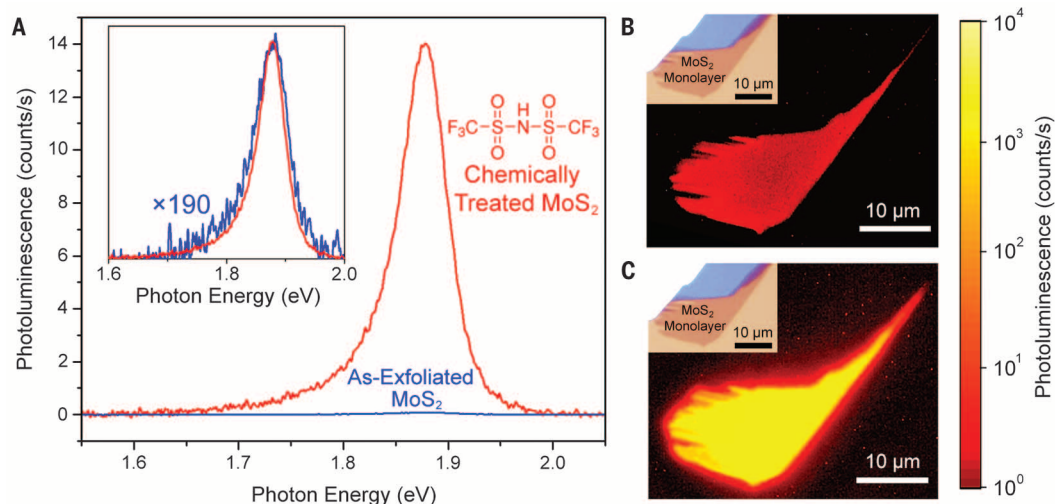
increased by more than two orders of magnitude, resulting in a QY > 95% and a characteristic lifetime of 10.8 ± 0.6 ns at low excitation densities.

In this study, we treated MoS_2 monolayers with a nonoxidizing organic superacid: bis(trifluoromethane) sulfonimide (TFSI). Superacids are strong protonating agents and have a Hammett acidity function (H_0) that is lower than that of pure sulfuric acid. [Details of the sample preparation and treatment procedure are discussed in the supplementary materials and methods (14).] The PL spectra of a MoS_2 monolayer measured before and after TFSI treatment (Fig. 1A) show a 190-fold increase in the PL peak intensity, with no change in the overall spectral shape. The magnitude of the enhancement depended strongly on the quality of the original as-exfoliated monolayer (14). (The term "as-exfoliated" indicates that the MoS_2 flakes were not processed after exfoliation.) PL images of a monolayer (Fig. 1, B and C, and fig. S4) (14), taken before and after treatment at the same illumination conditions, show that the enhancement from the superacid treatment is spatially uniform.

Calibrated steady-state PL measurements (14) showed that the spectral shape of the emission remained unchanged over a pump intensity dynamic range spanning six orders of magnitude (10^{-4} to 10^2 W cm^{-2}) (fig. S2) (14). From the pump-power dependence of the calibrated luminescence intensity (Fig. 2A), we extracted the QY (Fig. 2B). As-exfoliated samples exhibited low QY, with a peak efficiency of 1% measured at 10^{-2} W cm^{-2}. The absolute efficiency (12, 13) and observed power law (13) are consistent with previous reports for exfoliated MoS_2. After TFSI treatment, the QY reached a plateau at a low pump intensity (<10^{-2} W cm^{-2}), with a maximum value greater than 95%. The near-unity QY suggests that, within this range of incident power, there was negligible nonradiative recombination occurring in the sample. Although pure radiative recombination is commonly observed for fluorescent molecules that inherently have no dangling bonds, only a few semiconductors, such as GaAs double heterostructures (15) and surface-passivated
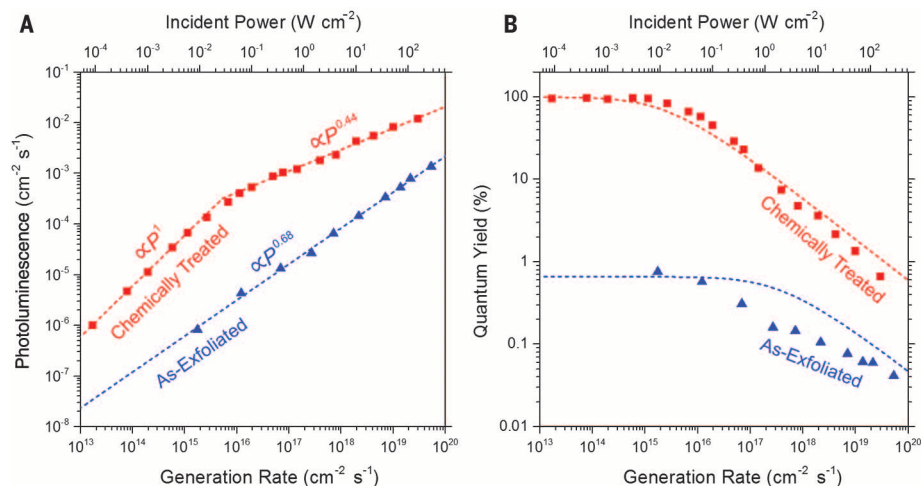
quantum dots (16), show this behavior at room temperature.

At high pump power, we observed a sharp dropoff in the QY, possibly caused by nonradiative biexcitonic recombination. We consider several models to explain the carrier density–dependent recombination mechanisms in MoS_2 before and after TFSI treatment. Here, n and p are the 2D electron and hole concentrations, respectively. At high-level injection, the dopant concentration is much less than the number of optically generated carriers, allowing $n = p$. The traditional interpretation without excitons (17) invokes a total recombination, R, as $R = An + Bn^2 + Cn^3$, where A is the Shockley-Read-Hall (SRH) recombination rate, B is the radiative recombination rate, and C is the Auger recombination rate. The QY is given as the radiative recombination rate over the total recombination. Auger processes dominate at high carrier concentrations, whereas SRH recombination dominates at low carrier concentrations. In the SRH regime (i.e., low pump power), QY increases with pump intensity. This behavior,

however, was not observed in previous MoS_2 studies (12, 13) or in this work.

The standard model poorly describes our QY data (fig. S10) (14), which are strongly influenced by bound excitons (9). As a result, the radiative rate is proportional to the total exciton population, $\langle N \rangle$ (18). At high exciton densities, nonradiative biexcitonic recombination can dominate, leading to a recombination rate proportional to $\langle N \rangle^2$ (18). Previous reports also suggest that the luminescence in as-exfoliated samples is limited by nonradiative defect-mediated processes (19, 20), resulting in low QY. Although the precise nature of the defect-mediated nonradiative recombination is unclear, a simple analytical model can be developed to describe our experimental results. The total excitation rate, R, in MoS_2 is balanced by recombination

$$R = B_{nr}n^2 + B_r n^2 \tag{1}$$

where B_{nr} is the nonradiative defect-mediated recombination rate and B_r is the formation rate of excitons. The generated excitons can then either

Fig. 2. Steady-state luminescence. (**A**) Pump-power dependence of the integrated PL for as-exfoliated and treated MoS_2. Dashed lines show power law fits for the three dominant recombination regimes. (**B**) Pump-power dependence of the QY for as-exfoliated and treated MoS_2. Dashed lines show the recombination model.
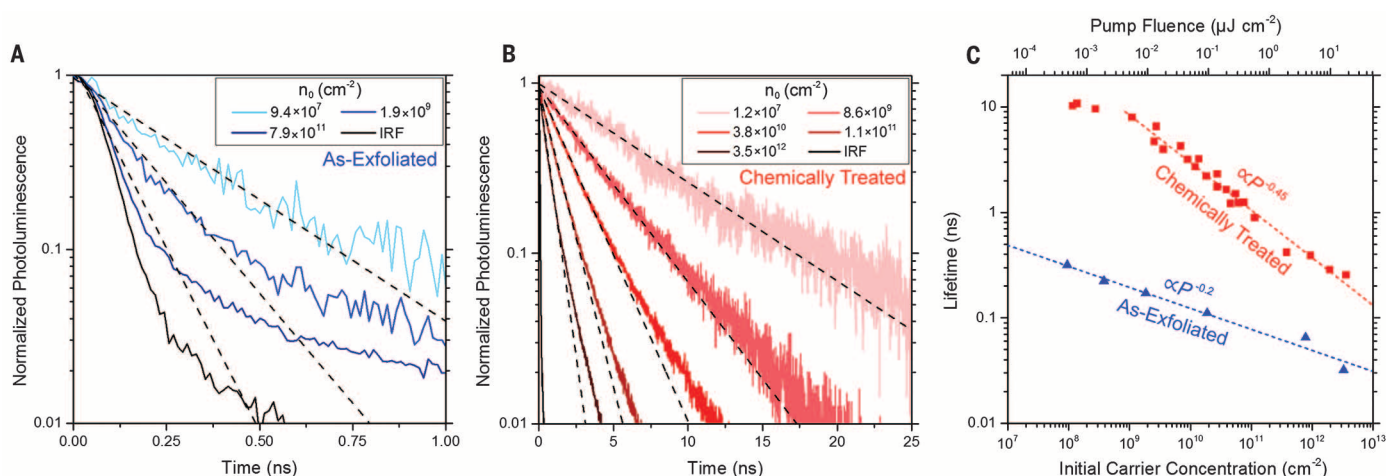
Fig. 3. Time-resolved luminescence. (**A**) Radiative decay of an as-exfoliated MoS$_2$ sample at various initial carrier concentrations (n$_0$), as well as the instrument response function (IRF). (**B**) Radiative decay of a treated MoS$_2$ sample plotted for several initial carrier concentrations (n$_0$), as well as the IRF. Dashed lines in (A) and (B) indicate single exponential fits. (**C**) Effective PL lifetime as a function of pump fluence. Dashed lines show a power law fit for the dominant recombination regimes.

undergo radiative recombination or nonradiatively recombine with a second exciton according to $B_r n^2 = \tau_r^{-1}\langle N \rangle + C_{bx}\langle N \rangle^2$ (*19*), where τ_r is the radiative lifetime and C_{bx} is the biexcitonic recombination rate. The QY is then given as

$$QY = \frac{\tau_r^{-1}\langle N \rangle}{\tau_r^{-1}\langle N \rangle + B_{nr}n^2 + C_{bx}\langle N \rangle^2} \quad (2)$$

For the case of the TFSI-treated sample, B_{nr} is negligible because the QY at low pump powers is >95%, allowing us to extract a biexcitonic recombination coefficient C_{bx} = 2.8 cm^2 s^{-1}. For the as-exfoliated sample, the defect-mediated nonradiative recombination can be fit to B_{nr} = 1.5 × 10^6 cm^2 s^{-1}, using the same C_{bx} value. The fitting results are plotted as the dashed curves in Fig. 2B.

To investigate the carrier recombination dynamics, we performed time-resolved measurements on both as-exfoliated and chemically treated samples. The luminescence decay was nonexponential, but not in the standard form known for bimolecular (Bn^2) recombination (*17*). As-exfoliated monolayers of MoS$_2$ had extremely short lifetimes on the order of 100 ps (Fig. 3A and fig. S6) (*14*), consistent with previous reports (*21*). After treatment, we saw a substantial increase in the lifetime, which is shown at several pump fluences in Fig. 3A. Fitting was performed with a single exponential decay that described only the initial characteristic lifetime for a given pump intensity. After the pump pulse, the exciton population decayed, which resulted in nonexponential decay through reduced nonradiative biexcitonic recombination. At the lowest measurable pump fluences, we observed a luminescence lifetime of 10.8 ± 0.6 ns in the treated sample, compared with ~0.3 ns in the untreated case at a pump fluence of 5 × 10^{-4} μJ cm^{-2} (Fig. 3C). The contrast between panels A and B of Fig. 3 is consistent with the QY trend.

Urbach tails, which depict the sharpness of the band edges (*22*), were derived from the steady-state PL spectra via the van Roosbroeck–Shockley equation and are plotted in fig. S8. After treatment with TFSI, a noticeable decrease in the Urbach energy (E_0) from 17.4 to 13.3 meV was observed, indicating a reduction in the overall disorder from potential fluctuations and improved band-edge sharpness (*22*). A spatial map showing Urbach energy (fig. S8) (*14*) further indicates that the treatment was highly uniform. To evaluate stability, the QY in air for chemically treated MoS$_2$ was measured daily at a constant pump power over the course of 1 week, during which the sample was stored without any passivation in ambient lab conditions (20° to 22°C, 40 to 60% relative humidity), as shown in fig. S9 (*14*). The QY remained above 80% during this period, indicating that the treatment resulted in samples that were relatively stable.

We then turned our attention to the effect of TFSI treatment on other properties of MoS$_2$. The monolayer surface was imaged by atomic force microscopy (AFM) before and after treatment (Fig. 4A). No visible change to the surface morphology was observed. We also investigated the effect of the treatment on the electrical properties of a back-gated MoS$_2$ transistor. The transfer characteristics of this majority carrier device before and after treatment showed a shift in the threshold voltage toward zero, indicating that the native n-type doping in the MoS$_2$ was removed while the same drive current was maintained (Fig. 4B). An improvement in the subthreshold slope indicated that the treatment reduces interface trap states. The Raman spectra of an as-exfoliated and treated monolayer (Fig. 4C) showed that there was no change in the relative intensity or peak position. Thus, the structure of MoS$_2$ was not altered during treatment, and the lattice was not subjected to any induced strain (*23*). Because absolute absorption was used in the calibration of QY, we performed careful absorption measurements using two different methods (*14*), both before and after treatment (Fig. 4D). At the pump wavelength (514.5 nm), no measurable change of the absolute absorption from the treatment was

observed. The strong resonances at 1.88 and 2.04 eV (corresponding to the A and B excitons, respectively) are consistent with previous reports (*12*). We then performed surface-sensitive x-ray photoelectron spectroscopy (XPS) on bulk MoS$_2$ from the same crystal used for micromechanical exfoliation. The Mo 3d and S 2p core levels (Fig. 4E) showed no observable change in oxidation state and bonding after treatment (*24*). Thus, an array of different techniques for materials characterization shows that the structure of the MoS$_2$ remains intact after TFSI treatment, with only the minority carrier properties (i.e., QY and lifetime) enhanced.

The effect of treatment on a wide variety of molecules is shown in table S1 and discussed in the supplementary text. Various polar, nonpolar, and fluorinated molecules, including strong acids and the solvents used for TFSI treatment (dichlorobenzene and dichloroethane), were explored. Treatment with the phenylated derivative of superacid TFSI was also performed (fig. S11) (*14*). These treatments all led to no or minimal (less than one order of magnitude) enhancement in PL QY.

The exact mechanism by which the TFSI passivates surface defects is not fully understood. Exfoliated MoS$_2$ surfaces contain regions with a large number of defect sites in the form of sulfur vacancies, adatoms on the surface, and numerous impurities (*25–27*). In fig. S12A (*14*), the calculated midgap energy is shown for several defect types, including a sulfur vacancy, adsorbed –OH, and adsorbed water. Deep-level traps—which contribute to defect-mediated nonradiative recombination, resulting in a low QY (*27*)—are observed for all of these cases. The strong protonating nature of the superacid can remove absorbed water, hydroxyl groups, oxygen, and other contaminants on the surface. Although these reactions will not remove the contribution of defects to nonradiative recombination, they will open the active defect sites to passivation by a second mechanism. One possibility is the protonation of the three dangling bonds at each sulfur vacancy site. However, density
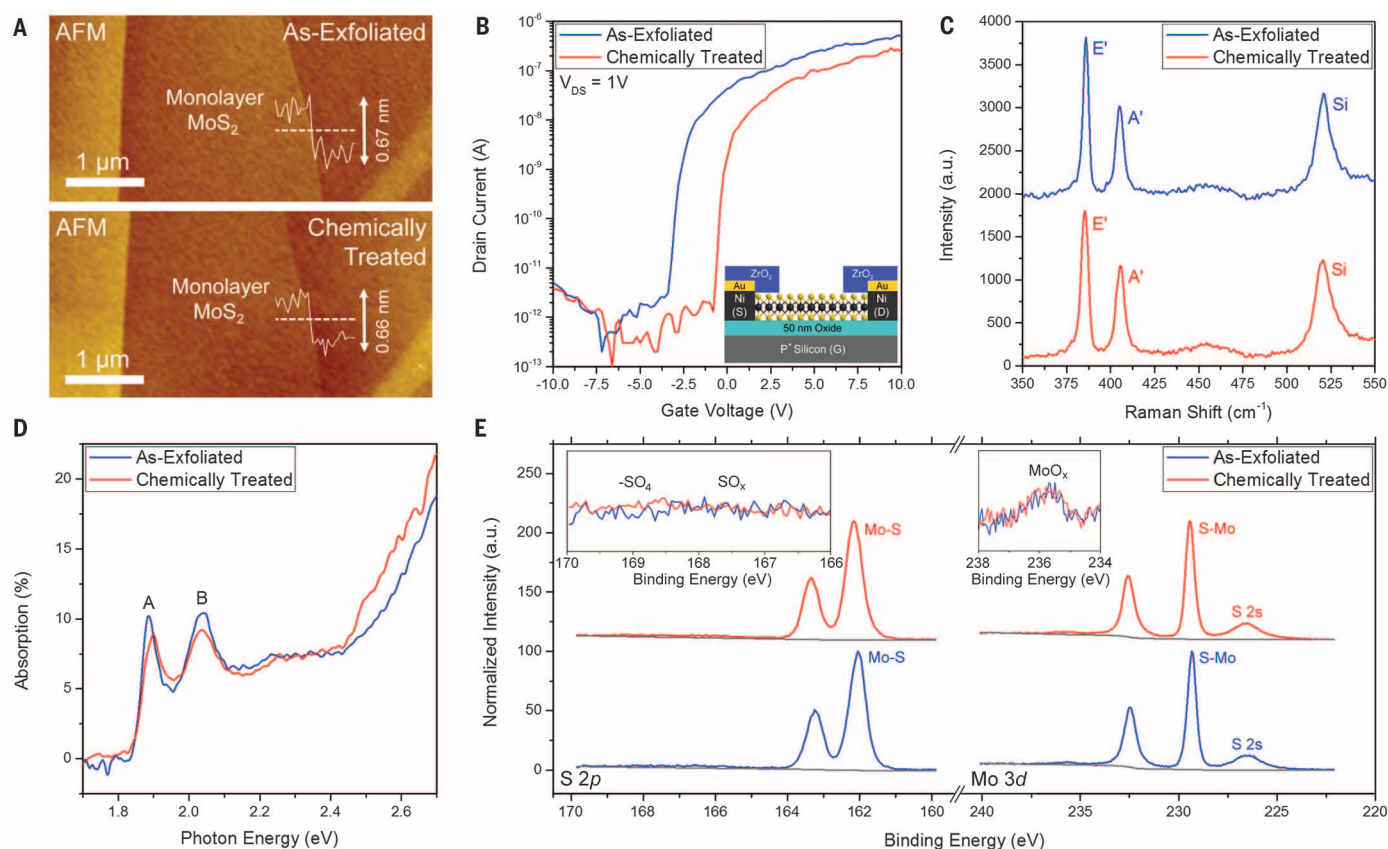
Fig. 4. Material and device characterization. (**A**) AFM images taken before and after TFSI treatment. (**B**) Transfer characteristics of a monolayer MoS₂ transistor, both before and after treatment. V_DS, drain-source voltage; S, source; D, drain; G, gate. (**C**) Raman spectrum of as-exfoliated and TFSI-treated MoS₂ samples. a.u., arbitrary units; E′, MoS₂ in-plane mode; A′, MoS₂ out-of-plane mode; Si, silicon Raman peak. (**D**) Absorption spectrum of the as-exfoliated and treated MoS₂ samples. A and B indicate the exciton resonances. (**E**) XPS spectrum of the S 2p and Mo 3d core levels before and after treatment. The insets show that there is no appearance of SO_x or change in the MoO_x peak intensity after treatment.

functional theory calculations (fig. S12C) (*14*) show that this reaction is energetically unfavorable. A second possibility is that the surface is restructured to reduce the sulfur vacancies through rearrangement of sulfur adatoms on the surface, which can be facilitated by hydrogenation via TFSI (*14*). The presence of sulfur adatom clusters has previously been confirmed by scanning tunneling microscopy and energy-dispersive x-ray spectroscopy (*27–30*). Careful examination of the XPS data over multiple spots before and after TFSI treatment (fig. S13) (*14*) reveals that the ratio of bonded sulfur to molybdenum (S/Mo) increased from 1.84 ± 0.04 in the as-exfoliated case to 1.95 ± 0.05 after treatment (table S2) (*14*).

We have demonstrated an air-stable process by which the PL of monolayer MoS₂ can be increased by more than two orders of magnitude, resulting in near-unity luminescence yield. This result sheds light on the importance of defects in limiting the performance of 2D systems and presents a practical route to eliminate their effect on optoelectronic properties. The existence of monolayers with near-ideal optoelectronic properties should enable the development of new high-performance light-emitting diodes, lasers, and solar cells. These devices can fulfill the revolutionary potential of the 2D semiconductors (*1*), which require interfacial passivation, as in all classic semiconductors.

REFERENCES AND NOTES

1. F. Xia, H. Wang, D. Xiao, M. Dubey, A. Ramasubramaniam, *Nat. Photonics* **8**, 899–907 (2014).
2. S. Wu *et al.*, *Nature* **520**, 69–72 (2015).
3. H. Fang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6198–6202 (2014).
4. L. Britnell *et al.*, *Science* **340**, 1311–1314 (2013).
5. K. F. Mak, K. He, J. Shan, T. F. Heinz, *Nat. Nanotechnol.* **7**, 494–498 (2012).
6. S. B. Desai *et al.*, *Nano Lett.* **14**, 4592–4597 (2014).
7. S. Tongay *et al.*, *Sci. Rep.* **3**, 2657 (2013).
8. A. Chernikov *et al.*, *Phys. Rev. Lett.* **113**, 076802 (2014).
9. H. M. Hill *et al.*, *Nano Lett.* **15**, 2992–2997 (2015).
10. T. C. Berkelbach, M. S. Hybertsen, D. R. Reichman, *Phys. Rev. B* **88**, 045318 (2013).
11. L. Yuan, L. Huang, *Nanoscale* **7**, 7402–7408 (2015).
12. K. F. Mak, C. Lee, J. Hone, J. Shan, T. F. Heinz, *Phys. Rev. Lett.* **105**, 136805 (2010).
13. H. Wang, C. Zhang, F. Rana, *Nano Lett.* **15**, 339–345 (2015).
14. Materials and methods are available as supplementary materials on *Science* Online.
15. I. Schnitzer, E. Yablonovitch, C. Caneau, T. J. Gmitter, *Appl. Phys. Lett.* **62**, 131–133 (1993).
16. Y. S. Park *et al.*, *Phys. Rev. Lett.* **106**, 187401 (2011).
17. P. T. Landsberg, *Phys. Status Solidi* **41**, 457–489 (1970).
18. F. Wang, Y. Wu, M. S. Hybertsen, T. F. Heinz, *Phys. Rev. B* **73**, 245424 (2006).
19. H. Wang *et al.*, *Phys. Rev. B* **91**, 165411 (2015).
20. C. Mai *et al.*, *Nano Lett.* **14**, 202–206 (2014).
21. T. Korn, S. Heydrich, M. Hirmer, J. Schmutzler, C. Schuller, *Appl. Phys. Lett.* **99**, 102109 (2011).
22. A. Iribarren, R. Castro-Rodriguez, V. Sosa, J. L. Pena, *Phys. Rev. B* **58**, 1907–1911 (1998).
23. Z. Liu *et al.*, *Nat. Commun.* **5**, 5246 (2014).
24. A. Azcatl *et al.*, *Appl. Phys. Lett.* **104**, 111601 (2014).
25. C. P. Lu, G. Li, J. Mao, L. M. Wang, E. Y. Andrei, *Nano Lett.* **14**, 4628–4633 (2014).
26. S. McDonnell, R. Addou, C. Buie, R. M. Wallace, C. L. Hinkle, *ACS Nano* **8**, 2880–2888 (2014).
27. R. Addou *et al.*, *ACS Nano* **9**, 9124–9133 (2015).
28. R. Addou, L. Colombo, R. M. Wallace, *Appl. Mat. Interfaces.* **7**, 11921–11929 (2015).
29. J. Y. Noh, H. Kim, Y. S. Kim, *Phys. Rev. B* **89**, 205417 (2014).
30. A. P. Nayak *et al.*, *Nat. Commun.* **5**, 3731 (2014).

GEOLOGY

Deformation-assisted fluid percolation in rock salt

Soheil Ghanbarzadeh,[1] Marc A. Hesse,[2,3]* Maša Prodanović,[1] James E. Gardner[2]

Deep geological storage sites for nuclear waste are commonly located in rock salt to ensure hydrological isolation from groundwater. The low permeability of static rock salt is due to a percolation threshold. However, deformation may be able to overcome this threshold and allow fluid flow. We confirm the percolation threshold in static experiments on synthetic salt samples with x-ray microtomography. We then analyze wells penetrating salt deposits in the Gulf of Mexico. The observed hydrocarbon distributions in rock salt require that percolation occurred at porosities considerably below the static threshold due to deformation-assisted percolation. Therefore, the design of nuclear waste repositories in salt should guard against deformation-driven fluid percolation. In general, static percolation thresholds may not always limit fluid flow in deforming environments.

R ock salt in sedimentary basins has long been considered to be impermeable and provides a seal for hydrocarbon accumulations in geological structures (1, 2). The low permeability of rock salt also has the potential to isolate nuclear waste from ambient groundwater and may provide a suitable deep geological waste repository (3, 4). This option is currently being reconsidered in the United States after the closure of the Yucca Mountain repository

in Nevada (3). However, field observations of oil-impregnated rock salt (5), geochemical evidence for the replacement of the in situ brines (6), and the drainage of brine from mining-induced fractures and dilatant microcracking (3, 7) demonstrate that the permeability of natural rock salt may not be negligible.

Brine-filled pore networks in rock salt approach textural equilibrium due to fast reaction kinetics of salt dissolution and repeciptation

(8). Percolation in these networks is controlled by the dihedral angle θ at the solid-solid-liquid triple junctions

$$\theta = 2\cos^{-1}[\gamma_{ss}/(2\gamma_{sl})] \qquad (1)$$

where γ_{ss} and γ_{sl} are the solid-solid and solid-liquid surface energies (9–12). The dihedral angle is therefore a thermodynamic property that changes with pressure P and temperature T. The static pore-scale theory shows that texturally equilibrated pore networks percolate at any porosity if $\theta \leq 60°$, whereas a finite porosity is required for percolation if $\theta > 60°$ (10–12).

The experimentally measured θ in salt-brine systems decreases with both increasing P and T (Fig. 1), suggesting that fluids at shallow depth must overcome a percolation threshold, whereas fluids at greater depth are likely to percolate at any porosity. The PT trajectory of multiple petroleum wells in the Gulf of Mexico crosses this transition and therefore provides an opportunity to test the static pore-scale theory in a realistic field setting.

We confirm the static pore-scale theory in undrained laboratory experiments on synthetic

[1]Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, 200 East Dean Keeton Street, Austin, TX 78712, USA. [2]Department of Geological Sciences, The University of Texas at Austin, 1 University Station, Austin, TX 78712, USA. [3]Institute for Computational Engineering and Sciences, The University of Texas at Austin, 201 East 24th Street, Austin, TX 78712, USA.
*Corresponding author. E-mail: mhesse@jsg.utexas.edu
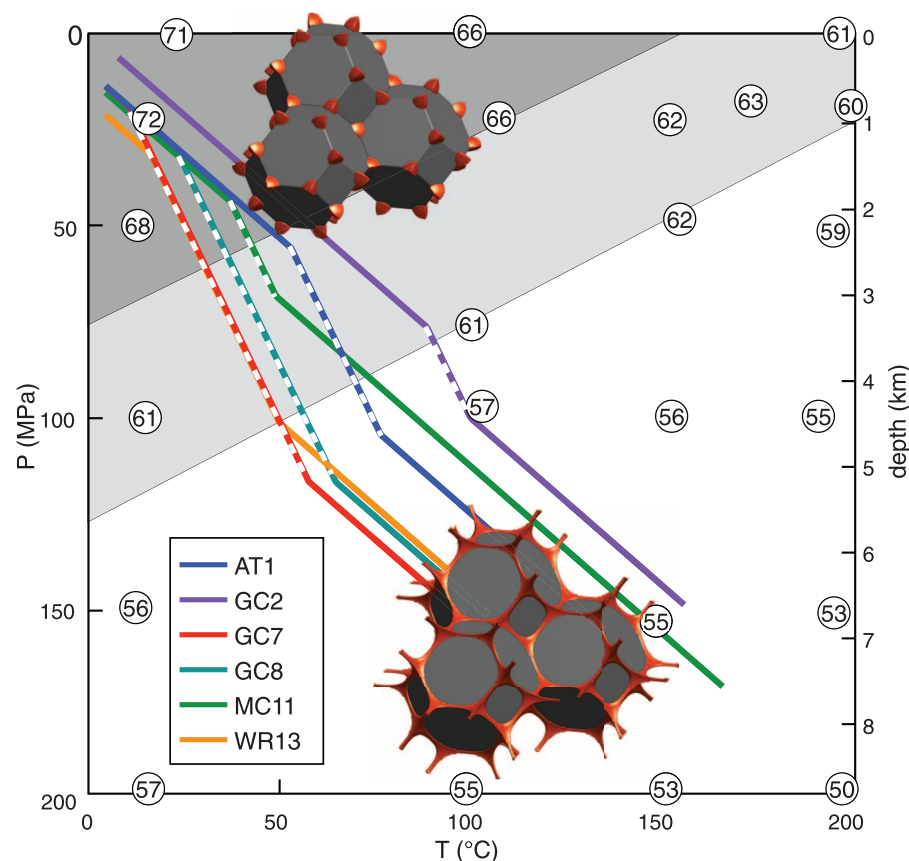
Fig. 1. Brine percolation in rock salt. PT trajectories of multiple subsalt petroleum wells are shown together with experimentally measured dihedral angles θ for the salt-brine system (8). The static theory predicts that fluid must overcome a percolation threshold in the gray area, whereas fluids are predicted to percolate at any porosity in the white area. The light gray area highlights the transition zone, 60° < θ < 65°, between percolating and disconnected pore space (8). The segment of each well that is located within the salt has a lower geothermal gradient due to the high conductivity of salt and is shown as a dashed line. The depth axis is only for illustration and assumes an overburden with constant density, ρ = 2300 kg/m³.

salt samples that have been imaged with non-destructive x-ray microtomography after quenching to ambient conditions (*13*). We present the results of two representative experiments (Fig. 2) performed at P = 20 MPa and T = 100°C (Exp-I) and P = 100 MPa and T = 275°C (Exp-II). The three-dimensional (3D) reconstruction (Fig. 2, A and B) and medial axis representation of the pore space (Fig. 2, C and D) show that the brine network is disconnected in Exp-I and is connected in Exp-II. This is confirmed by statistical analysis of the coordination number distributions that show that almost all nodes in Exp-I have coordination number 1, whereas the coordination numbers of 3 and 4 are most abundant in Exp-II (fig. S4). The distribution of the apparent dihedral angles has a median of 67 ± 5° for Exp-I and 52 ± 6° for Exp-II (Fig. 2E). Distributions with a single narrow peak, as well as similarity to previously reported values of dihedral angle (*8*), indicate that the experiments are approaching textural equilibrium. Comparison of experiments with the regime diagram for fluid percolation show that static pore-scale theory successfully predicts the connectivity of the pore space (Fig. 2F). These experimental results confirm the first-order control of the dihedral angle on brine percolation and serve as a baseline for the field observations of fluid distributions in deformed rock salt.

Commercial interest in the large hydrocarbon accumulations below extensive bodies of allochthonous salt in the deepwater Gulf of Mexico provides an opportunity to test the static pore-scale theory in slowly moving natural rock salt. We studied field data from the salt section of 48 wells crossing the predicted transition zone from disconnected to percolating pore space (Fig. 1) to constrain the brine and hydrocarbon connectivity. Typically, no intact core is recovered from the salt section of wells, and the available data sets consist of wireline well logs and mud logs (*13*). Wireline well logs, obtained by lowering a measurement tool into the well, characterize different properties of the formation rock and fluids (Fig. 3, A and B). Mud logs, which record the hydrocarbon gas content and observations from the drill cuttings brought to the surface, provide direct constraints on the presence of hydrocarbons in salt (Fig. 3, C to E). Hydrocarbon signs reported in mud logs include fluorescence, oil staining, oil cut, and dead oil embedded in the salt.

We chose only those salt sections for analysis that were free of other rock fragments, as indicated by low values of naturally occurring gamma radiation (Fig. 3A). In contrast to the uniform gamma-ray signature, all other logs (Fig. 3, B to E) show a distinct change in the bottom third of the salt. The very high electrical resistivity in

the upper two-thirds of the salt section implies that the conductive brine is not connected (Fig. 3B) (*14*). In this region, the porosity calculated from Archie's law is below 0.4% (Fig. 2G) (*13*). The reduction of electrical resistivity by an order of magnitude in the bottom third suggests that brine is connected at porosities below 0.8% (Fig. 2G). The salt-brine dihedral angle inferred from the *PT* trajectory of the well (fig. S6) (*13*) drops below 60° in the bottom third of the salt (Fig. 3F), consistent with the static pore-scale theory.

In addition to a connected brine phase, the total gas hydrocarbons and gas chromatography logs indicate a substantial increase in the amount of natural gas in the lower third of the salt (Fig. 3, C and D). We observe this general pattern also in the mud logs that contain no indications of hydrocarbons in the top two-thirds but show multiple signs of hydrocarbons in the bottom third (Fig. 3E). In the presence of brine, hydrocarbons are the nonwetting phase, so that the textural equilibration of the pore network occurs through brine-mediated dissolution and reprecipitation of the salt. The dihedral angle of the brine-salt system governs the connectivity of pore space, consistent with observations in wireline well logs and mud logs. Once hydrocarbons overcome the capillary entry pressure (*5*), they can enter the salt in regions where the brine network is
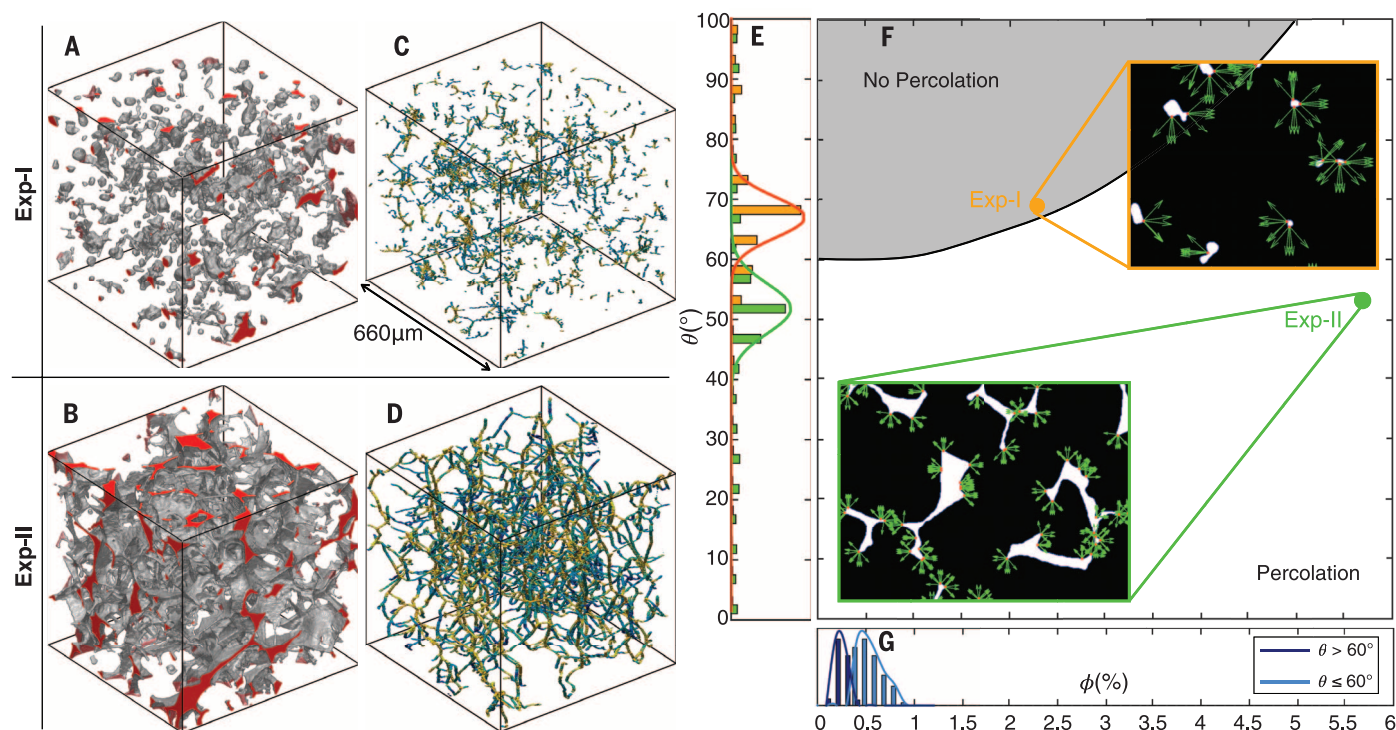
Fig. 2. Pore networks in rock salt. Hydrostatic experiments on synthetic rock salt have been performed at P = 20 MPa and T = 100°C (Exp-I) and P = 100 MPa and T = 275°C (Exp-II). (**A** and **B**) 3D reconstruction of the pore network at textural equilibrium; all edges of the 3D volumes correspond to 660 μm. (**C** and **D**) The skeletonized pore network extracted from the reconstructed 3D volume; colored according to local pore-space-inscribed radius, with warmer colors indicating larger radius. (**E**) Distribution of apparent dihedral angles in the experiments. (**F**) Exp-I and Exp-II in the θϕ space regime diagram with the percolation threshold obtained from the static pore-scale theory (*10*, *12*). Inserted images show the details of automated dihedral angle extraction from 2D images (*13*). We report the median value of dihedral angles and the estimated errors based on the 95% confidence interval. (**G**) Porosity of natural rock salt inferred from resistivity logs (Fig. 3B).

connected. Subsequent imbibition of the brine can trap the hydrocarbons in the pore space (fig. S7). The presence of hydrocarbons therefore indicates that a connected pore space existed during the entry of the hydrocarbons into the rock salt. This interpretation is consistent with previous work reporting direct observations of oil-stained salt cores recovered from conditions where θ < 60° (5).

High-quality resistivity logs (Fig. 3B) are only available in two wells due to technical difficulties and lack of commercial interest in the salt section of wells. Therefore, we rely on the logs that detect hydrocarbons to infer the connectivity of the brine in the remaining 46 wells. We group spatially associated wells to look at the distribution of hydrocarbons in salt sections (Fig. 4). The abundance of hydrocarbons is affected by the distance of the nearest hydrocarbon source from the bottom of the salt. For example, the first oil source is more than 2000 m below the base of salt in the wells of group WR13, justifying the sparsity of hydrocarbon signs.

We converted the depth to dihedral angle using available experimental data (Fig. 1 and fig. S6). All the wells that we considered show signs of connected pore space at depths where the dihedral angle is below 60°, except the shallow wells of group MC11. Using the two electrical resistivity logs and Archie's law, we estimate that the porosity of these connected regions is less than 1% (Fig. 2G). This provides direct field evidence that dihedral angles below 60° allow the percolation of texturally equilibrated pore networks at porosities below the transport limit in more typical porous media that originated as clastic sediments (15).

Nonetheless, field data also show evidence of percolating pore space at shallower depths, where the dihedral angle is substantially above 60° (Fig. 4). Under these conditions, the porosity must increase above a threshold to allow percolation. Static pore-scale theory requires porosities between 2 and 3% to allow percolation at dihedral angles between 65° and 70° (Fig. 2F). However, none of the porosities inferred from the available resistivity logs exceed 1%, and most are substantially lower (Fig. 2G), which is consistent with direct measurements of rock salt porosity (16, 17). The observation of percolating fluids at high dihedral angles and low porosities is not consistent with the static theory.

Viscous flow of rock salt due to the density contrast with the surrounding sediments may explain the failure of the static pore-scale theory to predict the percolation of pore space at high dihedral angles. At low effective mean stress, deformation-induced microcracking can lead to the formation of a percolating pore space (5). This microcracking-induced percolation is commonly observed in the zone of disturbed rock around openings in salt mines or nuclear waste repositories and under high overpressures in nature (3, 5). At the depth of petroleum wells considered here, the effective mean stress is sufficient that deformation occurs in the compac-

tion regime, where existing microcracks close and heal (18, 19).

However, deformation may induce permeability even in the absence of microcracking. At high effective mean stress and in the presence of small amounts of brine, the dislocation creep of salt is accompanied by fluid-assisted dynamic recrystallization and pressure solution creep (20–22). Both static and dynamic recrystallization are associated with transformation of the isolated grain-boundary fluid inclusions into grain-boundary fluid films (23, 24). The dynamic wetting of the grain boundaries and compaction have been observed in deformation experiments under conditions where θ ≈ 64° (22). This suggests that dynamic grain-boundary wetting induced fluid percolation and drainage at porosities below the percolation threshold.

These laboratory results must be extrapolated to natural conditions using appropriate microphysical models and suggest that fluid-assisted dynamic recrystallization becomes important at strain rates below 10^{-10} s^{-1} (21). This is consistent

with the recrystallized microstructures and x-ray microtomography of grain-boundary brine films in natural rock salt, as well as estimated natural strain rates between 10^{-15} and 10^{-11} s^{-1} (5, 25, 26). This confirms earlier suggestions that dynamic grain-boundary wetting associated with grain-boundary migration is a plausible mechanism in natural rock salt.

This conclusion is also supported by the comparison of the relative magnitude of shear stresses, Δσ, and the capillary pressure introduced by surface tension forces, Δp, given by capillary number

$$Ca = \frac{\Delta\sigma}{\Delta p} = \frac{\Delta\sigma}{2\gamma_{sl}/r} \qquad (2)$$

where r is the mean radius of disconnected pores. Microstructural evidence preserves records of differential stresses up to 1 MPa in subhorizontal bedded salts (27) and 2 MPa in salt domes (5, 28). In comparison, the capillary pressure for $r = 10^{-4}$ m and $\gamma_{sl} = 0.1$ N/m is on the
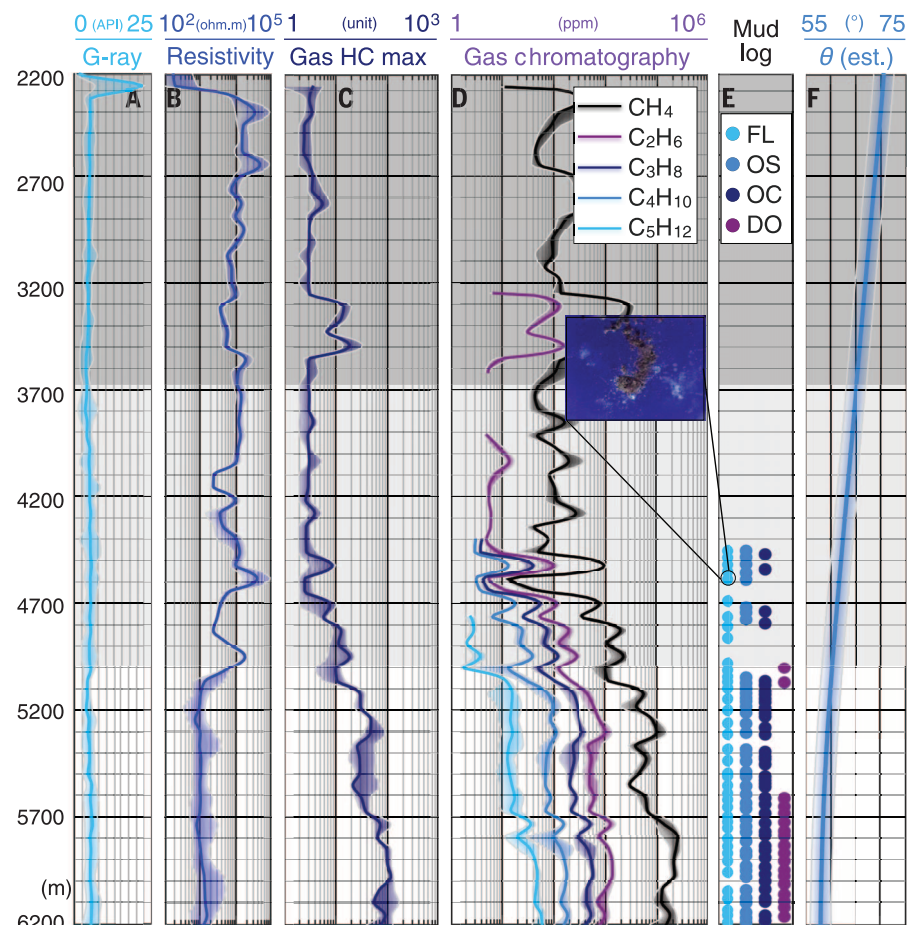
Fig. 3. Petrophysical observations. Wireline well logs and mud logs data constraining the fluid distribution and connectivity in the well GC8 from the deep water Gulf of Mexico (13). (**A**) Gamma-ray log, (**B**) electrical resistivity, (**C**) total hydrocarbons gas, (**D**) gas chromatography, (**E**) hydrocarbon signs (FL, fluorescence; OS, oil stain; DO, dead oil; and OC, oil cut) in mud logs, and (**F**) the dihedral angle inferred from experimental data (Fig. 1). Shading around each curve shows the measurement error and average fluctuations in data. The gray background corresponds to shaded areas in the experimental data (Fig. 1).

Fig. 4. Fluid distributions in salt wells. Hydrocarbons signs from mud logs of all 48 wells covering 150,000 m of salt are shown as a function of dihedral angle (13). Wells are divided into 14 groups based on spatial proximity. Salt extent is shown by an arrow in each region. Theoretical fluid connectivity is indicated by gray scale (Fig. 1). Abbreviations denote the following protraction areas in the Gulf of Mexico: AT, Atwater Valley; GC, Green Canyon; KC, Keathley Canyon; MC, Mississippi Canyon; and WR, Walker Ridge.
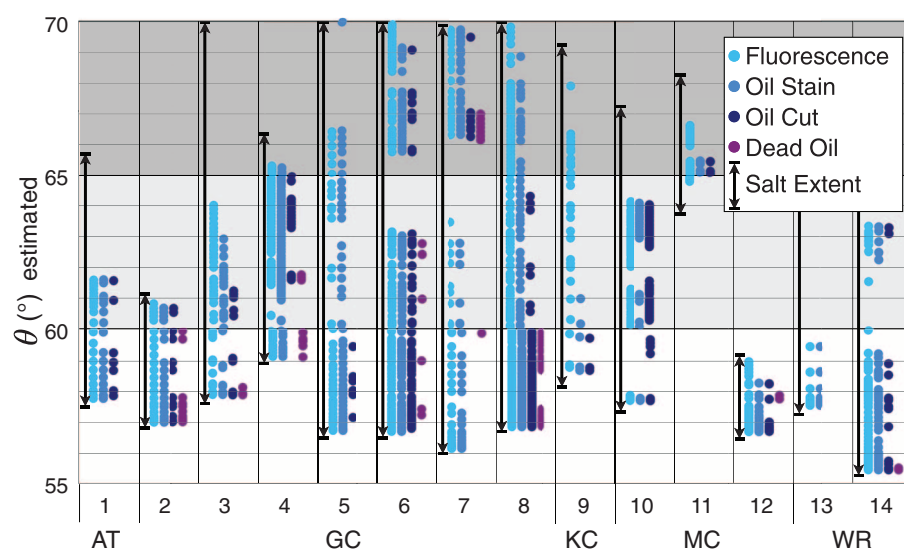
order of 10^3 Pa (29). Therefore, $Ca \approx 10^3$ and the shear stresses in rock salt may exceed capillary pressures and hence facilitate deformation-assisted percolation. This provides an explanation for the penetration of hydrocarbons into shallow regions of the salt, where $\theta > 60°$ and porosity is below the static percolation threshold (Fig. 2G and Fig. 4).

These field observations have implications for ensuring hydrological isolation of nuclear waste in rock salt. At the relatively shallow depth typically considered for geological storage, the dihedral angle is between $65°$ and $72°$ and should prevent brine percolation in rock salt, based on static pore-scale theory and experiments. However, field observations reported here show that such moderate dihedral angles do not guarantee hydrological isolation in deformed rock salt. The deformation-assisted percolation observed in salt sections of petroleum wells is not associated with human-made excavations, suggesting that this mechanism is not limited to the vicinity of the repository site and the duration of room closure around the waste. Lower differential stresses recorded in shallow bedded rock salt suggest that it is more likely to provide an impermeable barrier. However, tectonic forces and excavations can result in high stresses in shallow cold salt. Therefore, it is important to characterize the salt microstructure of potential repositories to determine the stress history, state of grain boundaries, and fluid distribution. Future work should also constrain the permeability that can be generated by deformation-assisted percolation and its persistence.

Beyond the direct application to salt-brine systems, the field observations reported here also provide an important test of a general theory that underlies our understanding of fluid percolation and flow in ductile regions of Earth. This is of particular interest to the debate about whether moderate dihedral angles can prevent the segregation of core-forming melts in the deforming lower mantle (30–32). The inaccessibility of

Earth's mantle to field observations has prevented the resolution of this debate. The observations of fluid distribution in rock salt reported here show that deformation-assisted percolation is possible and suggest that core formation by percolation may be a viable mechanism, even if the dihedral angle is above 60°.

REFERENCES AND NOTES

1. M. W. Downey, *AAPG Bull.* **68**, 1752–1763 (1984).
2. S. A. Stewart, *Geol. Soc. London Spec. Publ.* **272**, 361–396 (2007).
3. F. D. Hansen, C. D. Leigh, Technical Report SAND2011-0161 (Sandia National Laboratories, Albuquerque, NM, 2011).
4. U. Noseck, J. Wolf, W. Steininger, B. Miller, *Swiss J. Geosci.* **108**, 121–128 (2015).
5. J. Schoenherr *et al.*, *AAPG Bull.* **91**, 1541–1557 (2007).
6. L. S. Land, J. A. Kupecz, L. Mack, *Chem. Geol.* **74**, 25–35 (1988).
7. I. Davison, *J. Geol. Soc. London* **166**, 205–216 (2009).
8. S. Lewis, M. Holness, *Geology* **24**, 431–434 (1996).
9. S. Ghanbarzadeh, M. A. Hesse, M. Prodanović, *J. Comput. Phys.* **297**, 480–494 (2015).
10. N. von Bargen, H. S. Waff, *J. Geophys. Res.* **91**, 9261–9276 (1986).
11. D. A. Wark, E. Watson, *Earth Planet. Sci. Lett.* **164**, 591–605 (1998).
12. S. Ghanbarzadeh, M. Prodanović, M. A. Hesse, *Phys. Rev. Lett.* **113**, 048001 (2014).
13. Materials and methods are available as supplementary materials on *Science* Online.
14. T. Watanabe, C. J. Peach, *J. Geophys. Res. Solid Earth* **107**, ECV2-1–ECV2-12 (2002).
15. S. C. van der Marck, *Water Resour. Res.* **35**, 595–599 (1999).
16. U. Yaramanci, D. Flach, *Geophys. Prospect.* **40**, 85–100 (1992).
17. U. Yaramanci, *Geophys. Prospect.* **42**, 229–239 (1994).
18. N. D. Cristescu, U. Hunsche, in *Wiley Series in Materials, Modeling and Computation*, C. S. Desai, E. Krempl, Eds. (Wiley, New York, 1998), pp. 81–117.
19. O. Schulze, T. Popp, H. Kern, *Eng. Geol.* **61**, 163–180 (2001).
20. J. L. Urai, *Tectonophysics* **96**, 125–157 (1983).
21. J. L. Urai, C. J. Spiers, H. J. Zwart, G. S. Lister, *Nature* **324**, 554–557 (1986).
22. C. J. Peach, C. J. Spiers, P. W. Trimby, *J. Geophys. Res.* **106** (B7), 13315–13328 (2001).
23. G. Desbois *et al.*, *Contrib. Mineral. Petrol.* **163**, 19–31 (2012).
24. M. R. Drury, J. L. Urai, *Tectonophysics* **172**, 235–253 (1990).
25. N. Thiemeyer, J. Habersetzer, M. Peinl, G. Zulauf, J. Hammer, *J. Struct. Geol.* **77**, 92–106 (2015).
26. M. P. A. Jackson, C. J. Talbot, *Geol. Soc. Am. Bull.* **97**, 305–323 (1986).
27. Z. Schléder, J. L. Urai, *Int. J. Earth Sci.* **94**, 941–955 (2005).
28. N. L. Carter, S. T. Horseman, J. E. Russell, J. Handin, *J. Struct. Geol.* **15**, 1257–1271 (1993).
29. D. Tromans, J. A. Meech, *Miner. Eng.* **15**, 1027–1041 (2002).
30. D. Bruhn, N. Groebner, D. L. Kohlstedt, *Nature* **403**, 883–886 (2000).
31. M. C. Shannon, C. B. Agee, *Science* **280**, 1059–1061 (1998).
32. C. Y. Shi *et al.*, *Nat. Geosci.* **6**, 971–975 (2013).

ECONOMICS

Predicting poverty and wealth from mobile phone metadata

Joshua Blumenstock,[1]* Gabriel Cadamuro,[2] Robert On[3]

Accurate and timely estimates of population characteristics are a critical input to social and economic research and policy. In industrialized economies, novel sources of data are enabling new approaches to demographic profiling, but in developing countries, fewer sources of big data exist. We show that an individual's past history of mobile phone use can be used to infer his or her socioeconomic status. Furthermore, we demonstrate that the predicted attributes of millions of individuals can, in turn, accurately reconstruct the distribution of wealth of an entire nation or to infer the asset distribution of microregions composed of just a few households. In resource-constrained environments where censuses and household surveys are rare, this approach creates an option for gathering localized and timely information at a fraction of the cost of traditional methods.

R eliable, quantitative data on the economic characteristics of a country's population are essential for sound economic policy and research. The geographic distribution of poverty and wealth is used to make decisions about resource allocation and provides a foundation for the study of inequality and the determinants of economic growth (*1*, *2*). In developing countries, however, the scarcity of reliable quantitative data represents a major challenge to policy-makers and researchers. In much of Africa, for instance, national statistics on economic production may be off by as much as 50% (*3*). Spatially disaggregated data, which are necessary for small-area statistics and which are used by both the private and public sector, often do not exist (*4*, *5*).

In wealthy nations, novel sources of passively collected data are enabling new approaches to demographic modeling and measurement (*6*–*8*). Data from social media and the "Internet of Things," for instance, have been used to measure

unemployment (*9*), electoral outcomes (*10*), and economic development (*8*). Although most comparable sources of big data are scarce in the world's poorest nations, mobile phones are a notable exception: They are used by 3.4 billion individuals worldwide and are becoming increasingly ubiquitous in developing regions (*11*).

Here we examine the extent to which anonymized data from mobile phone networks can be used to predict the poverty and wealth of individual subscribers, as well as to create high-resolution maps of the geographic distribution of wealth. That this may prove fruitful is motivated by the fact that mobile phone data capture rich information, not only on the frequency and timing of communication events (*12*) but also reflecting the intricate structure of an individual's social network (*13*, *14*), patterns of travel and location choice (*15*–*17*), and histories of consumption and expenditure. Regionally aggregated measures of phone penetration and use have also been shown to correlate with regionally aggregated population statistics from censuses and household surveys (*8*, *18*, *19*).

Our approach is different from prior work that has examined the relation between regional wealth and regional phone use, as we focus on understanding how the digital footprints of a single individual can be used to accurately predict that same

individual's socioeconomic characteristics. This distinction is a scientific one, which also has several important implications: First, it allows for the method to be used in contexts for which recent census or household survey data are unavailable. Second, when an authoritative source of data does exist, it can be used to more objectively validate or refute the model's predictions. This limits the likelihood that the model is overfit on data from a single source, which is otherwise difficult to control, even with careful cross-validation (*20*). Third, our approach allows for a broad class of potential applications that require inferences about specific individuals instead of census tracts. As we discuss in the supplementary materials (section 6), future iterations of this approach could help to improve the targeting of humanitarian aid and social welfare, disseminate information to vulnerable populations, and measure the effects of policy interventions.

For this study, we used an anonymized database containing records of billions of interactions on Rwanda's largest mobile phone network and supplemented this with follow-up phone surveys of a geographically stratified random sample of 856 individual subscribers. Upon contacting and surveying each of these individuals, we received informed consent to merge their survey responses with the mobile phone transaction database. The surveys solicited no personally identifying information but contained questions on asset ownership, housing characteristics, and several other basic welfare indicators. From these data, we constructed a composite wealth index using the first principal component of several survey responses related to wealth (*21*, *22*) (supplementary materials section 1D). For each of the 856 respondents, we thus have ~75 survey responses, as well as the historical records of thousands of phone-based interactions such as calls and text messages (Table 1).

We use the merged data from this sample of 856 phone survey respondents to show that a mobile phone subscriber's wealth can be predicted from his or her historical patterns of phone use (Fig. 1A) (cross-validated correlation coefficient $r = 0.68$). Our approach to modeling combines feature engineering with feature selection by first transforming each person's mobile phone transaction logs into a large set of quantitative metrics and then winnowing out metrics

[1]Information School, University of Washington, Seattle, WA 98195, USA. [2]Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, USA. [3]School of Information, University of California, Berkeley, Berkeley, CA 94720, USA.
*Corresponding author. E-mail: joshblum@uw.edu

Table 1. Summary statistics for primary data sets. Phone survey data were collected by the authors in Kigali, in collaboration with the Kigali Institute of Science and Technology. Call detail records were collected by the primary mobile phone operator in Rwanda at the time of the phone survey. Demographic and Health Survey (DHS) data were collected by the Rwandan National Institute of Statistics. N/A, not applicable.

Summary statistic	Phone survey	Call detail records	DHS (2007)	DHS (2010)
Number of unique individuals	856	1.5 million	7377	12,792
Data collection period	July 2009	May 2008–May 2009	Dec. 2007–Apr. 2008	Sept. 2010–Mar. 2011
Number of questions in survey	75	N/A	1615	3396
Primary geographic units	30 districts	30 districts	30 districts	30 districts
Secondary geographic units	300 cell towers	300 cell towers	247 clusters	492 clusters
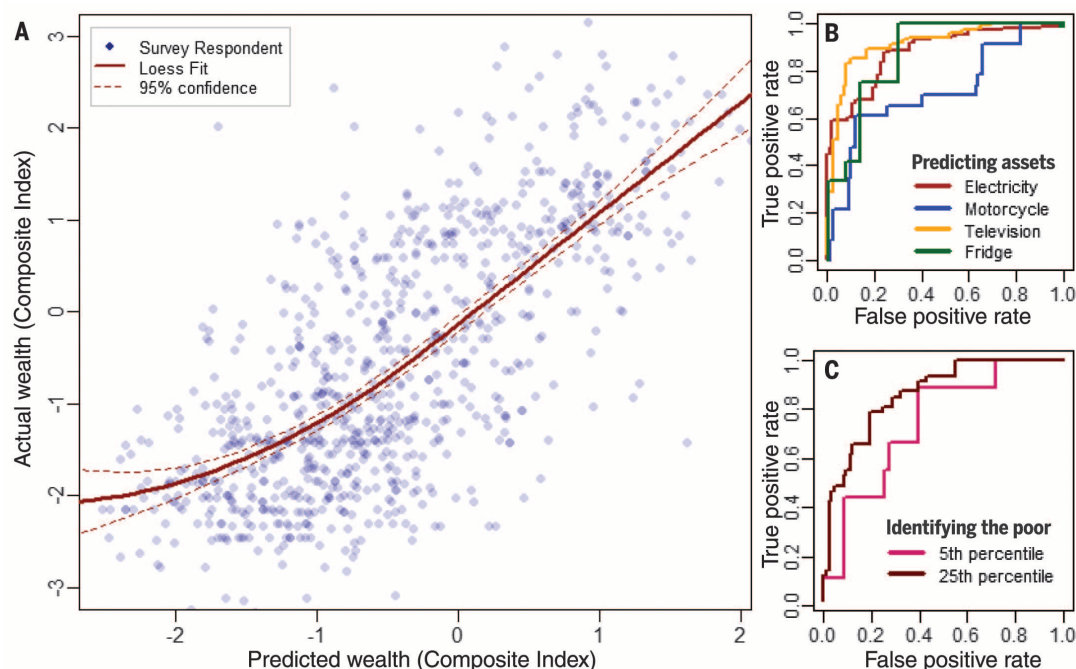
Fig. 1. Predicting survey responses with phone data. (**A**) Relation between actual wealth (as reported in a phone survey) and predicted wealth (as inferred from mobile phone data) for each of the 856 survey respondents. (**B**) Receiver operating characteristic (ROC) curve showing the model's ability to predict whether the respondent owns several different assets. AUC values for electricity, motorcycle, television, and fridge, respectively, are as follows: 0.85, 0.67, 0.84, and 0.88. (**C**) ROC curve illustrates the model's ability to correctly identify the poorest individuals. The poor are defined as those in the 5th percentile (AUC = 0.72) and the 25th percentile (AUC = 0.81) of the composite wealth index distribution.
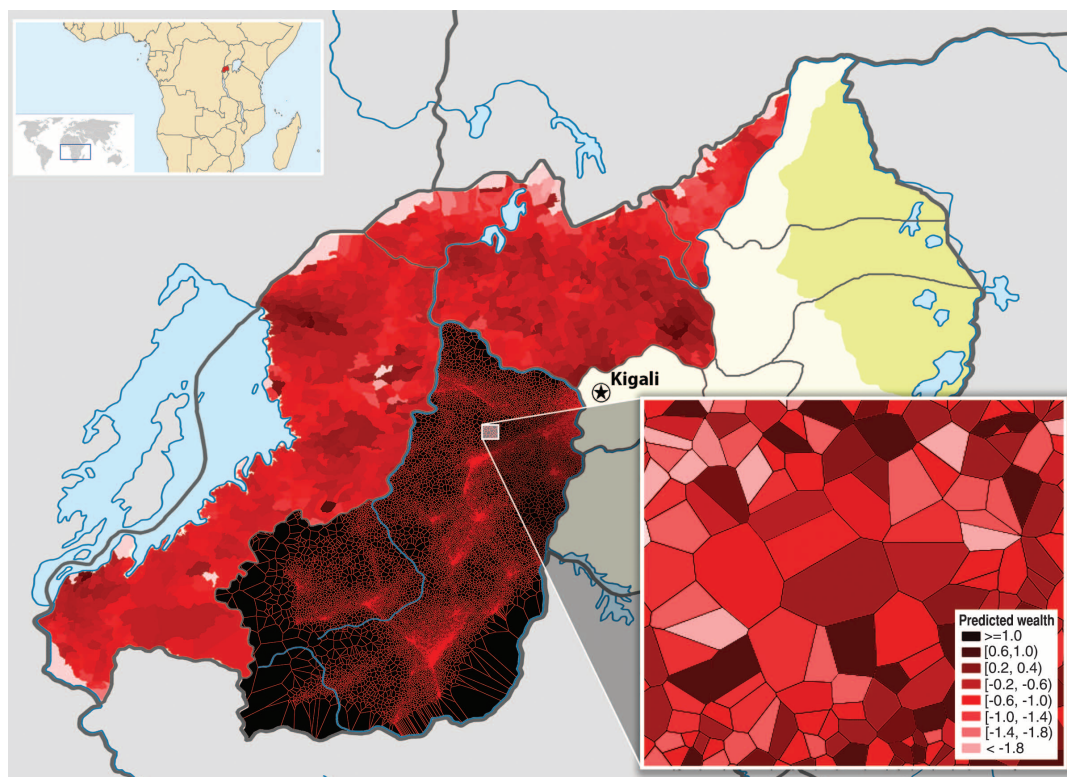
Fig. 2. Construction of high-resolution maps of poverty and wealth from call records. Information derived from the call records of 1.5 million subscribers is overlaid on a map of Rwanda. The northern and western provinces are divided into cells (the smallest administrative unit of the country), and the cell is shaded according to the average (predicted) wealth of all mobile subscribers in that cell. The southern province is overlaid with a Voronoi division that uses geographic identifiers in the call data to segment the region into several hundred thousand small partitions. (**Bottom right inset**) Enlargement of a 1-km² region near Kiyonza, with Voronoi cells shaded by the predicted wealth of small groups (5 to 15 subscribers) who live in each region.

that are not predictive of wealth. The first step employs a structured, combinatorial method to automatically generate several thousand metrics from the phone logs that quantify factors such as the total volume, intensity, timing, and directionality of communication; the structure of the individual's contact network; patterns of mobility and migration based on geospatial markers in the data; and so forth. The second step uses "elastic net" regularization to eliminate irrelevant phone metrics and select a parsimonious model that is more likely to generalize (23). We use cross-validation to limit the possibility that the model is overfit on the small sample on which it is trained. In the supplementary materials (section 3B), we provide details on these methods and show that comparable results are obtained under a variety of alternative supervised-learning

models, including tree-based ensemble regressors and classifiers (24). We also show that this two-step approach to feature engineering and model selection performs significantly better than a more intuitive approach based on a small number of hand-crafted metrics (table S1).

In addition to predicting composite wealth, this same approach can be used to estimate, with varying degrees of accuracy, how a phone survey participant will respond to any question, such as whether the respondent owns a motorcycle or has electricity in the household (Fig. 1B and table S1). Cross-validated area-under-the-curve (AUC) scores—which indicate the probability that the model will rank a randomly chosen positive response higher than a randomly chosen negative one—range from 0.50 (no better than random) to 0.88 (quite effective). An analogous method can

be used to accurately identify the individuals in the sample who are living below a relative poverty threshold (AUC = 0.72 to 0.81) (Fig. 1C). With further refinement, such methods could prove useful to policy-makers and organizations that target resources to the extreme poor (25) (supplementary materials section 6).

For each of these prediction tasks, we use the two-step procedure to select a different model with different metrics and parameters. Although not the focus of our analysis, we note discernible patterns in the set of features identified as the best joint predictors of these different response variables. For instance, features related to an individual's patterns of mobility are generally predictive of motorcycle ownership, whereas factors related to an individual's position within his or her social network are more useful in predicting
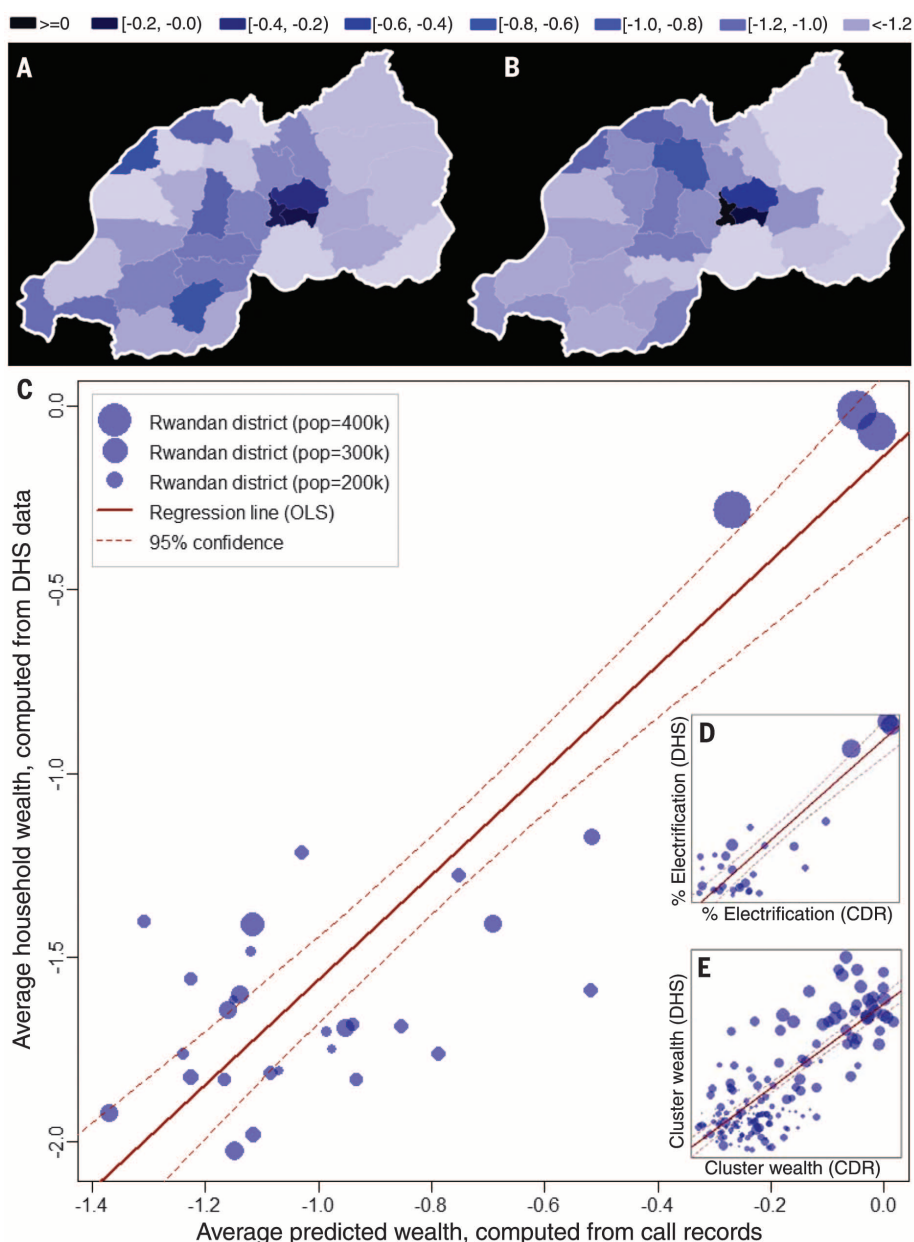
Fig. 3. Comparison of wealth predictions to government survey data. (**A**) Predicted composite wealth index (district average), computed from 2009 call data and aggregated by administrative district. (**B**) Actual composite wealth index (district average), as computed from a 2010 government DHS of 12,792 households. (**C**) Comparison of actual and predicted district wealth, for each of the 30 districts, with dots sized by population. (**D**) Comparison of actual and predicted rates of electrification, for each of the 30 districts. (**E**) Comparison of actual and predicted cluster wealth, for each of the 492 DHS clusters. CDR, call detail records.

poverty and wealth (fig. S3). These results suggest that our approach might be generalized to predict a broader class of survey responses, such as the subjective opinions and perceptions of mobile subscribers.

Having fit and cross-validated the model on the phone survey sample—a sample drawn to be representative of all active mobile phone users—we next generate out-of-sample predictions for the characteristics of the remaining 1.5 million Rwandan mobile phone users who did not participate in the survey. Combined with the rich geospatial markers in the phone data, the predicted attributes of millions of individual subscribers enable us to study the geographic distribution of subscriber wealth at an extremely fine degree of spatial granularity (Fig. 2). Whereas public data from Rwanda are only accurate at the level of the district (of which there are 30), the phone data can be used to infer characteristics of each of Rwanda's 2148 cells, as well as small microregions of just a few mobile subscribers (Fig. 2, bottom right inset).

The accuracy of these microregional wealth estimates cannot be directly verified, because no other data set provides wealth information with sufficient geographic resolution. However, when further aggregated to the district level, we can compare the distribution of wealth predicted from the call records of mobile subscribers (Fig. 3A) to the distribution of wealth measured with "ground truth" data collected by the Rwandan government (Fig. 3B). The former estimates are computed by averaging predicted wealth across the thousands of individual mobile phone–based predictions in each of Rwanda's 30 districts; the latter estimates are calculated using data from a nationally representative Demographic and Health Survey (DHS) of 12,792 households, conducted in person by the National Institute of Statistics of Rwanda (26). The strong correlation between these two predictions is evident in Fig. 3C and exists whether the ground truth is estimated from only those DHS households that report owning a mobile phone ($r = 0.917$) or from all households in the survey ($r = 0.916$). As we discuss in the supplementary materials (section 5A), the first correlation shows that the model's out-of-sample predictions are representative of the population of Rwandan mobile phone owners. The second correlation indicates that in countries like Rwanda, where patterns of mobile phone adoption are similar across regions, this method can provide a close approximation of the distribution of wealth of the full national population. Similar results are obtained when the analysis is disaggregated to the level of the DHS "cluster" ($r = 0.79$) (Fig. 3E), a geographic unit designed to be comparable to a village. These strong correlations are partially driven by the stark differences between urban and rural areas in Rwanda, but the correlations persist even when comparing clusters within urban or rural areas (fig. S6).

This same approach can be used to predict more than just the average wealth of a district. For instance, rates of district electrification estimated from phone records are comparable to those reported in the DHS survey ($r = 0.93$) (Fig. 3D). In the urban capital of Kigali, we also find a correlation ($r = 0.58$) between satellite estimates of night light intensity in 0.55-km^2 grid cells (fig. S7B) and the predicted distribution—based on phone data and the methods described earlier—of responses to the question "Does your household have electricity?" (fig. S7C).

How might such methods be used in practice? In addition to small-area estimation, one promising application is as a source of low-cost, interim national statistics. In many developing economies, long lag times typically occur between successive national surveys. In Angola, for instance, the most recent census before 2014 was conducted in 1970. In that 44-year period, the official population grew by more than 400%. Rwanda has better resources for data collection, and the DHS preceding the 2010 DHS was conducted in 2007. However, even in that relatively short period, the distribution of wealth in Rwanda shifted slightly. Thus, we find that the 2010 distribution of wealth is more accurately reflected in projections based on our analysis of phone data from 2009 than in estimates based on the 2007 DHS (fig. S8). This implies that a policy-maker tasked with targeting the poorest districts in Rwanda would obtain more accurate information from estimates based on mobile phone data than from estimates based on 2007 DHS data (supplementary materials section 6A).

In developing economies, where traditional sources of population data are scarce but mobile phones are increasingly common, these methods may provide a cost-effective option for measuring population characteristics. Whereas a typical national household survey costs more than $1 million and requires 12 to 18 months to complete (27), the phone survey we conducted cost only $12,000 and took 4 weeks to administer. Looking forward, the greatest challenge to such work lies in identifying protocols that enable analysis of similar data while respecting the privacy of individual subscribers and the commercial concerns of mobile operators (28, 29). With careful consideration, however, many compelling (and some speculative) applications are within reach, including population monitoring in remote and inaccessible regions, real-time policy evaluation, and the targeting of resources to those with the greatest need.

REFERENCES AND NOTES

1. S. Kuznets, *Am. Econ. Rev.* **45**, 1–28 (1955).
2. G. S. Fields, *World Bank Res. Obs.* **4**, 167–185 (1989).
3. M. Jerven, *Poor Numbers: How We Are Misled by African Development Statistics and What to Do About It* (Cornell Univ. Press, Ithaca, NY, 2013).
4. C. Elbers, J. O. Lanjouw, P. Lanjouw, *Econometrica* **71**, 355–364 (2003).
5. M. Ghosh, J. N. K. Rao, *Stat. Sci.* **9**, 55–76 (1994).
6. D. Lazer *et al.*, *Science* **323**, 721–723 (2009).
7. G. King, *Science* **331**, 719–721 (2011).
8. N. Eagle, M. Macy, R. Claxton, *Science* **328**, 1029–1031 (2010).
9. H. Choi, H. Varian, *Econ. Rec.* **88**, 2–9 (2012).
10. W. Wang, D. Rothschild, S. Goel, A. Gelman, *Int. J. Forecast.* **31**, 980–991 (2015).
11. "The mobile economy 2014" (GSMA Intelligence, 2014); www.gsmamobileeconomy.com/GSMA_ME_Report_2014_R2_WEB.pdf.
12. J. Candia *et al.*, *J. Phys. A* **41**, 224015 (2008).
13. J.-P. Onnela *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 7332–7336 (2007).
14. G. Palla, A. L. Barabási, T. Vicsek, *Nature* **446**, 664–667 (2007).
15. M. C. González, C. A. Hidalgo, A.-L. Barabási, *Nature* **453**, 779–782 (2008).
16. X. Lu, E. Wetter, N. Bharti, A. J. Tatem, L. Bengtsson, *Sci. Rep.* **3**, 2923 (2013).
17. J. E. Blumenstock, *Inf. Technol. Dev.* **18**, 107–125 (2012).
18. V. Frias-Martinez, J. Virseda, in *Proceedings of the Fifth International Conference on Information and Communication Technologies and Development* (Association for Computing Machinery, New York, 2012), pp. 76–84; http://doi.acm.org/10.1145/2160673.2160684.
19. P. Deville *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15888–15893 (2014).
20. G. C. Cawley, N. L. C. Talbot, *J. Mach. Learn. Res.* **11**, 2079–2107 (2010).
21. D. Filmer, L. H. Pritchett, *Demography* **38**, 115–132 (2001).
22. J. Blumenstock, N. Eagle, *Inf. Technol. Int. Dev.* **8**, 1–16 (2012).
23. H. Zou, T. Hastie, *J. R. Stat. Soc. Ser. B* **67**, 301–320 (2005).
24. L. Breiman, J. Friedman, C. J. Stone, R. A. Olshen, *Classification and Regression Trees* (Chapman and Hall/CRC Press, New York, ed. 1, 1984).
25. B. Abelson, K. R. Varshney, J. Sun, in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Association for Computing Machinery, New York, 2014), pp. 1563–1572; http://doi.acm.org/10.1145/2623330.2623335.
26. National Institute of Statistics of Rwanda (NISR) [Rwanda], Ministry of Health (MOH) [Rwanda], ICF International, "Rwanda Demographic and Health Survey 2010," *DHS Final Reports* (publication ID FR259, NISR, MOH, and ICF International, Calverton, MD, 2012).
27. M. Jerven, "Benefits and costs of the data for development targets for the post-2015 development agenda," in *Data for Development Assessment Paper* (Copenhagen Consensus Center, 2014).
28. Y.-A. de Montjoye, L. Radaelli, V. K. Singh, A. S. Pentland, *Science* **347**, 536–539 (2015).
29. A. Wesolowski *et al.*, *PLOS Curr.* 10.1371/currents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e (2014).

EVOLUTIONARY BIOLOGY

A pharyngeal jaw evolutionary innovation facilitated extinction in Lake Victoria cichlids

Matthew D. McGee,[1,2,3]* Samuel R. Borstein,[4] Russell Y. Neches,[1] Heinz H. Buescher,[5] Ole Seehausen,[2,3] Peter C. Wainwright[1]

Evolutionary innovations, traits that give species access to previously unoccupied niches, may promote speciation and adaptive radiation. Here, we show that such innovations can also result in competitive inferiority and extinction. We present evidence that the modified pharyngeal jaws of cichlid fishes and several marine fish lineages, a classic example of evolutionary innovation, are not universally beneficial. A large-scale analysis of dietary evolution across marine fish lineages reveals that the innovation compromises access to energy-rich predator niches. We show that this competitive inferiority shaped the adaptive radiation of cichlids in Lake Tanganyika and played a pivotal and previously unrecognized role in the mass extinction of cichlid fishes in Lake Victoria after Nile perch invasion.

Evolutionary innovations are adaptive traits that allow a lineage to cross a functional barrier and gain access to new niches (1). They are often framed as "key innovations" that can promote rapid diversification in the groups that evolve them (2, 3), and the search for key innovations has become a major component of modern macroevolutionary studies (4, 5). However, despite the obvious importance of evolutionary innovations in the history of life on Earth, innovative traits rarely show a direct link with increased diversification (6–15).

Evolutionary innovations are also traditionally thought to reduce extinction rates (2), but this may not be the case if innovation facilitates the evolution of specialist phenotypes sensitive to ecological disturbance (16, 17). Innovation may also exhibit niche-specific effects on extinction rates if the innovative trait involves a performance trade-off (13). Specifically, performance may increase in new niches at the cost of competitive exclusion and eventual extirpation from previously accessible niches.

We examined the potential cost of evolutionary innovation by using a classic example: pharyngognathy (18). Pharyngognathy involves multiple modifications of the jaw apparatus in the back of the throat that allow a fish to generate high bite force, which likely enables pharyngognathous fishes to exploit hard-shelled and processing-intensive prey items (19). However, these modifications reduce pharyngeal gape, which may alter the maximum size of prey that can be easily swallowed (20).

Several lineages within the spiny-finned fishes have independently evolved pharyngognathy, including wrasses, surfperches, damselfish, marine halfbeaks, flyingfishes, and cichlids (20). Most of these lineages live in shallow marine habitats, except for cichlids, which occur mostly in tropical freshwaters. Cichlids are especially well known for their tendency to undergo rapid speciation and accumulate exceptionally large species richness in spatially confined assemblages, particularly in Lakes Victoria, Malawi, and Tanganyika in eastern Africa (21, 22).

Each of these lineages has interacted with nonpharyngognathous spiny-finned lineages in different ways. In marine habitats, pharyngognathous lineages such as wrasses, parrotfishes and damselfishes have existed alongside closely related nonpharyngognathous spiny-finned fishes for tens of millions of years (20). In Lakes Victoria and Malawi, cichlids initially radiated in the complete absence of any nonpharyngognathous spiny-finned fish lineages. Unfortunately, in the 1950s, a nonpharyngognathous predatory fish, the Nile perch, *Lates niloticus*, was introduced into Lake Victoria, facilitating a cichlid mass extinction (23). In Lake Tanganyika, which hosts an older cichlid radiation than Victoria and Malawi, nonpharyngognathous *Lates* species and the pharyngognathous cichlids coexist, albeit with many fewer cichlid species and a lower speciation rate than the other two radiations (22, 24).

Comparative dietary data reveal that pharyngognathy has ecological consequences for the marine lineages that possess the trait. Unlike cichlids, which can sometimes evolve into predatory niches free from competition with predators like Nile perch, marine pharyngognaths always occur alongside typical nonpharyngognathous fish-eating lineages (20). We surveyed diet data across a phylogeny of marine spiny-finned fishes, including four marine transitions to pharyngognathy as well as other spiny-finned species occurring in the same environments as those four lineages, and measured rates of dietary evolution for fish and processing-intensive prey like plants and hard-shelled animals.
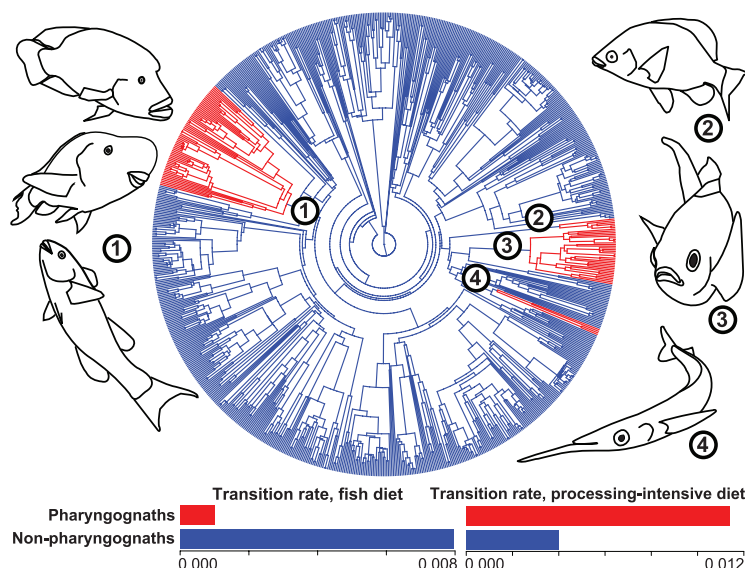
Fig. 1. Pharyngognathy affects dietary transitions in marine fishes. (**A**) Four transitions to pharyngognathy on a time-calibrated phylogeny of 851 marine spiny-finned fishes: (1) labroid fishes, including wrasses (Labridae), parrotfish (Scaridae), and weed whitings (Odacidae); (2) surfperches (Embiotocidae); (3) damselfishes (Pomacentridae); (4) marine halfbeaks (Hemirhamphidae). (**B**) Comparison of the transition rate for nonpharyngognathous and pharyngognathous fishes for fish prey and processing-intensive prey.

[1]Department of Evolution and Ecology and Center for Population Biology, University of California, Davis, CA 95616, USA. [2]Institute of Ecology and Evolution, University of Bern, CH-3012 Bern, Switzerland. [3]Department of Fish Ecology and Evolution, Eawag, Swiss Federal Institute for Aquatic Science and Technology, CH-6047 Kastanienbaum, Switzerland. [4]Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN 37996, USA. [5]Zoological Institute, University of Basel, CH-4051 Basel, Switzerland.
*Corresponding author. E-mail: mcgee.matthew@gmail.com

We analyzed diet both as a continuous character using Brownian motion and as a categorical variable using stochastic character mapping (*25*). In both cases, we examined whether fishes with pharyngognathy had different rates of dietary evolution or rates of transition to a specialist diet for fish prey and processing-intensive prey. Pharyngognathous marine fishes evolved into niches favoring processing-intensive prey items at a much higher rate than other spiny-finned fishes (Fig. 1). However, pharyngognaths evolved into fish-eating niches more slowly, suggesting that the evolution of the innovation may compromise access to this niche.

To assess the impact of pharyngognathous predators on competition with nonpharyngognathous predators, we investigated feeding performance and functional morphology of Nile perch and cichlids. We measured pharyngeal gape in Nile perch, which possess unfused pharyngeal jaws typical of nonpharyngognathous spiny-finned fishes, as well as in every major lineage of fish-eating cichlid (*25*). We found that pharyngognathy reduced cichlid pharyngeal gape to half that of Nile perch (Fig. 2A). The only exception to this pattern was in the South American genus *Cichla*, an old fish-eating cichlid lineage that has independently

lost pharyngognathy via loss of fusion of the lower pharyngeal jaw.

Feeding experiments indicate that pharyngognathy drastically increases handling time in cichlid predators relative to Nile perch. We measured handling time (*25*) in four predatory Lake Victoria cichlids and similarly-sized Nile perch by using fish prey of sizes and shapes comparable to those consumed in the wild by both groups (*26*). Cichlids were considerably slower, often taking many hours to process a prey item that a Nile perch could swallow in a few minutes (Fig. 2B). If processing time is examined with respect to pharyngeal gape (*25*), the difference between Nile perch and cichlids disappears, suggesting that the narrower pharyngeal gape of the cichlids is the primary cause of their long prey-processing times (*25*). Our results here are limited to predatory cichlids and *Lates*, but we suggest that similar analyses across marine lineages are likely to be of great interest for understanding the role of pharyngognathy in marine ecosystems.

If pharyngognathy hinders cichlid feeding performance when processing fish prey, fish-eating cichlids may have been particularly disadvantaged after Nile perch were introduced into Lake Victoria. We used conditional inference forests with

corrections for correlated variables to explore how ecological variables predict extinction in Victorian cichlids (*25*). A fish diet is the most important predictor of extinction (Fig. 3A), suggesting that competition played an important role in addition to known factors like predation (*23*) and eutrophication (*27, 28*).

Of the major functional morphological traits associated with the radiation, a large lower jaw length shows the strongest association with a fish diet in Victorian cichlids (*21, 25*). We reveal a large morphological shift in this character when comparing all fish-eating Victorian cichlids to a representative sample of fish-eaters in the relict fauna of Lake Victoria after Nile perch invasion (Fig. 3B). The preextinction fish-eater community was highly diverse and species rich, with many species possessing jaws of equal or greater size to Nile perch and often consuming large prey (*26*). However, the few relict fish-eating individuals collected postextinction all have a less predatory morphology than was typical for predatory cichlids of this radiation before the extinction events. The relict Victorian cichlid species now more closely resemble the less-extreme fish-eaters from Lake Tanganyika, where Nile perch and cichlids have coexisted for millions of years.

Fig. 2. Cichlids exhibit reduced pharyngeal gape and increased handling times relative to Nile perch. (**A**) Pharyngeal gape comparison of Nile perch, which possesses typical pharyngeal jaws, and fish-eating cichlids, including the one known loss of pharyngognathy in cichlids (genus *Cichla*). From top, *Lates niloticus*, *Harpagochromis* sp. "orange rock hunter," *Pyxichromis orthostoma*, *Harpagochromis cf serranus*, *Harpagochromis* sp. "two stripe white lip," *Lipochromis* sp. "matumbi hunter," *Lipochromis parvidens*, *Champsochromis caeruleus*, *Nimbochromis* sp., *Rhamphochromis longiceps*, *Boulengerochromis microlepis*, *Bathybates minor*, *Lepidiolamprologus profundicola*, *Cyphotilapia frontosa*, *Cichla ocellaris*, *Parachromis* sp., *Petenia splendida*. (**B**) Handling time comparison between Nile perch and four species of fish-eating Victorian cichlids with respect to the ratio of prey length:predator length. Colors as in (A).
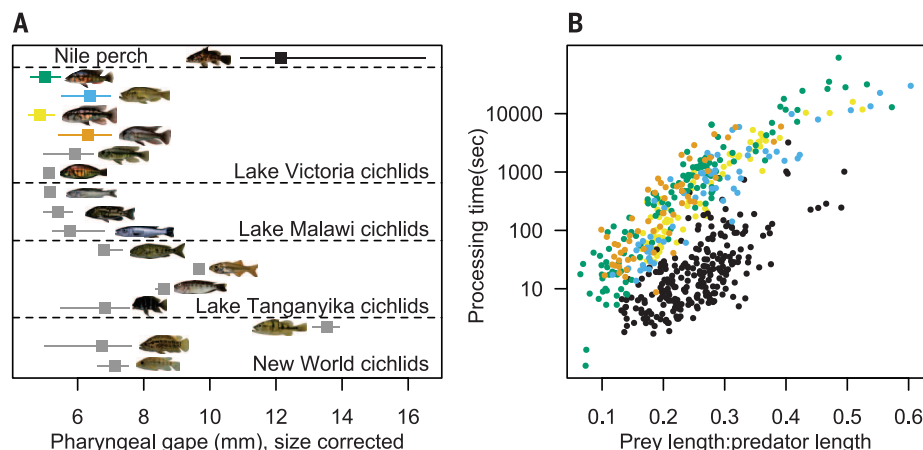
Fig. 3. Extinction of fish-eating Lake Victoria cichlids. (**A**) Relative importance of the four highest ecological variables predicting extinction in Victorian cichlids for the original 1992 data set (*23*) and an updated one (*25*). Pluses indicate increased risk; minuses indicate reduced risk. AUC, area under the curve. (**B**) Beeswarm plot of size-corrected lower-jaw length of Nile perch (NP), preinvasion Victorian cichlid fish-eaters (LV), postinvasion relict fish eaters (LVrelict), and eight transitions to fish-eating in Lake Tanganyika (LT) cichlids. Large bars indicate mean jaw length.
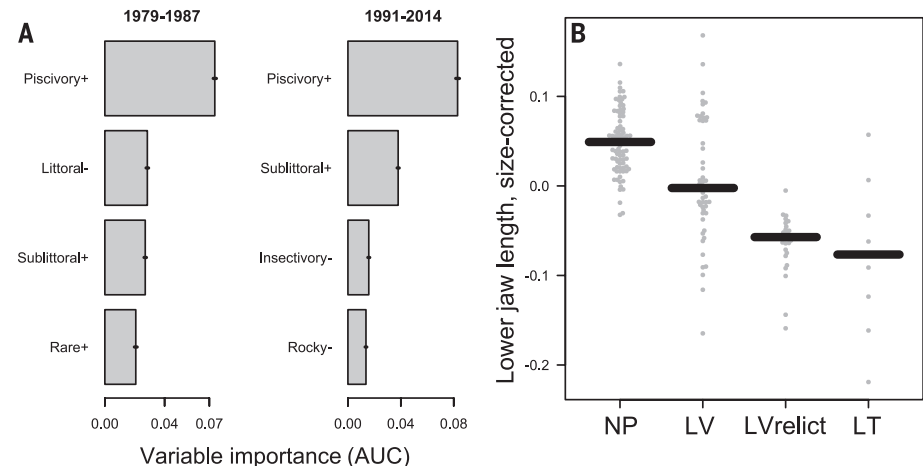
Interspecific competition is thought to be a pervasive force in evolution (29, 30), and we suggest that the pattern we observe across Lakes Victoria and Tanganyika is likely due to competition between Nile perch and cichlid predators.

For nearly half a century, the robust pharyngeal jaws of cichlids, wrasses, and other pharyngognathous fishes have been considered a classic example of evolutionary innovation that opened up new niches through increased trophic flexibility (18). Although this is almost certainly correct, our results suggest that the innovation involves a major trade-off that severely limits the size of prey that can be eaten, facilitating competitive inferiority in predatory niches and extinction in the presence of a predatory invader lacking the innovation. The evolutionary innovation of pharyngognathy is not a uniformly beneficial trait, but a specialization that can promote competitive exclusion and extinction depending on ecological context and community composition.

REFERENCES AND NOTES

1. E. Mayr, *Animal Species and Evolution* (Harvard Univ. Press Cambridge, MA, 1963).
2. S. B. Heard, D. L. Hauser, *Hist. Biol.* **10**, 151–173 (1995).
3. J. P. Hunter, *Trends Ecol. Evol.* **13**, 31–36 (1998).
4. R. Maia, D. R. Rubenstein, M. D. Shawkey, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10687–10692 (2013).
5. D. L. Rabosky, *PLOS ONE* **9**, e89543 (2014).
6. J. Cracraft, in *Evolutionary Innovations*, M. H. Nitecki, D. V. Nitecki, Eds. (Univ. of Chicago Press, Chicago, 1990), pp. 21–44.
7. G. J. Vermeij, *Biol. J. Linn. Soc. Lond.* **72**, 461–508 (2001).
8. G. J. Vermeij, *Paleobiology* **33**, 469–493 (2007).
9. G. J. Vermeij, *Evol. Ecol.* **26**, 357–373 (2012).
10. M. E. Alfaro, C. D. Brock, B. L. Banbury, P. C. Wainwright, *BMC Evol. Biol.* **9**, 255 (2009).
11. T. J. Givnish et al., *Evolution* **54**, 1915–1937 (2000).
12. S. A. Hodges, M. L. Arnold, *Proc. Biol. Sci.* **262**, 343–348 (1995).
13. D. Schluter, *The Ecology of Adaptive Radiation* (Oxford Univ. Press, Oxford, 2000).
14. D. Brawand et al., *Nature* **513**, 375–381 (2014).
15. C. Mitter, B. Farrell, B. Wiegmann, *Am. Nat.* **132**, 107–128 (1988).
16. D. J. Futuyma, G. Moreno, *Annu. Rev. Ecol. Syst.* **19**, 207–233 (1988).
17. T. J. Givnish, in *Evolution on Islands*, P. Grant, Ed. (Oxford Univ. Press, Oxford, 1998), pp. 281–304.
18. K. F. Liem, *Syst. Biol.* **22**, 425–441 (1973).
19. P. C. Wainwright, *J. Zool.* **213**, 283–297 (1987).
20. P. C. Wainwright et al., *Syst. Biol.* **61**, 1001–1027 (2012).
21. P. H. Greenwood, *The Haplochromine Fishes of the East African Lakes: Collected Papers on Their Taxonomy, Biology and Evolution* (Kraus International Publications, Munich, 1981).
22. O. Seehausen, *Proc. Biol. Sci.* **273**, 1987–1998 (2006).
23. F. Witte et al., *Environ. Biol. Fishes* **34**, 1–28 (1992).
24. G. W. Coulter, J.-J. Tiercelin, *Lake Tanganyika and Its Life* (Oxford Univ. Press, Oxford, 1991).
25. Materials and methods are available as supplementary materials on *Science* Online.
26. M. J. P. Van Oijen, *Neth. J. Zool.* **32**, 336–363 (1981).
27. O. Seehausen, J. J. Van Alphen, F. Witte, *Science* **277**, 1808–1811 (1997).
28. J. C. van Rijssel, F. Witte, *Evol. Ecol.* **27**, 253–267 (2013).
29. D. W. Pfennig, K. S. Pfennig, *Evolution's Wedge: Competition and the Origins of Diversity* (Univ. of California Press, CA, 2012).
30. D. L. Rabosky, *Annu. Rev. Ecol. Evol. Syst.* **44**, 481–502 (2013).

CANCER IMMUNOTHERAPY

Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota

Marie Vétizou,[1,2,3] Jonathan M. Pitt,[1,2,3] Romain Daillère,[1,2,3] Patricia Lepage,[4] Nadine Waldschmitt,[5] Caroline Flament,[1,2,6] Sylvie Rusakiewicz,[1,2,6] Bertrand Routy,[1,2,3,6] Maria P. Roberti,[1,2,6] Connie P. M. Duong,[1,2,6] Vichnou Poirier-Colame,[1,2,6] Antoine Roux,[1,2,7] Sonia Becharef,[1,2,6] Silvia Formenti,[8] Encouse Golden,[8] Sascha Cording,[9] Gerard Eberl,[9] Andreas Schlitzer,[10] Florent Ginhoux,[10] Sridhar Mani,[11] Takahiro Yamazaki,[1,2,6] Nicolas Jacquelot,[1,2,3] David P. Enot,[1,7,12] Marion Bérard,[13] Jérôme Nigou,[14,15] Paule Opolon,[1] Alexander Eggermont,[1,2,16] Paul-Louis Woerther,[17] Elisabeth Chachaty,[17] Nathalie Chaput,[1,18] Caroline Robert,[1,16,19] Christina Mateus,[1,16] Guido Kroemer,[7,12,20,21,22] Didier Raoult,[23] Ivo Gomperts Boneca,[24,25]* Franck Carbonnel,[3,26]* Mathias Chamaillard,[5]* Laurence Zitvogel[1,2,3,6]†

Antibodies targeting CTLA-4 have been successfully used as cancer immunotherapy. We find that the antitumor effects of CTLA-4 blockade depend on distinct *Bacteroides* species. In mice and patients, T cell responses specific for *B. thetaiotaomicron* or *B. fragilis* were associated with the efficacy of CTLA-4 blockade. Tumors in antibiotic-treated or germ-free mice did not respond to CTLA blockade. This defect was overcome by gavage with *B. fragilis*, by immunization with *B. fragilis* polysaccharides, or by adoptive transfer of *B. fragilis*–specific T cells. Fecal microbial transplantation from humans to mice confirmed that treatment of melanoma patients with antibodies against CTLA-4 favored the outgrowth of *B. fragilis* with anticancer properties. This study reveals a key role for *Bacteroidales* in the immunostimulatory effects of CTLA-4 blockade.

Ipilimumab is a fully human monoclonal antibody (Ab) directed against CTLA-4, a major negative regulator of T cell activation (1), approved in 2011 for improving the overall survival of patients with metastatic melanoma (MM) (2). However, blockade of CTLA-4 by ipilimumab often results in immune-related adverse events at sites that are exposed to commensal microorganisms, mostly the gut (3). Patients treated with ipilimumab develop Abs to components of the enteric flora (4). Therefore, given our previous findings for other cancer therapies (5), addressing the role of gut microbiota in the immunomodulatory effects of CTLA-4 blockade is crucial for the future development of immune checkpoint blockers in oncology.

We compared the relative therapeutic efficacy of the CTLA-4–specific 9D9 Ab against established MCA205 sarcomas in mice housed in specific pathogen–free (SPF) versus germ-free (GF) conditions. Tumor progression was controlled by Ab against CTLA-4 in SPF but not in GF mice (Fig. 1, A and B). Moreover, a combination of broad-spectrum antibiotics [ampicillin + colistin + streptomycin (ACS)] (Fig. 1C), as well as imipenem alone (but not colistin) (Fig. 1C), compromised the antitumor effects of CTLA-4–specific Ab. These results, which suggest that the gut microbiota is required for the anticancer effects of CTLA-4 blockade, were confirmed in the Ret melanoma and the MC38 colon cancer models (fig. S1, A and B). In addition, in GF or ACS-treated mice, activation of splenic effector CD4+ T cells and tumor-infiltrating lymphocytes (TILs) induced by Ab against CTLA-4 was significantly decreased (Fig. 1, D and E, and fig. S1, C to E).

We next addressed the impact of the gut microbiota on the incidence and severity of intestinal lesions induced by CTLA-4 Ab treatment. A "subclinical colitis" dependent on the gut microbiota was observed at late time points (figs. S2 to S5). However, shortly (by 24 hours) after the first administration of CTLA-4 Ab, we observed increased cell death and proliferation of intestinal epithelial cells (IECs) residing in the ileum and colon, as shown by immunohistochemistry using Ab-cleaved caspase-3 and Ki67 Ab, respectively (Fig. 2A and fig. S6A). The CTLA-4 Ab–induced IEC proliferation

was absent in RegIIIβ-deficient mice (fig. S6). Concomitantly, the transcription levels of *Il17a, Ifng, Ido1*, type 1 *Ifn*-related gene products and *Ctla4* (but not *Il6*), which indicate ongoing inflammatory processes, significantly increased by 24 hours in the distal ileum of CTLA-4 Ab–treated mice (fig. S6, B to D). Depletion of T cells, including intraepithelial lymphocytes (IELs) (by injection of Abs specific for CD4 and CD8), abolished the induction of IEC apoptosis by CTLA-4–specific Ab (Fig. 2A). When crypt-derived three-dimensional small intestinal enteroids (6) were exposed to Toll-like receptor (TLR) agonists (which act as microbial ligands in this assay) and subsequently admixed with IELs harvested from mice treated with Ab against CTLA-4 (but not isotype Ctl), IECs within the enteroids underwent apoptosis (Fig. 2B). Hence, CTLA-4 Ab compromises the homeostatic IEC-IEL equilibrium, favoring the apoptotic demise of IEC in the presence of microbial products.

To explore whether this T cell–dependent IEC death could induce perturbations of the microbiota composition, we performed high-throughput pyrosequencing of 16S ribosomal RNA (rRNA) gene amplicons of feces. The principal component analysis indicated that a single injection of CTLA-4 Ab sufficed to significantly affect the microbiome at the genus level (Fig. 2C). CTLA-4 blockade induced a rapid underrepresentation of both *Bacteroidales* and *Burkholderiales*, with a relative increase of *Clostridiales*, in

feces (Fig. 2C and table S1). Quantitative polymerase chain reaction (QPCR) analyses targeting the *Bacteroides* genus and species (spp.) in small intestine mucosa and feces contents showed a trend toward a decreased relative abundance of such bacteria in the feces, which contrasted with a relative enrichment in particular species [such as *B. thetaiotaomicron* (*Bt*) and *B. uniformis*] in the small intestine mucosa 24 to 48 hours after one CTLA-4 Ab injection (Fig. 2D and fig. S7). One of the most regulatory *Bacteroides* isolates, *B. fragilis* (*Bf*) (7–10), was detectable by PCR in colon mucosae

but was not significantly increased with CTLA-4 Ab (fig. S7).

Next, to establish a cause-and-effect relationship between the dominance of distinct *Bacteroides* spp. in the small intestine and the anticancer efficacy of CTLA-4 blockade, we recolonized ACS-treated and GF mice with several bacterial species associated with CTLA-4 Ab–treated intestinal mucosae as well as *Bf*. ACS-treated mice orally fed with *Bt, Bf, Burkholderia cepacia* (*Bc*), or the combination of *Bf* and *Bc*, recovered the anticancer response to CTLA-4 Ab, contrasting with all the other isolates that failed

[affiliations list omitted]
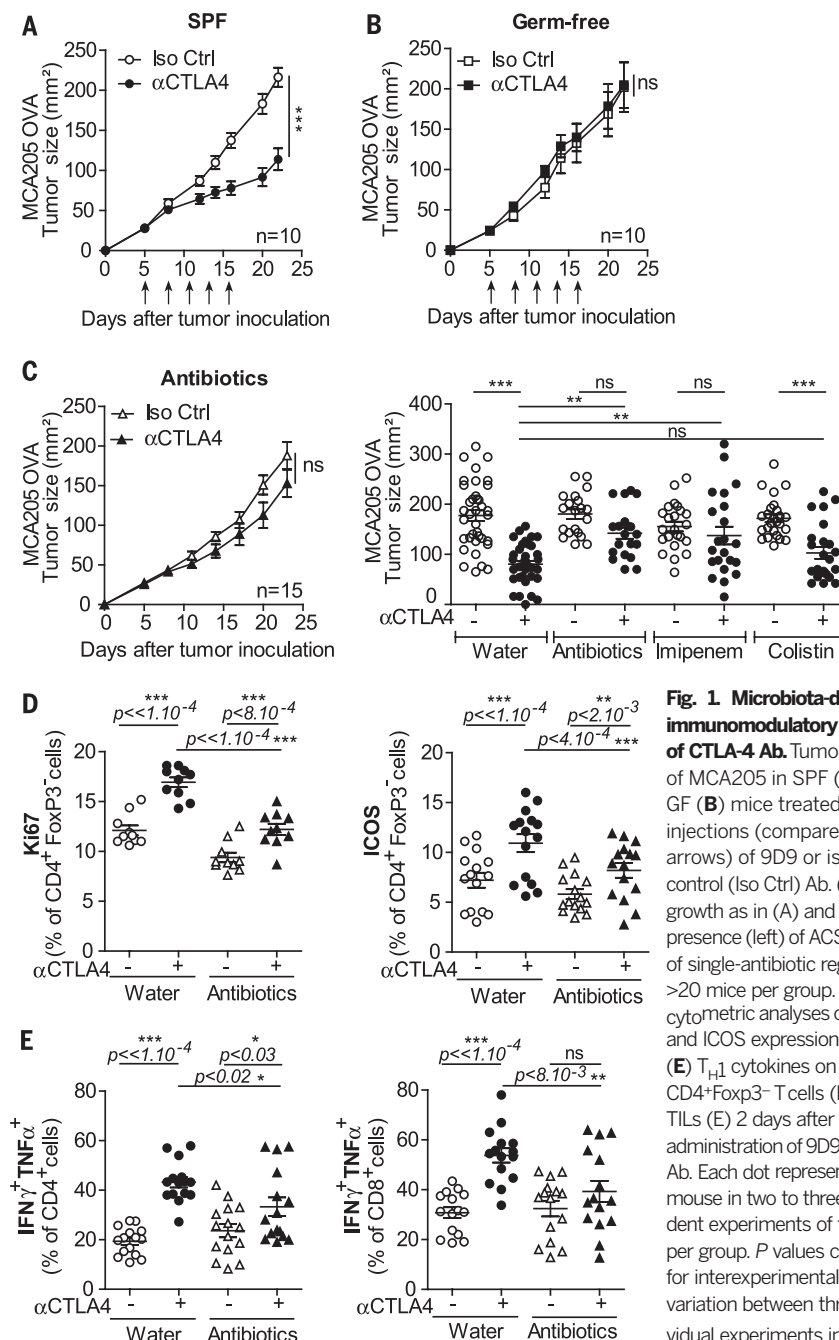
Fig. 1. Microbiota-dependent immunomodulatory effects of CTLA-4 Ab.

to do so (table S2 and Fig. 3A). Similarly, oral feeding with *Bf*, which colonized the mucosal layer of GF mice (fig. S8) (*11*), induced T helper 1 (T$_H$1) immune responses in the tumor-draining lymph nodes and promoted the maturation of intratumoral dendritic cells (DCs), which culminated in the restoration of the therapeutic response of GF tumor bearers to CTLA-4 Ab (Fig. 3B and fig. S9, A and B).

We analyzed the dynamics of memory T cell responses directed against distinct bacterial species in mice and humans during CTLA-4 blockade. CD4$^+$ T cells harvested from spleens of CTLA-4 Ab–treated mice (Fig. 3C) or from blood taken from individuals with MM or non–small cell lung carcinoma (NSCLC) patients after two administrations of ipilimumab (Fig. 3, D and E, and table S3) tended to recover a T$_H$1 phenotype (figs. S10 and S11). The functional relevance of such T cell responses for the anticancer activity

of CTLA-4 Ab was further demonstrated by the adoptive transfer of memory *Bf*-specific (but not *B. distasonis*-specific) T$_H$1 cells into GF or ACS-treated tumor bearers (Fig. 3F and fig. S12), which partially restored the efficacy of the immune checkpoint blocker.

The microbiota-dependent immunostimulatory effects induced by CTLA-4 blockade depended on the mobilization of lamina propria CD11b$^+$ DC that can process zwitterionic polysaccharides (*9*) and then mount interleukin-12 (IL-12)–dependent cognate T$_H$1 immune responses against *Bf* capsular polysaccharides (figs. S13 and S14). However, they did not appear to result from TLR2/TLR4-mediated innate signaling (*7, 8*) in the context of a compromised gut tolerance (figs. S15 to S19).

To address the clinical relevance of these findings, we analyzed the composition of the gut microbiome before and after treatment with

ipilimumab in 25 individuals with MM (table S4). A clustering algorithm based on genus composition of the stools (*12, 13*) distinguished three clusters (Fig. 4A and table S5) with *Alloprevotella* or *Prevotella* driving cluster A and distinct *Bacteroides* spp. driving clusters B and C (Fig. 4B). During ipilimumab therapy, the proportions of MM patients falling into cluster C increased, at the expense of those belonging to cluster B (Fig. 4B and fig. S20A). We next performed fecal microbial transplantation of feces harvested from different MM patients from each cluster, 2 weeks before tumor inoculation into GF mice that were subsequently treated with anti–CTLA-4 Ab. Tumors growing in mice that had been transplanted with feces from cluster C patients markedly responded to CTLA-4 blockade, contrasting with absent anticancer effects in mice transplanted with cluster B–related feces (Fig. 4C). QPCR analyses revealed
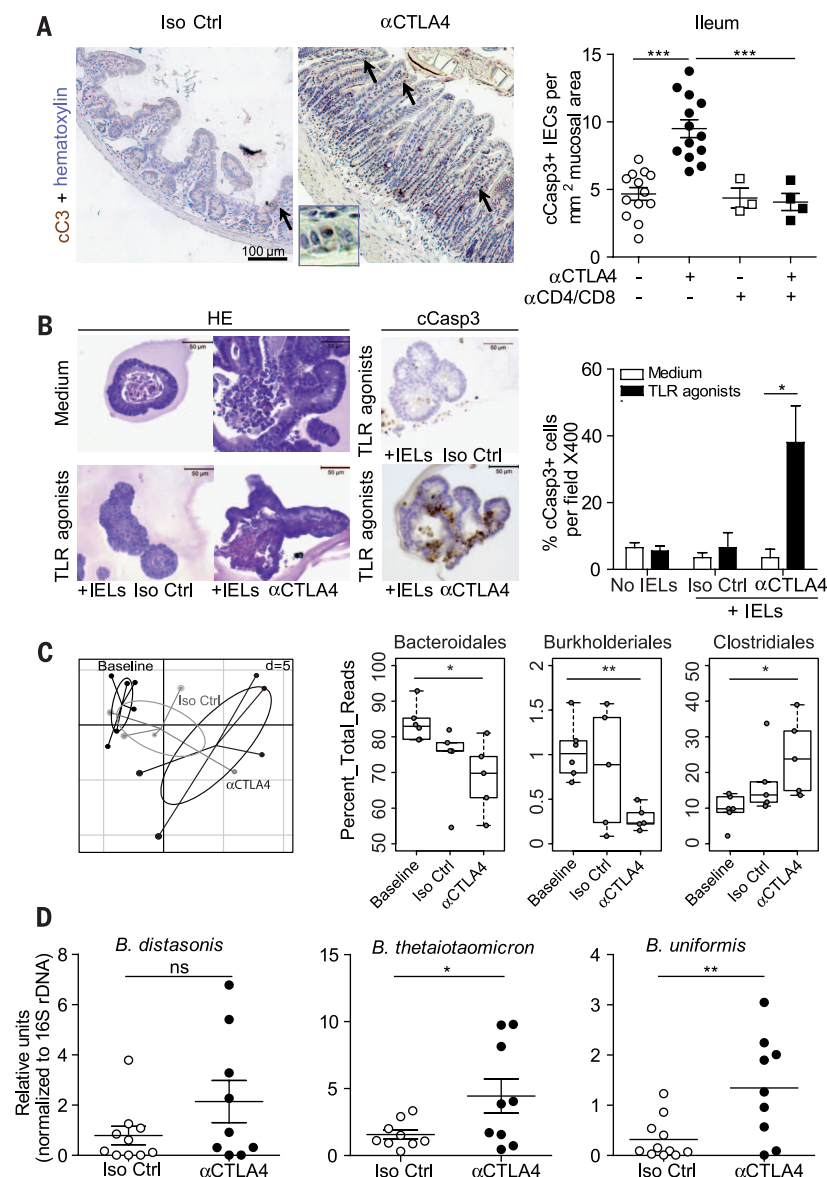
Fig. 2. IEC-IEL dialogue causes IEC apoptosis and intestinal dysbiosis after CTLA-4 Ab injection. (**A**) (left) Representative micrograph pictures of distal ileum after staining with Ab-cleaved caspase 3 (cCasp3) Ab 24 hours after one injection of 9D9 (or Iso Ctrl) Ab in naïve mice with or without prior depletion of CD4$^+$ and CD8$^+$ T cells. Inset enlarged 18-fold. (Right) Concatanated data of two experiments. (**B**) (left) Representative micrographs of 3D enteroid cocultures stimulated (or not) with TLR agonists and incubated with IELs harvested from 9D9 (or Iso Ctrl) Ab–treated mice in hematoxylin and eosin (H&E), then (middle) stained with cCasp3-specific Ab. (Right) Data concatenated from two experiments counting the means ± SEM percentages of apoptotic cells to organoid in 20 organoids. (**C**) Sequencing of 16S rRNA gene amplicons of feces from tumor bearers before and 48 hours after one administration of 9D9 or Iso Ctrl Ab. (Left) Principal component analysis (PCA) on a relative abundance matrix of genus repartition highlighting the clustering between baseline, Iso Ctrl Ab–, and 9D9 Ab–treated animals after one injection (five to six mice per group). Ellipses are presented around the centroids of the resulting three clusters. The first two components explain 34.41% of total variance (Component 1: 20.04%; Component 2: 14.35%) (Monte-Carlo test with 1000 replicates, *P* = 0.0049). (C) (right) Means ± SEM of relative abundance for each three orders for five mice per group are shown. (**D**) QPCR analyses targeting three distinct *Bacteroides* spp. in ileal mucosae performed 24 to 48 hours after Ab introduction. Results are represented as 2$^{-\Delta\Delta Ct}$ × 10^3, normalized to 16S rDNA and to the basal time point (before treatment). Each dot represents one mouse in two gathered experiments. *P* < 0.05; **P* < 0.01; ***P* < 0.001; ns, not significant.

Fig. 3. Memory T cell responses against *Bt* and *Bf* and anticancer efficacy of CTLA-4 blockade. (A and **B)** Tumoricidal effects of *Bf*, *Bt*, and/or *B. cepacia* (*Bc*) administered by oral feeding of ACS-treated or GF mice (also refer to fig. S8A). (A) (left) Tumor sizes at day 15 after 9D9 or Iso Ctrl Ab treatment are depicted. Each dot represents one tumor, and graphs depict two to three experiments of five mice per group. (Middle) Histopathological score of colonic mucosae in ACS-treated tumor bearers receiving 9D9 Ab after oral gavage with various bacterial strains, assessed on H&E-stained colons monitoring microscopic lesions as described in materials and methods at day 20 after treatment in five animals per group on at least six independent areas. (Right) Representative micrographs are shown; scale bar, 100 µm. (B) Tumoricidal effects of *Bf* in GF mice as indicated. (**C** to **E**) Recall responses of CD4[+] T cells in mice and patients to various bacterial strains after CTLA-4 blockade. DCs loaded with bacteria of the indicated strain were incubated with CD4[+] T cells, 2 days after three intraperitoneal (ip) CTLA-4 Ab in mice, and after at least two injections of ipilimumab (ipi) in patients. The graphs represent interferon-γ (IFN-γ) concentrations from coculture supernatants at 24 hours in mice (C) and 48 hours in MM patients (D). (E) IFN-γ/IL-10 ratios were monitored in DC–T cell cocultures of NSCLC patients at 48 hours. No cytokine release was observed in the absence of bacteria or T cells (fig. S11 with HV). Each dot represents one patient or mouse. Paired analyses are represented by linking dots pre- and post-ipi. (**F**) T cells harvested from spleens of mice exposed to CTLA-4 Ab and restimulated with *Bf* versus *B. distasonis* or bone marrow DCs alone (CD4[+] NT) were infused intravenously in day 6 MCA205 tumor-bearing GF mice. A representative experiment containing five to six mice per group is shown. *P < 0.05; **P < 0.01; ***P < 0.001; ns, not significant.
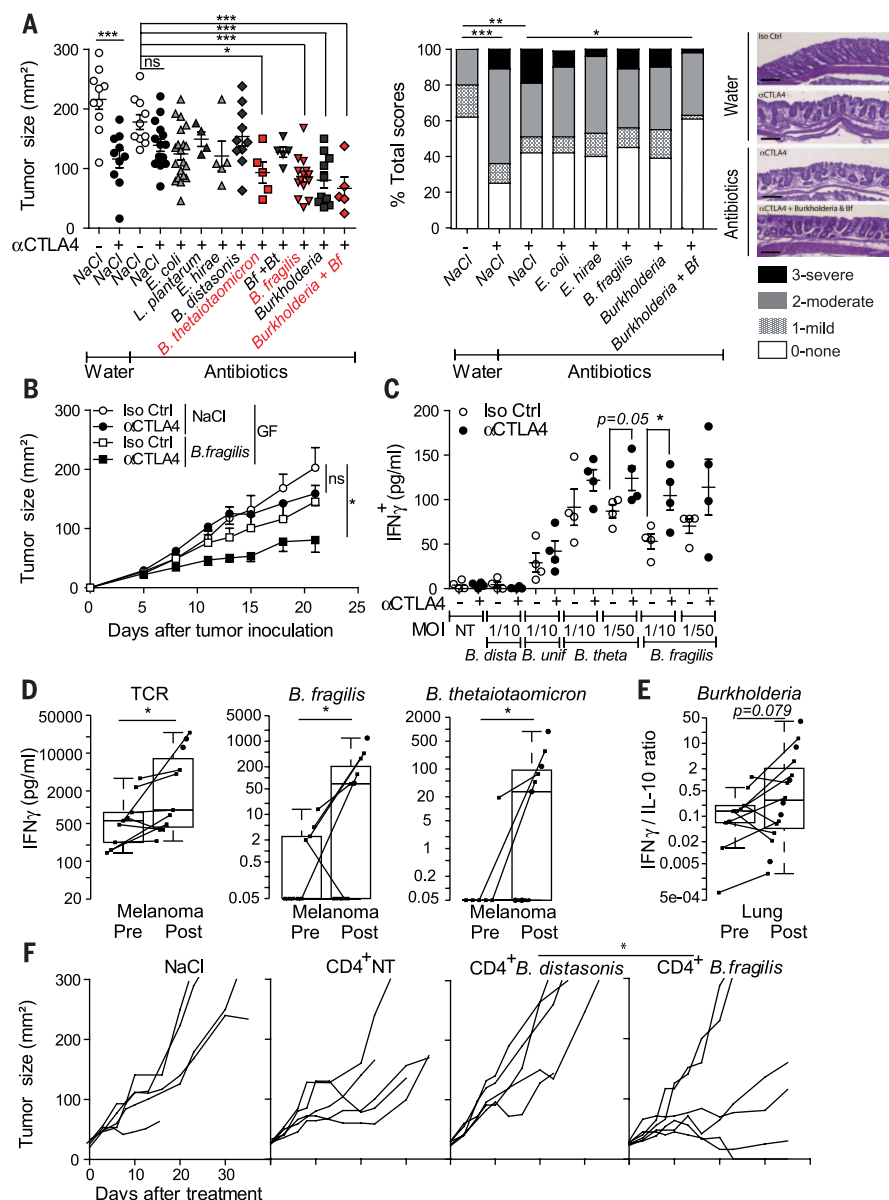
that, although bacteria from the *Bacteroidales* order equally colonized the recipient murine intestine, stools from cluster C (but not A or B) individuals specifically facilitated the colonization of the immunogenic bacteria *Bf* and *Bt* (7–10, 14, 15) (Fig. 4D). Moreover, after CTLA-4 Ab therapy, only cluster C (not A or B) recipient mice had outgrowth of *Bf* (fig. S20B). Note that the fecal abundance of *Bf* (but not *B. distasonis* or *B. uniformis*) negatively correlated with tumor size after CTLA-4 blockade in cluster C–recipient mice (Fig. 4E and fig. S20C). Hence, ipilimumab can modify the abundance of immunogenic *Bacteroides* spp. in the gut, which in turn affects its anticancer efficacy.

Finally, intestinal reconstitution of ACS-treated animals with the combination of *Bf* and *Bc* did not increase but rather reduced histopathological signs of colitis induced by CTLA-4 blockade (Fig. 3A). This efficacy-toxicity uncoupling effect was not

achieved with vancomycin, which could boost the antitumor effects of CTLA-4 blockade (presumably by inducing the overrepresentation of *Bacteroidales* at the expense of *Clostridiales*) but worsened the histopathological score (fig. S21). In support of this notion, *Bf* maintained its regulatory properties in the context of CTLA-4 blockade (fig. S22) (7).

Hence, the efficacy of CTLA-4 blockade is influenced by the microbiota composition (*B. fragilis* and/or *B. thetaiotaomicron* and *Burkholderiales*). The microbiota composition affects interleukin 12 (IL-12)–dependent T_H1 immune responses, which facilitate tumor control in mice and patients while sparing intestinal integrity. In accord with previous findings (16), colitis (observed in the context of IL-10 deficiency and CTLA-4 blockade) (fig. S17) could even antagonize anticancer efficacy. Several factors may dictate why such commensals could be suitable "anticancer

probiotics." The geodistribution of *Bf* in the mucosal layer of the intestine (fig. S8) and its association with *Burkholderiales*—recognized through the pyrin–caspase-1 inflammasome (17) and synergizing with TLR2/TLR4 signaling pathways (fig. S15)—may account for the immunomodulatory effects of CTLA-4 Ab. Future investigations will determine whether a potential molecular mimicry between distinct commensals and/or pathobionts and tumor neoantigens could account for the toxicity and/or efficacy of immune checkpoint blockers. Prospective studies in MM and/or NSCLC may validate the relevance of the enterotypes described herein in the long-term efficacy of immune checkpoint blockers, with the aim of compensating cluster B–driven patients with live and immunogenic or recombinant *Bacteroides* spp. (18) or fecal microbial transplantation from cluster C–associated stools to improve their antitumor immune responses.
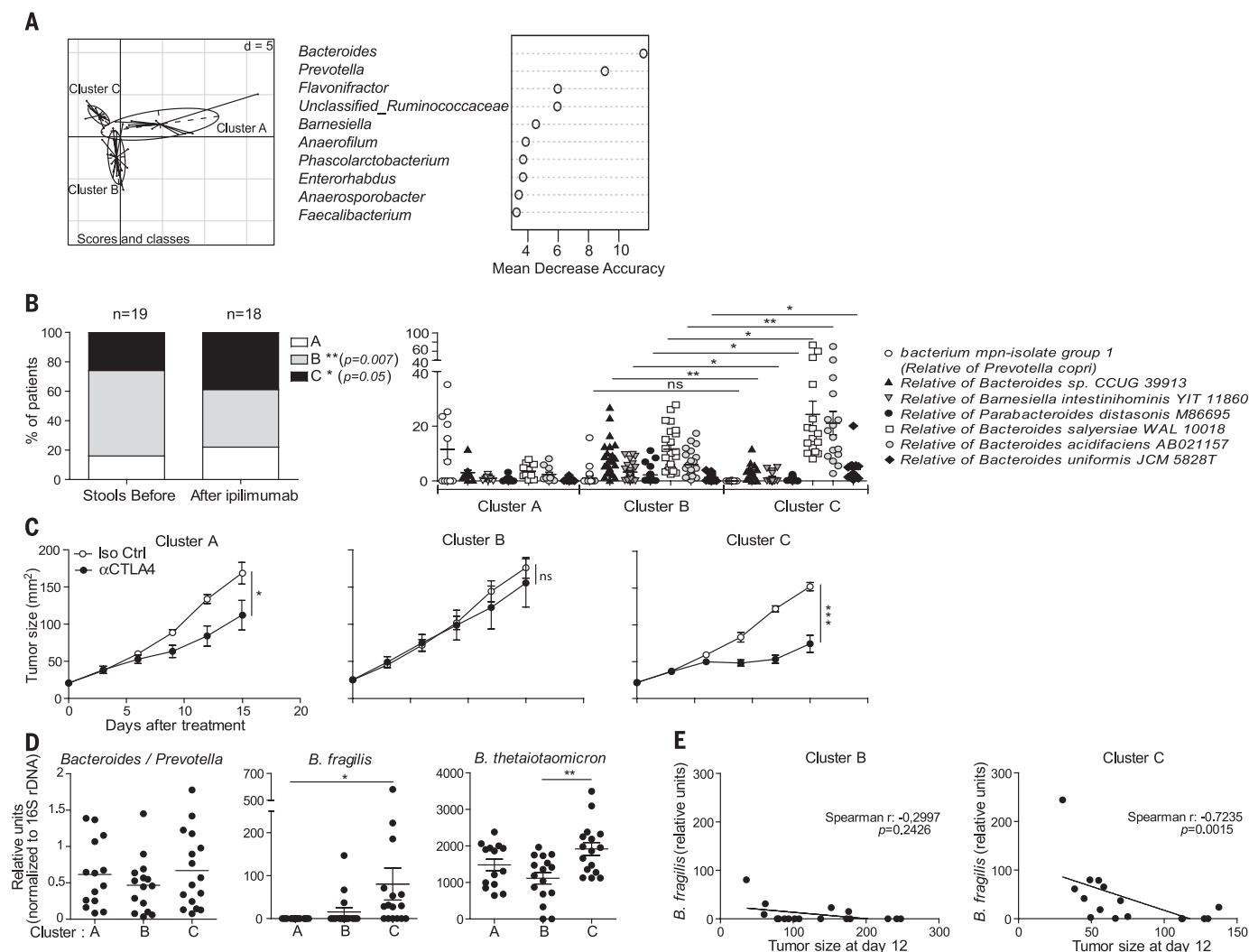
Fig. 4. Biological significance of ipilimumab-induced dysbiosis in patients. The k means clustering algorithm was applied on the basis of genus composition before and during ipilimumab treatment in 25 MM patients, validated using the Calinski-Harabasz index (14), and showed good performance in recovering three clusters before and after therapy (interclass PCA); (**A**) (left) (Monte-Carlo test, $P = 0.000199$). (A) (right) Random Forest analysis was applied to decipher the main genera responsible for this significant clustering. (**B**) (right) The relative abundance of main *Bacteroides* spp. significantly differed between clusters B and C. (B) (left) The proportions of patients falling into each cluster were analyzed in a nonpaired manner before versus after ipi injections regardless of the time point (fig. S20A). (**C**) Fecal microbial transplantation after introduction of ipilimumab from eight patients falling into each of the three clusters (stool selection for fecal microbial transplantation marked with an asterisk * in fig. S20A) into GF animals. One representative experiment out of three is shown with means ± SEM of tumor sizes depicted for each cluster over time. (**D**) QPCR analyses of feces DNA of the recipient before (2 weeks postcolonization) and 2 weeks after ipi, targeting *Bacteroidales* and *Bacteroides* spp. Results are represented as $2^{-\Delta Ct}$ x 10^3, normalized to 16S rDNA. No significant difference in the relative abundance of *Bf* was detectable in the donors of cluster B versus C before colonization (not shown). (**E**) Spearman correlations between the amount of *Bf* in stools 15 days after treatment with 9D9 Ab and tumor sizes across cluster B- and C-recipient mice. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

REFERENCES AND NOTES

1. K. S. Peggs, S. A. Quezada, A. J. Korman, J. P. Allison, *Curr. Opin. Immunol.* **18**, 206–213 (2006).
2. F. S. Hodi *et al.*, *N. Engl. J. Med.* **363**, 711–723 (2010).
3. K. E. Beck *et al.*, *J. Clin. Oncol.* **24**, 2283–2289 (2006).
4. D. Berman *et al.*, *Cancer Immun.* **10**, 11 (2010).
5. S. Viaud *et al.*, *Science* **342**, 971–976 (2013).
6. A. Rogoz, B. S. Reis, R. A. Karssemeijer, D. Mucida, *J. Immunol. Methods* **421**, 89–95 (2015).
7. S. Dasgupta, D. Erturk-Hasdemir, J. Ochoa-Reparaz, H. C. Reinecker, D. L. Kasper, *Cell Host Microbe* **15**, 413–423 (2014).
8. S. K. Mazmanian, C. H. Liu, A. O. Tzianabos, D. L. Kasper, *Cell* **122**, 107–118 (2005).
9. F. Stingele *et al.*, *J. Immunol.* **172**, 1483–1490 (2004).
10. A. O. Tzianabos *et al.*, *J. Biol. Chem.* **267**, 18230–18235 (1992).
11. J. Y. Huang, S. M. Lee, S. K. Mazmanian, *Anaerobe* **17**, 137–141 (2011).
12. M. Arumugam *et al.*, *Nature* **473**, 174–180 (2011).
13. J. Qin *et al.*, *Nature* **464**, 59–65 (2010).
14. A. Cebula *et al.*, *Nature* **497**, 258–262 (2013).
15. J. L. Sonnenburg, C. T. Chen, J. I. Gordon, *PLOS Biol.* **4**, e413 (2006).
16. W. Lam *et al.*, *Sci. Transl. Med.* **2**, 45ra59 (2010).
17. H. Xu *et al.*, *Nature* **513**, 237–241 (2014).
18. M. Mimee, A. C. Tucker, C. A. Voigt, T. K. Lu, *Cell Systems* **1**, 62–71 (2015).

CANCER IMMUNOTHERAPY

Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti–PD-L1 efficacy

Ayelet Sivan,[1]* Leticia Corrales,[1]* Nathaniel Hubert,[2] Jason B. Williams,[1] Keston Aquino-Michaels,[3] Zachary M. Earley,[2] Franco W. Benyamin,[1] Yuk Man Lei,[2] Bana Jabri,[2] Maria-Luisa Alegre,[2] Eugene B. Chang,[2] Thomas F. Gajewski[1,2]†

T cell infiltration of solid tumors is associated with favorable patient outcomes, yet the mechanisms underlying variable immune responses between individuals are not well understood. One possible modulator could be the intestinal microbiota. We compared melanoma growth in mice harboring distinct commensal microbiota and observed differences in spontaneous antitumor immunity, which were eliminated upon cohousing or after fecal transfer. Sequencing of the 16S ribosomal RNA identified *Bifidobacterium* as associated with the antitumor effects. Oral administration of *Bifidobacterium* alone improved tumor control to the same degree as programmed cell death protein 1 ligand 1 (PD-L1)–specific antibody therapy (checkpoint blockade), and combination treatment nearly abolished tumor outgrowth. Augmented dendritic cell function leading to enhanced CD8+ T cell priming and accumulation in the tumor microenvironment mediated the effect. Our data suggest that manipulating the microbiota may modulate cancer immunotherapy.

Harnessing the host immune system constitutes a promising cancer therapeutic because of its potential to specifically target tumor cells although limiting harm to normal tissue. Enthusiasm has been fueled by recent clinical success, particularly with antibodies that block immune inhibitory pathways, specifically CTLA-4 and the axis between programmed cell death protein 1 (PD-1) and its ligand 1 (PD-L1) (*1*, *2*). Clinical responses to these immunotherapies are more frequent in patients who show evidence of an endogenous T cell response ongoing in the tumor microenvironment before therapy (*3–6*). However, the mechanisms that govern the presence or absence of this phenotype are not well understood. Theoretical sources of interpatient heterogeneity include host germline genetic differences, variability in patterns of somatic alterations in tumor cells, and environmental differences.

The gut microbiota plays an important role in shaping systemic immune responses (*7–9*). In the cancer context, a role for intestinal microbiota in mediating immune activation in response to chemotherapeutic agents has been demonstrated (*10*, *11*). However, it is not known whether commensal microbiota influence spontaneous immune responses against tumors and thereby affect the therapeutic activity of immunotherapeutic interventions, such as anti–PD-1/PD-L1 monoclonal antibodies (mAbs).

To address this question, we compared subcutaneous B16.SIY melanoma growth in genetically similar C57BL/6 mice derived from two different mouse facilities, Jackson Laboratory (JAX) and Taconic Farms (TAC), which have been shown to differ in their commensal microbes (*12*). We found that JAX and TAC mice exhibited significant differences in B16.SIY melanoma growth rate, with tumors growing more aggressively in TAC mice (Fig. 1A). This difference was immune-mediated: Tumor-specific T cell responses (Fig. 1, B and C) and intratumoral CD8+ T cell accumulation (Fig. 1D) were significantly higher in JAX than in TAC mice. To begin to address whether this difference could be mediated by commensal microbiota, we cohoused JAX and TAC mice before tumor implantation. We found that cohousing ablated the differences in tumor growth (Fig. 1E) and immune responses (Fig. 1, F to H) between the two mouse populations, which suggested an environmental influence. Cohoused TAC

and JAX mice appeared to acquire the JAX phenotype, which suggested that JAX mice may be colonized by commensal microbes that dominantly facilitate antitumor immunity.

To directly test the role of commensal bacteria in regulating antitumor immunity, we transferred JAX or TAC fecal suspensions into TAC and JAX recipients by oral gavage before tumor implantation (fig. S1A). We found that prophylactic transfer of JAX fecal material, but not saline or TAC fecal material, into TAC recipients was sufficient to delay tumor growth (Fig. 2A) and to enhance induction and infiltration of tumor-specific CD8+ T cells (Fig. 2, B and C, and fig. S1B), which supported a microbe-derived effect. Reciprocal transfer of TAC fecal material into JAX recipients had a minimal effect on tumor growth rate and antitumor T cell responses (Fig. 2, A to C, and fig. S1B), consistent with the JAX-dominant effects observed upon cohousing.

To test whether manipulation of the microbial community could be effective as a therapy, we administered JAX fecal material alone or in combination with antibodies targeting PD-L1 (αPD-L1) to TAC mice bearing established tumors. Transfer of JAX fecal material alone resulted in significantly slower tumor growth (Fig. 2D), accompanied by increased tumor-specific T cell responses (Fig. 2E) and infiltration of antigen-specific T cells into the tumor (Fig. 2F), to the same degree as treatment with systemic αPD-L1 mAb. Combination treatment with both JAX fecal transfer and αPD-L1 mAb improved tumor control (Fig. 2D) and circulating tumor antigen–specific T cell responses (Fig. 2E), although there was little additive effect on accumulation of activated T cells within the tumor microenvironment (Fig. 2F). Consistent with these results, αPD-L1 therapy alone was significantly more efficacious in JAX mice compared with TAC mice (Fig. 2G), which paralleled improved antitumor T cell responses (fig. S1C). These data indicate that the commensal microbial composition can influence spontaneous antitumor immunity, as well as a response to immunotherapy with αPD-L1 mAb.

To identify specific bacteria associated with improved antitumor immune responses, we monitored the fecal bacterial content over time of mice that were subjected to administration of fecal permutations, using the 16S ribosomal RNA (rRNA) miSeq Illumina platform. Principal coordinate analysis revealed that fecal samples analyzed from TAC mice that received JAX fecal material gradually separated from samples obtained from sham- and TAC feces–inoculated TAC mice over time (*P* = 0.001 and *P* = 0.003, respectively, ANOSIM multivariate data analysis) and became similar

[1]Department of Pathology, University of Chicago, Chicago, IL 60637, USA. [2]Department of Medicine, University of Chicago, Chicago, IL 60637, USA. [3]Section of Genetic Medicine, University of Chicago, Chicago, IL 60637, USA.
*These authors contributed equally to this work. †Corresponding author. E-mail: tgajewsk@medicine.bsd.uchicago.edu

to samples obtained from sham- and JAX feces–inoculated JAX mice (Fig. 3A). In contrast, TAC-inoculated TAC mice did not change in community diversity relative to sham-inoculated TAC mice (P = 0.4, ANOSIM). Reciprocal transfer of TAC fecal material into JAX hosts resulted in a statistically significant change in community diversity (P = 0.003, ANOSIM), yet the distance of the microbial shift was smaller (Fig. 3A).

Comparative analysis showed that 257 taxa were of significantly different relative abundance in JAX mice relative to TAC mice [false discovery rate (FDR) < 0.05, nonparametric t test] (Fig. 3B and table S1). Members belonging to several of these groups were similarly altered in JAX-fed TAC mice relative to sham- or TAC-inoculated TAC mice (Fig. 3C and tables S1 and S2). To further identify functionally relevant bacterial taxa, we asked which genus-level taxa were significantly associated with accumulation of activated antigen-specific T cells within the tumor microenvironment across all permutations (Fig. 2C). The only significant association was *Bifidobacterium* (P =

5.7 × 10^{-5}, FDR = 0.0019, univariate regression) (table S3), which showed a positive association with antitumor T cell responses and increased in relative abundance over 400-fold in JAX-fed TAC mice (Fig. 3C). Stimulatory interactions between bifidobacteria and the host immune system, including those associated with interferon-γ (IFN-γ), have been described previously (*13–16*). We thus hypothesized that members of this genus could represent a major component of the beneficial antitumor immune effects observed in JAX mice.

At the sequence level, *Bifidobacterium* operational taxonomic unit OTU_681370 showed the largest increase in relative abundance in JAX-fed TAC mice (table S1) and the strongest association with antitumor T cell responses across all permutations (Fig. 3D and table S3). We further identified this bacterium as most similar to *B. breve*, *B. longum*, and *B. adolescentis* (99% identity). To test whether *Bifidobacterium* spp. may be sufficient to augment protective immunity against tumors, we obtained a commercially available

cocktail of *Bifidobacterium* species, which included *B. breve* and *B. longum* and administered this by oral gavage, alone or in combination with αPD-L1, to TAC recipients bearing established tumors. Analysis of fecal bacterial content revealed that the most significant change in response to *Bifidobacterium* inoculation occurred in the *Bifidobacterium* genus (P = 0.0009, FDR = 0.015, nonparametric t test), with a 120-fold increase in OTU_681370 (fig. S2A and table S4), which suggested that the commercial inoculum contained bacteria that were at least 97% identical to the taxon identified in JAX and JAX-fed TAC mice. An increase in *Bifidobacterium* could also be detected by quantitative polymerase chain reaction (PCR) (fig. S2B).

Bifidobacterium-treated mice displayed significantly improved tumor control in comparison with their non-*Bifidobacterium* treated counterparts (Fig. 3E), which was accompanied by robust induction of tumor-specific T cells in the periphery (Fig. 3F) and increased accumulation of antigen-specific CD8$^+$ T cells within the tumor

Fig. 1. Differences in melanoma outgrowth and tumor-specific immune responses between C57BL/6 JAX and TAC mice are eliminated when mice are cohoused. (**A**) B16.SIY tumor growth kinetics in newly arrived JAX and TAC mice. (**B**) IFN-γ enzyme-linked immunospot assay (ELISPOT) in tumor-bearing JAX and TAC mice 7 days after tumor inoculation. (**C**) Mean size of IFN-γ spots (10^{-3} mm^2). (**D**) Percentage of SIY$^+$ T cells of total CD8$^+$ T cells within the tumor of JAX and TAC mice as determined by flow cytometry 21 days after tumor inoculation. Representative plots (left), quantification (right). (**E**) B16.SIY tumor growth kinetics in JAX and TAC mice cohoused for 3 weeks before tumor inoculation. (**F**) Number of IFN-γ spots/10^6 splenocytes in tumor-bearing JAX and TAC mice cohoused

for 3 weeks before tumor inoculation. (**G**) Mean size of IFN-γ spots (10^{-3} mm^2). (**H**) Percentage of SIY$^+$ T cells of total CD8$^+$ T cells within the tumor of JAX and TAC mice cohoused for 3 weeks before tumor inoculation. Means ± SEM combined from six independent experiments, analyzed by two-way analysis of variance (ANOVA) with Sidak's correction for multiple comparisons (A) and (E), or individual mice with means ± SEM combined from four (B), (C), (F), (G) or three (D) and (H) independent experiments, analyzed by Student's t test; five mice per group per experiment; *P < 0.005, **P < 0.01; NS, not significant.
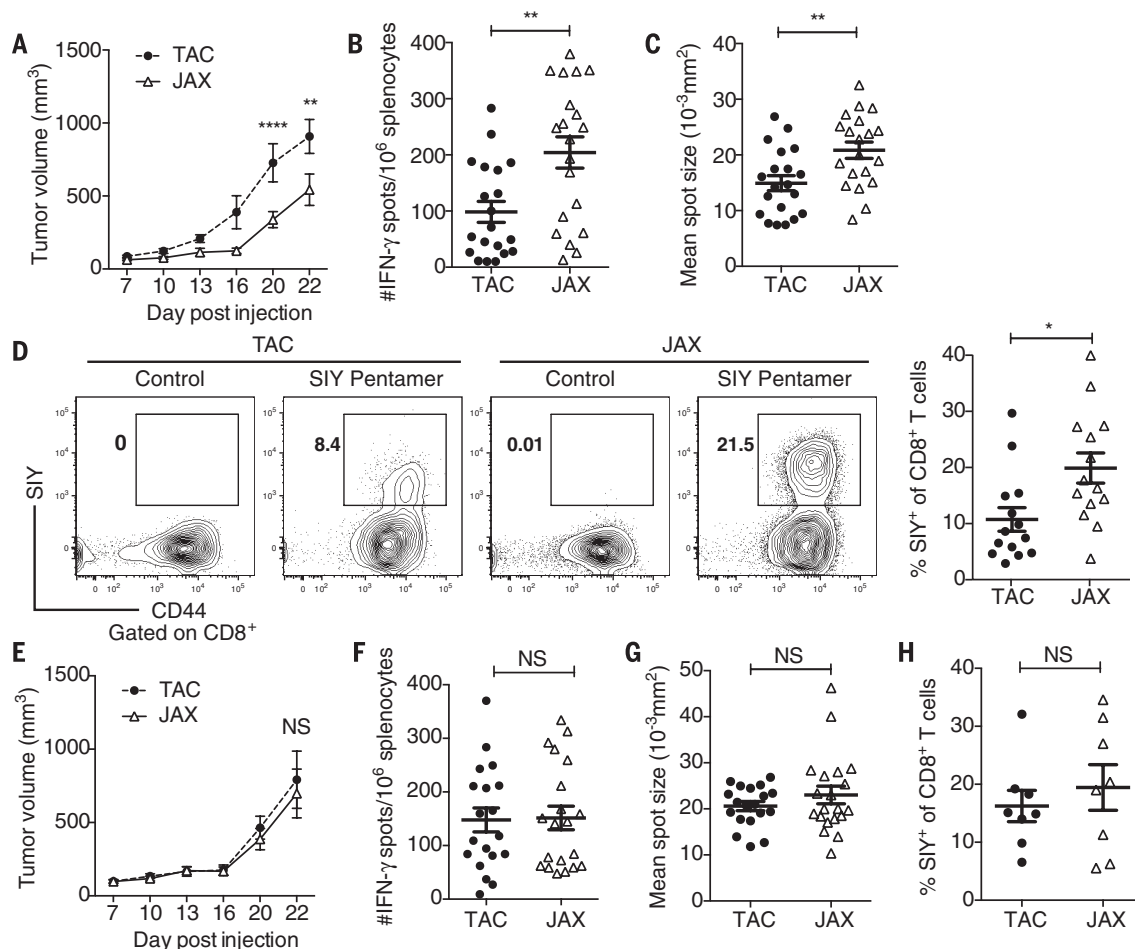
(Fig. 3G and fig. S2C). These effects lasted several weeks (fig. S2, D and E).

The therapeutic effect of *Bifidobacterium* feeding was abrogated in CD8-depleted mice (fig. S3A), which indicated that the mechanism was not direct but rather through host antitumor T cell responses. Heat inactivation of the bacteria before oral administration also abrogated the therapeutic effect on tumor growth and reduced tumor-specific T cell responses to baseline (fig. S3, B to D), which suggested that the antitumor effect requires live bacteria. As an alternative strategy, we tested the therapeutic effect of *B. breve* and *B. longum* strains obtained from the American Type Culture Collection, which also showed significantly improved tumor control (fig. S4A). Administration of *Bifidobacterium* to TAC mice inoculated

with B16 parental tumor cells or MB49 bladder cancer cells also resulted in delayed tumor outgrowth (fig. S4, B and C, respectively). Oral administration of *Lactobacillus murinus* to TAC mice, which was not among the overrepresented taxa in JAX-fed mice, had no effect on tumor growth (fig. S4D) or on tumor-specific T cell responses (fig. S4E), which suggested that modulation of antitumor immunity depends on the specific bacteria administered. Collectively, these data point to *Bifidobacterium* as a positive regulator of antitumor immunity in vivo.

Upon inoculation with *Bifidobacterium*, a small set of species were altered in parallel with *Bifidobacterium* (ANOSIM, $P = 0.003$) (fig. S5A and table S4), however, for the most part, they did not resemble the changes observed with JAX feces

administration. Although we observed reductions (~2- to 10-fold) in members of the order Clostridiales, as well as in butyrate-producing species, upon *Bifidobacterium* inoculation, which could point to an inhibitory effect on the regulatory T cell compartment (17–19), we did not observe any difference in the frequency of CD4+ Foxp3+ T cells in tumors isolated from JAX and TAC mice (fig. S5B). Thus, although we cannot definitively rule out an indirect effect, it is unlikely that *Bifidobacterium* is acting primarily through modulation of the abundance of other bacteria.

We next assessed whether translocation of *Bifidobacterium* was occurring into the mesenteric lymph nodes, spleen, or tumor; however, no *Bifidobacterium* was detected in any of the organs
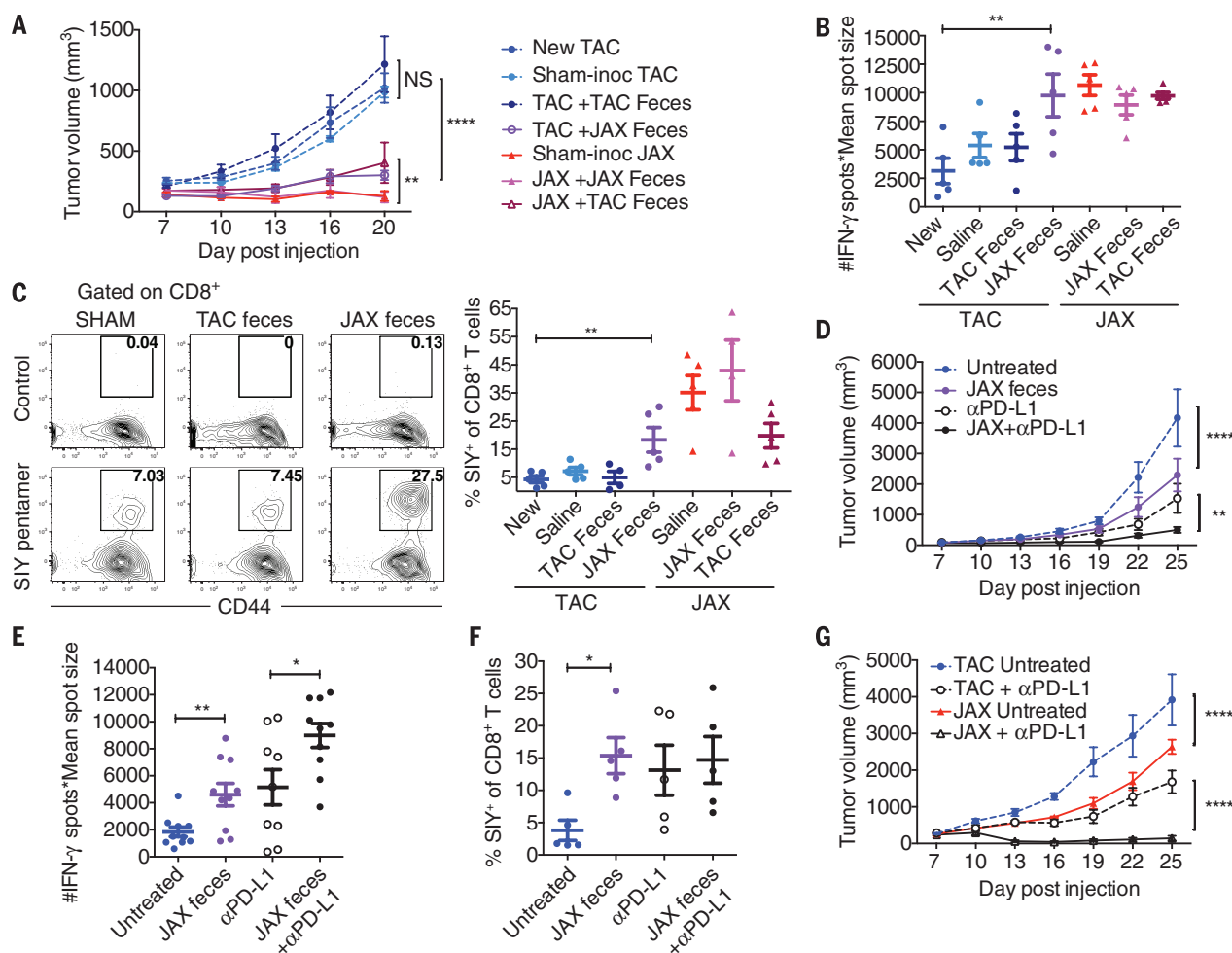
Fig. 2. Oral administration of JAX fecal material to TAC mice enhances spontaneous antitumor immunity and response to αPD-L1 mAb therapy. (**A**) B16.SIY tumor growth in newly arrived TAC mice, TAC and JAX mice orally gavaged with phosphate-buffered saline or TAC or JAX fecal material before tumor implantation. (**B**) Number of IFN-γ spots × mean spot size (10^{-3} mm²), determined by ELISPOT 7 days after tumor inoculation. (**C**) Percentage of SIY+ CD8+ T cells within the tumor of TAC and JAX mice treated as in (A), 21 days after tumor inoculation. Representative plots (left), quantification (right). (**D**) B16.SIY tumor growth in TAC mice, untreated or treated with JAX fecal material 7 and 14 days after tumor implantation, αPD-L1 mAb 7, 10, 13, and 16 days after tumor implantation, or both regimens. (**E**) IFN-γ ELISPOT as-

sessed 5 days after start of treatment. (**F**) Percentage of tumor-infiltrating SIY+ CD8+ T cells, determined by flow cytometry 14 days after start of treatment. (**G**) B16.SIY tumor growth kinetics in TAC and JAX mice, untreated or treated with αPD-L1 mAb 7, 10, 13, and 16 days after tumor implantation. Means ± SEM analyzed by two-way analysis of variance (ANOVA) with Dunnett's (A) or Tukey's (D) and (G) correction for multiple comparisons; or individual mice with means ± SEM analyzed by one-way ANOVA with Holm-Sidak correction for multiple comparisons (B), (C), (E), and (F); data are representative of (A) to (C), (F), and (G) or combined from (D) and (E) two to four independent experiments; five mice per group per experiment; *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$; NS, not significant.

Fig. 3. Direct administration of *Bifidobacterium* to TAC recipients with established tumors improves tumor-specific immunity and response to αPD-L1 mAb therapy. (A) Principal coordinate analysis plot of bacterial β-diversity over time in groups treated as in Fig. 2A, each group is made up of at least two cages, three or four mice per cage; data represent three independent experiments; **P < 0.01, ***P < 0.001 (ANOSIM). (**B**) Phylogenetic analysis of taxa that are of significantly different abundance in newly arrived JAX versus TAC mice FDR < 0.05 (nonparametric *t* test); bars represent log-transformed fold changes, inner circle, $\log_{10}(10)$; middle circle, $\log_{10}(100)$; outer circle, $\log_{10}(1000)$. (**C**) Heat map showing relative abundance over time of significantly altered genus-level taxa in JAX-fed TAC mice FDR < 0.05 (nonparametric *t* test); columns depict individual mice; each time point shows mice from two separate cages, three or four mice per cage. (**D**) Correlation plot of relative abundance of *Bifidobacterium* OTU_681370 in fecal material obtained from groups, as in (A), 14 days after arrival and frequency of SIY⁺ CD8⁺ T cells in tumor; $P = 1.4 \times 10^{-5}$, FDR =

0.0002, correlation $R^2 = 0.86$ (univariate regression). (**E**) B16.SIY tumor growth kinetics in TAC mice, untreated or treated with *Bifidobacterium* 7 and 14 days after tumor implantation, αPD-L1 mAb 7, 10, 13, and 16 days after tumor implantation, or both regimens. (**F**) IFN-γ ELISPOT assessed 5 days after start of treatment. (**G**) Percentage of tumor-infiltrating SIY⁺ CD8⁺ T cells, determined by flow cytometry 14 days after start of treatment. Means ± SEM analyzed by two-way ANOVA with Tukey's correction (E) or individual mice with means ± SEM analyzed by one-way ANOVA with Holm-Sidak correction (F) and (G), and are combined from two independent experiments; five mice per group per experiment: *P < 0.05, ***P < 0.001, ****P < 0.0001.

isolated from *Bifidobacterium*-gavaged tumor-bearing mice (fig. S5C). We thus concluded that the observed systemic immunological effects are likely occurring independently of bacterial translocation.

We subsequently interrogated the immunologic mechanisms underlying the observed differences in T cell responses between TAC, JAX, and *Bifidobacterium*-treated TAC mice (fig. S6A). CD8[+] SIY-specific 2C T cell receptor (TCR) Tg T cells exposed to tumors in JAX and *Bifidobacterium*-treated TAC mice exhibited greater expansion in the tumor-draining lymph node, as compared with their counterparts in TAC mice (fig. S6B).

However, they produced markedly greater IFN-γ in both the tumor-draining lymph node and the spleen of JAX and *Bifidobacterium*-fed TAC tumor-bearing mice (Fig. 4A), consistent with our analyses of the endogenous T cell response (Figs. 1C, 2E, and 3F). These data pointed to an improvement in immune responses upstream of T cells, at the level of host dendritic cells (DCs). Consistent with this hypothesis, we found an increased percentage of major histocompatibility complex (MHC) Class II[hi] DCs in the tumors of JAX and *Bifidobacterium*-treated TAC mice (Fig. 4B).

We therefore used genome-wide transcriptional profiling of early tumor-infiltrating DCs isolated from TAC, JAX, and *Bifidobacterium*-treated TAC mice (fig. S7A and table S5). Pathway analysis of 760 gene transcripts up-regulated in both JAX and *Bifidobacterium*-treated TAC-derived DCs relative to DCs from untreated TAC mice identified cytokine-cytokine receptor interaction, T cell activation, and positive regulation of mononuclear cell proliferation as significantly enriched pathways (Fig. 4C and fig. S7B). Many of these genes have been shown to be critical for antitumor responses, including those involved in CD8[+] T cell activation and costimulation [*H2-m2* (MHC-I), *Cd40*, *Cd70*, and *Icam1*] (20–22); DC maturation (*Relb* and *Ifngr2*) (23, 24); antigen processing and
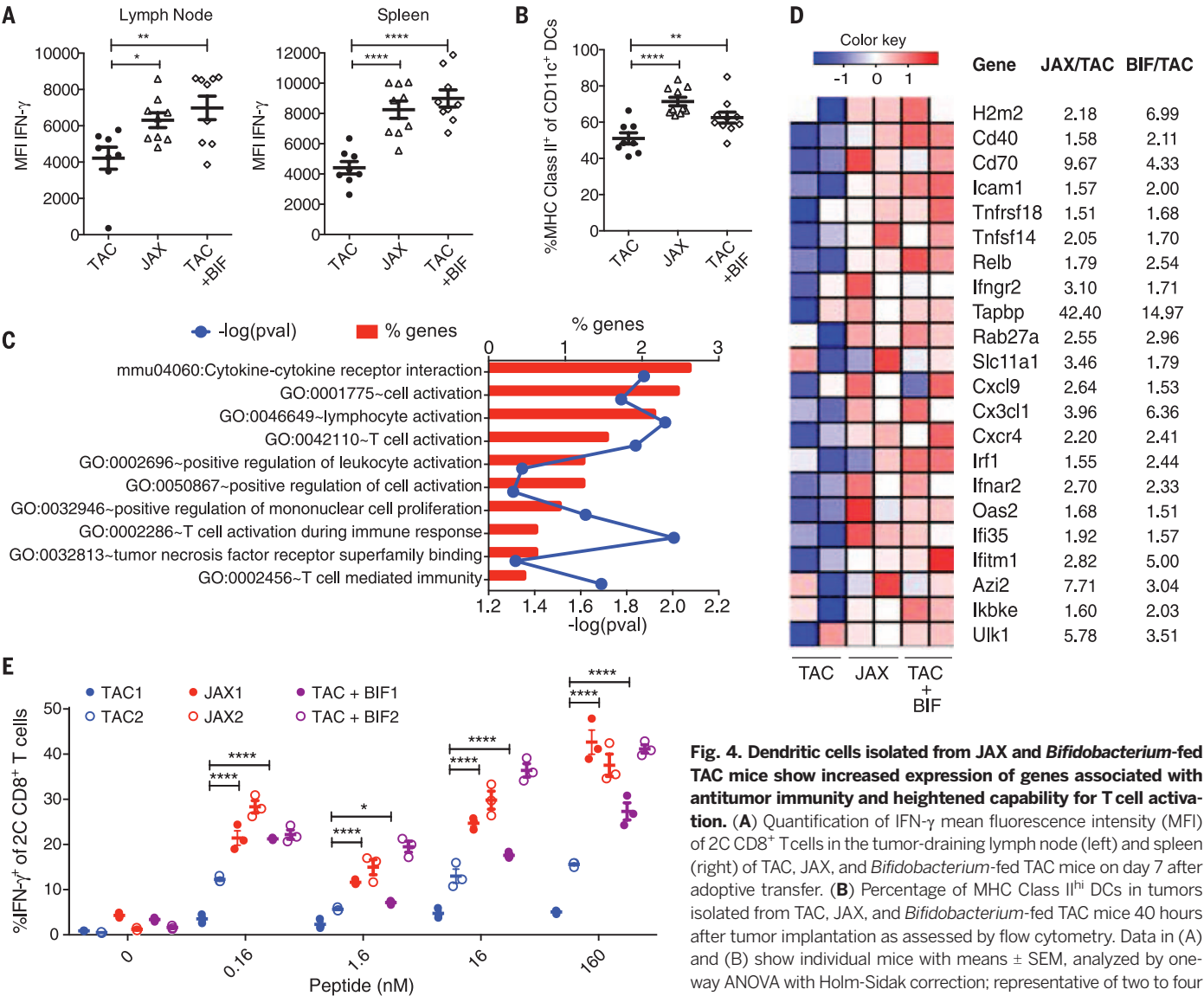
Fig. 4. Dendritic cells isolated from JAX and *Bifidobacterium*-fed TAC mice show increased expression of genes associated with antitumor immunity and heightened capability for T cell activation. (A) Quantification of IFN-γ mean fluorescence intensity (MFI) of 2C CD8[+] T cells in the tumor-draining lymph node (left) and spleen (right) of TAC, JAX, and *Bifidobacterium*-fed TAC mice on day 7 after adoptive transfer. (B) Percentage of MHC Class II[hi] DCs in tumors isolated from TAC, JAX, and *Bifidobacterium*-fed TAC mice 40 hours after tumor implantation as assessed by flow cytometry. Data in (A) and (B) show individual mice with means ± SEM, analyzed by one-way ANOVA with Holm-Sidak correction; representative of two to four independent experiments, eight or nine mice per group per experiment: *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$. (C) Enriched biological pathways and functions found within the subset of elevated genes in JAX and *Bifidobacterium*-treated TAC-derived DCs relative to untreated TAC DCs isolated from tumors 40 hours after tumor inoculation, as assessed by DAVID pathway analysis. Red bars indicate the percentage of genes in a pathway up-regulated in DCs isolated from JAX and *Bifidobacterium*-fed TAC mice. Blue line indicates P values calculated by Fisher's exact test. (D) Heat map of key antitumor immunity genes in DCs isolated from JAX, *Bifidobacterium*-treated TAC or untreated TAC mice. Mean fold-change for each gene transcript is shown on the right. (E) Quantification of IFN-γ[+] 2C TCR Tg CD8[+] T cells stimulated in vitro with DCs purified from peripheral lymphoid tissues of naïve TAC, JAX, and *Bifidobacterium*-treated TAC mice in the presence of different concentrations of SIY peptide. Analyses in (C) to (E) were performed on data combined from two independent experiments, five mice pooled per group per experiment. (E) Technical replicates of pooled samples from each experiment separately and were analyzed by fitting a linear mixed model, with Bonferroni correction for multiple comparisons: *$P < 0.05$, ****$P < 0.0001$.

cross presentation (*Tapbp*, *Rab27a*, and *Slc11a1*) (*25–27*); chemokine-mediated recruitment of immune cells to the tumor microenvironment (*Cxcl9*, *Cx3cl1*, and *Cxcr4*) (*28–30*); and type I interferon signaling (*Irf1*, *Ifnar2*, *Oas2*, *Ifi35*, and *Ifitm1*) (*31*, *32*) (Fig. 4D and fig. S7C). Expression of these genes was also increased in murine bone marrow–derived DCs stimulated with *Bifidobacterium* in vitro (table S6), consistent with previous reports that these species of *Bifidobacterium* can directly elicit DC maturation and cytokine production (*13*).

To test whether functional differences in DCs isolated from TAC, JAX, and *Bifidobacterium*-treated TAC mice could be sufficient to explain the differences in T cell priming observed in vivo, we purified DCs from lymphoid tissues of naïve TAC, JAX, and *Bifidobacterium*-treated TAC mice and tested their ability to induce carboxyfluorescein diacetate succinimidyl ester (CFSE)–labeled CD8$^+$ SIY-specific 2C TCR Tg T cell proliferation and acquisition of IFN-γ production in vitro. DCs purified from JAX and *Bifidobacterium*-treated TAC mice induced 2C T cell proliferation at lower antigen concentration than did DCs purified from naïve TAC mice (fig. S8, A and B). Furthermore, at all antigen concentrations, JAX-derived DCs elicited elevated levels of T cell IFN-γ production (Fig. 4E and fig. S8A). We observed similar effects upon oral administration of *Bifidobacterium* to TAC mice before DC isolation (Fig. 4E and fig. S8A). Taken together, these data suggest that commensal *Bifidobacterium*-derived signals modulate the activation of DCs in the steady state, which in turn supports improved effector function of tumor-specific CD8$^+$ T cells.

Our studies demonstrate an unexpected role for commensal *Bifidobacterium* in enhancing antitumor immunity in vivo. Given that beneficial effects are observed in multiple tumor settings and that alteration of innate immune function is observed, this improved antitumor immunity could be occurring in an antigen-independent fashion. The necessity for live bacteria may imply that *Bifidobacterium* colonizes a specific compartment within the gut that enables it to interact with host cells that are critical for modulating DC function or to release soluble factors that disseminate systemically and lead to improved DC function.

Our results do not rule out a contribution of other commensal bacteria species in having the capability to regulate antitumor immunity, either positively or negatively. Our data support the idea that one source of intersubject heterogeneity with regard to spontaneous antitumor immunity and therapeutic effects of antibodies targeting the PD-1/PD-L1 axis may be the composition of gut microbes, which could be manipulated for therapeutic benefit. These principles could apply to other immunotherapies, such as antibodies targeting the CTLA-4 pathway. Similar analyses can be performed in humans, by using 16*S* rRNA sequencing of stool samples from patients receiving checkpoint blockade or other immunotherapies, to identify commensals associated with clinical benefit.

REFERENCES AND NOTES

1. F. S. Hodi et al., *N. Engl. J. Med.* **363**, 711–723 (2010).
2. O. Hamid et al., *N. Engl. J. Med.* **369**, 134–144 (2013).
3. P. C. Tumeh et al., *Nature* **515**, 568–571 (2014).
4. S. Spranger et al., *Sci. Transl. Med.* **5**, 200ra116 (2013).
5. R. R. Ji et al., *Cancer Immunol. Immunother.* **61**, 1019–1031 (2012).
6. T. F. Gajewski, J. Louahed, V. G. Brichard, *Cancer J.* **16**, 399–403 (2010).
7. L. V. Hooper, D. R. Littman, A. J. Macpherson, *Science* **336**, 1268–1273 (2012).
8. I. I. Ivanov II, K. Honda, *Cell Host Microbe* **12**, 496–508 (2012).
9. J. P. McAleer, J. K. Kolls, *Immunity* **37**, 10–12 (2012).
10. N. Iida et al., *Science* **342**, 967–970 (2013).
11. S. Viaud et al., *Science* **342**, 971–976 (2013).
12. I. I. Ivanov II et al., *Cell* **139**, 485–498 (2009).
13. P. López, M. Gueimonde, A. Margolles, A. Suárez, *Int. J. Food Microbiol.* **138**, 157–165 (2010).
14. O. Ménard, M.-J. Butel, V. Gaboriau-Routhiau, A.-J. Waligora-Dupriet, *Appl. Environ. Microbiol.* **74**, 660–666 (2008).
15. P. Dong, Y. Yang, W. P. Wang, *Early Hum. Dev.* **86**, 51–58 (2010).
16. T. Kawahara et al., *Microbiol. Immunol.* **59**, 1–12 (2015).
17. N. Arpaia et al., *Nature* **504**, 451–455 (2013).
18. P. M. Smith et al., *Science* **341**, 569–573 (2013).
19. K. Atarashi et al., *Science* **331**, 337–341 (2011).
20. M. F. Mackey et al., *J. Immunol.* **161**, 2094–2098 (1998).
21. A. Scholer, S. Hugues, A. Boissonnas, L. Fetler, S. Amigorena, *Immunity* **28**, 258–270 (2008).
22. S. P. Bak et al., *J. Immunol.* **189**, 1708–1716 (2012).
23. J. Pan et al., *Immunol. Lett.* **94**, 141–151 (2004).
24. A. R. Pettit et al., *J. Immunol.* **159**, 3681–3691 (1997).
25. E. B. Compeer, T. W. H. Flinsenberg, S. G. van der Grein, M. Boes, *Front. Immunol.* **3**, 37 (2012).
26. C. Jancic et al., *Nat. Cell Biol.* **9**, 367–378 (2007).
27. C. B. Stober, S. Brode, J. K. White, J. F. Popoff, J. M. Blackwell, *Infect. Immun.* **75**, 5059–5067 (2007).
28. K. Kabashima et al., *Am. J. Pathol.* **171**, 1249–1257 (2007).
29. M. Nukiwa et al., *Eur. J. Immunol.* **36**, 1019–1027 (2006).
30. L. Zhang et al., *N. Engl. J. Med.* **348**, 203–213 (2003).
31. M. B. Fuertes et al., *J. Exp. Med.* **208**, 2005–2016 (2011).
32. S.-R. Woo et al., *Immunity* **41**, 830–842 (2014).

MALARIA

Malaria parasites target the hepatocyte receptor EphA2 for successful host infection

Alexis Kaushansky,[1]* **Alyse N. Douglass,[1]** **Nadia Arang,[1]** **Vladimir Vigdorovich,[1]** **Nicholas Dambrauskas,[1]** **Heather S. Kain,[1]** **Laura S. Austin,[1,2]** **D. Noah Sather,[1]** **Stefan H.I. Kappe[1,2]***

The invasion of a suitable host hepatocyte by mosquito-transmitted *Plasmodium* sporozoites is an essential early step in successful malaria parasite infection. Yet precisely how sporozoites target their host cell and facilitate productive infection remains largely unknown. We found that the hepatocyte EphA2 receptor was critical for establishing a permissive intracellular replication compartment, the parasitophorous vacuole. Sporozoites productively infected hepatocytes with high EphA2 expression, and the deletion of EphA2 protected mice from liver infection. Lack of host EphA2 phenocopied the lack of the sporozoite proteins P52 and P36. Our data suggest that P36 engages EphA2, which is likely to be a key step in establishing the permissive replication compartment.

M alaria infections place a tremendous burden on global health (*1*). Their causative agents, *Plasmodium* parasites, are transmitted to mammals as sporozoites by the bite of *Anopheles* mosquitoes. After entry into a capillary, sporozoites are carried to the liver, where they pass through multiple cells before recognizing and invading hepatocytes. During invasion, the sporozoite forms a protec- tive parasitophorous vacuole made of hepatocyte plasma membrane, which ensconces the parasite, establishes the intrahepatocytic replication

[1]Center for Infectious Disease Research (formerly Seattle Biomedical Research Institute), 307 Westlake Avenue North, No. 500, Seattle, WA 98109, USA. [2]Department of Global Health, University of Washington, Seattle, WA 98195, USA.
*Corresponding author. E-mail: alexis.kaushansky@cidresearch.org (A.K.); stefan.kappe@cidresearch.org (S.H.I.K.)
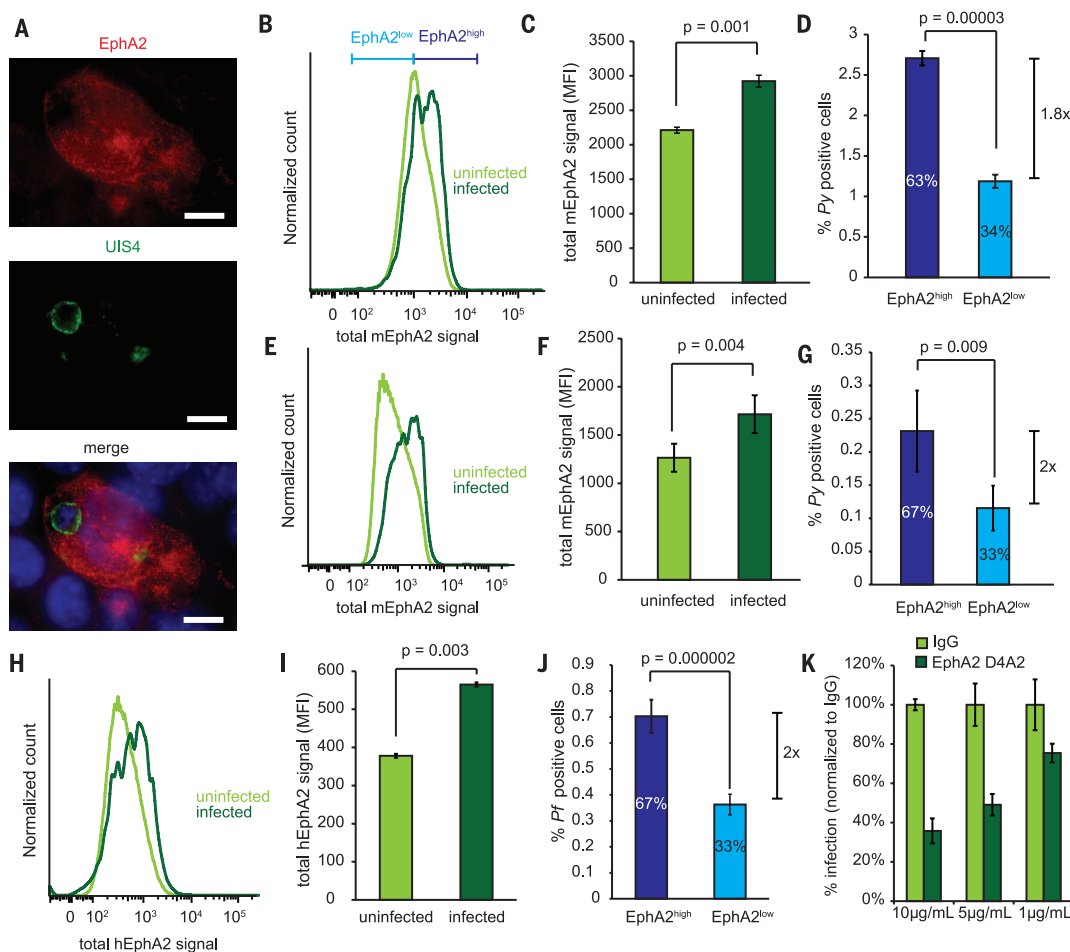
Fig. 1. *Plasmodium* sporozoites invade hepatocytes with high EphA2 expression. (**A**) Hepa1-6 cells were infected with *P. yoelii* sporozoites and visualized by immunofluorescence 24 hours after infection. The scale bar is 5 μm. (**B** to **D**) Hepa1-6 cells were infected with 10^5 *P. yoelii* sporozoites. (B) shows the distribution of EphA2 1.5 hours after infection (mEphA2, mouse EphA2). In (C), EphA2 levels are compared between parasite-infected and uninfected cells. (D) shows parasite-infection rates in EphA2^{high} and EphA2^{low} cells (*Py*, *P. yoelii*). The numbers in the bars are the percentages of infected cells within each sub-

set. (**E** to **G**) BALB/c mice were infected with 10^6 *P. yoelii* sporozoites by intravenous injection. Hepatocytes were analyzed as in (B) to (D). (**H** to **J**) HC-04 cells were infected with 10^5 *P. falciparum* sporozoites (hEphA2, human EphA2). Analyses were performed as in (B) to (D). (**K**) Hepa1-6 cells were incubated with EphA2-blocking antibody (D4A2) or immunoglobulin G (IgG) as a control 30 min before infection with 10^5 *P. yoelii* sporozoites. The infection rate was normalized to the infection rate in the presence of IgG. Each figure represents at least three independent experiments. The bar graphs show means with standard deviations.

niche, and supports successful infection. Highly sulfated proteoglycans are known to provide a signal to sporozoites to invade the liver parenchyma (*2*, *3*), and hepatocyte CD81 and scavenger receptor B1 are important for hepatocyte infection (*4*–*6*). Beyond this, the molecular mechanisms underlying infection remain poorly understood.

Hepatocytes exhibit differential susceptibility to infection. Sporozoites preferentially enter polyploid hepatocytes (*7*). Also, BALB/cByJ mice are more susceptible than BALB/cJ mice to *Plasmodium yoelii* sporozoite infection (*8*). To identify potential host receptors that might contribute to differential susceptibility, we used an antibody array to assess the levels of 28 activated receptors in the livers of BALB/cJ and BALB/cByJ mice. Nine receptors, including EphA2, were present at significantly ($P < 0.01$) and substantially elevated levels in highly susceptible BALB/cByJ mice (table S1). Polyploid hepatocytes also expressed higher levels of EphA2 (fig. S1).

In metazoans, Eph receptors and their cognate Ephrin ligands mediate cell-cell contact (*9*), making EphA2 a candidate to mediate the hepatocyte-sporozoite interaction. Furthermore, an Ephrin-like fold is present in the parasite's 6-Cys protein family (*10*). Although Hepa1-6 cells (a murine hepatocyte line) expressed EphA2 consistently across passages, variation within a culture was substantial (fig. S2). We therefore postulated that if EphA2 mediates sporozoite invasion, susceptibilities might vary within a culture of Hepa1-6 cells.

We infected Hepa1-6 cells with *P. yoelii* sporozoites; after 24 hours, we assessed parasites in hepatocytes that expressed high levels of EphA2 (Fig. 1A). We also observed this by flow cytometry 1.5 hours after infection (Fig. 1B and fig. S3A), and parasite-infected cells exhibited significantly increased levels of both total (Fig. 1C) and surface (fig. S3, B to D) EphA2. Similarly, the frequency of infection in cells in the top 50% of EphA2 expression (EphA2^{high}) was elevated compared

with infection frequency in cells in the bottom 50% (EphA2^{low}) (Fig. 1D). When we included only the top 40, 30, 20, or 10% of EphA2-expressing cells in the EphA2^{high} gate, the preference was even more pronounced (fig. S3E).

We next challenged BALB/c mice with 10^6 *P. yoelii* sporozoites and isolated hepatocytes after 3 hours. We again observed a strong parasite preference for EphA2^{high} hepatocytes (Fig. 1, E to G). Finally, we tested whether the preference for infection of EphA2^{high} hepatocytes is also present in the human parasites by infecting HC-04 hepatocytes with *P. falciparum*. We observed elevated levels of EphA2 in infected cells and a higher proportion of sporozoite-containing cells in the EphA2^{high} population (Fig. 1, H to J).

EphA2 has an extracellular ligand-binding region and an intracellular kinase domain, which mediates downstream signaling. To assess whether interaction with the extracellular portion of EphA2 is critical for *Plasmodium* infection, we infected

Fig. 2. EphA2 affects PVM formation. (**A**) Time sequence showing the maintenance of *P. yoelii* infection in Hepa1-6 cells. Maintenance of infection was defined as the infection rate at a given point divided by the infection rate at 1.5 hours after infection. Error bars indicate the standard deviation of biological replicates. (**B**) $EphA2^{(-/-)}$ or age-matched wild-type (WT) mice were infected with 10^5 *P. yoelii* sporozoites. Infection was assessed by quantitative polymerase chain reaction 42 hours after infection (*Py18s*, *P. yoelii* 18S ribosomal RNA; mGAPDH, mouse glyceraldehyde-3-phosphate dehydrogenase). Error bars show SEM. (**C**) $EphA2^{(-/-)}$ or strain-matched WT mice were infected with 10^2 *P. yoelii* sporozoites. Patency was monitored daily by thin smear. The horizontal bar indicates the median. (**D**) A *Py*UIS4-Myc parasite in a Hepa1-6 cell 24 hours after infection. The scale bar is 5 μm. (**E** and **F**) *Py*UIS4-Myc parasites were used to infect Hepa1-6 cells, and the cells were analyzed for the presence of the PVM (UIS4^{pos}) after 24 hours (MFI, median florescence intensity). (**G**) *Py*UIS4-Myc infected Hepa1-6 cells were assessed for permeability after 48 hours. Each figure represents at least three independent experiments. (E) to (G) show means with standard deviations.
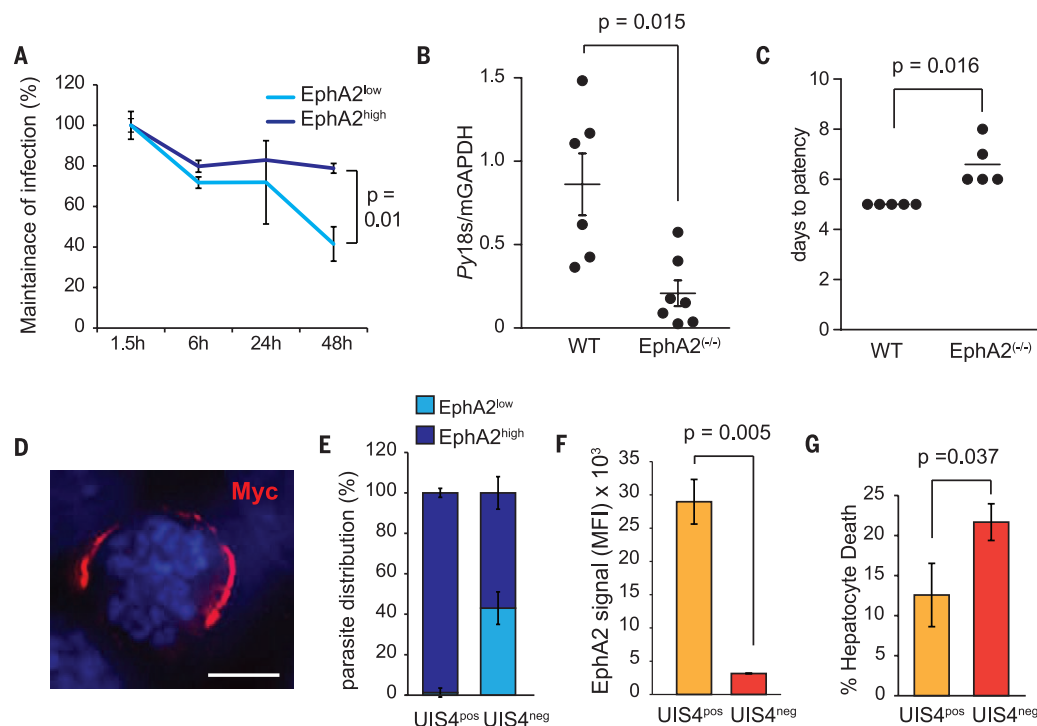
Fig. 3. P36 interacts with EphA2. (**A**) 2×10^5 WT or $p52^-/p36^-$ *P. yoelii* parasites were used to infect Hepa1-6 cells. Levels of EphA2 were monitored in infected and uninfected cells. (**B**) *P. yoelii* sporozoites were used to infect Hepa1-6 cells 30 min after treatment with IgG or EphA2-blocking antibody (αEphA2). Infection levels were normalized to the infection rate in the presence of IgG. (**C** and **D**) Hepa1-6 cells were incubated with recombinant P52 or P36, alone or in combination with 10 min of EphrinA1 treatment. Immunoblots show levels of pEphA2 (pY772). (**E**) 2×10^5 WT or $p52^-/p36^-/sap1^-$ *P. falciparum* sporozoites were used to infect 6×10^5 HC-04 cells. EphA2 levels were measured in infected and uninfected cells. Each figure represents at least three independent experiments. The bar graphs show means with standard deviations.
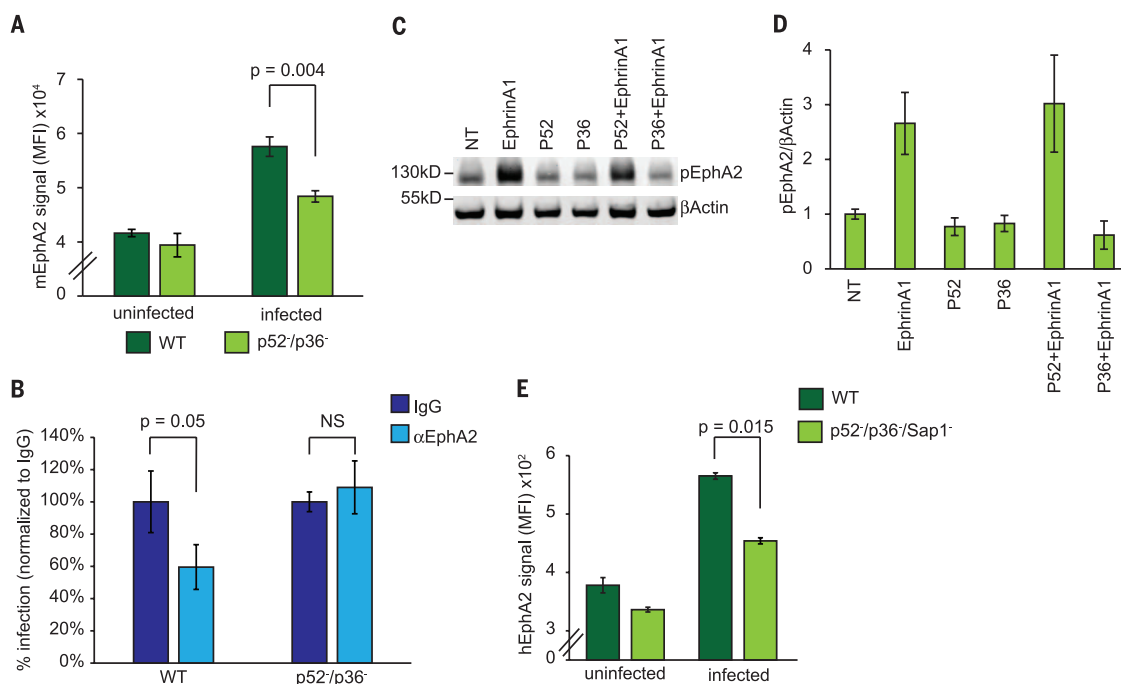
hepatocytes in the presence of an antibody that binds extracellular EphA2. The presence of the antibody reduced sporozoite infection in a dose-dependent manner (Fig. 1K). In contrast, inhibiting the kinase domain of EphA2 did not inhibit infection (fig. S4). Thus, the extracellular portion of EphA2 facilitates *Plasmodium* invasion of hepatocytes.

To test whether EphA2 levels are important for liver-stage parasite survival and development, we measured infection rates in EphA2^{high} and EphA2^{low} cells over the course of 48 hours, normalizing each infection rate to the rate at 1.5 hours after infection. Whereas the number of EphA2^{high} infected cells was maintained throughout the course

of infection, the number of EphA2[low] infected cells decreased over time (Fig. 2A). This difference could not be accounted for by division rates, because we observed lower levels of host cell division among EphA2[low] cells. Thus, our results may in fact underestimate the impact of EphA2 on infected cell survival (fig. S5). When we infected EphA2[(−/−)] and wild-type mice with 10^5 P. yoelii sporozoites, we observed a large decrease in liver-stage burden after 42 hours in EphA2[(−/−)] mice (Fig. 2B). EphA2[(−/−)] mice also exhibited a delay in the onset of blood-stage infection by 1 to 3 days (Fig. 2C). Thus, without EphA2, the host is far less susceptible to productive parasite liver infection.

The parasitophorous vacuole membrane (PVM) is critical for liver-stage development. One liver-stage PVM-resident protein, UIS4, is highly expressed after invasion when it is exported to the PVM (11), making it a useful marker. We constructed a P. yoelii parasite line, PyUIS4-Myc, which expressed a UIS4-Myc fusion protein driven by the endogenous UIS4 promoter (Fig. 2D). This allowed us to monitor PVM prevalence (UIS4[pos]) in infected cells by flow cytometry. Most of the UIS4[pos] infected host cells were in the EphA2[high] category (Fig. 2E). Similarly, the level of EphA2 expression was higher in UIS4[pos] infected cells than in UIS4[neg] infected cells (Fig. 2F). Thus, sporozoites not only preferentially entered EphA2[high] cells, but invasion accompanied by PVM formation was far more effective in these cells. UIS4[neg] infected hepatocytes suffered a higher frequency of cell death (Fig. 2G).

Two members of the 6-Cys family of parasite proteins (12, 13), P52 and P36, are expressed in sporozoites, are important for the invasion of hepatocytes (14–16), and are critical for PVM formation (14). In mouse livers, parasites without P52 or P36 were almost entirely eliminated within 3 hours after infection (fig. S6). We tested whether the lack of P52 and P36 phenocopies the lack of host EphA2 and found that p52[−]/p36[−] P. yoelii sporozoites exhibited a reduced preference for EphA2[high] cells (Fig. 3A). The related 6-Cys protein P12 shows structural similarity to the mammalian ligand for EphA2, EphrinA1 (10).

We showed that an interaction in the extracellular region of EphA2 was required for sporozoite entry using an EphA2-blocking antibody (Fig. 1K). Therefore, we next asked whether the presence of P36 and P52 was required for the antibody to block sporozoite entry. The EphA2 antibody blocked infection for wild-type P. yoelii sporozoites, but p52[−]/p36[−] sporozoite entry was not affected (Fig. 3B). These data suggest that P36 or P52 engages EphA2 at the point of host cell invasion. We next tested whether P52 or P36 could directly impede the interaction between EphrinA1 and EphA2 on the hepatocyte surface, which results in EphA2 activation. When we added EphrinA1 in the presence of P36 to Hepa1-6 cells, P36 blocked the activation of EphA2 (Fig. 3, C and D). P52, however, did not block EphrinA1-mediated activation of EphA2 (Fig. 3, C and D). To determine whether the interaction between EphA2 and P36 also occurs in human parasites, we assessed levels of EphA2 in P. falciparum

wild-type or p52[−]/p36[−]/sap1[−] parasite-infected HC-04 cells. The P52-P36–deficient P. falciparum sporozoites exhibited partially reduced selectivity for EphA2[high] HC-04 cells compared with P. falciparum wild-type sporozoites (Fig. 3E). Thus, P36 engages EphA2 but does not trigger its activation in rodent and human parasites.

We have shown that both host EphA2 and parasite 6-Cys proteins have a role in sporozoite invasion of hepatocytes and the establishment of the growth-permissive intracellular niche. Without either component, the parasite can still enter hepatocytes, but it does so without a PVM, which can result in death of the infected hepatocyte. The convergence of infection-permissive phenotypes is best explained by an interaction between parasite P36 and hepatocyte EphA2 when the PVM is formed. This role for EphA2 in hepatocyte infection does not preclude the possibility that additional hepatocyte receptors may be critical for infection. Interventional strategies aimed at either EphA2 or sporozoite 6-Cys proteins might block parasite infection before the onset of clinical malaria.

REFERENCES AND NOTES

1. S. H. Kappe, A. M. Vaughan, J. A. Boddey, A. F. Cowman, Science 328, 862–866 (2010).
2. A. Coppi et al., Cell Host Microbe 2, 316–327 (2007).
3. U. Frevert et al., J. Exp. Med. 177, 1287–1298 (1993).
4. O. Silvie et al., Nat. Med. 9, 93–96 (2003).
5. S. Yalaoui et al., Cell Host Microbe 4, 283–292 (2008).
6. C. D. Rodrigues et al., Cell Host Microbe 4, 271–282 (2008).
7. L. S. Austin, A. Kaushansky, S. H. Kappe, Cell. Microbiol. 16, 784–795 (2014).
8. A. Kaushansky et al., Infect. Immun. 83, 39–47 (2015).
9. E. M. Lisabeth, G. Falivelli, E. B. Pasquale, Cold Spring Harb. Perspect. Biol. 5, a009159 (2013).
10. S. A. Arredondo et al., Proc. Natl. Acad. Sci. U.S.A. 109, 6692–6697 (2012).
11. A. K. Mueller et al., Proc. Natl. Acad. Sci. U.S.A. 102, 3022–3027 (2005).
12. T. Annoura et al., FASEB J. 28, 2158–2170 (2014).
13. M. R. van Dijk et al., PLOS Pathog. 6, e1000853 (2010).
14. M. Labaied et al., Infect. Immun. 75, 3758–3768 (2007).
15. M. R. van Dijk et al., Proc. Natl. Acad. Sci. U.S.A. 102, 12194–12199 (2005).
16. K. M. VanBuskirk et al., Proc. Natl. Acad. Sci. U.S.A. 106, 13004–13009 (2009).

GENOMICS

Gene essentiality and synthetic lethality in haploid human cells

Vincent A. Blomen,[1]* Peter Májek,[2]* Lucas T. Jae,[1] Johannes W. Bigenzahn,[2] Joppe Nieuwenhuis,[1] Jacqueline Staring,[1] Roberto Sacco,[2] Ferdy R. van Diemen,[1] Nadine Olk,[2] Alexey Stukalov,[2] Caleb Marceau,[3] Hans Janssen,[1] Jan E. Carette,[3] Keiryn L. Bennett,[2] Jacques Colinge,[2,4]† Giulio Superti-Furga,[2,5] Thijn R. Brummelkamp[1,2,6]†

Although the genes essential for life have been identified in less complex model organisms, their elucidation in human cells has been hindered by technical barriers. We used extensive mutagenesis in haploid human cells to identify approximately 2000 genes required for optimal fitness under culture conditions. To study the principles of genetic interactions in human cells, we created a synthetic lethality network focused on the secretory pathway based exclusively on mutations. This revealed a genetic cross-talk governing Golgi homeostasis, an additional subunit of the human oligosaccharyltransferase complex, and a phosphatidylinositol 4-kinase β adaptor hijacked by viruses. The synthetic lethality map parallels observations made in yeast and projects a route forward to reveal genetic networks in diverse aspects of human cell biology.

S ingle-cell organisms can often tolerate inactivating mutations in the majority of genes (1–3), but it is unclear whether human cells require more essential genes because of increased complexity, or fewer because of added redundancy. To study this, we used mutagenesis in the near-haploid chronic myeloid leu-

kemia (CML) cell line KBM7 (karyotype 25, XY, +8, Ph+), and its nonhematopoietic derivative HAP1, which is haploid for all chromosomes (fig. S1A) (4). More than 34.3 million and 65.9 million gene-trap integrations were identified in KBM7 and HAP1 cells, respectively. The gene-trap vector we used was unidirectional by design (fig. S1B),

and for most genes, the number of intronic integrations in the sense direction was similar to that in the antisense direction (for example, pro-apoptotic factor *BBC3*) (Fig. 1, A and B) (*5–7*). For a fraction of genes, however, disruptive muta-

[1]Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, Netherlands. [2]CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria. [3]Department of Microbiology and Immunology, Stanford University School of Medicine, 299 Campus Drive, Stanford, CA 94305, USA. [4]University of Montpellier, Institut de Recherche en Cancérologie de Montpellier Inserm U1194, Institut régional du Cancer Montpellier, 34000 Montpellier, France. [5]Center for Physiology and Pharmacology, Medical University of Vienna, Vienna, Austria. [6]Cancer Genomics Center (CGC.nl), Plesmanlaan 121, 1066CX, Amsterdam, Netherlands.
*These authors contributed equally to this work. †Corresponding author. E-mail: jacques.colinge@inserm.fr (J.C.); gsuperti@cemm.at (G. S.-F.); t.brummelkamp@nki.nl (T.R.B.)

tions were underrepresented, which is indicative of impaired fitness: Some genes (such as *STAT5B*) appeared essential in one cell type (fig. S1C), whereas others appeared essential in both (such as *RPL13A*) (Fig. 1, A and B, and tables S1 to S3).

In KBM7 cells, 2054 genes (table S1 and figs. S2 and S3) and in HAP1, 2181 genes (table S2 and figs. S2 and S3) appeared to be needed for viability or optimal fitness under the experimental growth conditions (referred to as "essential," although the approach does not distinguish between the two). The 1734 genes identified in both cell lines were designated as "core essentialome" (table S3). In KBM7 cells, genes on chromosome 8 (present in two copies) tolerated disruptive mutations, underscoring the specificity of the approach (fig. S1D). Furthermore, nearly all subunits of the proteasome were identified as es-

sential (fig. S4). In general, essential genes are overrepresented in categories such as translation or transcription but not signaling (Fig. 1C and figs. S5 and S6).

Many genes required for fitness in yeast were also essential in human cells. Exceptions were largely explained by paralogs in the human genome or by yeast-specific requirements (fig. S7A and table S4) (*1*). We estimated the evolutionary age of essential genes and found that 77% emerged in premetazoans ("old" essential genes) (fig. S7B). Essential genes had fewer paralogs and higher protein abundance and contained fewer single-nucleotide polymorphisms (SNPs) predicted to impair function (Fig. 1D). Proteins encoded by essential genes displayed more protein-protein interactions (fig. S8, A to D), and these occurred more frequently with other essential proteins (49.8%) (fig. S9A) and within the same functional
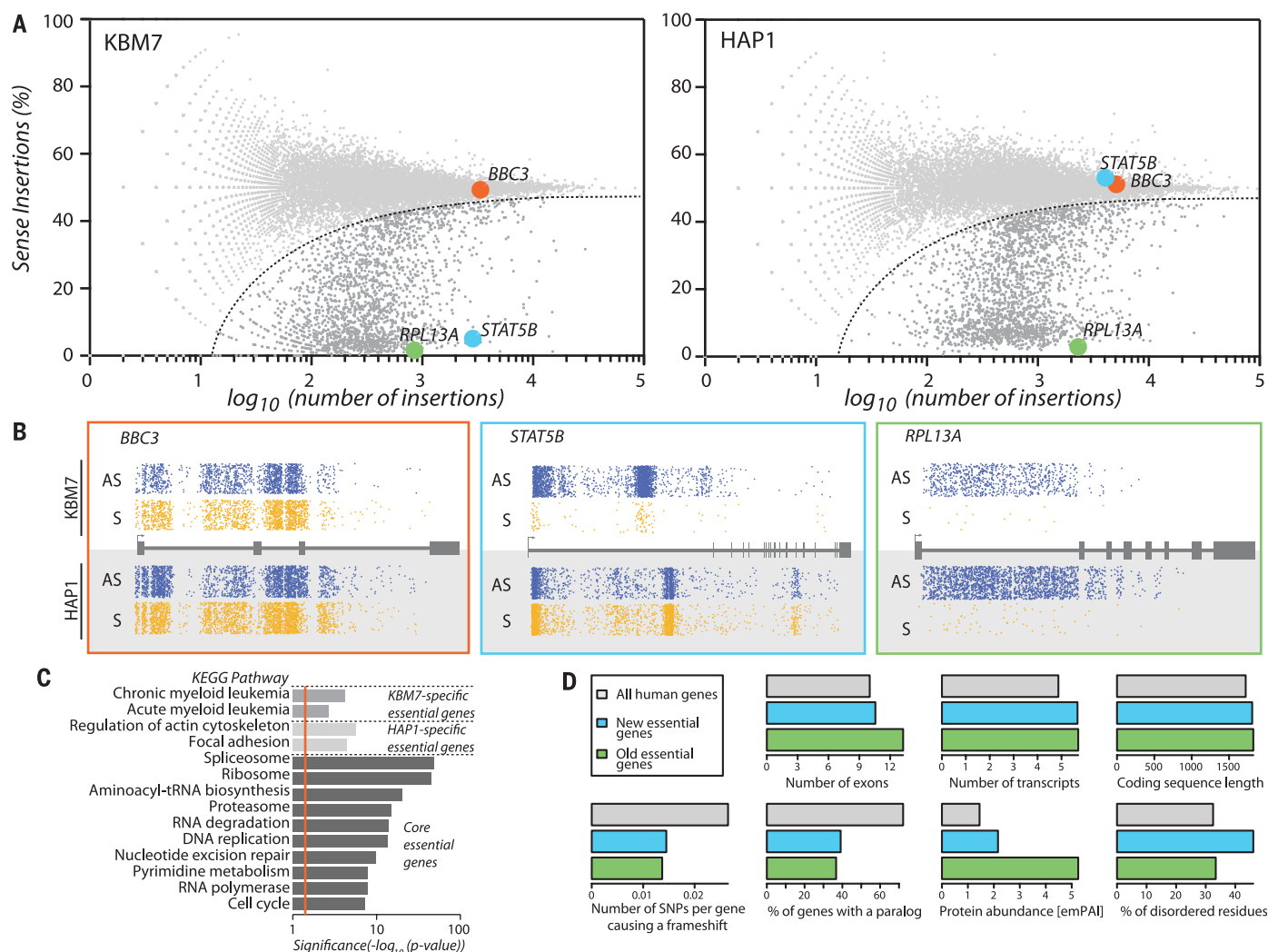
Fig. 1. Identification of genes required for fitness in KBM7 and HAP1 cells through insertional mutagenesis. (**A**) Distinct gene-trap insertions were mapped in KBM7 and HAP1 cells, and their orientation relative to the affected genes was counted. Per gene, the percentage of sense orientation gene-trap insertions (*y* axis) and the total number of insertions in a particular gene (*x* axis) are plotted. (**B**) Gene-trap insertions identified in the sense (S, yellow) or antisense orientation (AS, blue) in a nonessential gene

(*BBC3*), a gene essential only in KBM7 cells (*STAT5B*), and a gene essential in both cell lines (*RPL13A*). (**C**) KEGG pathway enrichment analysis of essential genes shared between or specific to KBM7 or HAP1 cells. (**D**) Properties of new and old essential genes compared with the human genome. Averages for the sets are displayed, except for protein abundance, in which median Exponentially Modified Protein Abundance Index (emPAI) values are shown.
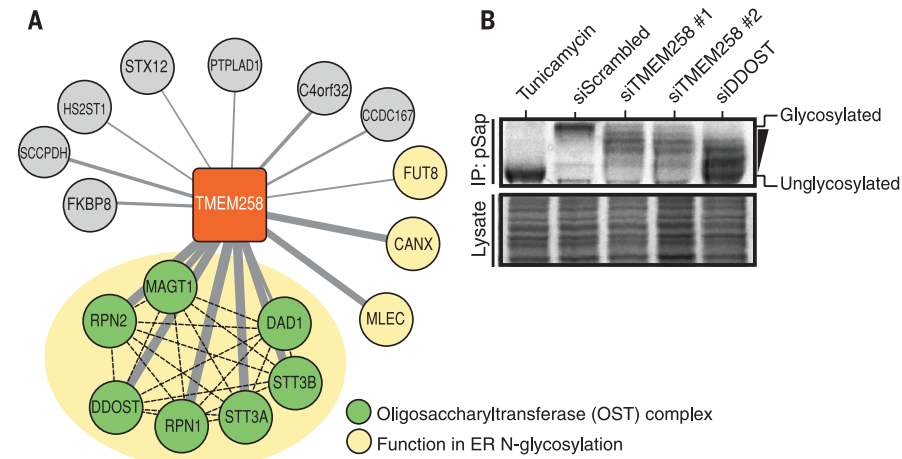
Fig. 2. The essential gene *TMEM258* encodes a component of the OST complex. (**A**) High-confidence protein-protein interactions associated with TMEM258. Green proteins indicate members of the OST complex. Dashed lines indicate the OST complex subnetwork. (**B**) Effects of depletion of TMEM258 with small interfering RNAs on the glycosylation of endogenous prosaposin. Cells were pulsed with ^{35}S-methionine/cysteine, lysed, and subjected to immunoprecipitation by using antibodies to prosaposin. Precipitated proteins were detected by means of phosphorimaging, and hypoglycosylated prosaposin species are indicated. Tunicamycin treatment and depletion of the established OST subunit DDOST served as positive controls.

category (fig. S5B). Remarkably, the products of "new" essential genes are more often connected with old rather than other new essential gene products, suggesting that they largely function within ancient molecular machineries (fig. S9, B and C).

To identify proteins interacting with products of 18 uncharacterized essential genes, we used tandem affinity purification coupled to mass spectrometry (fig. S10). Interactors were frequently essential proteins (52.4%, $P < 2.5 \times 10^{-36}$, hypergeometric test) involved in processes such as splicing, translation, and trafficking (fig. S11 and table S5). The small transmembrane protein TMEM258 associated with components of the conserved oligosacharyltransferase (OST) complex (Fig. 2A and fig. S12A) that are essential for protein N-glycosylation (8). TMEM258 localized to the endoplasmic reticulum (fig. S12B), and depletion (fig. S12, C and D) impaired OST catalytic activity as monitored by means of hypoglycosylation of prosaposin (Fig. 2B) (9). This also rationalizes the observed clustering of TMEM258 with OST complex subunits in a recent genetic screen (10). Thus, TMEM258 constitutes a subunit of the
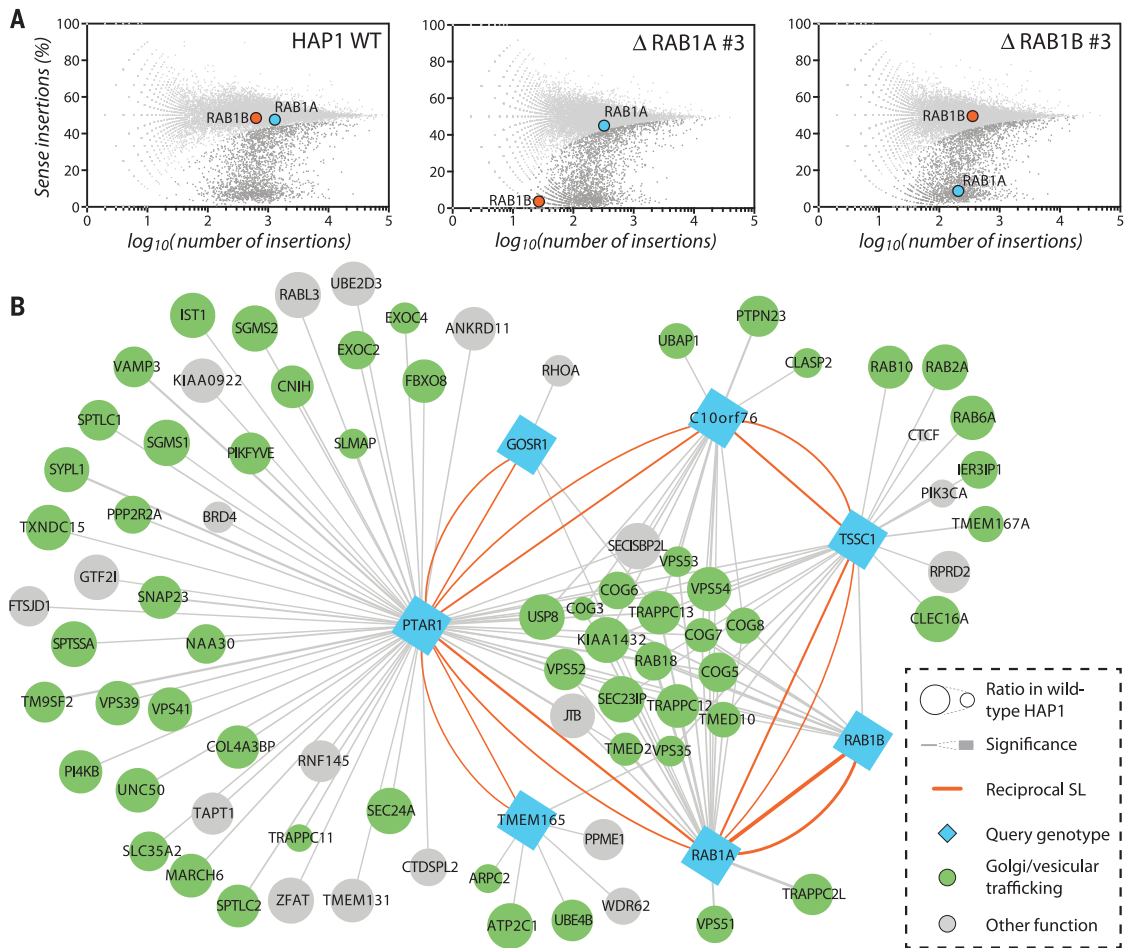
Fig. 3. Synthetic lethality network generated based on mutations. (**A**) Essentiality of *RAB1A* and *RAB1B* in wild-type HAP1 cells and cells deficient for *RAB1A* or *RAB1B*. (**B**) Genetic interaction network indicating synthetic lethal/sick interactions that were identified by scoring genes for fitness reduction in three nuclease-generated knockout clones per genotype. This revealed an interconnected network, with many genes that could be functionally assigned to the secretory pathway (labeled in green). Reciprocal interactions, scored in either query genotype, are indicated by orange edges. Edge thickness reflects the effect size of the interaction (compared with wild-type cells).

human OST complex, and although homology searches (fig. S12E) do not identify a yeast ortholog, TMEM258 may relate to the similarly sized yeast transmembrane protein OST5 (*11*).

Whereas most genes appear nonessential, their function may be buffered by other genes so that only simultaneous disruption is lethal (*12*–*15*). The frequency of such synthetic lethal interactions between human genes is debated and challenging to address experimentally (*16, 17*). We studied the small guanosine triphosphatases (GTPases) RAB1A and RAB1B by creating individual knockout lines and assessing the genes needed for fitness in these backgrounds (Fig. 3A and fig. S13A). Whereas neither *RAB1A* nor *RAB1B* were essential in wild-type cells, *RAB1A* became indispensable in *RAB1B* knockout cells and vice versa (Fig. 3A and fig. S13B). To explore the breadth of synthetic lethality, we probed the secretory pathway using three independent knockout cell lines (fig. S14) for *RAB1A, RAB1B, GOSR1* [a subunit of the Golgi soluble *N*-ethylmaleimide–sensitive factor attachment protein (SNAP) receptor] (*18*), and *TMEM165* (a Golgi-resident Ca^{2+}/H^+ antiporter whose deficiency impairs glycosylation) (Fig. 3B, figs. S15 and S16, and table S6) (*19*). Most of their genetic interactions impinged on the secretory pathway (Fig. 3B and table S7), and many were found synthetic lethal with *PTAR1*. Synthetic lethality screens in PTAR1-deficient cells confirmed these genetic interactions and additionally identified the uncharacterized gene *C10orf76* (Fig. 3B and fig. S17A). Validation by use of *C10orf76* as query gene confirmed synthetic lethality with *PTAR1* and (reciprocally) identified

TSSC1, which was recently reported to interact with the Golgi-associated retrograde protein complex (GARP) (Fig. 3B) (*20*). The human genes we studied display on average ~20 synthetic lethal interactions, a number comparable with that in yeast (*12*), although this varies between genes, with *PTAR1* (causing a fitness defect when deleted alone) having close to 60 interactions (fig. S17B). This illustrates that synthetic lethal interactions can be identified and validated by using reciprocal haploid screens and that, similarly to yeast, interactions frequently occur between genes whose products act in related processes (fig. S17B) (*13, 16*). However, we acknowledge a caveat that this approach cannot readily distinguish between synthetic lethal or synthetic "sick" interactions.

The impaired growth of *PTAR1*-deficient cells (table S2) was suppressed by loss of the Golgi factor GOLGA5 (Fig. 4A and fig. S18) (*21*). *PTAR1*-deficient cells had an abnormally dilated Golgi morphology (fig. S19A), which was partially corrected by codeletion of *GOLGA5* (Fig. 4B and fig. S19B). Functionally, *PTAR1* deficiency impaired glycosylation (fig. S19C) (*7*), possibly owing to dysregulation of RAB proteins (*22*). Indeed, *PTAR1*-deficient cells showed attenuated geranylgeranylation of RAB1A and RAB1B (fig. S19D). Partial correction of the Golgi morphology in cells lacking both *PTAR1* and *GOLGA5* could relate to the effect of GOLGA5, itself a RAB effector, on Golgi fragmentation (*21, 23*). Thus, the interaction map reveals *PTAR1* and *GOLGA5* as opposing handles tuning Golgi morphology and homeostasis.

Genetic analysis suggested a link between the unstudied gene *C10orf76* and *PI4KB*, which were

both synthetic lethal with *PTAR1* (Fig. 3B). A host factor screen with coxsackievirus A10 also identified *C10orf76* as well as *PI4KB* (fig. S20, A and B), and a proteomics survey (*24*) suggested association between *C10orf76* and *PI4KB*. We confirmed this interaction in immunoprecipitation experiments with cells expressing FLAG-tagged C10orf76 (Fig. 4C). Phosphatidylinositol 4-kinase β (PI4KB) regulates abundance of phosphatidylinositol 4-phosphate [PI(4)P] (*25*) and has a role in genome replication of various RNA viruses, including coxsackieviruses (*26*). Infection studies confirmed that cells in which *C10orf76* was knocked out were particularly resistant to coxsackievirus A10 (fig. S20C). Although virus entry occurred normally, replication of viral RNA was decreased in cells in which *C10orf76* was knocked out (Fig. 4D and fig. S20D). Enteroviruses hijack PI4KB activity to construct "replication factories," which were abundant in wild-type cells but rare in *C10orf76*-deficient cells (fig. S20E). Amounts of PI(4)P were decreased in these cells, and Golgi retention of PI4KB after chemical inhibition (*27*) was largely dependent on C10orf76, which also localized to this compartment under these conditions (fig. S21, A and B). Thus, C10orf76 is a PI4KB-associated factor hijacked by specific picornaviruses for replication.

This study identifies ~2000 genes required for optimal fitness of cultured haploid human cells. Despite technical limitations, the identification of gene essentiality shows high concordance with the gene-trap and clustered regularly interspaced short palindromic repeats (CRISPR) data reported in the accompanying manuscript of Wang *et al.*
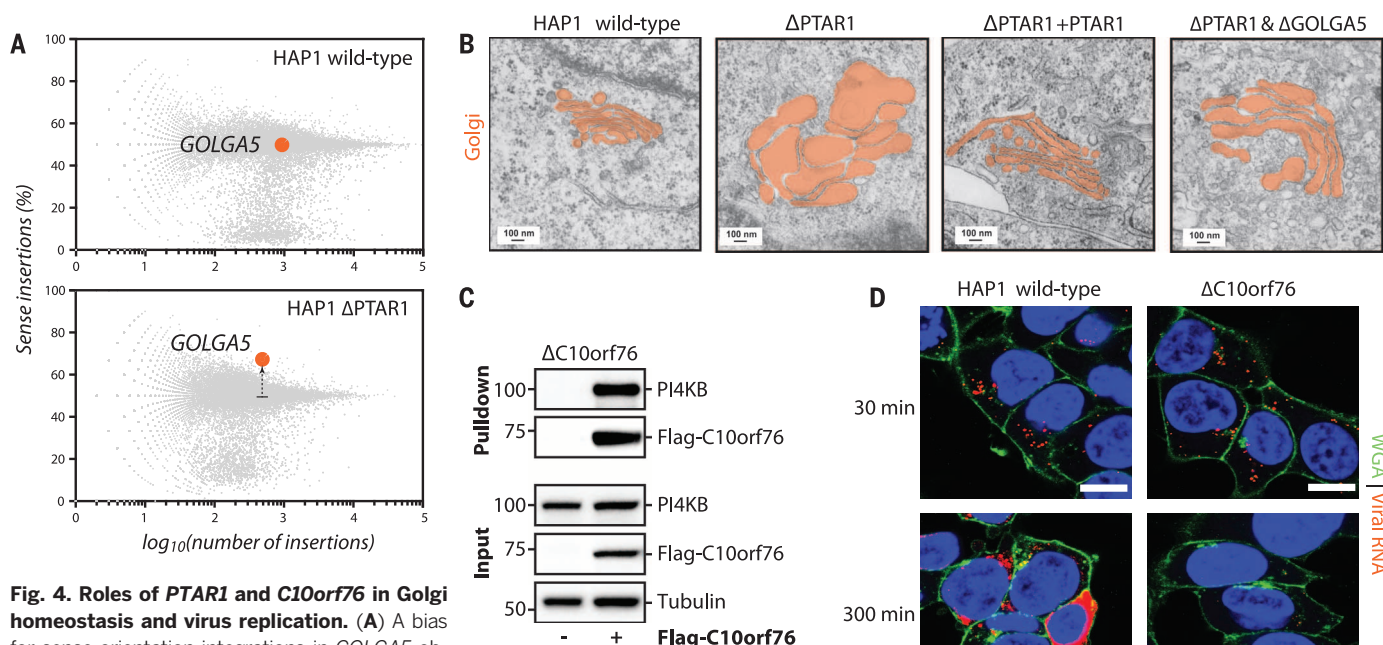
Fig. 4. Roles of *PTAR1* and *C10orf76* in Golgi homeostasis and virus replication. (**A**) A bias for sense-orientation integrations in *GOLGA5* observed in *PTAR1*-deficient HAP1 cells but not wild-type cells. (**B**) Electron micrographs of the Golgi apparatus (orange highlight) in the indicated genotypes. (**C**) Interaction of FLAG-tagged C10orf76 with PI4KB in HAP1 cells detected with immunoprecipitation by using antibodies to FLAG. (**D**) Coxsackievirus A10 amplification in wild-type and *C10orf76*-deficient cells measured by means of single molecule fluorescent in situ hybridization (smFISH) to localize individual viral genomes (red). Intracellular viral RNA was first detected after 30 min. Increased RNA signal after 300 min indicates RNA replication.

(supplementary text and fig. S22) (*28*). This suggests that the increase in total number of genes in humans as compared with that in yeast yielded a system of higher complexity rather than more robustness through added redundancy. Nonessential human genes appear to frequently engage in synthetic lethal interactions. Our studies start to reveal an interconnected module of genetic interactions affecting the secretory pathway and link it to uncharacterized genes. The experimental strategy is applicable to various cellular processes and may help unravel the genetic network encoding a human cell.

REFERENCES AND NOTES

1. G. Giaever et al., Nature **418**, 387–391 (2002).
2. A. H. Tong et al., Science **294**, 2364–2368 (2001).
3. S. Mnaimneh et al., Cell **118**, 31–44 (2004).
4. J. E. Carette et al., Nature **477**, 340–343 (2011).
5. J. E. Carette et al., Nat. Biotechnol. **29**, 542–546 (2011).
6. J. E. Carette et al., Science **326**, 1231–1235 (2009).
7. L. T. Jae et al., Science **340**, 479–483 (2013).
8. D. J. Kelleher, R. Gilmore, Glycobiology **16**, 47R–62R (2006).
9. A. Dumax-Vorzet, P. Roboti, S. High, J. Cell Sci. **126**, 2595–2606 (2013).
10. O. Parnas et al., Cell **162**, 675–686 (2015).
11. G. Reiss, S. te Heesen, R. Gilmore, R. Zufferey, M. Aebi, EMBO J. **16**, 1164–1172 (1997).
12. A. H. Tong et al., Science **303**, 808–813 (2004).
13. J. L. Hartman 4th, B. Garvik, L. Hartwell, Science **291**, 1001–1004 (2001).
14. C. Boone, H. Bussey, B. J. Andrews, Nat. Rev. Genet. **8**, 437–449 (2007).
15. M. Schuldiner et al., Cell **123**, 507–519 (2005).
16. S. M. Nijman, S. H. Friend, Science **342**, 809–811 (2013).
17. J. Tischler, B. Lehner, A. G. Fraser, Nat. Genet. **40**, 390–391 (2008).
18. V. N. Subramaniam, F. Peter, R. Philp, S. H. Wong, W. Hong, Science **272**, 1161–1163 (1996).
19. D. Demaegd et al., Proc. Natl. Acad. Sci. U.S.A. **110**, 6859–6864 (2013).
20. E. L. Huttlin et al., Cell **162**, 425–440 (2015).
21. A. Diao, D. Rahman, D. J. Pappin, J. Lucocq, M. Lowe, J. Cell Biol. **160**, 201–212 (2003).
22. W. L. Charng et al., PLOS Biol. **12**, e1001777 (2014).
23. A. Rejman Lipinski et al., PLOS Pathog. **5**, e1000615 (2009).
24. A. L. Greninger, G. M. Knudsen, M. Betegon, A. L. Burlingame, J. L. DeRisi, MBio **4**, e00098–e00013 (2013).
25. T. R. Graham, C. G. Burd, Trends Cell Biol. **21**, 113–121 (2011).
26. N. Y. Hsu et al., Cell **141**, 799–811 (2010).
27. H. M. van der Schaar et al., Cell Res. **22**, 1576–1592 (2012).
28. T. Wang et al., Science **350**, 1096–1101 (2015).

Identification and characterization of essential genes in the human genome

Tim Wang,[1,2,3,4] Kıvanç Birsoy,[1,2,3,4]* Nicholas W. Hughes,[3] Kevin M. Krupczak,[2,3,4] Yorick Post,[2,3,4] Jenny J. Wei,[1,2] Eric S. Lander,[1,3,5]†‡ David M. Sabatini[1,2,3,4,6]†‡

Large-scale genetic analysis of lethal phenotypes has elucidated the molecular underpinnings of many biological processes. Using the bacterial clustered regularly interspaced short palindromic repeats (CRISPR) system, we constructed a genome-wide single-guide RNA library to screen for genes required for proliferation and survival in a human cancer cell line. Our screen revealed the set of cell-essential genes, which was validated with an orthogonal gene-trap–based screen and comparison with yeast gene knockouts. This set is enriched for genes that encode components of fundamental pathways, are expressed at high levels, and contain few inactivating polymorphisms in the human population. We also uncovered a large group of uncharacterized genes involved in RNA processing, a number of whose products localize to the nucleolus. Last, screens in additional cell lines showed a high degree of overlap in gene essentiality but also revealed differences specific to each cell line and cancer type that reflect the developmental origin, oncogenic drivers, paralogous gene expression pattern, and chromosomal structure of each line. These results demonstrate the power of CRISPR-based screens and suggest a general strategy for identifying liabilities in cancer cells.

The systematic identification of essential genes in microorganisms has provided critical insights into the molecular basis of many biological processes (*1*). Similar studies in human cells have been hindered by the lack of suitable tools. Moreover, little is known about how the set of cell-essential genes differs across cell types and genotypes. Differentially essential genes are likely to encode tissue-specific modulators of key cellular processes and important targets for cancer therapies. We used two independent approaches for inactivating genes at the DNA level to define the cell-essential genes of the human genome.

The first approach uses the clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9–based gene editing system, which has emerged as a powerful tool to engineer the genomes of cultured cells and whole organisms (*2*, *3*). We and others have shown that lentiviral single-guide RNA (sgRNA) libraries can enable pooled loss-of-function screens and have used the technology to uncover mediators of drug resistance and pathogen toxicity (*4*–*6*). To systematically identify cell-essential genes, we constructed

a library, which was optimized for high cleavage activity, and performed a proliferation-based screen in the near-haploid human KBM7 chronic myelogenous leukemia (CML) cell line (Fig. 1, table S1, and supplementary text S1).

The unusual karyotype of these cells also allows for an independent method of genetic screening. In this approach, null mutants are generated at random through retroviral gene-trap mutagenesis, selected for a phenotype, and monitored by sequencing the viral integration sites to pinpoint the causal genes (*7*). Positive selection–based screens by use of this method have identified genes underlying processes such as epigenetic silencing and viral infection (*7*–*9*). We extended this technique by developing a strategy for negative selection and conducted a screen for cell-essential genes (Fig. 1 and supplementary text S2).

For both methods, we computed a score for each gene that reflects the fitness cost imposed by inactivation of the gene. We defined the CRISPR score (CS) as the average $\log_2$ fold-change in the abundance of all sgRNAs targeting a given gene, with replicate experiments showing a high degree of reproducibility [correlation coefficient (r) = 0.90] (Fig. 2A, fig. S1A, and table S2). Of the 18,166 genes targeted by the library, 1878 scored as essential for optimal proliferation in our screen, although this precise number depends on the cutoff chosen (Fig. 2A and tables S2 and S3). Overall, this fraction represents ~10% of genes within our data set or roughly 9.2% of the entire genome (many of the genes not targeted by our library encode olfactory receptors that are unlikely to be cell-essential). Gene products that act in a non-cell-autonomous manner are not expected to score as essential in this pooled setting (fig. S1B).

[1]Department of Biology, Massachusetts Institute of Technology, Cambridge MA 02139, USA. [2]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA. [3]Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142, USA. [4]David H. Koch Institute for Integrative Cancer Research at MIT, Cambridge, MA 02139, USA. [5]Department of Systems Biology, Harvard Medical School, Boston, MA 02115, USA. [6]Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
*Present address: Laboratory of Metabolic Regulation and Genetics, The Rockefeller University, New York, NY 10065, USA.
†These authors contributed equally to this work.
‡Corresponding author. E-mail: lander@broadinstitute.org (E.S.L.); sabatini@wi.mit.edu (D.M.S.).

We defined the gene-trap score (GTS) as the fraction of insertions in a given gene occurring in the inactivating orientation. Because the accuracy of this score depends on the depth of insertional coverage, we set a requirement on the minimum number ($n = 65$) of antisense inserts in a gene needed for inclusion in our analysis by measuring the concordance between replicate experiments (Fig. 2B; fig. S1, C and D; table S4; and supplementary text S2). For the 7370 genes on the haploid chromosomes that exceeded this threshold, the GTS was well-correlated with the CS and with results from a copublished study that used a similar gene-trap approach ($r = 0.68$) (Fig. 2C; fig. S1, E and F; and supplementary text S3). The strong correspondence between the overlapping sets of cell-essential genes defined by the two methods provides support for the accuracy of the CRISPR scores for the full set of 18,166 genes.

The two methods differed with respect to the diploid chromosome 8. Whereas the gene-trap screen failed to detect any cell-essential genes on this chromosome, the CRISPR screen uncovered a similar proportion of cell-essential genes on all the autosomes (Fig. 2, A to C). These observations indicate that (i) the vast majority of cell-essential genes are haplosufficient and (ii) biallelic inactivation occurs at high frequency in our CRISPR screen (4).

To assess the accuracy of our scores with other measures of gene essentiality, we relied on functional profiling experiments conducted in yeast *Saccharomyces cerevisiae* as a benchmark (1, 10). Specifically, we ranked genes common to all data sets by their scores in each data set—CRISPR, gene trap, scores from similar loss-of-function RNA interference (RNAi) screens (11), and as a naïve proxy for gene essentiality, gene expression levels determined by means of RNA sequencing (RNA-Seq)—and compared these rankings with the essentiality of yeast homologs. The CRISPR and gene-trap methods had significantly stronger correlations with the yeast results than did the RNAi screens or gene expression, which performed similarly to each other (both methods, $P < 10^{-4}$, permutation test) (Fig. 2D). On the basis of additional comparisons with yeast gene essen-

tiality, we also found that (i) our new optimized sgRNA library gave better results than those from screens using older unoptimized libraries (4, 5) and (ii) the coverage of this library (~10 constructs per gene) approaches saturation, as evidenced by down-sampling (decreasing the coverage by randomly eliminating subsets of data) (fig. S2, A and B). Together, our results suggest that scores from the CRISPR and gene-trap screens both provide accurate measures of the cell-essentiality of human genes.

Essential genes should be under strong purifying selection and should thus show greater evolutionary constraint than that of nonessential genes (12). Consistent with this expectation, the essential genes found in our screens were more broadly retained across species, showed higher levels of conservation between closely related species, and contain fewer inactivating polymorphisms within the human species, as compared with their dispensable counterparts (Fig. 2, E to G). Essential genes also tend to have higher expression and encode proteins that engage in more protein-protein interactions (13–15). These
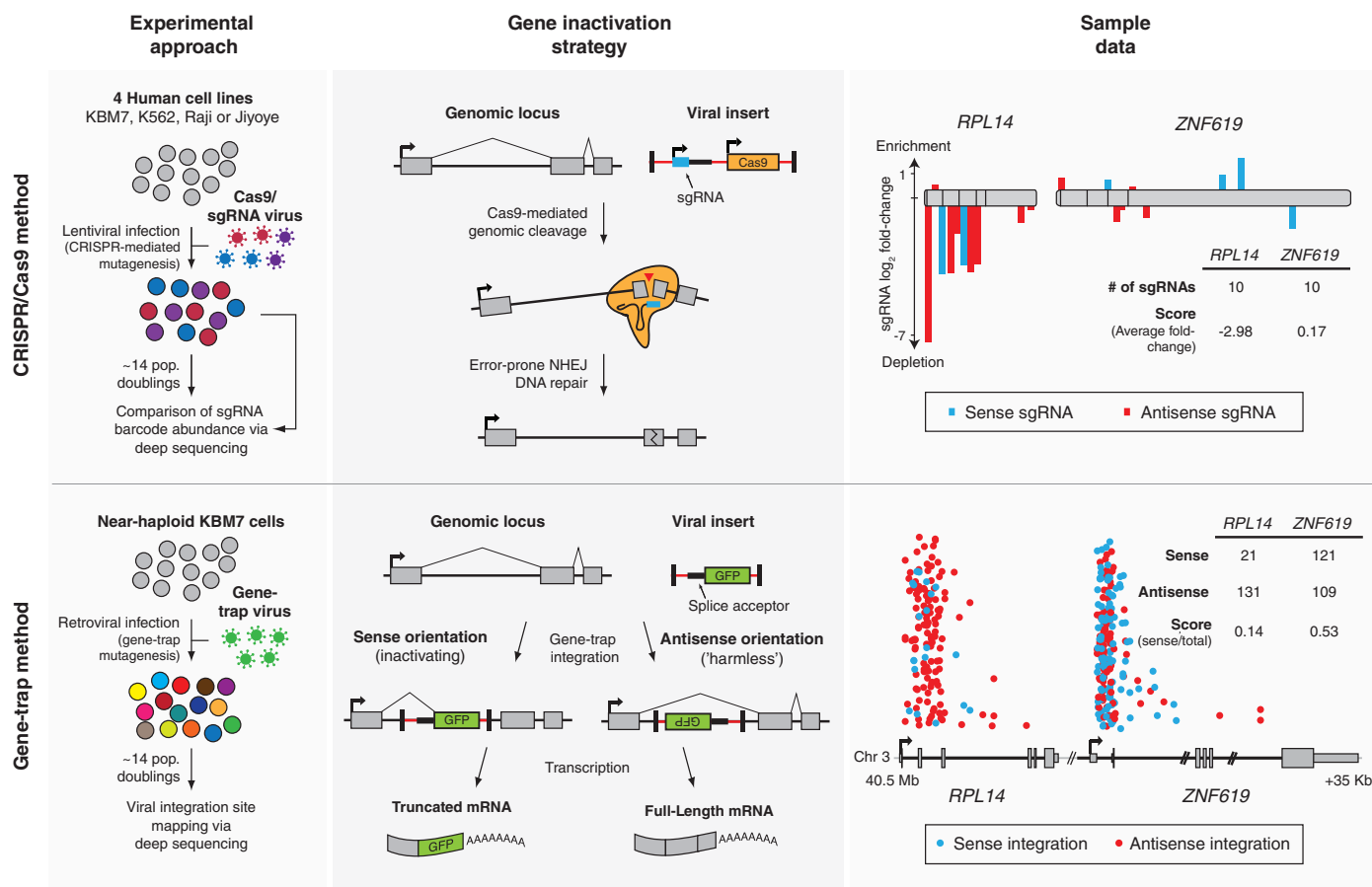
Fig. 1. Two approaches for genetic screening in human cells. (**Top**) CRISPR/Cas9 method. Cells are transduced with a genome-wide sgRNA lentiviral library. Gene inactivation via Cas9-mediated genomic cleavage is directed by the 20–base pair (bp) sequence at the 5′ end of the sgRNA. Cells bearing sgRNAs targeting essential genes are depleted in the final population. (**Bottom**) Gene-trap method. KBM7 cells are transduced with a gene-trap retrovirus that integrates in an inactivating or "harmless" orientation at random genomic loci. Essential genes contain fewer insertions in the inactivating orientation. Sample data for two neighboring genes—*RPL14*, encoding an essential ribosomal protein, and *ZNF619*, encoding a dispensable zinc finger protein—are displayed. For CRISPR/Cas9, sgRNAs are plotted according to their target position along each gene, with the height of each bar indicating the level of depletion. Boxes indicate individual exons. For gene trap, the intronic insertion sites in each gene are plotted according to their orientation and genomic position. The height of each point is randomized.

SCIENCE sciencemag.org

27 NOVEMBER 2015 • VOL 350 ISSUE 6264 **1097**

patterns were also observed in our CRISPR data set (Fig. 2, H and I).

In *S. cerevisiae*, genes with paralogous copies in the genome show a lower degree of essentiality, presumably because of at least partial functional overlap (*16*). Surprisingly, meta-analysis of knockout mouse collections has suggested that there is no such correlation in mammals (*17*, *18*). However, others have challenged this interpretation because the genes analyzed were far from a random sample (*19*). Using the results from our genome-wide screens, we revisited this question and observed that genes with paralogs are indeed less likely to be essential, which is consistent with the idea that paralogs can pro-

vide functional redundancy at the cellular level (Fig. 2J).

To examine the functions of the cell-essential genes, we used gene set enrichment analysis (GSEA) and found strong enrichment for many fundamental biological processes, such as DNA replication, RNA transcription, and mRNA translation (fig. S3A) (*20*). Whereas most of the genes could be assigned to such well-defined pathways, no function has been ascribed to ~330 of the cell-essential genes (18%) (Fig. 3A). For this set of uncharacterized genes, an analysis of the domains within their encoded gene products and comparisons with proteomic data sets from organellar purifications revealed substantial enrichment in

proteins found in the nucleolus and those containing domains associated with RNA processing (fig. S3, B and C) (*21*).

We characterized three such genes—*C16orf80*, *C3orf17*, and *C9orf114*—whose mRNA expression patterns across the Cancer Cell Line Encyclopedia (CCLE) were correlated with that of genes involved in RNA processing (Fig. 3B). We validated the essentiality of these genes in short-term proliferation assays and detected localization of their products to the nucleus (C16orf80) or nucleolus (C3orf17 and C9orf114) (Fig. 3, C and D) (*22*). Additionally, mass spectrometric analyses of anti-FLAG-immunoprecipitates prepared from KBM7 cells expressing FLAG-tagged C16orf80,
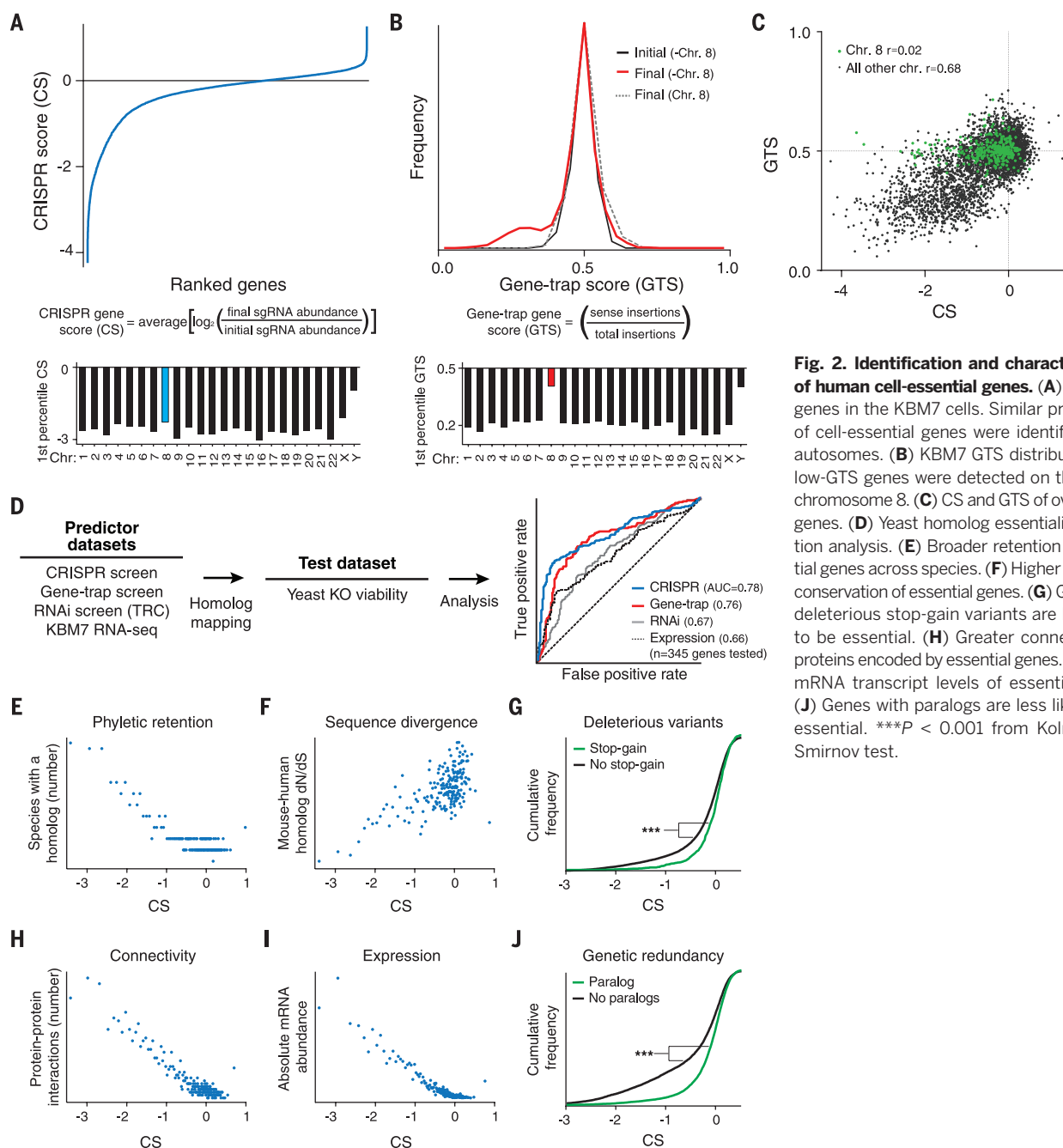
Fig. 2. Identification and characterization of human cell-essential genes. (**A**) CSs of all genes in the KBM7 cells. Similar proportions of cell-essential genes were identified on all autosomes. (**B**) KBM7 GTS distributions. No low-GTS genes were detected on the diploid chromosome 8. (**C**) CS and GTS of overlapping genes. (**D**) Yeast homolog essentiality prediction analysis. (**E**) Broader retention of essential genes across species. (**F**) Higher sequence conservation of essential genes. (**G**) Genes with deleterious stop-gain variants are less likely to be essential. (**H**) Greater connectivity of proteins encoded by essential genes. (**I**) Higher mRNA transcript levels of essential genes. (**J**) Genes with paralogs are less likely to be essential. ***P < 0.001 from Kolmogorov-Smirnov test.

C3orf17, and C9orf114 revealed interactions with multiple subunits of the spliceosome, ribonuclease (RNase) P/MRP, and H/ACA small nucleolar ribonucleoprotein (snoRNP) complexes, respectively (Fig. 3E). These results implicate *C16orf80* in splicing, which is consistent with its association with mRNAs; *C3orf17* in ribosomal RNA/tRNA processing; and *C9orf114* in RNA modification (*23*). More broadly, our results indicate that the molecular components of many critical cellular processes, especially RNA processing, have yet to be fully defined in mammalian cells.

To determine how the set of essential genes differs among cell lines, we screened another CML cell line (K562) and two Burkitt's lymphoma cell lines (Raji and Jiyoye) using the CRISPR system (tables S2 and S3). Overall, the sets of essential genes in the four cell lines showed a high degree of overlap (Fig. 4A). Out of these

four cell lines, the KBM7 CRISPR results showed the highest correlation with the KBM7 gene-trap data set, suggesting that the few differences observed are likely to be biologically meaningful (fig. S4A).

We focused first on genes found to be essential in only one of the four cell lines. The Raji, Jiyoye, and KBM7 cell lines had 6, 7, and 19 such genes, respectively (fig. S4B and table S5). One example was *DDX3Y*, which resides in the non-pseudoautosomal region of the Y chromosome and was required only in Raji cells (Fig. 4B). Its X-linked paralog, *DDX3X*, was essential in KBM7 and K562 cells (Fig. 4E). Both genes encode DEAD-box helicases that likely have similar cellular functions (*24*). Thus, the dependence on one paralog might reflect functional absence of the other paralog. Indeed, DNA sequencing of *DDX3X* in Raji cells revealed hemizygous muta-

tions in the 5′-splice site of intron 8 that resulted in the production of a truncated mRNA transcript (Fig. 4, C and D). Conversely, *DDX3Y* was not expressed in KBM7 cells and was not present in K562 cells, which are of female origin (Fig. 4E). Introduction of wild-type *DDX3X* cDNA into Raji cells fully rescued the proliferation defect resulting from *DDX3Y* loss, indicating that the paralogous genes are essential and functionally overlapping (Fig. 4F). Essential paralogous gene pairs, involved in glucose metabolism (*HK1/2* and *SLC2A1/3*) and cell-cycle regulation (*CDK4/6*), were also observed in the Jiyoye line (fig. S4C and supplementary text S4). Vulnerabilities due to the loss of a paralogous partner may serve as targets for highly personalized antitumor therapies (*25*).

In some cases, cell line–specific essentiality of paralogous genes did not reflect differential expression. For example, the transcription factors *GATA1* and *GATA2* are expressed in both K562 and KBM7 cells, but the first is specifically essential in K562 cells and the second in KBM7 cells (fig. S4D). These master regulators are known to promote proliferation and survival during distinct developmental stages in the hematopoietic lineage; *GATA1* is required for the survival of erythroid progenitors, and *GATA2* is required for the maintenance and proliferation of immature hematopoietic progenitors (*26*). These two cell types likely correspond to the cells of origin of the two CML lines (*27, 28*). We also identified similar instances of genes required for cell line–specific functions such as nuclear factor κB (NF-κB) pathway regulation and homology-directed DNA repair in Raji and KBM7 cells, respectively (fig. S4, E and F, and supplementary text S5).

Whereas the other three cell lines showed only a few cell line–specific essential genes, the K562 cell line was an outlier, with 63 such genes (fig. S4B). Oddly, these genes showed no discernible common functions, and many were not even expressed in the K562 cell line. (Additionally, a few encoded secreted factors whose loss would not be expected to be lethal in a pooled screen.) This mystery was resolved when we examined the chromosomal location of the genes: The majority (39 of the 63 genes) reside near 9q34 and 22q11. These two regions are translocated to produce the *BCR-ABL* oncogene and are present in a high-copy tandem amplification in K562 cells (fig. S5, A and B) (*29*). All 61 contiguous genes within the amplicon on 22q11 scored as essential, suggesting that sgRNA-mediated cleavage in this repeated region induces cytotoxicity in a manner unrelated to the function of the target genes themselves (Fig. 4G and fig. S5F). Indeed, two sgRNAs targeting nongenic sites within the amplicon were toxic to K562 but not KBM7 cells. They increased the abundance of phosphorylated histone H2AX (γH2AX), a marker of DNA damage, and induced erythroid differentiation, which occurs upon DNA damage in this cell line (Fig. 4, H to J, and fig. S5C) (*30*). We obtained similar results in another cell line, HEL, which contains a highly amplified
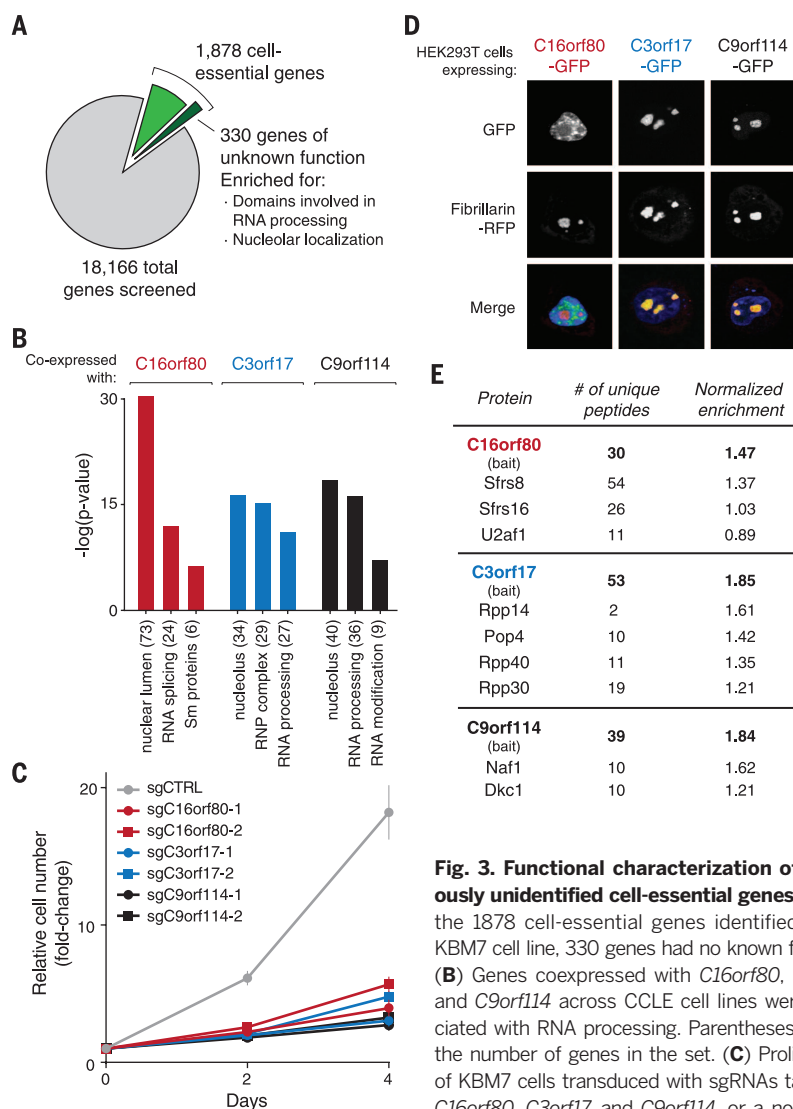
Fig. 3. Functional characterization of previously unidentified cell-essential genes. (**A**) Of the 1878 cell-essential genes identified in the KBM7 cell line, 330 genes had no known function. (**B**) Genes coexpressed with *C16orf80*, *C3orf17*, and *C9orf114* across CCLE cell lines were associated with RNA processing. Parentheses denote the number of genes in the set. (**C**) Proliferation of KBM7 cells transduced with sgRNAs targeting *C16orf80*, *C3orf17*, and *C9orf114*, or a nontargeting control. Error bars denote SD (n = 4 experiments per group). (**D**) C16orf80 localized to the nucleus, and C3orf17 and C9orf114 localized to the nucleolus. Fibrillarin–red fluorescent protein (RFP) was used as a nucleolar marker. GFP, green fluorescent protein. (**E**) Multiple subunits of the spliceosome, RNase P/MRP, and H/ACA ribonucleoprotein complexes interact with C16orf80, C3orf17, and C9orf114, respectively.
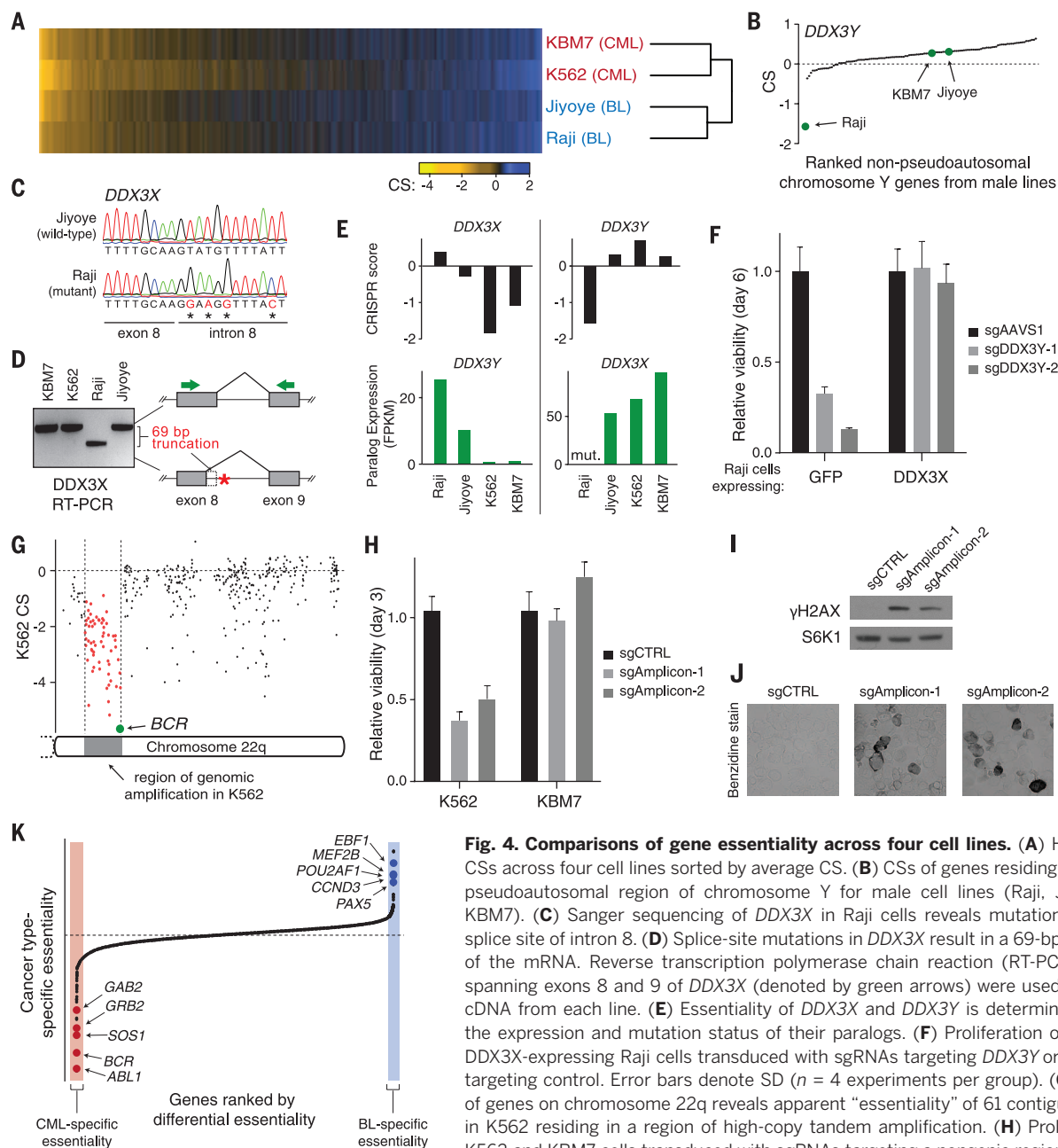
Fig. 4. Comparisons of gene essentiality across four cell lines. (**A**) Heatmap of CSs across four cell lines sorted by average CS. (**B**) CSs of genes residing in the non-pseudoautosomal region of chromosome Y for male cell lines (Raji, Jiyoye, and KBM7). (**C**) Sanger sequencing of *DDX3X* in Raji cells reveals mutations in the 5′ splice site of intron 8. (**D**) Splice-site mutations in *DDX3X* result in a 69-bp truncation of the mRNA. Reverse transcription polymerase chain reaction (RT-PCR) primers spanning exons 8 and 9 of *DDX3X* (denoted by green arrows) were used to amplify cDNA from each line. (**E**) Essentiality of *DDX3X* and *DDX3Y* is determined through the expression and mutation status of their paralogs. (**F**) Proliferation of GFP- and DDX3X-expressing Raji cells transduced with sgRNAs targeting *DDX3Y* or an AAVS1-targeting control. Error bars denote SD (*n* = 4 experiments per group). (**G**) Analysis of genes on chromosome 22q reveals apparent "essentiality" of 61 contiguous genes in K562 residing in a region of high-copy tandem amplification. (**H**) Proliferation of K562 and KBM7 cells transduced with sgRNAs targeting a nongenic region within the BCR-ABL amplicon in K562 cells or a nontargeting control. Error bars denote SD (*n* = 4 experiments per group). (**I**) γH2AX (phospho-S139 H2AX) immunoblot analysis of K562 transduced with sgRNAs as in (H). S6K1 was used as a loading control. (**J**) Cleavage within amplified region induced erythroid differentiation in K562 cells as assessed by means of 3,3′-dimethoxybenzidine hemoglobin staining. (**K**) Comparison of gene essentiality between the two cancer types reveals oncogenic drivers and lineage specifiers. Genes were ranked by the difference between the average CS of each cancer type.

region surrounding the *JAK2* tyrosine kinase (fig. S5, D and E). Together, these findings indicate that lethality upon Cas9-mediated cutting may also reflect chromosome structure and therefore should be evaluated in light of copy-number information.

Last, we looked for consistent differences in essential genes between the two CML and two Burkitt's lymphoma lines. Such genes might represent attractive targets for antineoplastic therapies because their inhibition is less likely to be broadly cytotoxic. Overall, we identified 33 genes that were specifically essential in the CML lines

and 15 genes in the Burkitt's lymphoma lines (Fig. 4K and table S6). As a control, permuted comparisons—that is, a set containing of one CML and one Burkitt's line versus the complementary sets—showed roughly half as many "set-specific" essential genes (fig. S6, A to C).

In the CML lines, the top two genes were *BCR* and *ABL1*, which is consistent with the known essentiality of the BCR-ABL translocation product and the therapeutic effect of BCR-ABL inhibitors such as imatinib (*31*). Additional members of the BCR-ABL signaling pathway—*SOS1*,

GRB2, and *GAB2*—scored strongly as well (ranked 3, 4, and 7, respectively). Network analysis of the other top hits also uncovered several genes encoding assembly factors for the electron transport chain, as well as enzymes involved in folate-mediated one-carbon metabolism. These results suggest additional potential targets for CML therapy (table S6).

In the B cell–derived Burkitt's lymphoma cell lines, the top genes included three B cell–lineage transcription factors *EBF1*, *POU2AF1*, and *PAX5* (ranked 3, 6, and 8, respectively). Each of these

genes is the target of recurrent translocations in lymphoma (*32–34*). Enhancers of the corresponding three gene loci all show a high level of bromodomain containing 4 (BRD4) occupancy in Ly1 cells, a related diffuse large B cell lymphoma cell line, suggesting bromodomain inhibitors such as JQ1 as potential treatments (*35*). Other selectively essential genes included *MEF2B*—a transcriptional activator of *BCL6*—and *CCND3*, both of which are frequently mutated and implicated in the pathogenesis of various lymphomas (*36*). Intriguingly, the top two hits, *CHM* and *RPP25L*, do not appear to have specific roles in B cells; rather, their differential essentiality is likely explained by the lack of expression of their paralogs, *CHML* and *RPP25*, in both of the Burkitt's lymphoma cell lines studied (fig. S6D).

We used two complementary and concordant approaches, CRISPR and gene trap, to define the cell-essential genes in the human genome. Although the gene-trap method is suitable only for loss-of-function screening in rare haploid cell lines, the CRISPR method is broadly applicable. Extending our analysis across different cell lines and tumor types, we developed a framework to assess differential gene essentiality and identify potential drivers of the malignant state. The method can be readily applied to more cell lines per cancer type so as to eliminate idiosyncrasies particular to a given cell line and to more cancer types so as to systematically uncover tumor-specific liabilities that might be exploited for targeted therapies.

REFERENCES AND NOTES

1. G. Giaever et al., *Nature* **418**, 387–391 (2002).
2. L. Cong et al., *Science* **339**, 819–823 (2013).
3. H. Wang et al., *Cell* **153**, 910–918 (2013).
4. T. Wang, J. J. Wei, D. M. Sabatini, E. S. Lander, *Science* **343**, 80–84 (2014).
5. O. Shalem et al., *Science* **343**, 84–87 (2014).
6. H. Koike-Yusa, Y. Li, E. P. Tan, M. C. Velasco-Herrera, K. Yusa, *Nat. Biotechnol.* **32**, 267–273 (2014).
7. J. E. Carette et al., *Science* **326**, 1231–1235 (2009).
8. J. E. Carette et al., *Nature* **477**, 340–343 (2011).
9. I. A. Tchasovnikarova et al., *Science* **348**, 1481–1485 (2015).
10. J. P. Kastenmayer et al., *Genome Res.* **16**, 365–373 (2006).
11. G. S. Cowley, B. A. Weir, W. C. Hahn, *Sci. Data* 10.1038/data.2014.35 (2014).
12. A. C. Wilson, S. S. Carlson, T. J. White, *Annu. Rev. Biochem.* **46**, 573–639 (1977).
13. Y. Ishihama et al., *BMC Genomics* **9**, 102 (2008).
14. H. Jeong, S. P. Mason, A. L. Barabási, Z. N. Oltvai, *Nature* **411**, 41–42 (2001).
15. T. Hart, K. R. Brown, F. Sircoulomb, R. Rottapel, J. Moffat, *Mol. Sys. Biol.* 10.15252/msb.20145216 (2014).
16. Z. Gu et al., *Nature* **421**, 63–66 (2003).
17. H. Liang, W.-H. Li, *Trends Genet.* **23**, 375–378 (2007).
18. B.-Y. Liao, J. Zhang, *Trends Genet.* **23**, 378–381 (2007).
19. T. Makino, H. Hokamp, A. McLysaght, *Trends Genet.* **25**, 152–155 (2009).
20. A. Subramanian et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 15545–15550 (2005).
21. Y. Ahmad, F.-M. Boisvert, E. Lundberg, M. Uhlen, A. I. Lamond, *Mol. Cell. Proteomics* **11**, 013680 (2012).
22. J. Barretina et al., *Nature* **483**, 603–607 (2012).
23. A. G. Baltz et al., *Mol. Cell* **46**, 674–690 (2012).
24. S. Sekiguchi, H. Iida, J. Fukumura, T. Nishimoto, *Exp. Cell Res.* **300**, 213–222 (2004).
25. F. L. Muller et al., *Nature* **488**, 337–342 (2012).
26. K. Ohneda, M. Yamamoto, *Acta Haematol.* **108**, 237–245 (2002).
27. L. C. Andersson, K. Nilsson, C. G. Gahmberg, *Int. J. Cancer* **23**, 143–147 (1979).
28. B. S. Andersson et al., *Leukemia* **9**, 2100–2108 (1995).
29. S. Q. Wu et al., *Leukemia* **9**, 858–862 (1995).
30. A. Constantinou, K. Kiguchi, E. Huberman, *Cancer Res.* **50**, 2618–2624 (1990).
31. B. Scappini et al., *Cancer* **100**, 1459–1471 (2004).
32. S. Iida et al., *Blood* **88**, 4110–4117 (1996).
33. H. Bouamar et al., *Blood* **122**, 726–733 (2013).
34. S. Galiègue-Zouitina et al., *C. R. Acad. Sci. III* **318**, 1125–1131 (1995).
35. B. Chapuy et al., *Cancer Cell* **24**, 777–790 (2013).
36. R. D. Morin et al., *Nature* **476**, 298–303 (2011).

GENOME EDITING

Genome-wide inactivation of porcine endogenous retroviruses (PERVs)

Luhan Yang,[1,2,3]*† Marc Güell,[1,2,3]† Dong Niu,[1,4]† Haydy George,[1]† Emal Lesha,[1] Dennis Grishin,[1] John Aach,[1] Ellen Shrock,[1] Weihong Xu,[6] Jürgen Poci,[1] Rebeca Cortazio,[1] Robert A. Wilkinson,[5] Jay A. Fishman,[5] George Church[1,2,3]*

The shortage of organs for transplantation is a major barrier to the treatment of organ failure. Although porcine organs are considered promising, their use has been checked by concerns about the transmission of porcine endogenous retroviruses (PERVs) to humans. Here we describe the eradication of all PERVs in a porcine kidney epithelial cell line (PK15). We first determined the PK15 PERV copy number to be 62. Using CRISPR-Cas9, we disrupted all copies of the PERV *pol* gene and demonstrated a >1000-fold reduction in PERV transmission to human cells, using our engineered cells. Our study shows that CRISPR-Cas9 multiplexability can be as high as 62 and demonstrates the possibility that PERVs can be inactivated for clinical application of porcine-to-human xenotransplantation.

Pig genomes contain from a few to several dozen copies of PERV elements (*1*). Unlike other zoonotic pathogens, PERVs cannot be eliminated by biosecure breeding (*2*). Prior strategies for reducing the risk of PERV transmission to humans have included small interfering RNAs (RNAi), vaccines (*3–5*), and PERV elimination using zinc finger nucleases (*6*) or TAL effector nucleases (*7*), but these have had limited success. Here we report the successful use of the CRISPR-Cas9 RNA-guided nuclease system (*8–10*) to inactivate all copies of the PERV *pol* gene and effect a 1000-fold reduction of PERV infectivity of human cells.

To design Cas9 guide RNAs (gRNAs) that specifically target PERVs, we analyzed the sequences of publicly available PERVs and other endogenous retroviruses in pigs (methods). Using droplet digital polymerase chain reaction (PCR), we identified a distinct clade of PERV elements (Fig. 1A) and determined that there were 62 copies of PERVs in PK15 cells (a porcine kidney epithelial cell line) (Fig. 1B). We then designed two Cas9 gRNAs that targeted the highly conserved catalytic center (*11*) of the *pol* gene on PERVs (Fig. 1C and fig. S1). The *pol* gene product functions as a reverse transcriptase (RT) and is thus essential for viral replication and infection. We determined that these gRNAs targeted all PERVs but no other endogenous retrovirus or other sequences in the pig genome (methods).

Initial experiments showed inefficient PERV editing when Cas9 and the gRNAs were transiently transfected (fig. S2). Thus, we used a PiggyBac transposon (*12*) system to deliver a doxycycline-inducible Cas9 and the two gRNAs

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]Wyss Institute for Biologically Inspired Engineering, Harvard University, Cambridge, MA, USA. [3]eGenesis Biosciences, Boston, MA 02115, USA. [4]College of Animal Sciences, Zhejiang University, Hangzhou 310058, China. [5]Transplant Infectious Disease and Compromised Host Program, Massachusetts General Hospital, Boston, MA 02115, USA. [6]Department of Surgery, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA.
*Corresponding author. E-mail: gchurch@genetics.med.harvard.edu (G.C.); luhan.yang@egenesisbio.com (L.Y.)
†These authors contributed equally to this work.

Fig. 1. CRISPR-Cas9 gRNAs were designed to specifically target the *pol* gene in 62 copies of PERVs in PK15 cells. (**A**) Phylogenetic tree representing endogenous retroviruses present in the pig genome. PERVs are highlighted in blue. (**B**) Copy number determination of PERVs in PK15 cells via digital droplet PCR. The copy number of *pol* elements was estimated to be 62, using three independent reference genes: *ACTB*, *GAPDH*, and *EB2*. n = 3 independent reference genes, mean ± SEM. (**C**) We designed two CRISPR-Cas9 gRNAs to target the catalytic region of the PERV *pol* gene. The two gRNA targeting sequences are shown below a schematic of PERV gene structure. Their PAM (protospacer adjacent motif) sequences are highlighted in red.
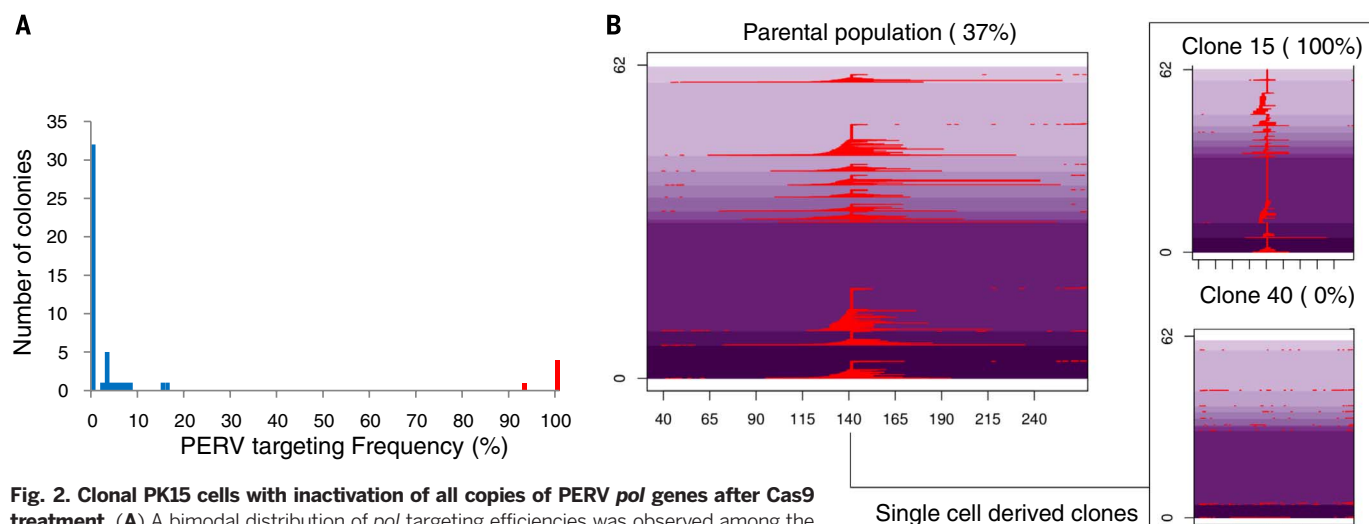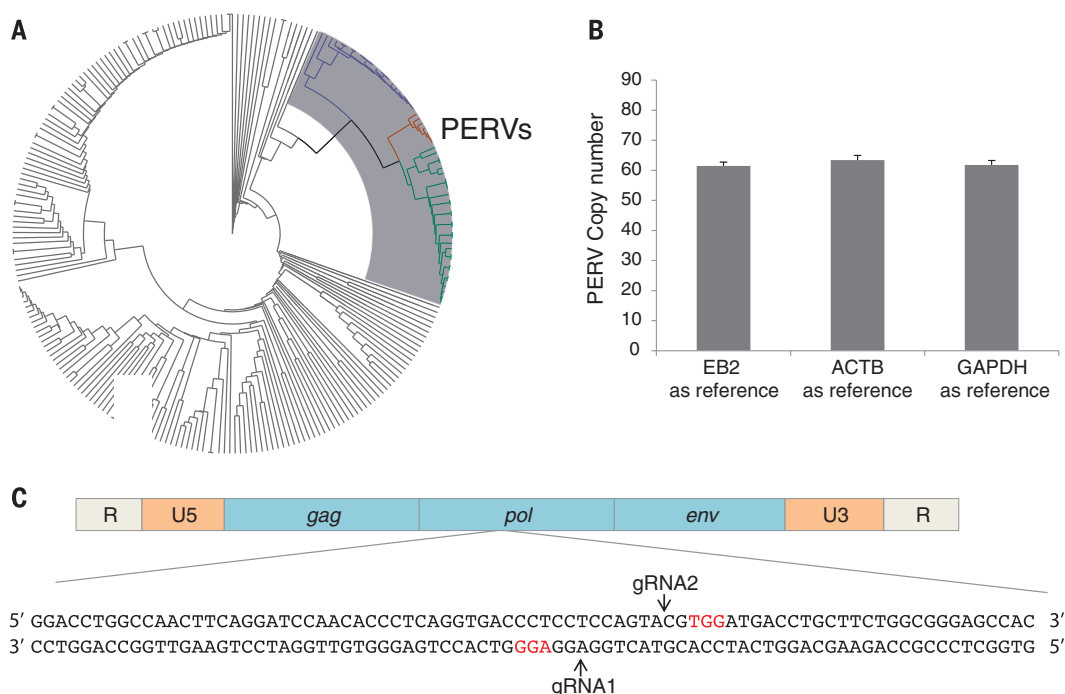
Fig. 2. Clonal PK15 cells with inactivation of all copies of PERV *pol* genes after Cas9 treatment. (**A**) A bimodal distribution of *pol* targeting efficiencies was observed among the single-cell–derived PK15 clones after 17 days of Cas9 induction. 45 out of 50 exhibited <16% targeting efficiency; 5 out of 50 clones exhibited >93% targeting efficiency. (**B**) PK15 haplotypes at PERV *pol* loci after CRISPR-Cas9 treatment. Indel events in the PERV *pol* sequence are represented in red. Shades of purple indicate endogenous PERVs.

into the genome of PK15 cells (figs. S2 and S3). Continuous induction of Cas9 led to increased targeting frequency of the PERVs (fig. S5), with a maximum targeting frequency of 37% (~23 PERV copies per genome) observed on day 17 (fig. S5). Neither higher concentrations of doxycycline nor prolonged incubation increased targeting efficiency (figs. S4 and S5), possibly because of the toxicity of nonspecific DNA damage by CRISPR-Cas9. Similar trends were observed when Cas9 was delivered using lentiviral constructs (fig. S6). We then genotyped the cell lines that exhibited maximal PERV targeting efficiencies. We observed 455 different insertion and deletion

(indel) events centered at the two gRNA target sites (Fig. 2B). Indel sizes ranged from 1 to 148 base pairs (bp); 80% of indels were small deletions (<9 bp). We validated the initial deep sequencing results with Sanger Sequencing (fig. S7).

We next sorted single cells from PK15 cells with high PERV targeting efficiency using flow cytometry, and we genotyped the *pol* locus of the resulting clones via deep sequencing (*13*, *14*). A repeatable bimodal (Fig. 2A and figs. S8 and S9) distribution was observed, with ~10% of the clones exhibiting high levels of PERV disruption (97 to 100%) and the remaining clones exhibiting

low levels of editing (<10%). We then examined individual indel events in the genomes of these clones (Fig. 2B and figs. S10 and S11). For the highly edited clones (clone 20, 100%; clone 15, 100%; clone 29, 100%; clone 38, 97.37%), we observed only 16 to 20 unique indel patterns in each clone (Fig. 2B and fig. S11). In addition, there was a much higher degree of repetition of indels within each clone than across the clones (fig. S25), suggesting a mechanism of gene conversion in which previously mutated PERV copies were used as templates to repair wild-type PERVs cleaved by Cas9 (Fig. 2B and fig. S25). Mathematical modeling of DNA repair during PERV

Fig. 3. PK15 cells with all PERV pol genes inactivated lose the infection capacity of human HEK 293 cell lines. (**A**) Detection of PERV *pol*, *gag*, and *env* DNA in the genomes of HEK 293–green fluorescent protein (GFP) cells after coculturing with PK15 cells for 5 and 7 days (293G5D and 393G7D, respectively). A pig GGTA1 primer set was used to detect pig cell contamination of the purified human cells. (**B**) Quantitative PCR (qPCR) analysis of the number of PERV elements in 1000 293G cells derived from a population cocultured with WT PK15 cells using specific primer sets (*n* = 3 qPCR experiment replicates, mean ±

SEM). (**C**) qPCR quantification of the number of PERV elements in PK15 clones 15, 20, 29, and 38, with high levels of PERV *pol* modification and minimally modified clones 40 and 41 (*n* = 3 qPCR experiment replicates, mean ± SEM). (**D**) Results of PCR on PERV *pol* on genomic DNA from various numbers of HEK 293–GFP cells (0.1, 1, 10, and 100) isolated from populations previously cultured with highly modified PK15 clone 20 and minimally modified clone 40 (see figs. S18 to S21 for a full panel of PCR reactions).
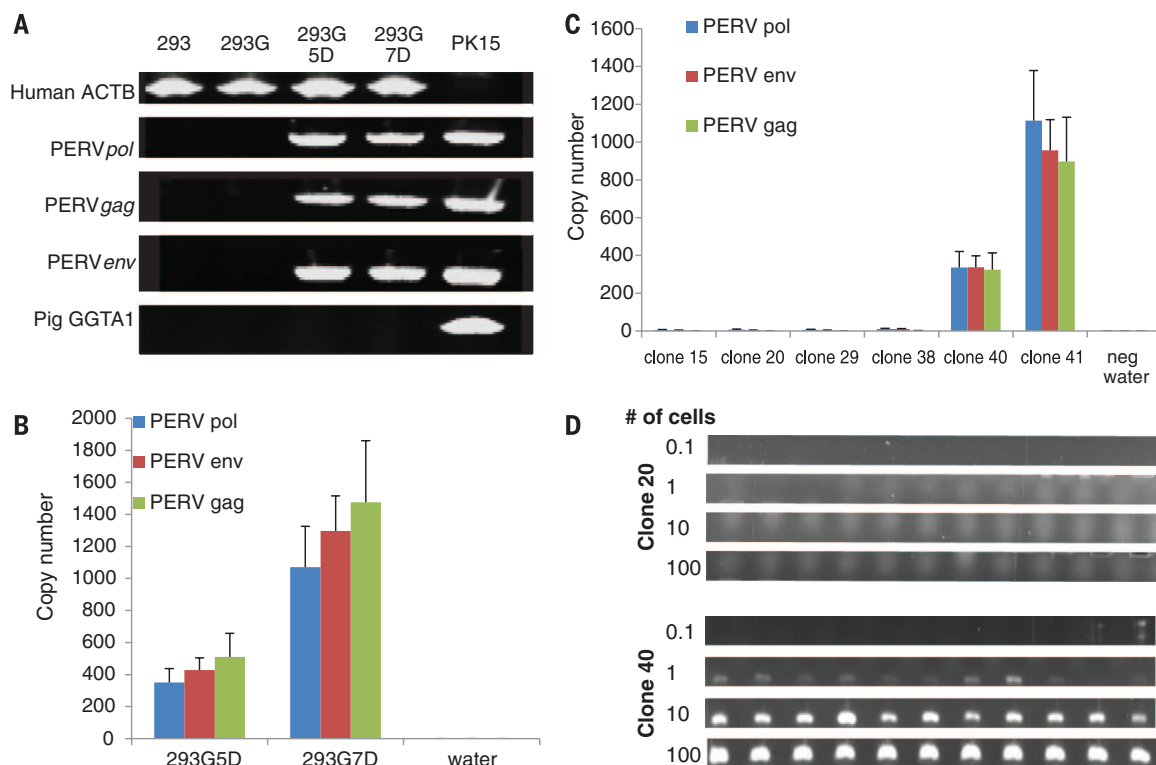
elimination (fig. S26) and analysis of expression data (figs. S22 to S24) supported this hypothesis and suggested that highly edited clones were derived from cells in which Cas9 and the gRNAs were highly expressed.

Next, we examined whether unexpected genomic rearrangements had occurred as a result of the multiplexed genome editing. Karyotyping of individual modified clones (figs. S12 to S14) indicated that there were no observable genomic rearrangements. We also examined 11 independent genomic loci with at most 2 bp mismatches to each of the intended gRNA targets and observed no nonspecific mutations (fig. S27). This suggests that our multiplexed Cas9-based genome engineering strategy did not cause catastrophic genomic instability.

Last, we examined whether our disruption of all copies of PERV *pol* in the pig genome could eliminate in vitro transmission of PERVs from pig to human cells. We could not detect RT activity in the cell culture supernatant of the highly modified PK15 clones (fig. S15), suggesting that modified cells only produced minimal amounts of PERV particles. We then cocultured wild-type (WT) and highly modified PK15 cells with human embryonic kidney (HEK) 293 cells to test directly for transmission of PERV DNA to human cells (*15*). After coculturing PK15 WT and HEK 293 cells for 5 days and 7 days (figs. S16 to S17), we detected PERV *pol*, *gag*, and *env* sequences in

the HEK 293 cells (Fig. 3A). We estimated the frequency of PERV infection to be approximately 1000 PERVs per 1000 human cells (Fig. 3B). However, PK15 clones with >97% PERV *pol* targeting exhibited up to 1000-fold reduction of PERV infection, producing results that were similar to background levels (Fig. 3C). We validated these results with PCR amplification of serial dilutions of HEK 293 cells that had a history of contact with PK15 clones (Fig. 3D and figs. S18 to S21). We could consistently detect PERVs in single HEK 293 cells isolated from the population cocultured with minimally modified clone 40, but we could not distinctly detect PERVs in 100 human cells from the population cocultured with highly modified clone 20. Thus, we concluded that the PERV infectivity of the engineered PK15 cells had been reduced by up to 1000-fold.

We successfully targeted the 62 copies of PERV *pol* in PK15 cells and demonstrated greatly reduced in vitro transmission of PERVs to human cells. Although in vivo PERV transmission to humans has not been demonstrated (*16*, *17*), PERVs are still considered risky (*18*, *19*), and our strategy could completely eliminate this liability. Because no porcine embryonic stem cells exist, this system will need to be recapitulated in primary porcine cells and cloned into animals by means of somatic cell nuclear transfer. Moreover, we achieved simultaneous Cas9 targeting of 62 loci

in single pig cells without salient genomic rearrangement. To our knowledge, the maximum number of genomic sites previously reported to be simultaneously edited has been six (*20*). Our methods thus open the possibility of editing other repetitive regions of biological significance.

REFERENCES AND NOTES

1. D. Lee *et al.*, *Anim. Biotechnol.* **22**, 175–180 (2011).
2. H.-J. Schuurman, *Xenotransplantation* **16**, 215–222 (2009).
3. J. Ramsoondar *et al.*, *Xenotransplantation* **16**, 164–180 (2009).
4. M. Semaan, D. Kaulitz, B. Petersen, H. Niemann, J. Denner, *Xenotransplantation* **19**, 112–121 (2012).
5. U. Fiebig, O. Stephan, R. Kurth, J. Denner, *Virology* **307**, 406–413 (2003).
6. M. Semaan, D. Ivanusic, J. Denner, *PLOS ONE* **10**, e0122059 (2015).
7. D. Dunn, M. DaCosta, M. Harris, R. Idriss, A. O'Brien, *FASEB J.* **29**, LB761 (2015).
8. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
9. P. Mali *et al.*, *Science* **339**, 823–826 (2013).
10. L. Cong *et al.*, *Science* **339**, 819–823 (2013).
11. S. G. Sarafianos *et al.*, *J. Mol. Biol.* **385**, 693–713 (2009).
12. M. H. Wilson, C. J. Coates, A. L. George Jr., *Mol. Ther.* **15**, 139–145 (2007).
13. L. Yang *et al.*, *Nucleic Acids Res.* **41**, 9049–9061 (2013).
14. M. Güell, L. Yang, G. M. Church, *Bioinformatics* **30**, 2968–2970 (2014).
15. C. Patience, Y. Takeuchi, R. A. Weiss, *Nat. Med.* **3**, 282–286 (1997).
16. W. Heneine *et al.*, *Lancet* **352**, 695–699 (1998).
17. J. H. Dinsmore, C. Manhart, R. Raineri, D. B. Jacoby, A. Moore, *Transplantation* **70**, 1382–1389 (2000).
18. D. Butler, *Nature* **391**, 320–324 (1998).
19. *ILAR J.* **38**, 49–51 (1997).
20. H. Wang *et al.*, *Cell* **153**, 910–918 (2013).

PROTEIN FOLDING

Cotranslational protein folding on the ribosome monitored in real time

Wolf Holtkamp,[1]* Goran Kokic,[1]* Marcus Jäger,[1] Joerg Mittelstaet,[1]†
Anton A. Komar,[2] Marina V. Rodnina[1]‡

Protein domains can fold into stable tertiary structures while they are synthesized on the ribosome. We used a high-performance, reconstituted in vitro translation system to investigate the folding of a small five-helix protein domain—the N-terminal domain of *Escherichia coli* N5-glutamine methyltransferase HemK—in real time. Our observations show that cotranslational folding of the protein, which folds autonomously and rapidly in solution, proceeds through a compact, non-native conformation that forms within the peptide tunnel of the ribosome. The compact state rearranges into a native-like structure immediately after the full domain sequence has emerged from the ribosome. Both folding transitions are rate-limited by translation, allowing for quasi-equilibrium sampling of the conformational space restricted by the ribosome. Cotranslational folding may be typical of small, intrinsically rapidly folding protein domains.

I n living cells, folding of many proteins begins cotranslationally as soon as the N-terminal part of a given protein emerges from the peptide exit tunnel of the ribosome (*1–3*); secondary structure elements, such as α helices, can form within the exit tunnel (*4–7*). Whereas small- and medium-size protein domains can acquire their native structures in less than a second (*8*), the bacterial ribosome requires 5 to 10 s to synthesize a protein domain 100 amino acids in length. Cotranslational folding may be attenuated by interactions of the nascent peptide with the ribosome (*9*) and by auxiliary proteins, such as chaperones or other protein biogenesis factors (*10–12*). Changes in local translational kinetics, such as those caused by rare codons or transfer RNA (tRNA) abundance, can influence the conformation of newly synthesized proteins (*13, 14*). Little is known about the exact timing of cotranslational protein folding in relation to protein synthesis or the conformation of the polypeptide emerging from the ribosome.

To monitor cotranslational protein folding during ongoing translation, we used a reconstituted in vitro translation system in combination with the selective site-specific labeling of nascent proteins with fluorescent probes (*15*). We studied the N-terminal domain (NTD) of *Escherichia coli* N5-glutamine methyltransferase HemK. The HemK NTD consists of five helices (residues 3 to 12, 20 to 29, 35 to 42, 48 to 65, and 67 to 73) (Fig. 1) connected by a linker (residues 72 to 96) to the C-terminal domain (CTD) (*16*). The isolated HemK NTD (residues 1 to 73) forms a stable α-helical structure independent of the CTD (figs. S1 and S2). Free NTD in solution folds on a (sub)millisecond time scale in a predominantly two-state fashion (fig. S3 and table S1). For in vitro translation, we used an mRNA construct coding for the N-terminal 112 amino acids of HemK (HemK112) comprising the NTD (73 amino acids) plus 39 amino acids from the linker and the CTD. This length should allow the NTD to fully emerge from the ribosome exit tunnel when HemK112 is synthesized (Fig. 1).

We rapidly mixed synchronized initiation complexes containing BodipyFL (BOF)–Met-tRNA^{fMet} with elongation factors (EF-Tu, EF-Ts, and EF-G) and purified aminoacyl-tRNAs. Instead of Lys-tRNA, we added εNH$_2$-Bodipy576/589 (BOP)–Lys-tRNA^{Lys} to introduce the second fluorescence label at position 34 of the nascent peptide (Fig. 1). The two reporters provide a donor-acceptor pair to monitor Förster resonance energy transfer (FRET). Donor- and acceptor-only controls served to correct the time courses measured in the presence of both donor and acceptor (fig. S4A). The appearance of FRET after 10 s (Fig. 2A and fig. S4C) implies the formation of a structure in which positions 1 and 34 of HemK112 come into close proximity to one another. When compared to the time course of translation, the band representing the full-length protein appears after 40 s (Fig. 2B and fig. S5). Thus, chain compaction begins earlier than the full-length domain emerges from the ribosome.

We translated HemK peptides of different lengths ranging from 42 to 112 residues (Figs. 1A and 2B) and derived a translation velocity of 3.6 ± 0.1 amino acids/s (Fig. 2B and fig. S5B). Translation of HemK42 results in a low final FRET efficiency, reflecting an extended conformation of short nascent chains within the exit tunnel (Fig. 2, B and C). Synthesis of HemK56 leads to the formation of a high-FRET state, which suggests that the two probes come close—presumably by compaction of the nascent peptide—even though BOP-Lys on the nascent HemK56 chain should be occluded inside the tunnel, even when the peptide is in a fully extended conformation (Fig. 1A). Thus, the structure manifested by high FRET must form largely inside the ribosome, close to the exit of the peptide tunnel. With the HemK70 construct, the FRET efficiency is even higher, and the time of FRET appearance coincides with the synthesis of the 70–amino acid product (Fig. 2B). For all constructs, FRET rapidly increases after ~10 s of translation, whereas the formation of full-length HemK84, 98, and 112 is markedly slower than the appearance of the high-FRET state.

The final FRET efficiency is somewhat lower for HemK84, 98, and 112 relative to HemK70, indicating that there are structural rearrangements when the peptide chain becomes longer than 70 amino acids (Fig. 2C). With fully translated HemK112, the entire 73–amino acid NTD is likely to emerge from the exit tunnel, and thus the lower FRET level reflects the end state of folding. In contrast, the conformation captured by using the HemK70 construct does not represent a native-like fold, but rather a compact state that is stabilized by the ribosome until the next part of the protein sequence emerges. When we released the nascent HemK70 peptide from the ribosome by puromycin treatment, the FRET efficiency decreased to the level measured with the longer constructs (Fig. 2C and fig. S6), reflecting spontaneous domain folding (fig. S3). The decrease of FRET in HemK112 is attributable to the stabilization of the N terminus: BOP-Lys^{34} is in close proximity to Tyr^3, which can quench BOP fluorescence (*17, 18*). When we substituted Tyr^3 with Phe, we recovered the high FRET signal (Fig. 2C). These data indicate the existence of two states along the folding pathway of the NTD: a compact state, formed early during translation, which rearranges into a near-native fold upon further translation.

We probed the folding of HemK nascent peptides of different lengths by limited proteolysis with thermolysin, which cleaves at sites with bulky and aromatic residues (*19*), and monitored the

[1]Department of Physical Biochemistry, Max Planck Institute for Biophysical Chemistry, 37077 Göttingen, Germany. [2]Center for Gene Regulation in Health and Disease and Department of Biological, Geological and Environmental Sciences, Cleveland State University, Cleveland, OH 44115, USA.
*These authors contributed equally to this work. †Present address: Miltenyi Biotec GmbH, 51429 Bergisch Gladbach, Germany.
‡Corresponding author. E-mail: rodnina@mpibpc.mpg.de

cleavage of the BOF-labeled N-terminal peptide. In addition to HemK wild-type constructs, we used a mutated NTD in which four conserved Leu residues, comprising the hydrophobic core of the NTD, were simultaneously replaced with Ala (4×A) (Fig. 3A). The NTD with even a single Ala mutation was almost completely unfolded at 37°C (Fig. 3B). HemK70 nascent chains (wild-type or 4×A) were largely protected from protease digestion (Fig. 3C and fig. S7), because the N-terminal BOF resides inside or in close proximity to the peptide exit tunnel, which protects it from proteolytic cleav-

age. When released from the ribosome by puromycin treatment, HemK70 4×A peptides were rapidly digested, whereas wild-type peptides were more resistant to proteolysis (Fig. 3C and fig. S7B). HemK84 nascent peptides (wild-type or 4×A) were rapidly digested, indicating that the N-terminal parts of the peptides are exposed and do not adopt a protease-resistant conformation. In contrast, longer wild-type nascent peptides were significantly more protease-resistant than the respective 4×A mutants (figs. S7 and S8). The kinetics of proteolysis was similar for HemK112

nascent chains and HemK70 peptides released by puromycin (Fig. 3C and fig. S7), supporting the notion that HemK112 attains a native-like structure similar to the one that forms after release of nascent chains from the ribosome.

Together, the results of time-resolved FRET and limited proteolysis suggest that cotranslational folding proceeds through a compact state that is formed early during peptide elongation, when the nascent peptides are still confined in the exit tunnel. The latter intermediate state converts to the native-like fold upon emergence of the entire
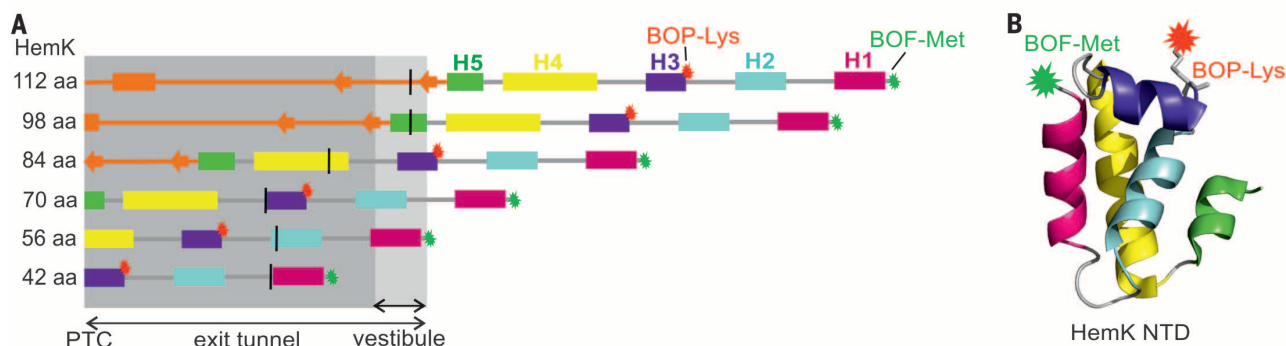
Fig. 1. Schematics of the model protein HemK NTD. (**A**) Secondary structure elements of HemK; helices H1 to H5 are shown as bars assuming fully α-helical conformation (1.5 Å per residue). Green and red stars indicate, respectively, the positions of the BOF-Met and BOP-Lys dyes in the proteins. Dark and light gray shaded areas indicate the positions of the ribosome exit tunnel and the vestibule relative to the nascent peptide if all secondary structures were formed. Vertical bars delineate boundaries of the exit tunnel assuming a fully extended conformation of the nascent peptide (3.5 Å per residue). Numbers of amino acids (42 to 112 aa) correspond to lengths of the nascent peptides. PTC, peptidyl transferase center. (**B**) Crystal structure of the HemK NTD (PDB ID 1T43). Secondary structure elements H1 to H5 are color-coded as in (A).
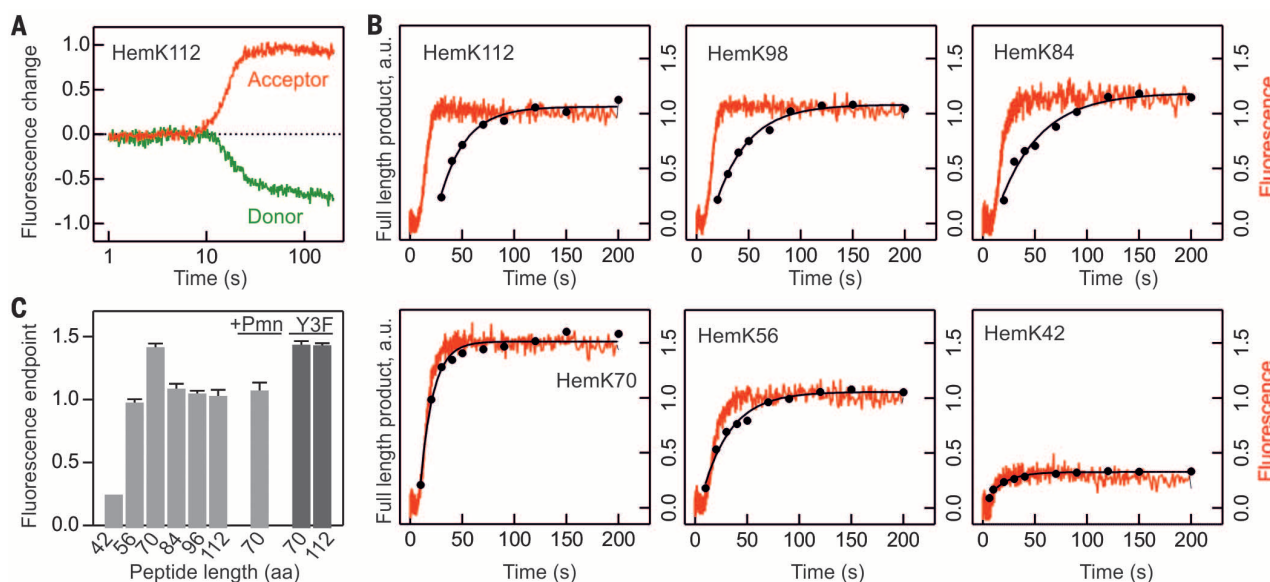
Fig. 2. Cotranslational folding monitored by FRET. (**A**) Changes in BOF-Met[1] (donor, green) and BOP-Lys[34] (acceptor, red) fluorescence corrected for bleed-through and fluorescence changes unrelated to folding (see supplementary materials). Translation was carried out in HiFi buffer at 37°C. (**B**) Time courses of translation, derived from SDS–polyacrylamide gel electrophoresis experiments (circles) and normalized FRET changes (acceptor; red lines), for HemK constructs of different lengths. FRET traces are normalized to the maximum fluorescence change (endpoint) measured for HemK112; a.u., arbitrary units. (**C**) Endpoints of FRET change (acceptor; light gray bars) for HemK nascent peptides of different lengths. +Pmn, FRET in HemK70 peptide released from the ribosome by puromycin treatment; dark gray bars, nascent peptides (chain lengths 70 or 112) with replacement Y3F. Error bars indicate SD ($n = 12$ replicates).

domain from the peptide exit tunnel or after the release of the nascent chain from the ribosome.

We used photoinduced electron transfer (PET) between the N-terminal BOF-Met and Trp residues within the nascent chain, which is ideally suited to study conformational rearrangements on short length scales (20). We introduced single Trp residues in different positions of HemK (6W, F38W, D49W) (Fig. 4A) and monitored the change in BOF-Met fluorescence during translation and movement of the nascent chain through the exit tunnel. We extracted the changes reflecting nascent chain folding by correcting for the fluorescence change of BOF-Met in a HemK construct lacking any intrinsic Trp residue (fig. S9).

Upon synthesis of the peptide that contains Trp at position 6 (6W), the PET efficiency changes in a multiphasic way (Fig. 4B). The initial phases are independent of the chain length or mutations in the protein core (4×A). After about 20 s of translation, the PET efficiency for the various constructs starts to deviate. Nascent HemK70 (wild-type and

4×A) or the 4×A variants of HemK112 adopt a high-PET conformation. In contrast, native HemK112 adopts a low-PET state. Comparison of the fluorescence endpoints for different peptide lengths (Fig. 4C) shows that the PET efficiency is low when the nascent chain is short (35 and 42 amino acids), probably because the peptide exit tunnel restricts chain dynamics, thereby inhibiting BOF-Trp interactions. Upon arrival at the end of the exit tunnel (56 to 84 amino acids), the PET efficiency peaks at 13 to 18% and then decreases with increasing peptide length. With the further increase in length and extrusion of the folded domain from the exit tunnel, BOF-Met becomes shielded from fluorescence-quenching interactions with the Trp residue. In contrast, in the intermediate or unfolded state, Trp is solvent-exposed and accessible, consistent with the high PET observed with the 4×A mutants of any length (Fig. 4C). These data demonstrate the existence of a transient compact state that rearranges into the native state when the nascent chain reaches an appropriate length.

To better define the timing of native-state formation, we carried out PET experiments using constructs carrying Trp engineered at positions 38 or 49 of the HemK NTD (Fig. 4, A and D, and fig. S2). With Trp at position 38, a high-PET intermediate emerges as the nascent chain length reaches 98 amino acids, somewhat later than with Trp at position 6 (Fig. 4, C and E). Further translation results in a rearrangement (PET decrease) that leads to the final state. With Trp at position 49, the transient folding state is not observed; rather, PET reports the formation of the final native-like structure (Fig. 4E). In that case, the time course of folding coincides with the synthesis of the full-length HemK112 (Fig. 4D), which should be sufficient to extrude all helices required to fold into the NTD just outside the ribosome exit tunnel, and does not change further with the increase of the chain length up to 154 amino acids.

Assuming that residues 74 to 112 that do not belong to the NTD adopt a fully extended conformation (~3.5 Å per residue) within the ribosome exit tunnel (length ~100 Å), the folded nascent domain would reside ~33 Å away from the ribosome exit tunnel. However, for HemK98, which shows a fold similar to that of nascent HemK112, the same calculation suggests that the native-like fold may assemble even before the domain is fully released from the tunnel, as the distance within the ribosome covered by residues 74 to 98 is only 84 Å—shorter than the tunnel length, but within the area attributed to the tunnel vestibule (21, 22) (Fig. 1A). The difference in the environment or local folding of HemK98 and HemK112 is reported by PET from BOF to Trp[38] (Fig. 4E). Thus, the compact state may rearrange to the native-like fold already in the tunnel vestibule, just before the full domain emerges from the exit tunnel. When HemK70 folding was followed after the release of the compacted nascent chains into solution by puromycin treatment, PET changed in exactly
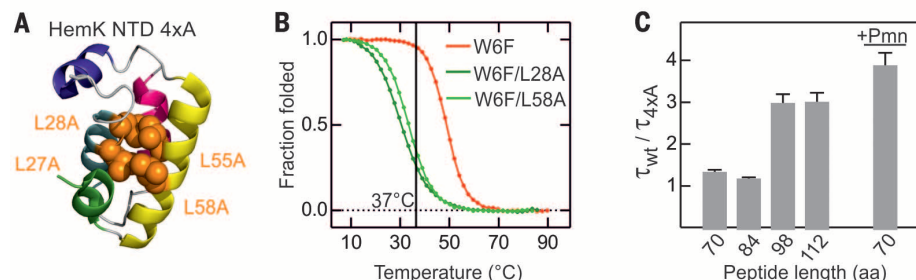
Fig. 3. Probing the folding status by proteolysis. (A) Positions of the Leu → Ala replacements in the protein core of the HemK 4×A construct. (B) The stability of the isolated NTD in solution measured by far-ultraviolet circular dichroism spectra at different temperatures. (C) Relative stability against thermolysin digestion of the native (τ_{wt}) and 4×A ($\tau_{4×A}$) NTD on the ribosome and after release into solution by puromycin treatment (+Pmn). Error bars indicate SEM of the fits.
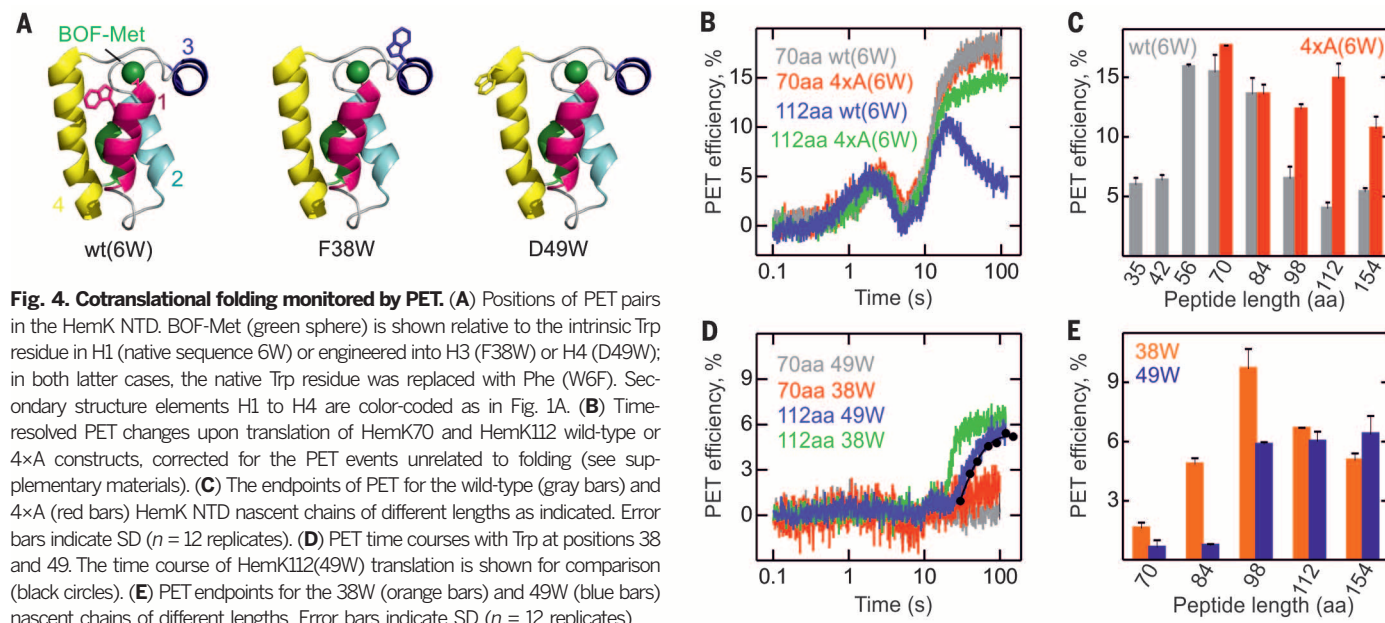
Fig. 4. Cotranslational folding monitored by PET. (A) Positions of PET pairs in the HemK NTD. BOF-Met (green sphere) is shown relative to the intrinsic Trp residue in H1 (native sequence 6W) or engineered into H3 (F38W) or H4 (D49W); in both latter cases, the native Trp residue was replaced with Phe (W6F). Secondary structure elements H1 to H4 are color-coded as in Fig. 1A. (B) Time-resolved PET changes upon translation of HemK70 and HemK112 wild-type or 4×A constructs, corrected for the PET events unrelated to folding (see supplementary materials). (C) The endpoints of PET for the wild-type (gray bars) and 4×A (red bars) HemK NTD nascent chains of different lengths as indicated. Error bars indicate SD (n = 12 replicates). (D) PET time courses with Trp at positions 38 and 49. The time course of HemK112(49W) translation is shown for comparison (black circles). (E) PET endpoints for the 38W (orange bars) and 49W (blue bars) nascent chains of different lengths. Error bars indicate SD (n = 12 replicates).

the same way as upon protein folding on the ribosome (Fig. 4 and fig. S10). In contrast, release of HemK112 did not lead to any additional PET changes, which suggests that the final conformation of HemK112 on the ribosome and in solution after folding is very similar.

Our results provide an insight into nascent protein folding on the ribosome in real time (fig. S11). When the peptide reaches a length of about 56 to 70 amino acids, the nascent chain becomes compact, at a stage of translation when parts of the nascent peptide are still enclosed in the exit tunnel of the ribosome—possibly in the exit tunnel vestibule, which can accommodate a substantial degree of protein structure (e.g., tertiary hairpins) (7, 22). If such a state formed off the ribosome, it was too short-lived to be captured by stopped-flow experiments. The ribosome may induce an alternative folding pathway, or it may stabilize an arrangement that is hardly sampled in solution.

Retention of compact or intermediate states may represent a fundamental feature of cotranslational folding that acts to prevent the chain from falling into kinetic traps, such as stably misfolded non-native conformations that may appear when only part of a protein has been synthesized. Owing to the slow pace of translation, folding of intrinsically rapidly folding domains appears to exhibit equilibrium-like properties (21) with a landscape of accessible conformations restricted by the environment of the exit tunnel. Our findings show how the ribosome can, in principle, define the pathway for cotranslational folding.

REFERENCES AND NOTES

1. Y. Han et al., Proc. Natl. Acad. Sci. U.S.A. 109, 12467–12472 (2012).
2. S. T. Hsu et al., Proc. Natl. Acad. Sci. U.S.A. 104, 16516–16521 (2007).
3. C. Eichmann, S. Preissler, R. Riek, E. Deuerling, Proc. Natl. Acad. Sci. U.S.A. 107, 9111–9116 (2010).
4. C. A. Woolhead, P. J. McCormick, A. E. Johnson, Cell 116, 725–736 (2004).
5. J. Lu, C. Deutsch, Nat. Struct. Mol. Biol. 12, 1123–1129 (2005).
6. S. Bhushan et al., Nat. Struct. Mol. Biol. 17, 313–317 (2010).
7. O. B. Nilsson et al., Cell Rep. 12, 1533–1540 (2015).
8. J. Kubelka, T. K. Chiu, D. R. Davies, W. A. Eaton, J. Hofrichter, J. Mol. Biol. 359, 546–553 (2006).
9. C. M. Kaiser, D. H. Goldman, J. D. Chodera, I. Tinoco Jr., C. Bustamante, Science 334, 1723–1727 (2011).
10. Y. E. Kim, M. S. Hipp, A. Bracher, M. Hayer-Hartl, F. U. Hartl, Annu. Rev. Biochem. 82, 323–355 (2013).
11. F. Gloge, A. H. Becker, G. Kramer, B. Bukau, Curr. Opin. Struct. Biol. 24, 24–33 (2014).
12. S. Pechmann, F. Willmund, J. Frydman, Mol. Cell 49, 411–421 (2013).
13. A. A. Komar, Trends Biochem. Sci. 34, 16–24 (2009).
14. G. Zhang, Z. Ignatova, Curr. Opin. Struct. Biol. 21, 25–31 (2011).
15. J. Mittelstaet, A. L. Konevega, M. V. Rodnina, J. Am. Chem. Soc. 135, 17031–17038 (2013).
16. Z. Yang et al., J. Mol. Biol. 340, 695–706 (2004).
17. H. Chen, S. S. Ahsan, M. B. Santiago-Berrios, H. D. Abruña, W. W. Webb, J. Am. Chem. Soc. 132, 7244–7245 (2010).
18. N. Marmé, J. P. Knemeyer, M. Sauer, J. Wolfrum, Bioconjug. Chem. 14, 1133–1139 (2003).
19. B. Keil, Specificity of Proteolysis (Springer-Verlag, Berlin, 1992).
20. S. Doose, H. Neuweiler, M. Sauer, ChemPhysChem 10, 1389–1398 (2009).
21. E. P. O'Brien, J. Christodoulou, M. Vendruscolo, C. M. Dobson, J. Am. Chem. Soc. 133, 513–526 (2011).
22. A. Kosolapov, C. Deutsch, Nat. Struct. Mol. Biol. 16, 405–411 (2009).

myIDP: A career plan customized for you, by you.

For your career in science, there's only one **Science**

my IDP
Science Careers

Recommended by leading professional societies and the NIH

Features in myIDP include:

- Exercises to help you examine your skills, interests, and values.

- A list of 20 scientific career paths with a prediction of which ones best fit your skills and interests.

- A tool for setting strategic goals for the coming year, with optional reminders to keep you on track.

- Articles and resources to guide you through the process.

- Options to save materials online and print them for further review and discussion.

- Ability to select which portion of your IDP you wish to share with advisors, mentors, or others.

- A certificate of completion for users that finish myIDP.

Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers In partnership with:

FASEB
advancing the life sciences
1912-2012
Federation of American Societies for Experimental Biology

UCSF
University of California
San Francisco

MEDICAL COLLEGE OF WISCONSIN

BURROUGHS WELLCOME FUND

Join AAAS. Get instant access to *Science*. Support all of the sciences.

When you subscribe to *Science*, you become part of the American Association for the Advancement of Science (AAAS), a nonprofit community of more than 100,000 members worldwide who believe in the power of science to make the world a better place. AAAS is hard at work promoting science in government, schools, and in the public commons around the globe.

AAAS's award-winning journal *Science* offers the top peer-reviewed research across multiple disciplines. With your subscription, you'll get:

- 51 weeks of home delivery of *Science*
- Instant online retrieval of every *Science* article ever published, dating back to 1880
- Full access to the *Science* mobile site and apps
- Career advice, webinars, blogs and fascinating features exclusively for AAAS members
- Members-only newsletters, and much more

With increasing public skepticism about science–and public funding for research more uncertain than ever–our work has never been more important. Join hands with us today!

Visit promo.aaas.org/joinaaas. Together, we can make a difference.

Science
AAAS

Protein Analysis Kits

A new tool for protein digestion is designed to substantially reduce preparation times and associated errors, facilitating fast, reliable analytical results. Current protein digestion techniques comprise complex procedures requiring determination of protein concentration followed by sample reduction and alkylation prior to digestion with trypsin. To improve this operation, the SMART Digest kit uses a simple three-step, easy-to-implement process that can be automated. Its immobilized trypsin design is formulated for high-quality data with a significantly reduced sample preparation time compared to traditional in-solution digestion methods. The increased reproducibility of digestion possible with this kit should enable users to implement generic, rapid, and robust methods for high-throughput processing of samples. Furthermore, the SMART Digest kit is designed to provide higher sequence coverage while reducing both chemically induced posttranslational modifications and the number of chemicals in the final sample.

Thermo Fisher Scientific
For info: 800-556-2323
www.thermoscientific.com/smartdigest

SPE Columns

New polymer-based solid phase extraction (SPE) columns have been added to the tried and tested EVOLUTE EXPRESS line. EVOLUTE EXPRESS is a novel family of plates, and now columns, that simplify polymer-based SPE by employing a streamlined three-step protocol: load-wash-elute. By using this approach and removing the conditioning and equilibration steps necessary with traditional bioanalytical SPE procedures, processing and method development times are dramatically reduced without loss of analyte recovery or method robustness. The EVOLUTE EXPRESS design provides a uniform flow in the processing of aqueous samples and eliminates the need for reruns due to clogging. This clear-flow technology makes EVOLUTE EXPRESS particularly efficient and effective when used with SPE automated systems (such as the Biotage Extrahera), and enables consistent, rapid, and reliable manual processing. EVOLUTE EXPRESS columns are packed with polymer-based EVOLUTE chemistries (ABN, CX, WCX, AX, and WAX) to address most application needs in a wide range of laboratories.

Biotage
For info: 800-446-4752
www.biotage.com

Liquid Chromatograph System

The new ACQUITY Arc System is a quaternary liquid chromatograph that gives analytical laboratories running established liquid chromatography (LC) methods a clear choice for replicating or improving their separations performance. The system is specifically engineered to respond to the needs of analytical scientists for a single LC platform that can enable them to efficiently transfer, adjust, or improve their methods regardless of the LC platform on which they were developed. To date, scientists working with established methods haven't had an LC platform versatile enough to bridge the gap between high-performance liquid chromatography (HPLC) and ultra-high-performance liquid chromatography (UHPLC). With the introduction of the ACQUITY Arc System and its enabling Arc Multi-flow path technology, scientists now have the ability to emulate the gradient dwell volume and mixing behavior of various LC systems. When the appropriate fluidic path is selected, the ACQUITY Arc System can easily emulate a variety of HPLC systems without altering the method's gradient table, or provide UHPLC performance with the flip of a switch.

Waters
For info: 800-252-4752
www.waters.com/arc

Biomarker Services

Our new Biomarker Services include biomarker screening and assay development services, providing the customer a clear path from early discovery and exploratory studies to validated products. Following biomarker identification, the assays are easily transferred to a customer's own lab or a contract research organization (CRO) to continue validation studies, allowing for consistency across the entire study. This ensures the reproducibility and accuracy of results and speeds up the biomarker research cycle. These biomarker services leverage Meso Scale Diagnostic's Multi-Array technology, which has been widely adopted by researchers worldwide due to its high sensitivity, excellent reproducibility, and wide dynamic range. By utilizing the U-PLEX, V-PLEX, and S-PLEX platforms, each custom service can deliver rapid multiplex development, rigorous assay validation, and ultimate sensitivity to measure previously unmeasurable biomarkers, depending on user requirements. With the addition of over 80 human, mouse, and nonhuman primate assays, the U-PLEX platform will continue to help customers quickly and easily design and build personalized multiplex assays.

Meso Scale Diagnostics
For info: 240-314-2795
www.mesoscale.com

Single-Cell T-Cell Receptor Assay

The single cell T-cell receptor (TCR) panel for Precise assays enables a fast, simple, and high-throughput method for sequencing thousands of single cells. It is the first commercial assay to enable the sequencing of the alpha and beta chain within the same cell, 96 or 384 cells at a time. The TCR panel may be combined with other standard or custom gene expression panels to enable the correlation of TCR sequences to function in individual T cells. Based on patented Molecular Indexing technology, the Precise assays offer absolute quantization and bias-free expression information. The assays allow the examination of large numbers of standard or low-input messenger RNA (mRNA) samples from precious samples or whenever absolute quantitation is required. The Precise assays combine molecular and sample indexing in 96- and 384-sample formats, enabling customers to sequence up to 4,608 samples on one sequencing run without investment in new equipment or extensive training.

Cellular Research
For info: 650-752-6144
www.cellular-research.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.

IALS

Five Associate/Full Professor Positions at the Institute for Applied Life Sciences and College of Natural Sciences, University of Massachusetts Amherst

OPPORTUNITIES

The University of Massachusetts Amherst Institute for Applied Life Sciences (IALS) is a fast-growing enterprise dedicated to developing the next generation of life science products and technologies to improve human health.

We seek candidates for a number of tenure track positions in the areas listed below. Candidates should have a passion for research in drug delivery technologies, therapeutic candidate development, personalized health monitoring and have what it takes to move from exciting idea to product innovation. With a new state-of-the-art building, more than 20 core equipment facilities, and over 150 interdisciplinary collaborators, IALS provides exciting opportunities for faculty interested in conducting novel academic research in the life sciences. Specific positions open are:

Human Magnetic Resonance Center Director (Focus: Neuroimaging Research)
We seek candidates with relevant expertise in areas such as: anatomical or functional neuroimaging; development of new neuroimaging methodologies; or neuroimaging of cognition. Successful candidates will have a strong track record of research publications and federal funding focused on the development or use of state-of-the-art magnetic resonance techniques and approaches for the study of human brain structure and function. Preference will be given to a candidate with demonstrated leadership and extensive experience and involvement in a successful MR center. A Ph.D. in Neuroscience, Psychology, Physics, Biology or related field is required. **Application Materials for the HMR Director position must be submitted to:** http://umass.interviewexchange.com/jobofferdetails.jsp?JOBID=66077

Computational Biology We seek interdisciplinary candidates who bring expertise in computational biology and internationally recognized research programs that complement IALS strengths, which include protein homeostasis, cancer biology, immunology, neurodevelopment, neurodegeneration, cellular dynamics, and drug delivery. The ideal candidate will utilize computational approaches to exploit large databases, analyze complex molecular architectures, and/or interrogate interconnected cellular pathways. A Ph.D. in any of the Life Sciences, Math, Chemistry or a related field is required. **Application Materials for the Computational Biology position must be submitted to:** http://umass.interviewexchange.com/jobofferdetails.jsp?JOBID=66083

Biosensors We seek candidates with expertise in designing and applying biosensors for addressing key issues in areas of food safety, biomedical diagnostics, tremor detection, or neuroscience. The candidate's research should have a strong emphasis on using engineering and/or nanotechnology to design and apply biosensors that can be used in food, agriculture, environmental, and/or biomedical settings. Candidates with research expertise in food science, health monitoring, bio-engineering, neuroscience, or related areas that aim to improve human health or safety of foods using biosensing techniques are strongly encouraged to apply. A Ph.D. in Food Science, Psychology, Neuroscience, Biology, Chemistry, Microbiology or a related field is required. **Application Materials for the Biosensors position must be submitted to:** http://umass.interviewexchange.com/jobofferdetails.jsp?JOBID=66078

Disease Models Stem Cell We seek candidates who work in areas that complement IALS strengths in biomaterials chemistry, protein homeostasis, cancer biology, immunology, neurodevelopment, neurodegeneration, and cellular dynamics, and bring expertise in animal models of disease. Those utilizing stem cells to explore disease etiology and develop therapeutic strategies are also of great interest. The ideal candidate will sustain a vigorous research program linking basic chemical, biochemical, and cellular approaches with organism-level disease models. A Ph.D. in any of the Life Sciences, Chemistry, Psychology, Veterinary & Animal Sciences or a related field is required. **Application Materials for the Disease Models Stem Cell position must be submitted to:** http://umass.interviewexchange.com/jobofferdetails.jsp?JOBID=66082

Animal Model of Disease We seek candidates who direct internationally recognized research programs using animal models of disease and work in areas that complement IALS strengths in protein homeostasis, cancer biology, immunology, neurodevelopment, neurodegeneration, and cellular dynamics. The ideal candidate will sustain a vigorous research program linking basic cellular biochemistry with an organism-level understanding of disease. A Ph.D. in any of the Life Sciences, Chemistry, Psychology, Veterinary & Animal Sciences or a related field is required. **Application Materials for the Animal Model of Disease position must be submitted to:** http://umass.interviewexchange.com/jobofferdetails.jsp?JOBID=66080

APPLICATION PROCESS: The search committee will begin reviewing applications on Juanuary 6, 2016. Searches will continue until the positions are filled. Applicants must send curriculum vitae, a description of research interests, and contact information for three references. Applicants must apply through the respective Interview Exchange Links. The complete position announcements can be found on Interview Exchange. For questions about these positions, contact John McCarthy, search committee co-chair at **jmccarthy@grad.umass.edu**.

中国西北高校招聘

Recruitment for Universities in Northwest China

一帶一路：中國夢與世界夢的交滙橋梁

One Belt and One Road:The bridge between the Chinese dream and the world dream

西安交通大學
www.xjtu.edu.cn

陝西師範大學
www.snnu.edu.cn

西安電子科技大學
www.xidian.edu.cn

西北工業大學
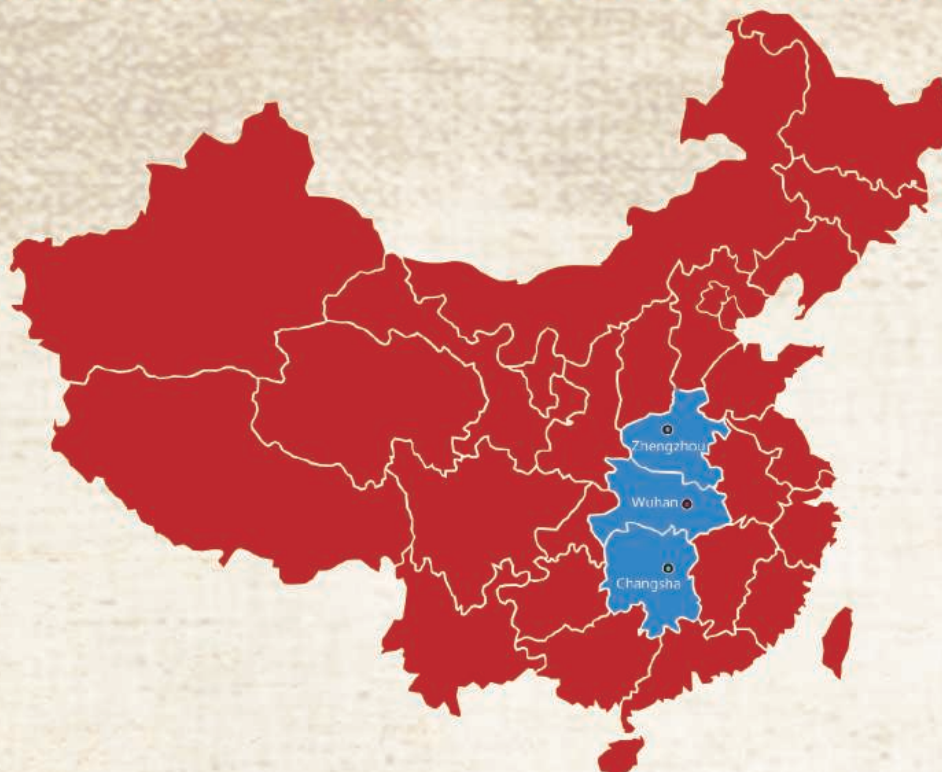www.nwpu.edu.cn

Looking for more positions? Please send your CV to acabridge@163.com or call the direct line: 86-10-62603334

www.edu.cn/xbzp

中国西南高校招聘

Recruitment for Universities in Southwest China

钟灵毓秀，杰归西南

Welcome your joining in southwest China's colleges and universities

www.swu.edu.cn

www.cqu.edu.cn

www.uestc.edu.cn

www.sicau.edu.cn

Looking for more positions? Please send your CV to acabridge@163.com or call the direct line: 86-10-62603334

www.edu.cn/xnzp

By **Kamal J. K. Gandhi**

The best decision I ever made

What prompts a person to make a decision that will change their life forever? For me, it was being born a female in a developing country, India. I had a happy childhood filled with warm memories and a supportive family, but because I was a girl, it seemed that I had fewer rights than boys and couldn't have the freedoms and education they were offered. My future was to marry a man as arranged by my parents, rather than have a full career. But I wanted something different. After spending a glorious summer at a youth World Wildlife Fund camp in Mudumalai National Park in the mountains of the Western Ghats when I was 13 years old, I decided that I wanted to be an ecologist to study and preserve nature.

I undertook a bachelor's degree in zoology at the University of Delhi, with the intention of going to graduate school in North America. But my family couldn't understand why I wanted to live in the Western world as a poor graduate student when I could have a comfortable life in India. Elders clicked their tongues in disapproval, and friends thought that I had lost my mind when I told them about my dream to become a scientist. After I completed my undergraduate degree, I had the choice to either get married or leave home for a career. I took the latter option, and now, more than 20 years later, I know that it was the best decision I ever made.

But the path wasn't easy, especially at the start. I moved to Canada with only a few hundred dollars and a suitcase full of books and clothes. I worked in retail, as a waitress, and other odd jobs for a few years while trying to get into graduate school. And once I began my master's program, I found myself chronically behind my peers. I didn't even know how to turn on a computer, let alone use one! Only after many years of hard work was I finally up to speed. Luckily, along the way I met wonderful friends and colleagues who gave me a sense of family and belonging, and helped me get settled in a new and foreign culture. I always had a place to go for Thanksgiving and Christmas, and when I married an American husband of my own choice, we started hosting at our home.

But developing confidence was a major challenge. Both as a graduate student and young faculty member, I caught myself trivializing my accomplishments when I was complimented. Slowly, I learned to take ownership of my successes. Yet it took decades to truly feel confident about myself as a scientist, and to not dwell on negative messages or the

"I decided that I would move anywhere to pursue my dream."

failures that are more common than successes in science. I owe a huge debt of gratitude to my colleagues and mentors, who guided and taught me along the way. My mentors were patient; they believed in me and gave me opportunities to thrive in science. This is not an easy task, as I understand better now that I am a mentor myself.

Life is full of choices, and those dictated by passion, dedication, and perseverance are the ones that will make us successful. As I was establishing my career, I decided that I would move anywhere to pursue my dream. I wanted to see the world, pursue opportunities as they arose, and learn widely about ecological problems. Before we got married, I asked my husband if he was willing to move with me, and he agreed to follow me across the continent. So, I worked with various state and federal agencies and forest companies in Alberta, British Columbia, California, Michigan, Minnesota, and Ohio before becoming a professor and settling in Georgia. Working in many ecosystems gave me a unique perspective and appreciation for ecological problems, which brought depth and diversity to my research program.

Many times during graduate school I wondered if, by leaving home, I had ruined my life. But I found a supportive and nurturing community that made the challenges lighter and easier, and I am glad I stayed the course. After all, the biggest adventures in our lives are the ones we allow ourselves to have. ∎

Kamal J. K. Gandhi is an associate professor of forest entomology at the Daniel B. Warnell School of Forestry and Natural Resources at the University of Georgia, Athens. Send your story to SciCareerEditor@aaas.org.