Can billions buy a cure for Alzheimer's? *p. 838*

Eight carbonyl ligands on calcium *pp. 849 & 912*

CRISPR barcodes elucidate mouse development *p. 893*

Science

$15
31 AUGUST 2018
sciencemag.org

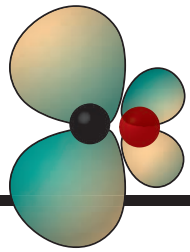
AAAS

SPECIAL ISSUE

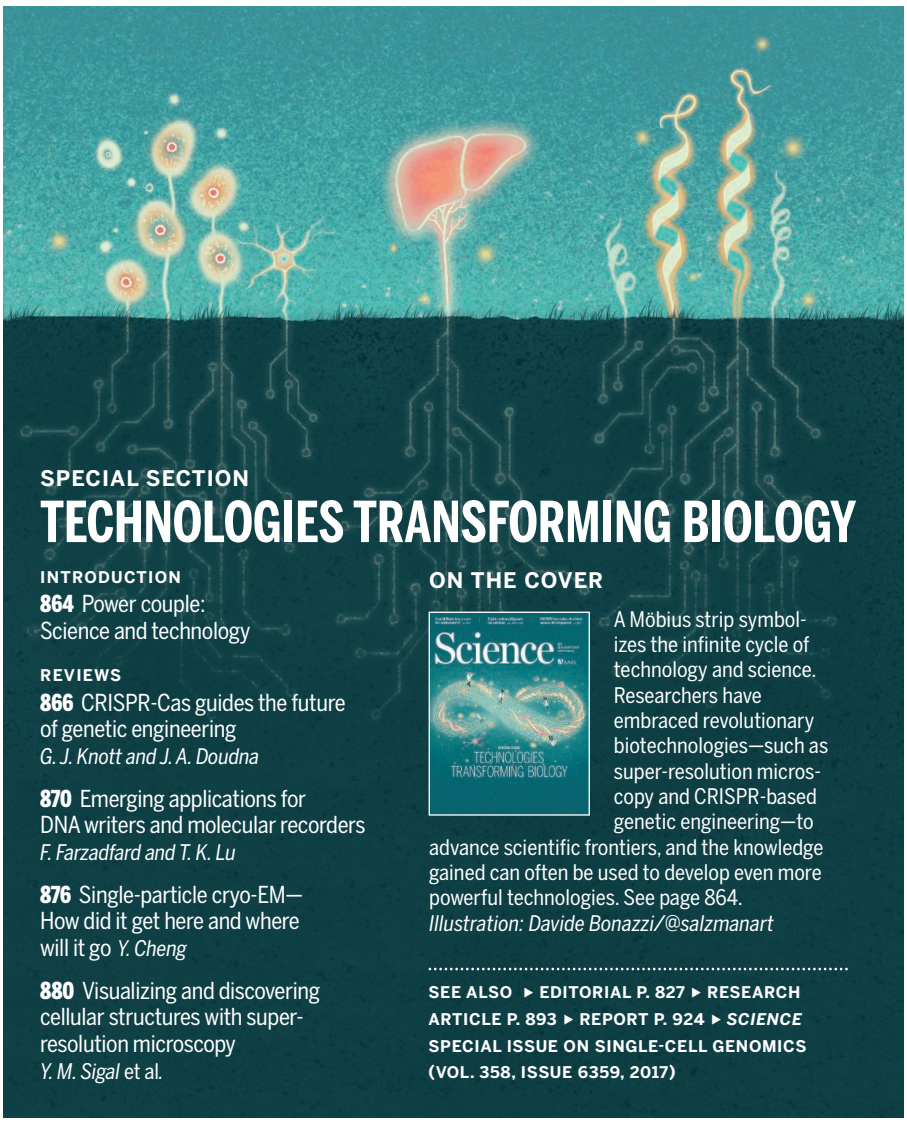
TECHNOLOGIES TRANSFORMING BIOLOGY

CONTENTS

31 AUGUST 2018 • VOLUME 361 • ISSUE 6405



849 & 912
Backbonding in group 2



SPECIAL SECTION

TECHNOLOGIES TRANSFORMING BIOLOGY

INTRODUCTION

REVIEWS

ON THE COVER

A Möbius strip symbolizes the infinite cycle of technology and science. Researchers have embraced revolutionary biotechnologies—such as super-resolution microscopy and CRISPR-based genetic engineering—to advance scientific frontiers, and the knowledge gained can often be used to develop even more powerful technologies. See page 864.
Illustration: Davide Bonazzi/@salzmanart

SEE ALSO ▸ EDITORIAL P. 827 ▸ RESEARCH ARTICLE P. 893 ▸ REPORT P. 924 ▸ *SCIENCE* SPECIAL ISSUE ON SINGLE-CELL GENOMICS (VOL. 358, ISSUE 6359, 2017)

NEWS

IN BRIEF

IN DEPTH

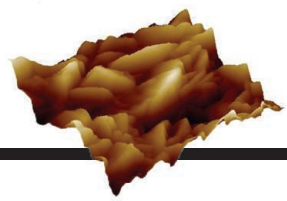
FEATURE

INSIGHTS

PERSPECTIVES

Published by AAAS

CONTENTS



31 AUGUST 2018 • VOLUME 361 • ISSUE 6405

RESEARCH

CREDITS: (GRAPHIC) Q. HAN *ET AL.*; (ILLUSTRATION) ROBERT NEUBECKER

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005. Periodicals mail postage (publication No. 484460) paid at Washington, DC, and additional mailing offices.

Domestic individual membership, including subscription (12 months): $165 ($74 allocated to subscription). Domestic institutional subscription (51 issues): $1808; Foreign postage extra: Mexico, Caribbean (surface mail) $55; other countries (air assist delivery): $89. First class, airmail, student, and emeritus rates on request. Canadian rates with GST available upon request, GST #125488122. Publications Mail Agreement Number 1069624. **Printed in the U.S.A.**
Change of address: Allow 4 weeks, giving old and new addresses and 8-digit account number. **Postmaster:** Send change of address to AAAS, P.O. Box 96178, Washington, DC 20090–6178. **Single-copy sales:** $15 each plus shipping and handling; bulk rate on request. **Authorization to reproduce** material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and others who use Copyright Clearance Center (CCC) Pay-Per-Use services provided that $35.00 per article is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923. The identification code for *Science* is 0036-8075. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Science
AAAS
1200 New York Avenue, NW, Washington, DC 20005
Clarendon House, Clarendon Road, Cambridge, UK CB2 8FH

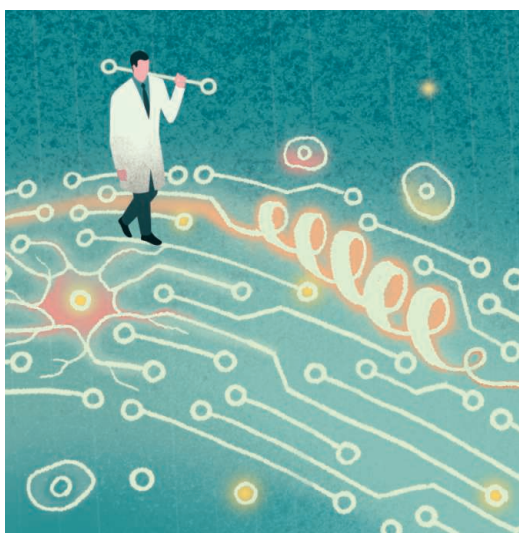
Published by AAAS

EDITORIAL

Revolutionary technologies

Editor-in-Chief, Science *Journals.*
jberg@aaas.org

In this issue of *Science*, we present reviews of four technologies whose power and rapid growth across biological research communities make them revolutionary (see page 864). New technology is one of the most powerful drivers of scientific progress. For example, the earliest microscopes magnified images only 50-fold at most. When the Dutch fabric merchant and amateur scientist Antonie van Leeuwenhoek developed microscopes with more than 200-fold magnifications (likely to examine cloth), he used them to study many items, including pond water and plaque from teeth. His observations of "animalcules" led to fundamental discoveries in microbiology and cell biology, and spurred the elaboration of improved microscopes. Today, various light microscopes remain prime tools in modern biology. This example embodies two characteristics of a revolutionary technology: a capability for addressing questions better than extant technologies, and the possibility of being utilized and adapted by many other investigators.



"New technology is one of the most powerful drivers of scientific progress."

The discovery of x-rays in 1895 ushered in a multifaceted revolution in imaging. As scientists sought to understand the nature of these electromagnetic waves, they realized that they were diffracted by crystals, establishing that the wavelengths of x-rays were comparable to the separation between atoms in crystals. In 1913, William Henry Bragg and his son William Lawrence Bragg found that diffraction patterns could be interpreted to reveal the arrangement of atoms in a crystal. The Braggs determined the structures of many simple substances, including table salt and diamond. Others began using similar techniques to reveal more complex structures of inorganic and organic compounds. In the late 1950s, these methods were extended to determine the structure of proteins, and eventually to larger proteins and protein complexes. Thousands of structures are now reported each year and are foundational to our understanding of biochemistry and cell biology. Technical innovations, improved commercial and shared-facility instrumentation, and powerful software continue to drive the x-ray crystallography revolution.

As the field of recombinant DNA technology was evolving, a revolutionary technique in the form of the polymerase chain reaction (PCR) was developed by Kary Mullis (from 1983 to 1985) at the biotechnology company Cetus. PCR allows tremendous amplification of specific DNA sequences. It had an almost immediate revolutionary impact on many fields, including gene cloning and DNA analysis, and forms the foundation of many methods in modern molecular biology. PCR depends on several underlying technologies, including the chemical synthesis of short sequences of DNA and the availability of appropriate enzymes, but also machines for programmable temperature cycling. The method was invented shortly before I was setting up my first independent laboratory. Cetus had partnered with one company to sell PCR machines, although other devices with similar capabilities were available. I remember calling one of the other companies and asking if its machine would work for PCR. Concerned about patent issues, the sales representative said, "I can't say, but no one has said that it didn't work for their particular application!" My lab joined the PCR revolution.

The reviews in this issue of *Science* focus on two imaging methods that are extending and complementing the powers of traditional light microscopy and x-ray crystallography, and two methods for manipulating DNA to drive a range of discoveries and potentially powerful applications. Such technologies can help to resolve long-standing questions and can open up new vistas, revealing new phenomena and allowing the formulation of questions previously unimagined.

–Jeremy Berg

10.1126/science.aav1775

NEWS

> "I do not want to lie to myself anymore."
>
> **Nicolas Hulot**, who resigned his position as France's environment minister live on France Inter radio to protest the "ministeps" taken by his government and others to slow global warming.

IN BRIEF

Edited by **Catherine Matacic**

POPULAR CULTURE

Wisecracker doesn't crack NIH chief

Comedian Sacha Baron Cohen (left) quizzes National Institutes of Health Director Francis Collins (right) on HIV/AIDS and trans fats.

British comedian Sacha Baron Cohen has duped many public figures into making fools of themselves with his spoof shows and interviews. But Francis Collins, director of the U.S. National Institutes of Health in Bethesda, Maryland, quickly caught on when he was interviewed for an episode of Baron Cohen's *Who is America?*, which aired 19 August on Showtime. The trickster—dressed as wheelchair-driving southerner Billy Wayne Ruddick, Jr. Ph.D.—asks Collins why "big agriculture" is spiking food with trans fats to make people transgender. Collins gamely replies that trans fats contribute to heart attacks and stroke, but have nothing to do with gender. When Baron Cohen describes an experiment to prove whether AIDS exists—one that involved taking a sample of his own blood using a needle from an HIV-positive homeless person—Collins gives him a stern talking-to: "You might want to wait 6 weeks and then have a test and see if you turned positive." Collins, who was told he was doing a Showtime interview with a nonexpert, says he knew "a few minutes in" that it was a ruse. "I was pretty irritated from having been misled." But ever the dutiful public servant, he decided to keep going—and get across whatever public health messages would stick.

STD cases set record

PUBLIC HEALTH | Cases of the sexually transmitted diseases (STDs) chlamydia, gonorrhea, and syphilis hit a record-setting 2.3 million in the United States in 2017, a 10% increase over 2016, the Centers for Disease Control and Prevention (CDC) in Atlanta reported on 28 August. The 2017 data, which are preliminary, mark the fourth consecutive year of sharp increases in these STDs. Since 2013, gonorrhea diagnoses have increased 67% and nearly doubled among men; diagnoses of the early stages of syphilis have grown by 76%; and cases of chlamydia grew from 1.4 million to 1.7 million—nearly half of them in women 15 to 24 years old. "The systems that identify, treat, and ultimately prevent STDs are strained to the near-breaking point," says Jonathan Mermin, director of the CDC center responsible for STD prevention. "We are sliding backward."

Former CDC director arrested

#METOO | Tom Frieden, the physician who directed the U.S. Centers for Disease Control and Prevention in Atlanta from 2009 to 2017, was arrested on 24 August in New York City and charged with forcible touching of intimate parts, sexual abuse in the third degree, and harassment in the second degree. He is accused of grabbing the buttocks of a female friend without permission at his home in October 2017, making her "alarmed and annoyed." Frieden, 57, pleaded not guilty, surrendered his passport, and was released without bail by Judge Michael Yavinsky of the Kings County Criminal Court. José Castro, CEO of the New York City nonprofit Vital Strategies, where Frieden last year launched the public health initiative Resolve to Save Lives (*Science*, 22 September 2017, p. 1217), wrote in a statement that the accuser is a 30-year friend of the Frieden family, and that the former director "has the highest ethical standards both personally and professionally." Frieden's next court appearance is scheduled for 11 October.

Global wind mapper launched

METEOROLOGY | After nearly 2 decades of preparation, the European Space Agency (ESA) launched its €480 million

PHOTO: SHOWTIME

Corrected 29 August 2018. See full text.

Published by AAAS

Pollution could cause trees to break more easily in storms.

FORESTRY

Climate change is making trees bigger, but weaker

As global temperatures rise, trees around the world are experiencing longer growing seasons, sometimes as much as three extra weeks a year. All that time helps trees grow faster. But a study of the forests of Central Europe suggests higher temperatures—combined with pollution—are making wood weaker. Scientists analyzed core samples from four species of century-old trees, including the European beech (above), and discovered that their wood density has dropped between 8% and 12% since 1900, they report in *Forest Ecology and Management*. Rising temperatures, and the faster growth they spur, probably account for some of the drop. But another factor, scientists say, is more nitrogen in the soil from agricultural fertilizers and vehicle exhaust. The result? Less durable lumber and trees that may be less efficient at soaking up the greenhouse gas carbon dioxide.

wind-measuring satellite, Aeolus, into space last week from French Guiana. Named for the "keeper of the winds" in Greek mythology, Aeolus is the first satellite to measure winds directly. It will blast Earth's atmosphere with a laser and catch photons reflected off air molecules, allowing the winds' height, speed, and direction to be detected. The mission should improve climate models and weather forecasts, especially in the tropics, where few weather balloons are deployed. ESA selected Aeolus for flight in 1999, but engineering its high-power laser was challenging and caused delays.

A go for magic mushroom study

BIOMEDICINE | A phase II trial to treat depression with psilocybin, the active ingredient in magic mushrooms, just got the go-ahead from the U.S. Food and Drug Administration. Preliminary studies have suggested the psychedelic substance, a schedule 1 drug in the United States, can help alleviate depressive symptoms. The new trial, set to start next week in the United Kingdom, will give 216 patients who have not responded to other therapies a single dose of the compound: either 1, 10, or 25 milligrams. A spokesperson for the company running the trial, Compass Pathways, says results are expected toward the end of 2019. If positive, the spokesperson adds, Compass will likely conduct a phase III trial, the last step before licensing. Psilocybin is also being tested in clinical trials to treat migraines, nicotine addiction, and obsessive-compulsive disorder.

Carbon plan sinks Australian PM

CLIMATE SCIENCE | Disagreement over Australia's efforts to cut carbon emissions led the ruling coalition to dump Prime Minister Malcolm Turnbull last week. Turnbull, leader of the Liberal and National parties' coalition since 2015, recently proposed an energy bill calling for Australia to cut greenhouse gas emissions by 26% from 2005 levels by 2030, in line with the country's Paris agreement target. Conservative coalition members, many of whom want Australia to withdraw from the climate treaty, pushed Turnbull into dropping the target; but they still forced a leadership vote that ended in a victory for Scott Morrison, a friend of Australia's coal industry. This is the fourth time in 10 years that squabbles over energy and emissions policies have contributed to the toppling of a prime minister.

Nominee ducks climate query

SCIENCE POLICY | Kelvin Droegemeier sidestepped the only hardball question last week from the Senate panel reviewing his nomination for director of the White House Office of Science and Technology Policy. Senator Ted Cruz (R–TX), who says the planet is not warming and that Democrats have demagogued the issue, asked about "empirical" satellite data that "show no statistically significant warming over the past 18 years." Droegemeier, an expert in severe storm prediction at the University of Oklahoma in Norman, gave a nonanswer: "I'm familiar with some of those studies … but I don't study climate." His predecessor, John Holdren, repeatedly dismissed the supposed evidence for a "hiatus," but some climate scientists are worried that Droegemeier has decided to remain mum on the issue as the price of becoming the president's science adviser. He revealed some of his views in a 2014 talk to researchers, when he described the planet's resiliency by saying: "You can kick it in the butt really, really hard and it will come back."

SCIENCEMAG.ORG/NEWS
Read more news from *Science* online.

Corrected 29 August 2018. See full text.

Published by AAAS

American chestnuts were once a dominant tree, and a major source of food, in the forests of eastern North America.

BIOTECHNOLOGY

Can a transgenic chestnut restore a forest icon?

Researchers seek permit to release American chestnut engineered to resist a deadly blight

By **Gabriel Popkin**, *in Syracuse, New York*

Two deer-fenced plots here contain some of the world's most highly regulated trees. Each summer researchers double-bag every flower the trees produce. One bag, made of breathable plastic, keeps them from spreading pollen. The second, an aluminum mesh screen added a few weeks later, prevents squirrels from stealing the spiky green fruits that emerge from pollinated flowers. The researchers report their every move to regulators with the U.S. Department of Agriculture (USDA). "We tell them when we plant and where we plant and how many we plant," says Andrew Newhouse, a biologist at the nearby State University of New York College of Environmental Science and Forestry (SUNY ESF).

These American chestnut trees (*Castanea dentata*) are under such tight security because they are genetically modified (GM) organisms, engineered to resist a deadly blight that has all but erased the once widespread species from North American forests. Now, Newhouse and his colleagues hope to use the GM chestnuts to restore the tree to its former home. In the coming weeks, they plan to formally ask U.S. regulators for approval to breed their trees with nonengineered relatives and plant them in forests.

If the regulators approve the request, it would be "precedent setting"—the first use of a GM tree to try to restore a native species in North America, says Doria Gordon, lead senior scientist at the Environmental Defense Fund (EDF) in Washington, D.C. But deciding whether to unleash a GM tree into the wild could take years.

American chestnuts, towering 30 meters or more, once dominated forests throughout the Appalachian Mountains. But in the early 1900s, a fungal infection appeared on trees at the Bronx Zoo in New York City, and then spread rapidly. The so-called chestnut blight—an accidental import from Asia—releases a toxin that girdles trees and kills everything above the infection site, though still-living roots sometimes send up new shoots. By midcentury, large American chestnuts had all but disappeared.

Researchers seal off the flowers of a chestnut carrying a wheat gene that neutralizes a fungal toxin.

In 1990, SUNY ESF tree geneticists William Powell and Charles Maynard (now retired) decided to try to create resistant chestnuts with the then-new technology of genetic engineering. Eventually, they inserted into the tree's genome a wheat gene that codes for an enzyme called oxalate oxidase, or OxO. It breaks down the oxalic acid the pathogen releases, which is what kills the trees. "We're basically taking the weapon away from the fungus," Powell says.

It didn't work at first. Then, the scientists changed the wheat gene's promoter sequence to cause OxO to be expressed at high levels. In 2014, they reported that a GM tree named Darling 58 both resisted blight infection and transmitted resistance to its offspring. Subsequent tests showed that it produces nuts indistinguishable from those of native trees, Newhouse says. And its pollen, flowers, and decaying leaves don't harm bees, beneficial soil fungi, or tadpoles that hatch in pools on the forest floor.

But the request to release it is likely to face a lengthy regulatory road. The United States, China, and Brazil have approved some transgenic trees for use in fruit orchards, biofuel plantations, and afforestation projects. But like GM crops and animals, GM trees are controversial, and ethical and ecological concerns are heightened because

PHOTOS: ANDREW NEWHOUSE

the chestnut trees would grow wild. Regulators from three federal agencies are likely to take a close look at those concerns. USDA officials, for instance, will seek to determine whether the tree could become a weed or otherwise threaten existing plants. The Food and Drug Administration will study whether the tree's fruit is safe to eat, and the Environmental Protection Agency will consider whether the trees' blight-blocking enzyme should be regulated as a fungicide.

Regulators also "need a really clear process for transparently incorporating … cultural and spiritual values into the decision-making," says Gordon, who serves on a committee convened by the National Academies of Sciences, Engineering, and Medicine to examine issues surrounding GM trees. The American chestnut was a culturally important tree and important food source for many Native Americans, and some are wary of genetically altering a species with which they have a long relationship, says Neil Patterson, a member of the Tuscarora Nation and assistant director of the Center for Native Peoples and the Environment at SUNY ESF.

If the tree survives the regulatory gauntlet, researchers not directly involved in its development are cautiously optimistic that it could help with restoration. It seems to fend off blight better than hybrids produced so far through traditional breeding methods, says Jared Westbrook, chief scientist of the Asheville, North Carolina–based American Chestnut Foundation, which has spent 35 years attempting to breed a blight-resistant chestnut and helped fund the GM tree research. But to maximize survival, the GM trees—which are all descended from clones of one "founder" tree—will need to be crossed with trees adapted to local climate and diseases, Westbrook says. "We're not going to restore a species with a clone."

All that work could be undone if the fungus evolves a way around the defense, says Richard Sniezko, a tree geneticist with the U.S. Forest Service in Cottage Grove, Oregon. Powell and Newhouse doubt that kind of natural selection will occur, because their tree does not actually kill the fungus. Still, "None of us wants something to be put out there … and it fails after 10 years," Sniezko says.

The continuing influx of insects and pathogens from abroad also could present a new chestnut killer, says Gary Lovett, an ecologist at the Cary Institute of Ecosystem Studies in Millbrook, New York. Creating resistant varieties "is a good thing," he says. "But it doesn't do any good if we keep introducing new pests." ■

Gabriel Popkin is a journalist in Mount Rainier, Maryland.

BIOMEDICINE

New pain drugs may lower overdose and addiction risk

By slowing action or targeting different receptors, altered opioids or alternatives aim to sidestep abuse

By **Robert F. Service**, *in Boston*

As the opioid crisis continues to ravage U.S. communities, scientists and drug companies have intensified their efforts to develop safer and less addictive pain medications. Now, multiple research groups are claiming progress in devising novel opioids—or alternatives—that seem to offer pain relief with far less risk of addiction or of the opioid-induced respiratory depression that all too commonly leads to death.

Most of these studies, reported at a meeting here and in a paper released this week, have only been done in animals, so the experimental compounds face significant hurdles before they can become approved medications. Yet they are raising tentative hopes among researchers. "It's encouraging," says Laura Bohn, a biochemist at Scripps Research in Jupiter, Florida. "There has been a really big push to develop nonopioid pain relievers. But it has been really hard."

> ***"There has been a really big push to develop nonopioid pain relievers. But it has been really hard."***
>
> **Laura Bohn**, Scripps Research

A record 72,000 people in the United States died last year from overdoses, up nearly 10% from 2016, according to an estimate this month by the Centers for Disease Control and Prevention. That rise was driven primarily by an increase in overdoses from highly potent synthetic opioids such as fentanyl and carfentanil. Another 2.1 million Americans are believed to regularly abuse opioids, including natural ones like morphine, semisynthetic compounds such as oxycodone, and the synthetics, and have signs of addiction, such as withdrawal symptoms, if they try to quit.

Opioids are powerful pain relievers because they bind to a key cell membrane protein, known as the μ-opioid receptor (MOR), on neurons in the brain and spinal cord. Once activated, the MOR triggers an intracellular "G protein" to initiate a molecular cascade that leads to pain relief. But traditional opioids also activate another intracellular protein, β-arrestin2, which produces respiratory depression and constipation, the most common opioid side effects for such drugs. Several "biased opioids," including one now under review by the Food and Drug Administration, offer pain relief while reducing β-arrestin2 activation, but it's not clear whether they are less addictive than conventional opioids (*Science*, 17 November 2017, p. 847).

At last week's meeting of the American Chemical Society here, Neel Anand, a senior director for medicinal chemistry at Nektar Therapeutics, a biotech firm in South San Francisco, California, described an approach that might help. Nektar's drug, called NKTR-181, is a version of oxycodone to which researchers have linked a molecular tail called polyethylene glycol, a common pharmaceutical strategy for extending the life span of medicines in the blood. Anand reported that in animal studies, NKTR-181 crosses the blood-brain barrier 70 times more slowly than oxycodone. Instead of a sharp spike in both pain relief and euphoria, caused by an upsurge of the neurotransmitter dopamine in brain regions tied to addiction, NKTR-181 triggers a slower release of dopamine that produces flatter, more sustained pain relief and less euphoria. In clinical studies of more than 600 patients taking the compound, Nektar researchers found far fewer signs of addiction than in patients given oxycodone, as well as fewer side effects.

"It clearly works" as a painkiller, says Steven McKerrall, a medicinal chemist with Genentech in South San Francisco. "They've built [a timed release] into the drug itself." But McKerrall and others caution that opioid addicts have devised strategies to defeat other abuse-resistant formulations, for example, by crushing pills that have timed-release coatings. "Addicts will always find a way," Bohn says.

They might have a tougher time with a compound developed by Astraea Therapeutics, a biotech company in Mountain View, California, that hits two brain molecules at once. AT-121 stimulates not only MOR, but also a close cousin known as the nociceptin opioid receptor (NOR). When activated in the brain, NOR appears to counteract MOR. At the same time, it reinforces MOR's pain relieving activity elsewhere in the central nervous system, says Nurulain Zaveri, Astraea's founder and chief scientific officer. The drug isn't the first to target both receptors—another one is already in phase III trials for diabetic nerve pain, among other uses, but that compound targets other receptors as well, and animal studies suggest it may have addictive properties.

In this week's issue of *Science Translational Medicine*, Zaveri and academic colleagues in the United States and Japan report that rhesus monkeys given AT-121 experienced 100-fold greater pain relief than the same dose of morphine provided. Yet the drug did not trigger respiratory depression, addictivelike behaviors, or even tolerance, where more of a compound is needed over time to produce the same desirable effects such as pain relief. AT-121 even appears to counteract addiction to standard opioids, such as oxycodone, Zaveri says. Monkeys hooked on oxycodone and trained to self-administer the drug sharply reduced further drug seeking when given AT-121. "It looks very promising," Bohn says of the new compound.

Avoiding opioid receptors altogether is another appealing strategy for relieving pain with a reduced risk of addiction, says Roger Kroes, senior director for discovery science at Aptinyx, a biotech firm in Evanston, Illinois, who described one of his company's compounds at the meeting. Called NYX-2925, it activates the NMDA receptor, which helps strengthen neural synapses involved in learning and memory. Although acute pain doesn't involve a learned component, chronic pain is thought to bring about long-term neural changes orchestrated, in part, by NMDA receptors.

Many well-known drugs that block these receptors—among them ketamine and methadone—can relieve pain and can be less addictive than opioids. But these compounds hit other targets as well and have widespread side effects. NYX-2925, however, is more selective, data show. At the meeting, Kroes reported that in preclinical studies on mice and rats, the compound reduced pain and led to a remodeling of synapses involved in learning and memory, essentially rewiring neural circuitry away from being habituated to pain.

The results "were pretty exciting," says Ben Milgram, a medicinal chemist with Amgen in Cambridge, Massachusetts, who attended the meeting. Aptinyx is now testing NYX-2925 in two phase II clinical studies in people with diabetic nerve pain and fibromyalgia, a disease marked by widespread muscle and skeletal pain.

Drugs designed to deliver the benefits of opioids without the deadly risks can easily falter. At the meeting, researchers from Genentech, Merck & Co., and Amgen described compounds designed to tamp down yet another nonopioid receptor target, a protein called $Na_v1.7$. Although all found their target and reduced pain in animals, they proved weaker on other scores; for example, some were poorly absorbed in the blood or blocked other Na_v proteins, causing side effects. Still, with the opioid crisis taking an ever-larger toll, even preliminary good news is welcome. ■

Synthetic opioids, such as this fentanyl captured in a drug raid, have caused an alarming rise in overdose deaths.

PHOTO: CLIFF OWEN/AP PHOTO

INFECTIOUS DISEASES

Hybridization may give some parasites a leg up

Genomic study helps explain how schistosomiasis gained a foothold in Europe

By **Elizabeth Pennisi**, *in Montpellier, France*

Infecting an estimated 230 million people, schistosomiasis is the world's most widespread parasitic disease after malaria. But temperate latitudes were thought to be spared: Schistosome flatworms are common only in warm places in Africa, India, and South America. So parasitologist Jerome Boissier was surprised when, in a single week in 2014, physicians in France and Germany called him to report that two families who had never left Europe had developed the disease, which can cause fever, chills, muscle aches, and bloody urine.

Epidemiologists later traced the cases to the Cavu River on Corsica, a French island in the Mediterranean Sea, where the patients had swum during a vacation. Scientists found that a local freshwater snail was serving as the intermediate host that's essential to the flatworms' complicated life cycle. The river is still infested: At least 120 people have become infected. And the disease is turning up elsewhere on Corsica.

In earlier work, Boissier, who's at the University of Perpignan Via Domitia in France, had shown that the culprit is no ordinary schistosome parasite, but rather a hybrid of two species. Now, his team has uncovered the hybrid's advantage: It appears to be better than the parent species at infecting both the snails and its unfortunate mammalian hosts. Such hybrids, discovered in other parasitic species as well, could widen a parasite's range of host mammals, complicating efforts to control it. Presented here last week at the Second Joint Congress on Evolutionary Biology by Boissier's grad student Julien Kincaid-Smith, the work "is changing the way we think about disease transmission," says Christina Faust, a disease ecologist at the University of Glasgow in the United Kingdom.

Humans and other mammals infected with schistosomiasis shed eggs in their feces

At least 120 people contracted schistosomiasis in the Cavu River; another river nearby has become infested as well.

or urine, which hatch if they reach freshwater in time. The hatchlings then take up residence in snails, where they mature and reproduce asexually, yielding tiny larvae that exit the snail. If those larvae encounter another swimming or wading mammal, they burrow into its skin and settle in blood vessels, completing the life cycle. Five species infect humans; the most common one, *Schistosoma haematobium*, causes urogenital schistosomiasis. It often resides in veins in the bladder wall or the reproductive tract and can damage organs or impair fertility. Although the antiparasitic drug praziquantel is effective, patients in developed countries can go undiagnosed for years.

S. haematobium probably reached Europe after a patient infected elsewhere traveled to Corsica and urinated in the Cavu River, Boissier says. An intermediate host was waiting: The river is home to the snail *Bulinus truncatus*—one of a few *Bulinus* species that can support schistosomes—which also occurs in some African and Middle Eastern countries. The outbreak "is a wake-up call that this disease can establish itself wherever the right [conditions] exist," says immunologist Daniel Colley of the University of Georgia in Athens, who notes that global travel makes such introductions more likely.

Two years ago, Boissier's team reported that DNA tests on the parasite eggs suggested the new arrival was a hybrid of *S. haematobium* and *S. bovis*, a schistosome species that infects livestock; on the basis of the hybrid's DNA, Senegal was the most likely source. The hybrids themselves were not news; Tine Huyse, a parasitologist at the Catholic University of Leuven in Belgium, and a colleague had found them in Senegal in 2008. But Kincaid-Smith traveled to Senegal and Cameroon to collect the parent strains, and the team bred them in the lab to re-create the hybrid. The researchers then tested the ability of the parents and hybrid to infect snails and—as a stand-in for humans—hamsters.

The human parasites found in Africa didn't infect the Corsican snails, Boissier reported. *S. bovis*, the animal variety, did infect the snails, but the hybrid did so even more readily, and it thrived not only in Corsican snails but also in *B. truncatus* snails from Spain and a related snail species from Portugal. The hybrid also developed faster in hamsters and made them sicker.

Hybrids have emerged in other parasites, including the agents that cause malaria, leishmaniasis, and Chagas disease. Just how important they are in epidemiology is still unclear, but their existence is worrisome, Huyse says, and they seem set to become more common as travel and migration expand. Hybrids are more likely to infect multiple hosts, allowing some of them to "hide" in nonhuman animals, out of reach of the drugs given to people. And combining two genomes gives a parasite more genetic variation with which to adapt to new places and hosts, Faust says.

When Kincaid-Smith and his colleagues teamed up with the Wellcome Sanger Institute in Hinxton, U.K., to fully sequence the hybrid, they found that three-quarters of its DNA came from the human parasite and the rest from *S. bovis*. That mixture may boost the ability of the hybrid to infect the Corsican snail, but with a quarter of the genes from *S. haematobium* missing, "it's a wonder how the parasite can still infect humans," said Kincaid-Smith, who with his colleagues posted a preprint about the genome study on bioRxiv on 11 August. The fact that DNA from the two parent species was quite mixed up—sections of *S. bovis* chromosomes appeared at various places along the *S. haematobium* chromosomes—indicates that hybrids have been around long enough to mate with parents and with each other over multiple generations.

"The level of genomic information [in the study] is impressive," Colley says. But he's cautious about extrapolating the findings about the infectious superpowers of the labmade hybrid to what happens in nature. "We do not know how it will play out in the long run in terms of worsening the spread of or impeding the control of schistosomiasis," he says.

Schistosomiasis seems set to stay on Corsica. Although no human cases occurred in 2017—after a total of seven cases in the two preceding years—the worms still occur in snails in the Cavu River; they also have surfaced in the nearby Solenzara River, Boissier says. Whether they overwinter in snails or take refuge in rodents or some other mammalian host isn't clear, Kincaid-Smith told the meeting: "That's also something that needs to be investigated." ■

A European foothold

Schistosomiasis was discovered on the French island of Corsica in 2014; a DNA analysis suggests it originated in Senegal.

CREDITS: (PHOTO) BORGES SAMUEL/ALAMY STOCK PHOTO; (MAP) A. CUADRA/*SCIENCE*

Corrected 29 August 2018. See full text.

Published by AAAS

ILLUSTRATION: ERHUI1979/GETTY IMAGES. ADAPTED BY C. AYCOCK/*SCIENCE*

RESEARCH FUNDING

Amid fears of idea theft, NIH targets foreign funding links

Agency reminds applicants to report all grant sources, warns against violating confidentiality of peer review

By **Jocelyn Kaiser** *and* **David Malakoff**

Fears that foreign governments are tapping U.S.-funded research for valuable information have reached the nation's largest research funder, the National Institutes of Health (NIH) in Bethesda, Maryland. Last week it sent a letter to more than 10,000 research institutions, urging them to ensure that NIH grantees are properly reporting their foreign ties. The agency also said it is investigating about a half-dozen cases in which NIH-funded investigators may have broken reporting rules, and it reminded researchers who review grant applications that they should not share proposal information with outsiders.

At a Senate committee hearing on NIH oversight last week, NIH Director Francis Collins said "the robustness of the biomedical research enterprise is under constant threat" and "the magnitude of these risks is increasing," although he did not mention specific incidents. He added that in addition to sending the 20 August letter asking institutions to help curb "unacceptable breaches of trust and confidentiality," NIH has established a new advisory group to help the agency tighten procedures.

NIH is feeling pressure from Congress. At the hearing, Senate health committee chair Senator Lamar Alexander (R–TN) lauded the contributions of foreign-born scientists to the United States but worried about "bad actors." His comments reflect a resurgence of concern about foreign competitors to the United States—especially China, Russia, and Iran—attempting to harvest the fruits of federal investments in academic science. In March, federal prosecutors indicted nine Iranians on charges of hacking into the accounts of nearly 4000 professors at 144 U.S. universities and stealing data that cost $3.4 billion to develop. In another case, a professor at Duke University in Durham, North Carolina, has alleged that a Chinese doctoral student working in his laboratory on materials for "cloaking" objects from electromagnetic waves returned to China with sensitive, government-funded findings that he used to start a succesful tech company. Such incidents have prompted the Federal Bureau of Investigation to begin to meet with university officials to brief them on information security issues.

Adding to the worries is the growing number of researchers who receive funding from—and run laboratories in—the United States and another nation, potentially opening a conduit for the transfer of data and technology. U.S.-based scientists are also being targeted by so-called talent recruitment programs run by China and others, which offer lucrative funding.

In general, NIH and other federal research funders don't bar U.S.-based grantees from receiving foreign funding, often sponsor research conducted jointly with scientists in other nations, and encourage grantees to freely share the results of funded research unless the government stamps it as classified. But grantees do have to inform the government if they patent research discoveries as well as disclose all sources of funding when applying for a grant.

Collins told *Science* that NIH's interest in the issue was prompted not by "some big explosive episode" involving a violation of such rules, but "just a gathering sense that it's time to take action." Agency officials have spotted NIH-funded papers noting foreign support that had not been properly disclosed to NIH itself, for instance. The agency is also concerned about NIH-funded scientists who spend several months a year in their home country at what Collins called shadow labs, making it hard to tell which country is funding their discoveries. NIH won't name the six or so institutions that it is investigating, but Collins told STAT that the agency is concerned that some researchers have hidden their foreign ties because they intend to share intellectual property or private information with other nations. But it "may all turn out to be fine—they forgot to tell us something," he says.

NIH also is moving to defend its peer-review system, which annually uses thousands of volunteer reviewers to evaluate more than 80,000 applications—many of which include unpublished findings. A particular concern, the letter states, is "sharing of confidential information on grant applications by NIH peer reviewers with others, including foreign entities."

Academic groups say they share NIH's concerns. Schools "look forward to working with NIH to identify opportunities to mitigate breaches and help ensure accurate reporting," says Lisa Nichols of the Council on Governmental Relations in Washington, D.C., which tracks regulatory issues for universities. They can't always uncover problematic foreign links on their own, adds Tobin Smith of the Association of American Universities, also in Washington, D.C. "We're going to have to do a better job of making sure that faculty are being honest," he says.

Some lawmakers in Congress would like to see stricter oversight of foreign-funded projects on U.S. campuses. A draft amendment to a recent defense spending bill, for example, would have allowed the Pentagon to bar funding for U.S.-based researchers who received support from talent recruitment programs funded by foreign governments. The provision—perceived as targeting China—was ultimately shelved in favor of language asking the Department of Defense to work with universities to examine the risks and benefits of such arrangements. ■

BIOMEDICINE

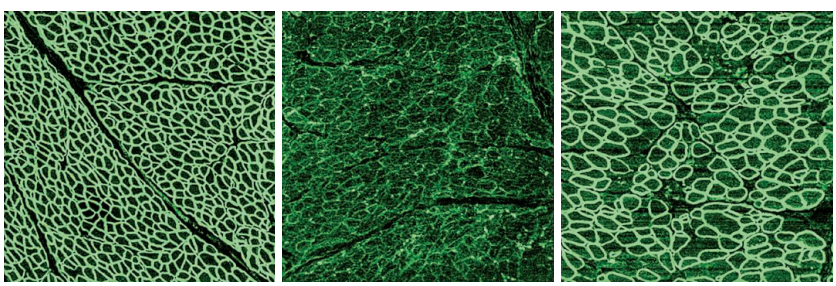
In dogs, CRISPR fixes a muscular dystrophy

Treatment repairs gene in beagles by further mutating it, but human trials are far off

By **Jon Cohen**

Fighting fire with fire, researchers working with dogs have fixed a genetic glitch that causes Duchenne muscular dystrophy (DMD) by further damaging the DNA. The unusual approach, using the genome editor CRISPR, allowed a mutated gene to again make a key muscle protein. The feat—achieved for the first time in a large animal—raises hopes that such genetic surgery could one day prevent or treat this crippling and deadly disease in people. An estimated 300,000 boys around the world are currently affected by DMD.

The study monitored just four dogs for less than 2 months; more animal experiments must be done to show safety and efficacy before human trials can begin. Even so, "I can't help but feel tremendously excited," says Jennifer Doudna of the University of California, Berkeley, who heard the results last week at a CRISPR meeting she helped organize. "This is really an indication of where the field is heading, to deliver gene-edited molecules to the tissues that need them and have a therapeutic benefit. Obviously, we're not there yet, but that's the dream."



Dystrophin (green) is abundant in normal dog muscle cells (left), almost absent in those of a beagle with a muscular dystrophy (middle), and partially restored in an affected beagle treated with CRISPR (right).

IMAGES: UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER

The study, which also appears online this week in *Science*, was led by molecular biologist Eric Olson of the University of Texas (UT) Southwestern Medical Center in Dallas, whose team earlier had similar results in mice. "We wanted to put this to the ultimate test and see if we could do it in a large animal," Olson says. The positive findings—CRISPR quickly restored the protein dystrophin in critical body muscles, including the heart—"brought tears to the eyes and were jaw-dropping," he says.

The study offers little evidence that dogs regained muscle function, however, and that, coupled with the short duration of the study and the small number of animals studied, left some scientists less enthusiastic. One researcher in the tight-knit DMD field who asked not to be named wonders whether the study was rushed to help draw investment in Exonics Therapeutics, a Boston-based company Olson launched last year to develop the potential treatment.

Olson says his team worked quickly not because of corporate ambitions, but rather to prove the concept before expanding to longer, more thorough dog experiments that ultimately are needed to launch human trials. The few animals initially studied, he adds, reflects sensitivities about experimenting with dogs. "We're very mindful of ethical concerns and have done our best to keep our use of dogs to an absolute minimum."

The *dystrophin* gene, the largest in the human body, contains 79 separate coding regions, or exons, that work together to create a protein that has 3500 amino acids. That much DNA offers a lot of opportunities for mutations that can cause DMD. But only one functional copy of the gene is needed, and because it sits on the X chromosome, girls have a backup copy. Boys with their one copy disabled develop walking problems early in life and die on average in their mid-20s from heart and respiratory failure.

About 13% of boys with DMD have mutations in a region between exon 45 and 50, which bumps exon 51 "out of frame" and throws a wrench into the cellular machinery that reads the gene's instructions, stopping production of dystrophin. In 2009, a team led by Richard Piercy at the Royal Veterinary College in London identified a spaniel with signs of DMD that had a spontaneous mutation deleting exon 50, which similarly moves exon 51 out of frame. They later bred a relative of that dog with beagles, which have long been used in biomedical research, to create a colony with DMD symptoms.

Together with Piercy's group, Olson and colleagues designed CRISPR's molecular scissors to make a cut at the beginning of exon 51 in the diseased beagles. The team hoped that when the cell tried to repair the slice, it would accidentally introduce errors to exon 51, leading its proteinmaking machinery to skip the exon altogether and produce a shortened but still functional dystrophin. (A newly approved drug for DMD, eteplirsen, promotes such exon-skipping as well, but its efficacy remains hotly debated.)

Another challenge was to alter billions of muscle cells throughout a living animal. So the team enlisted a helper: a harmless adeno-associated virus that preferentially infects skeletal muscle and heart tissue. Two 1-month-old dogs received intramuscular injections of the virus, engineered to carry CRISPR's molecular components. Six weeks later, those muscles were making dystrophin again. Those results led the researchers to give an intravenous infusion to two more dogs, also 1 month old, to see whether the CRISPR-carrying viruses could add the genome editor to muscles throughout the body. By 8 weeks, Olson told the meeting, dystrophin levels climbed to relatively high levels in several muscles, reaching 58% of normal in the diaphragm and 92% in the heart. But because the dogs were euthanized, Olson could show little evidence that they had avoided DMD symptoms, save for a dramatic video of a treated dog walking and jumping normally.

"There are a lot of questions that have to be addressed," acknowledges Leonela Amoasii, who works in Olson's lab at UT Southwestern and is director of gene editing at Exonics. Skeletal muscle is constantly being replaced, so the treatment would have to reach its stem cells to avoid the need for repeated injections. Longer studies will be needed to make sure that the CRISPR treatment does not introduce cancer-causing mutations. Even if it safely restores the ability to make dystrophin, the treatment likely will only help boys who receive it early in life because the muscle damage is irreversible. And ultimately the treatment would have to target many other DMD-related mutations to help most boys with the disease. "We have to make sure that we dot all the i's and cross all the t's because the implications for both DMD and CRISPR therapy are immense," Olson says. ■

REPRODUCIBILITY

Social science studies get a 'generous' test

New replication effort aimed to detect effects overstated in the original reports

By **Kelly Servick**

In 2013, social psychologist David Kidd, then a graduate student at The New School in New York City, learned that his very first paper as lead author had passed peer review and would be published in *Science*. Now, Kidd's paper, which suggested that reading literary fiction improves a person's ability to intuit the mental states of others, has come under scrutiny again—with a less gratifying outcome. It is among eight studies called into question by a painstaking effort to replicate all 21 experimental social science papers published in *Science* and *Nature* between 2010 and 2015.

Called the Social Sciences Replication Project, it is the latest bid by the nonprofit Center for Open Science (COS) in Charlottesville, Virginia, and far-flung collaborators to quality check the scientific literature. Like its predecessors, the new effort found that a large fraction of published studies don't yield the same results when done a second time. But this time, the five independent research teams that did the replications strove to give the studies the benefit of the doubt: They increased the statistical power of the studies by enlisting, on average, five times as many participants as the originals. "This is an effort to be very generous," says Brian Nosek, a psychologist at the University of Virginia in Charlottesville who co-founded COS and whose lab conducted five of the new replications.

That may help explain why the new project successfully replicated 62% of the experiments, compared with 39% in a much larger study of papers in three psychology journals, which COS and collaborators released in 2015 (*Science*, 28 August 2015, p. 910). A similar project scrutinizing economics studies reported in 2016 that it had replicated 61% of experiments.

The current findings, published this week in *Nature Human Behavior*, seem to contradict the claim that studies in high-profile journals, which put a premium on groundbreaking or surprising results, are less reproducible than those in more specialized journals. But cognitive psychologist Hal Pashler at the University of California, San Diego, cautions that differences in replication rates between various projects probably aren't statistically significant. And the 62% figure "certainly is consistent with there being a problem" in the field, he says. "It seems funny that there's been a drift in standards to the point where 62% seems very respectable."

The teams aimed to test the notion that many studies are hard to reproduce because the claimed effect, though real, is inflated. If so, a replication effort would need to be more sensitive to find that smaller positive effect, Nosek says. "We didn't want low power to be an explanation for why some effects didn't replicate."

"We didn't want low power to be an explanation for why some effects didn't replicate."

Brian Nosek, Center for Open Science

The replication efforts, almost all of them designed in collaboration with original authors, were large enough to be sensitive to an effect only 75% of the size originally reported. If an initial replication attempt failed, the researchers added even more participants. The approach made a difference: Two experiments made the cut only after the sample sizes ballooned.

Kidd, now a postdoctoral researcher at Harvard University, says the extra rigor makes it easier to accept the project's negative verdict on his study (*Science*, 18 October 2013, p. 377). "I can't imagine a reason why one would privilege the original findings over this replication." But in commentaries published alongside the study, Kidd and other authors defend the underlying hypotheses of their papers. Kidd, for example, points out that the project repeated only one experiment from each paper, and in his case, it wasn't the strongest or the most important.

Even repeating just a fraction of the work in the papers took years and cost more than $200,000, with plenty of donated labor and lab time. But the new project also highlights a cheaper way to gauge a paper's replicability. The authors asked roughly 200 scientists and students, most of them psychologists and economists, to guess how likely each study was to be replicated. These experts also participated in an online "prediction market," trading shares that corresponded to studies, which paid out only if the given study was replicated.

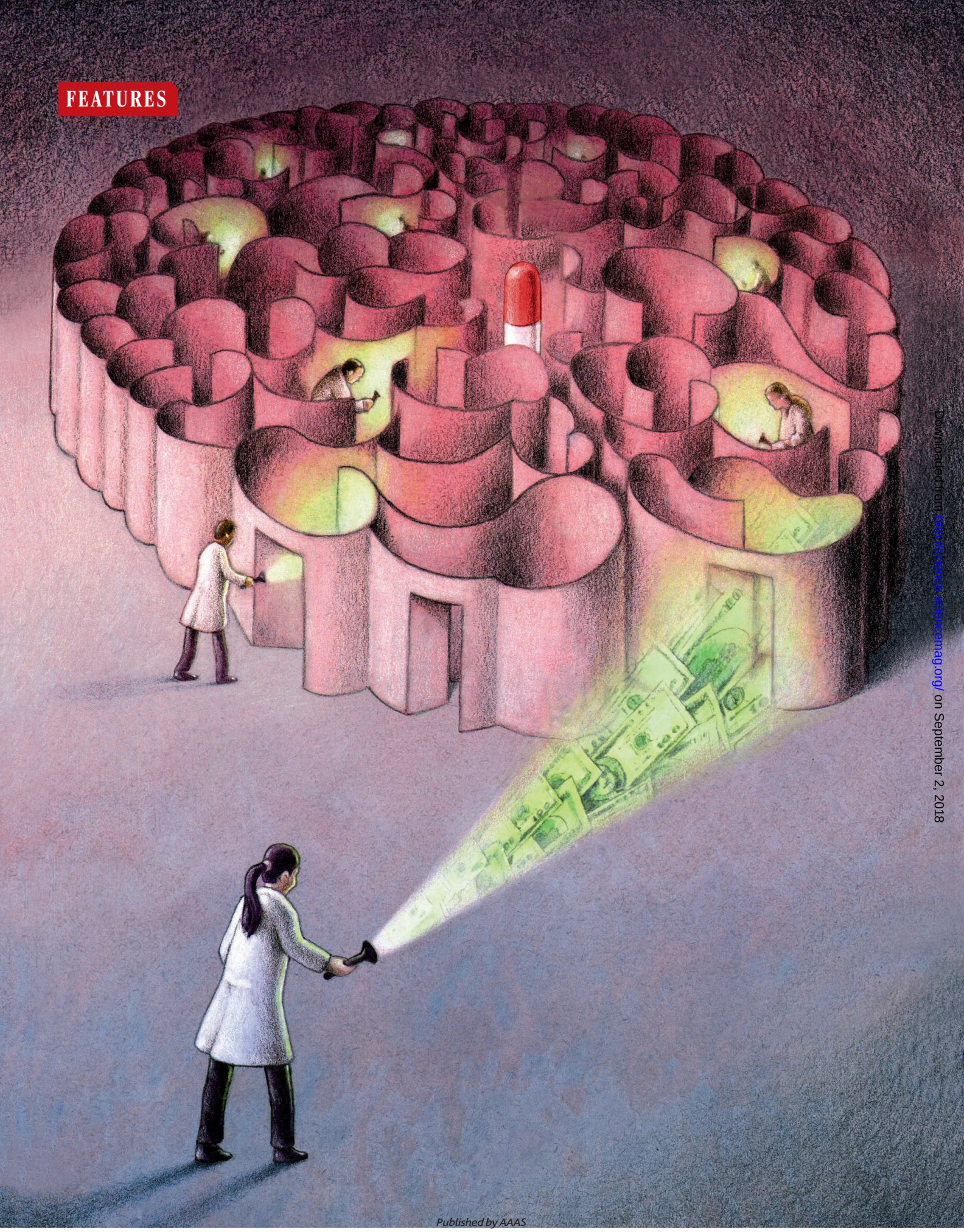
Both approaches did well at predicting the outcome for individual studies, and they predicted an overall replication rate very close to the actual figure of 62%. The finding echoes others suggesting expert judgment is a highly accurate proxy for replication. "There is definitely some wisdom of crowds going on here," economist Anna Dreber of the Stockholm School of Economics, a member of the replication research team, said in a press conference. Anecdotal feedback from the expert evaluators shows they had higher confidence in the replicability of studies with larger sample sizes, and were more dubious of those with surprising or counterintuitive findings.

If experts can instinctively spot an irreproducible finding, "that kind of begs the question of why that doesn't seem to be happening in peer review," says Fiona Fidler, a philosopher of science at The University of Melbourne in Australia. But if future studies can identify and weigh the best predictors of replicability, reviewers might be given a rubric to help them weed out problematic work before it's published.

Another trend may also help tame the problem of irreproducible studies: the push in many fields for authors to share the design of their studies ahead of time, to keep them from changing their approach midstream in search of a flashy, statistically significant result. The studies analyzed here mostly predate that shift, Nosek says. Whether it will really boost social science's track record, he says, is "the next big question." ■

PHOTO: DAN ADDISON UNIVERSITY COMMUNICATIONS

FEATURES



THE ALZHEIMER'S GAMBLE

Can the National Institute on Aging turn a funding windfall into a treatment for the dreaded brain disease?

By **Jocelyn Kaiser**

When molecular biologist Darren Baker was winding up his postdoc studying cancer and aging a few years ago at the Mayo Clinic in Rochester, Minnesota, he faced dispiritingly low odds of winning a National Cancer Institute grant to launch his own lab. A seemingly unlikely area, however, beckoned: Alzheimer's disease. The U.S. government had begun to ramp up research spending on the neurodegenerative condition, which is the sixth-leading cause of death in the United States and will afflict an estimated 14 million people in this country by 2050. "There was an incentive to do some exploratory work," Baker recalls.

Baker's postdoc studies had focused on cellular senescence, the cellular version of aging, which had not yet been linked to Alzheimer's. But when he gave a drug that kills senescent cells to mice genetically engineered to develop an Alzheimer's-like illness, the animals suffered less memory loss and fewer of the brain changes that are hallmarks of the disease. Last year, those data helped Baker win his first independent National Institutes of Health (NIH) research grant—not from NIH's National Cancer Institute, which he once expected to rely on, but from the National Institute on Aging (NIA) in Bethesda, Maryland. He now has a six-person lab at the Mayo Clinic, working on senescence and Alzheimer's.

Baker is the kind of newcomer NIH hoped to attract with its recent Alzheimer's funding bonanza. For years, patient advocates have pointed to the growing toll and burgeoning costs of Alzheimer's as the U.S. population ages. Spurred by those projections and a controversial national goal to effectively treat the disease by 2025, Congress has over 3 years tripled NIH's annual budget for Alzheimer's and related dementias, to $1.9 billion. The growth spurt isn't over: Two draft 2019 spending bills for NIH would bring the total to $2.3 billion—more than 5% of NIH's overall budget.

Such a dramatic increase in research funding for a disease has no precedent at NIH aside from the War on Cancer, an effort launched in 1971, and an explosion of AIDS funding in the late 1980s. With the largesse come logistical challenges. Overworked NIH staff are scrambling to review and process thousands of grant proposals, including those for this year's $414 million bolus—a sum that equals the entire budget of some smaller NIH institutes—which Congress approved in March.

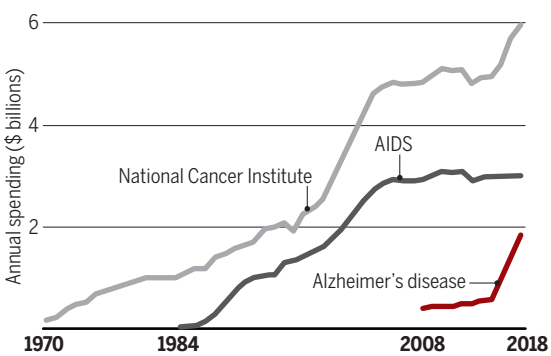
NIA, which oversees the new funds, doesn't just want to plump up existing Alzheimer's labs, says Director Richard Hodes. The institute is also luring investigators, such as Baker, from other fields to bring in fresh ideas. Many are answering the call. "Nearly everyone I know is putting the words 'Alzheimer's disease' in their grants in an effort to tap into the money," says Matt Kaeberlein of the University of Washington in Seattle, who studies aging.

The funding blitz targets a problem that looks more intractable than ever. The only approved drugs for Alzheimer's don't stop the neurodegeneration, but merely treat symptoms—and not very well. In the past year, several major clinical trials based on the field's leading hypothesis—that reducing the level of β-amyloid plaques that riddle the brains of Alzheimer's patients would halt disease progression—have flopped. An antibody that targets β-amyloid recently delivered seemingly promising results in a phase II trial. Yet given past failures for other eagerly watched compounds, many researchers remain skeptical and want to see a larger phase III trial.

Those setbacks have amplified concerns that U.S. officials and some scientists have oversold the plan for a treatment by the middle of the next decade. "I am convinced that we are destined to fail to make the 2025 goal and therefore look like we have failed at our promise," says Alzheimer's researcher Samuel Gandy of the Icahn School of Medicine at Mount Sinai in New York City. Some researchers also worry about focusing so much money on just Alzheimer's. The biomedical community "has mixed feelings" about such targeted funding, says biogerontologist Judy Campisi of the Buck Institute for Research on Aging in Novato, California, who wonders whether more should go to basic research.

Catching up

The National Institutes of Health (NIH) has dramatically ramped up funding for only three specific disease priorities: cancer, AIDS, and, most recently, Alzheimer's.



*Alzheimer's disease funding, which NIH began to track in 2008, does not include related dementias.

CREDITS: (ILLUSTRATION, OPPOSITE PAGE) PAWEL KUCZYNSKI; (GRAPHIC) N. DESAI/*SCIENCE*; (DATA) NIH APPROPRIATIONS DATA, OFFICE OF AIDS RESEARCH, NIH RCDC

Even Baker has qualms. "I think it is great that there's all of this funding. I just hope it's not at the expense of something interesting in the cancer realm."

But naysayers are few. "Overall, what is wrong with it? Nothing," says biochemist Rozalyn Anderson of the University of Wisconsin in Madison, who studies caloric restriction in monkeys to slow aging and is now tying that work to Alzheimer's. "It's a great experiment underway: By increasing funding and access to resources, can we bring on a game-changer in research in a particular area?"

A "CONFLUENCE OF FACTORS" unleashed the funding surge, says Sue Peschin, president and CEO of the Alliance for Aging Research in Washington, D.C. Families became more open about the once-hidden disease, and advocates became savvier. In the late 1990s, the Alzheimer's Association in Chicago, Illinois, and later other groups began to frame care for Alzheimer's patients as a financial crisis looming as the large baby boomer population ages. Alzheimer's already costs Medicare and Medicaid $186 billion per year, and the figure will balloon to $750 billion by 2050, according to the Alzheimer's Association.

Advocates also argued that Alzheimer's is underfunded in the United States in comparison with major killers such as cancer and heart disease. That's especially true for AIDS, which until recently received a fixed 10% of NIH's overall budget—it now gets $3 billion per year—yet affects far fewer Americans. "Neurodegenerative diseases have historically never really had the same funding. In a sense this is a correction," says Alzheimer's researcher John Hardy of University College London.

Those messages resonated with U.S. lawmakers, including Senator Susan Collins (R–ME) and then-Representative (now Senator) Edward Markey (D–MA), who in 1999 co-founded the Congressional Task Force on Alzheimer's Disease. In 2011, they co-sponsored the National Alzheimer's Project Act, which called for a U.S. plan to improve research and care for people with Alzheimer's and related dementias. After Congress passed the bill, the Department of Health and Human Services (HHS), NIH's parent department, outlined ambitious goals, the most striking being to "prevent and effectively treat Alzheimer's disease by 2025." Some Alzheimer's researchers had misgivings about the deadline, says David Holtzman of Washington University School of Medicine in St. Louis, Missouri: "I don't think most thought it was realistic."

The Mayo Clinic's Ronald Petersen, who chaired the advisory board that drafted the HHS plan, defends the 2025 goal: "We wanted to make a bold statement. Not 'We hoped to make progress.' That isn't going to inspire anybody."

As more lawmakers joined the cause, Congress in 2015 mandated that NIH prepare a "professional judgment" budget on Alzheimer's research, a wish list of needs to meet the 2025 target that would bypass the federal budget process and go directly to the president and Congress. Until then, only cancer and AIDS had enjoyed that special treatment. Alzheimer's advocates also lobbied the administration of former President Barack Obama to include extra funding in the White House budget request, Peschin says.

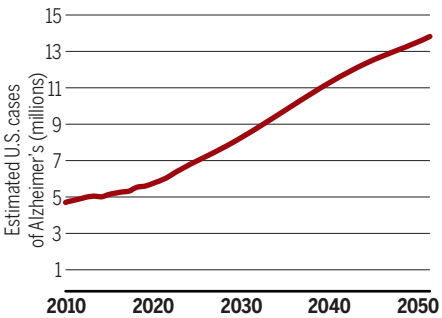
The lobbying began to pay off as early as 2012 when then–HHS Secretary Kathleen Sebelius held a press conference to announce modest increases in funding for Alzheimer's research. That gained the attention of some scientists, including Baker, who submitted his grant proposal to NIA in 2015. However, the big ramp up began only in 2016 after Obama and lawmakers struck a deal to lift federal spending caps and Congress boosted NIH's overall budget after a decade of stagnation. That fiscal year, the share of NIH money going to Alzheimer's shot up 56% to $986 million, including $57 million for separate research on three related dementias, such as vascular dementia. By now, 3 years of such funding boosts have transformed NIA—once a midsize NIH institute and "almost a backwater," as one official put it on a blog—to the fifth-largest of NIH's 27 institutes and centers with a $2.6 billion overall budget. "Our continued investment will pay dividends for the millions of families affected by Alzheimer's," Collins said in a statement to *Science*.

The windfall is incredible, says Eliezer Masliah, director of NIA's Division of Neuroscience. "I've been in this field for over 30 years, and I've never seen anything like this. This is really a golden era for [studying] Alzheimer's disease."

NOW, THE ONUS IS ON NIA and the research community not to waste the money. Under the national plan, NIH holds summits every 3 years to guide its Alzheimer's efforts, targeting the most promising lines of research. Some 140 treatment or prevention trials are underway, testing both drugs and preventive interventions such as exercise. The funding has supported a consortium working on novel mouse models, genetically engineered to mimic the common, late-onset form of the disease. Other money goes to modeling the disease by editing Alzheimer's risk genes in neural cells derived from stem cells.

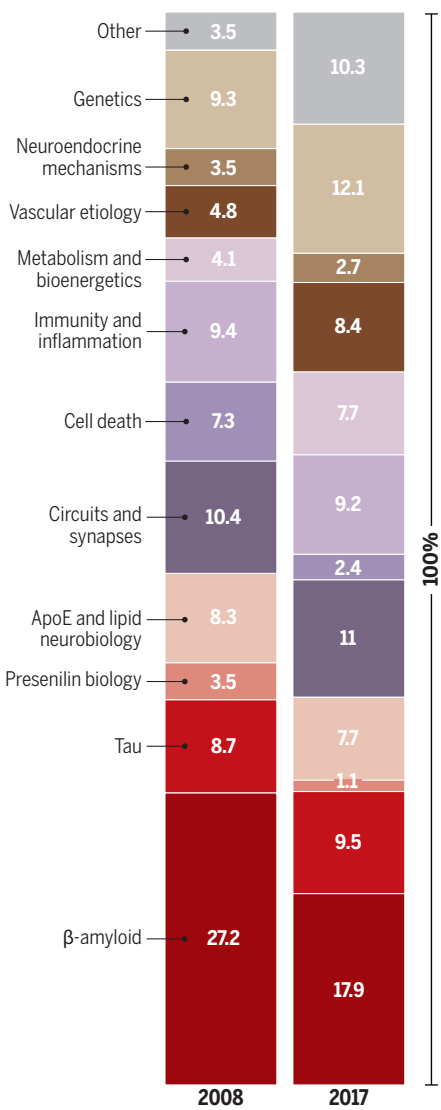
Deadly rise

The number of people in the United States with Alzheimer's disease may reach nearly 14 million by the middle of this century.



Shifting priorities

Researchers seeking Alzheimer's drugs are choosing targets other than β-amyloid and tau, the proteins long thought to be the key to treatments. The bars below reflect the percentage of National Institute on Aging grants for basic research devoted to various topics in 2008 and 2017.



CREDITS: (GRAPHIC) N. DESAI/*SCIENCE*; (DATA, TOP TO BOTTOM) L. E. HEBERT *ET AL. NEUROLOGY* 1778 (2013); NATIONAL INSTITUTE ON AGING AND THE ALZHEIMER'S ASSOCIATION; THE INTERNATIONAL ALZHEIMER'S AND RELATED DEMENTIAS

Basic researchers are exploring new hypotheses. Some of NIA's recent funding opportunities invite research on alternatives to the long-dominant idea that β-amyloid deposits outside brain cells and "tangles" of the protein tau inside neurons are the key drivers of Alzheimer's disease and the best treatment targets. The announcements call for proposals in less-explored areas, such as the role of protective genes, how neurodegeneration affects other animal species, and how metabolic changes might contribute to Alzheimer's. "This brought in many people who were reluctant to submit an Alzheimer's application in part because they thought, 'We're never going to do well, we're going to be outsiders,'" Hodes says. At a recent Senate hearing, he pointed out that of 452 investigators who won new Alzheimer's and related dementia grants from 2015 to 2017, 27% were receiving their first independent NIH grant, like Baker, and 36% were established researchers who had never had NIH support for Alzheimer's. (Some had funding from Alzheimer's foundations, however.) "We're not just repeating the things that failed and hoping we get a different result," Hodes says.

Masliah says that compared with a few years ago, when less than half of NIH's portfolio in Alzheimer's was devoted to areas other than β-amyloid or tau, it's now more than 60% for translational studies and about 70% for basic research. "I do believe there is more money available for us to explore these other ideas," says Carol Colton of Duke University in Durham, North Carolina, who studies inflammation as a possible cause of Alzheimer's. She and others add, however, that the academics called on to review NIH grant proposals are sometimes less open-minded than NIA staff and nix proposals in new areas. They "need to catch up," Colton says.

To cast an even wider net, NIA is offering 1-year funding supplements to researchers already funded by NIH in other areas who want to add an Alzheimer's component to their research. The hope is that the extra money will lead to full-fledged proposals.

Alzheimer's grants are now much easier to get than other NIA grants: For most Alzheimer's proposals this year, those ranked in the top 28th percentile by peer-review panels get money. For non-Alzheimer's grants, that pay line is the 19th percentile. The competition for grants is still stiff, Hodes stresses. After all, he notes, high-quality applications for the Alzheimer's pool of money have "increased dramatically" in the last couple years "as word got out."

NIA grantees in fields with scarcer funding aren't complaining, so far. Some recipients even suggest they're benefiting because competitors in the field of aging are shifting into Alzheimer's. "Paradoxically, the new funding injection could improve everyone's chances of funding," says Duke psychologist Terrie Moffitt, a member of NIA's advisory council.

NIA HAS HAD TO BE CREATIVE to cope with the tide of applications for the Alzheimer's bounty, agency officials say. After a crushing scramble to process grant proposals last summer, this year NIA called early for proposals and scheduled peer-review panels even before it knew its final 2018 budget. Adding to the pressure, President Donald Trump's administration imposed a federal hiring freeze last year that was only recently lifted at NIH. "I think our staff has managed heroically to still be doing an extremely conscientious job. … Where we've compromised probably is the quality of life of a lot of our staff," Hodes says.

At the NIH Center for Scientific Review in Bethesda, which arranges peer-review panels for much of the funding, "We're handling the load as best we can," says acting Director Noni Byrnes. The pool of potential reviewers—U.S. Alzheimer's researchers who aren't applying for the new funding themselves and so don't have a conflict of interest—is limited. So, for NIA-organized review panels, the institute is also using Alzheimer's experts in Canada and Europe, says Ramesh Vemuri, NIA's chief of scientific review.

Senator Susan Collins (right), visiting a retirement home specializing in dementia care, co-sponsored a bill that made research on Alzheimer's disease a national priority.

PHOTO: PORTLAND PRESS HERALD/CONTRIBUTOR/GETTY IMAGES

Clinical trials won't be easy to staff either. Clinical researchers and neuropathologists focused on dementia are in short supply, says Alzheimer's Association Chief Science Officer Maria Carrillo. NIA is trying to attract them by funding fellowships. Another huge problem is finding enough subjects for trials—especially those who are at high risk for the disease but still without symptoms, the population on which some researchers think amyloid-busting drugs could yet work. NIA plans to launch a national recruitment strategy that includes raising awareness about trials.

Looming over the massive research push is the 2025 goal. It was set when optimism ran high that drug trials based on the β-amyloid hypothesis would pan out, Carrillo and others say. But if patients must begin antiamyloid treatments well before symptoms set in, seeing clinical benefits could take decades, Gandy notes. And the chances of meeting the deadline by targeting a different disease mechanism are small; such treatments remain far off. Still, Holtzman hopes for good news from an antiamyloid treatment trial. "Something is likely to be approved by 2025. It won't be the be all, end all," he says, but he hopes it will keep everyone motivated. "Because we don't just need money from the NIH, we need the pharmaceutical industry to not drop out"—as Pfizer did this year when it announced it was abandoning Alzheimer's research.

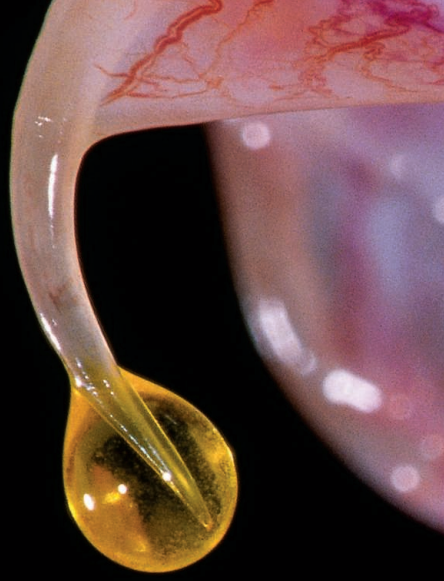
Some researchers point to the mixed success of NIH's other disease "wars": AIDS funding hasn't led to a cure or a vaccine, though it has yielded drugs that allow people infected with HIV to lead nearly normal lives. The war on cancer has led to treatments that are improving survival, but cancer remains the second-leading cause of death in the United States.

That history makes former NIH Director Harold Varmus cautious about the 2025 goal. "No one denies the enormous need to make progress against Alzheimer's," he says. But, "I wish a date were not attached."

Hodes concedes that, like real wars, disease wars can last far longer than anyone imagined—or feared. But that doesn't mean it was a mistake to launch an all-out offensive against Alzheimer's disease, he says. "If 2025 comes and we haven't achieved all we wanted, I'm not going to stop there and declare failure." ■

INSIGHTS

PERSPECTIVES



EVOLUTION

Venoms to the rescue

Insights into the evolutionary biology of venoms are leading to therapeutic advances

By **Mandë Holford[1], Marymegan Daly[2], Glenn F. King[3], Raymond S. Norton[4]**

Venomous animals have been admired and feared since prehistoric times, and their venoms have been used to both benefit and impair human health. In 326 BCE, Alexander the Great encountered lethal arrowheads in India that, based on the symptoms of dying soldiers, were most likely laced with venom from the deadly Russell's viper. By contrast, snake venom has been used in Ayurvedic medicine since the 7th century BCE to prolong life and treat arthritis and gastrointestinal ailments, while tarantulas are used in the traditional medicine of indigenous populations of Mexico and Central and South America. The modern era of venom research has so far yielded six venom-derived drugs (*1*). Recent work has elucidated the evolutionary biology of venoms and provided an impressive diversity of new therapeutic drug candidates.

Venomous organisms are ubiquitous. All known animal phyla contain venomous species. There are more than 220,000 known venomous animal species, or ~15% of all described animal biodiversity on Earth. Venomous animals inhabit virtually all marine and terrestrial habitats, ranging from desert snakes and scorpions to Antarctic sea anemones and jellyfish. However, most of their venoms have not been studied. For example, invertebrates make up more than 90% of all extant species, yet we know very little about their venoms (*2*). In large part, this neglect has been due to the lack of appropriate technologies for studying the tiny amounts of venom that can be extracted from small animals. However, the recent revolution in omics technologies (genomics, transcriptomics, proteomics) has enabled the study of venoms from animals that are small, rare, or hard to maintain in the lab (*3*).

These recent investigations of a broad range of animals have provided new perspectives on venom diversity, evolution, and pharmacology (see the figure). For example, studies of venom in Cnidaria, the most ancient venomous lineage that evolved ~500 million years ago, suggest that there are differences in the tempo, mode, and nature of evolution in old and young venoms, with evolutionarily older venoms experiencing relatively more purifying selection (*4*). This observation necessitates reconsideration of commonly held ideas about toxin evolution based on studies of snakes and cone snails. The latter animals are relatively young in an evolutionary sense, and their venom toxins are still undergoing strong diversifying selection.

Another example of how studies of neglected or understudied venomous organisms are proving beneficial is the recent examination of centipede venoms. The results show both toxin radiations within gene families and convergent recruitment of toxin genes, highlighting the multiplicity of processes by which venom toxins arise and diversify even within a single lineage (*5*). The discovery of venom in remipede crustaceans, the sister lineage to Hexapoda (Arthropoda), provides context for understanding and comparing the similarities and differences in the venom of centipedes, spiders, and insects (*6*). Extensive research on a broad range of organisms is imperative to derive and test robust hypotheses about venom as it relates to species diversification and predator-prey interactions, and to describe the immense biodiversity of animals on Earth.

The organs that produce venom add yet another dimension to the evolutionary pressures that govern venom phenotype and genotype, and as we investigate a broad swathe of venomous organisms, we clarify more of the story. For instance, in centipedes, the nature of the venom production facilities probably constrains venom diversity (*5*), whereas it may enhance it in cone snails (*7*) and assassin bugs (*8*). Centipedes modified the first pair of walking legs into appendages (forcipules) able to deliver venom and evolved venom glands from cuticular dermal glands (*5*), whereas assassin bugs have two morphologically distinct venom glands that each produce entirely different venoms with contrasting ecological roles—one venom is used for predation, and the other for defense against predators (*8*). Similarly, the distal and proximal portions of cone snail venom glands can be subdivided to deliver venom toxins needed for predation or defense (*7*). Understanding the bifurcation of venom use has important implications for deciphering how venom has evolved.

The acquisition of venom is a transformative event in the evolution of an animal, because it remodels the predator-prey interaction from a physical to a biochemical battle, enabling venomous animals to prey on, and defend themselves against, much larger animals. The ongoing evolutionary arms race between venomous animals and their prey, which can evolve resistance to venoms, has resulted in venoms that are extremely complex, with toxins that seek out physiological molecular targets with exquisite potency and selectivity (*9*).

Just as venom frequently evolves by convergent evolution from a limited suite of common genes, mammalian resistance to venom relies, at least in part, on independent instances of modification of a limited suite of genes (*10*). There is substantial similarity throughout the animal kingdom in the basic molecular structure of venom toxins and

[1]Hunter College and CUNY Graduate Center, New York, NY, USA. [2]Ohio State University, Columbus, OH, USA. [3]The University of Queensland, Brisbane, Australia. [4]Monash University, Melbourne, Victoria, Australia. Email: mholford@hunter.cuny.edu

PHOTO: JOE MCDONALD/GETTY IMAGES

Over 200,000 animal species, such as this prairie rattlesnake (*Crotalus viridis*), produce venom.

their targets. The latter include neuronal ion channels and receptors in the case of neurotoxic venoms (most invertebrate venoms) and components of the blood coagulation cascade in the case of hemotoxic venoms (leech, snake, and lizard venoms). Indeed, the definitive pharmacological markers for many ion channels are venom toxins, such as tetrodotoxin in the case of voltage-gated sodium channels and α-conotoxins for nicotinic acetylcholine receptors. Ion channels and receptors drive physiological processes involved in everything from seeing to breathing, and venom peptides are providing the tools with which to investigate, manipulate, and modify these macromolecular machines (*11*).

The specificity, potency, stability, and speed with which venom peptides manipulate their molecular targets make them ideal candidates for therapeutics. Examples of the value of venom peptides in guiding the development of human therapeutics include the U.S. Food and Drug Administration (FDA)–approved drugs exenatide, an antidiabetic peptide from the venomous Gila monster (*Heloderma suspectum*), and ziconotide, an analgesic peptide from the venomous cone snail (*Conus magus).* Promising new developments for venom peptide therapeutics and insecticides include monomeric insulins found in the venom of cone snails (*12*); the sea anemone venom peptide ShK for treatment of autoimmune diseases (*13*); chlorotoxin from the deathstalker scorpion for imaging brain tumors during surgery (*14*); and spider toxins for use as eco-friendly insecticides (*15*).

Venom research is a highly interdisciplinary enterprise, requiring studies of the biology and ecology of venomous organisms, the structure and function of venom deployment, the biochemistry and pharmacology of venoms, the pathophysiological effects that venom induces in prey and predators, and the translational development of venom components for biomedical and biotechnological applications. As research on neglected or poorly studied venomous organisms gains momentum with advanced omics techniques, a substantial database of new molecules with novel mechanisms of action will be produced. A major challenge facing this emerging field will be the development of robust high-throughput assays for determining the molecular targets of these new compounds.

Recent progress has greatly expanded the known molecular targets of venom toxins and highlighted the value of venoms as a source of pharmacological tools and drug leads. Similarly, the field will have to develop new computational methods for modeling the interaction of toxins and their molecular targets to reduce cost, labor, and guesswork in identifying selective venom peptides. Unlike small molecules, venom peptides have substantial barriers that must be addressed when considering in silico procedures (*11*). An evolutionarily informed perspective will help to focus venom research to leverage the extraordinary biochemical warfare nature has created to yield transformative developments for evolutionary biology, chemical biology, and the discovery of therapeutics and bioinsecticides. ■

REFERENCES AND NOTES

1. G. F. King, *Expert Opin. Biol. Ther.* **11**, 1469 (2011).
2. B. M. von Reumont, L. I. Campbell, R. A. Jenner, *Toxins* **6**, 3488 (2014).
3. A. Verdes *et al.*, *Toxins* **8**, 117 (2016).
4. M. Jouiaei *et al.*, *Mol. Biol. Evol.* **32**, 1598 (2015).
5. E. A. B. Undheim *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 4026 (2015).
6. B. M. von Reumont, E. A. B. Undheim, R.-T. Jauss, R. A. Jenner, *Toxins* **9**, 234 (2017).
7. S. Dutertre *et al.*, *Nat. Commun.* **5**, 3521 (2014).
8. A. A. Walker *et al.*, *Nat. Commun.* **9**, 755 (2018).
9. K. Arbuckle, R. C. Rodríguez de la Vega, N. R. Casewell, *Toxicon* **140**, 118 (2017).
10. M. L. Holding, D. H. Drabeck, S. A. Jansa, H. L. Gibbs, *Integr. Comp. Biol.* **56**, 1032 (2016).
11. A. E. Leffler *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E8100 (2017).
12. J. G. Menting *et al.*, *Nat. Struct. Mol. Biol.* **23**, 916 (2016).
13. K. G. Chandy, R. S. Norton, *Curr. Opin. Chem. Biol.* **38**, 97 (2017).
14. A. N. Mamelak *et al.*, *J. Clin. Oncol.* **24**, 3644 (2006).
15. G. F. King, M. C. Hardy, *Annu. Rev. Entomol.* **58**, 475 (2013).

ACKNOWLEDGMENTS

M.H. acknowledges support from the U.S. National Science Foundation and the Camille Dreyfus Teacher-Scholar Award. M.D. acknowledges support from the U.S. National Science Foundation. G.F.K. and R.S.N. are supported by a Principal Research Fellowship from the Australian National Health and Medical Research Council.

10.1126/science.aau7761

An evolutionary tree of venomous animals

Animal venoms are an important source of potential therapeutic compounds, but traditional methods only allowed study of venoms from a few species. Genomic and proteomic advances are now opening the range of animal species to scientific study.

SIGNALING

From oncogenic mutation to dynamic code

Oncogenic BRAF mutations can distort downstream signaling outcomes

By **Walter Kolch**[1,2,3] *and* **Christina Kiel**[1,2,4]

Signal transduction pathways (STPs) convert biochemical reactions into precise and reproducible biological outcomes. These functions are performed reliably and reproducibly against a background of noise (variation) arising from the stochastic nature of biochemical reactions (*1*). Indeed, information theory analysis of STPs indicates that they have a limited capacity to discriminate information, including different ligands or different activation states of components (*2*). However, this discrimination is dramatically enhanced by adding dynamic information, such as signal rise time, signal duration, amplitude, and decay rate (*3*). The observation that differential activation dynamics of the extracellular signal-regulated kinase (ERK) pathway can determine whether rat pheochromocytoma cells proliferate or differentiate was reported more than 20 years ago (*4*), and evidence has since accumulated that STP dynamics control cell fate decisions. However, we are still struggling to understand how signaling dynamics is encoded and decoded and how pathological changes, such as the expression of mutant proteins, affect the dynamic STP code. On page 892 of this issue, Bugaj *et al.* (*5*) make use of new tools with which to decipher this code and reveal how certain cancer-associated BRAF mutations can corrupt the dynamic STP code and trick cells into unlicensed proliferation.

BRAF is a pivotal kinase in the biochemical circuitry that controls cell proliferation and transformation. It links the activation of RAS, a group of small G proteins that are activated by many growth factor receptors, to mitogen-activated extracellular

[1]*Systems Biology Ireland, University College Dublin, Belfield, Dublin 4, Ireland.* [2]*School of Medicine, University College Dublin, Belfield, Dublin 4, Ireland.* [3]*Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland.* [4]*Charles Institute, University College Dublin, Belfield, Dublin 4, Ireland.*
Email: walter.kolch@ucd.ie

GRAPHIC: N. DESAI/*SCIENCE*

kinase (MEK) and ERK (*6*). ERK has >650 substrates that regulate cell proliferation, survival, differentiation, metabolism, and many other biochemical processes (*7*). ERK activation dynamics control both the regulation of gene transcription (*8*) and peripheral biochemical processes that determine whether a cell undergoes proliferation or differentiation (*9*). The RAS-RAF-MEK-ERK signaling module is altered in >50% of human cancers because of activating mutations or overexpression of growth factor receptors or mutations in RAS and BRAF (*10*). Potent inhibitors of BRAF and MEK kinases have been developed to treat various cancer types, but their clinical deployment has resulted in several surprises. For example, RAF inhibitors cause a paradoxical activation of ERK because they promote the dimerization of BRAF with CRAF, subverting a normally transient element of physiological RAF activation into a long-lasting change of ERK activation dynamics, which leads to resistance to RAF inhibitors (*10*). This has led to the hypothesis that drug mechanisms of action have to be considered in the context of STP topologies and dynamics, rather than only being optimized for inhibiting a specific target.

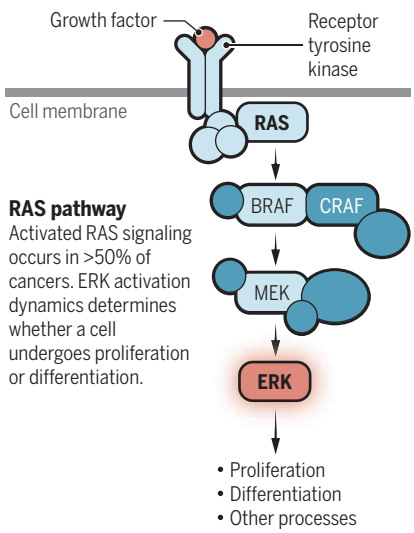
Bugaj *et al.* used optogenetic tools to switch on and off the expression of either RAS or BRAF with precise temporal dynamics. By modulating the frequency of activation, they could investigate the effects of common RAS and BRAF mutations on ERK activation dynamics. Interestingly, they found that in cells with different oncogenic RAS and BRAF mutations, ERK activity was still dependent on optogenetic RAS activation. Recent data indicate that oncogenic RAS mutants can still transition between active and inactive conformations and thus could maintain responsiveness to growth factor stimulation (*11*), whereas the activity of the BRAF mutants studied by Bugaj *et al.* was previously deemed to be RAS-independent (*12*). If confirmed, this suggests that in general, oncogenically activated proteins are not deadlocked into a constitutively active state but are still subject to regulation. This could change the way we design combination therapies by introducing the simple principle that inhibiting upstream activators of an oncogenic protein will augment the efficacy of drugs targeting the oncogenic protein itself.

Additionally, the common BRAF-V600E mutant produced fast-responding pulses of ERK activation, as occurs when BRAF is not mutated, in response to optogenetic activation of RAS. By contrast, oncogenic mutations in the adenosine triphosphate (ATP)–binding loop (P-loop) of BRAF induced delays in both activation and deactivation of ERK, causing pulses to merge into sustained ERK activity at higher frequencies of RAS stimulation. Using various RAF and MEK inhibitors and an optogenetic BRAF activation system, the authors mapped the source of this reduction of kinetic resolution to BRAF itself. They suggest that the requirement for P-loop mutants to dimerize with and signal via activated CRAF blurs and broadens the ERK response kinetics and leads to changes in early gene expression and proliferation. Importantly, this finding ties an oncogenic mutation to alterations in signaling dynamics. This replaces the traditional view that mutations are off-on switches with a more nuanced picture in which mutations distort the dynamic STP code (see the figure).

It will be interesting to examine how the frequency of RAS activation pulses affect the kinetics of BRAF-CRAF heterodimerization, and how P-loop mutations selectively modulate these kinetics. P-loop mutations diminish MEK binding to BRAF (*13*), which may retard ERK activation. The P-loop is also the target for autoinhibitory phosphorylation, which is abolished by P-loop mutations (*14*). The lack of autoinhibition could plausibly explain the slower signaling decay of BRAF P-loop mutants. Understanding the mechanism underlying the signal distortion in cells expressing these mutants will be important to gauge the role of such dynamic alterations in malignant transformation. BRAF-V600E does not affect the resolution of activated RAS signaling in terms of ERK activation dynamics, yet it is a potent oncogene, suggesting that altered signaling dynamics may contribute to transformation by certain mutations but not by others.

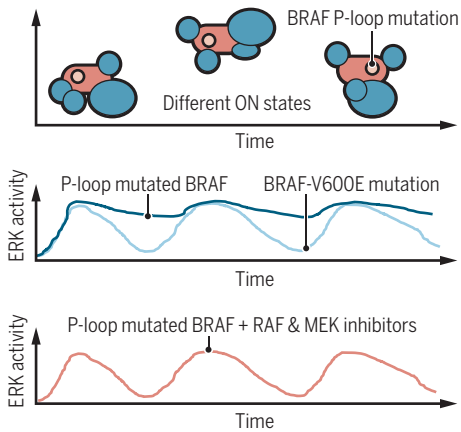
Cancer mutations can distort the dynamic signaling code

Oncogenic mutations were thought to lock oncoproteins in an active state. It is now appreciated that protein-protein interactions (PPIs) and network context cause dynamic changes in the outcome of signaling.



RAS pathway
Activated RAS signaling occurs in >50% of cancers. ERK activation dynamics determines whether a cell undergoes proliferation or differentiation.

PPIs change activation states
Certain oncogenic BRAF mutations can alter ERK activation dynamics by altering PPIs, which can be rectified by pathway inhibitory drugs.



These intricacies highlight that mutations do not just statically affect STPs; they can have wide repercussions on functional features of STPs, including signaling dynamics and protein-protein interactions that change pathway configurations. This view has wide ramifications for understanding drug resistance. For instance, the mechanism of kinase dimerization conveying drug resistance is a general principle if the kinases can allosterically activate each other (*15*). In a broader context, mutations in proteins such as RAS that can connect to many STPs may cause a profound rewiring of both the topologies and signaling dynamics of downstream effector pathways. This viewpoint also helps rationalize why targeting mutated proteins with potent inhibitors has frequently resulted in unexpected responses or poor clinical efficacy. It is time to move the crosshairs of drug discovery away from single molecules and toward considering the functional dynamic context that potential targets are embedded in. ■

GRAPHIC: C. BICKEL/*SCIENCE*

REFERENCES

1. V. Shahrezaei, P. S. Swain, *Curr. Opin. Biotechnol.* **19**, 369 (2008).
2. R. Cheong *et al.*, *Science* **334**, 354 (2011).
3. J. Selimkhanov *et al.*, *Science* **346**, 1370 (2014).
4. C. J. Marshall, *Cell* **80**, 179 (1995).
5. L. J. Bugaj *et al.*, *Science* **361**, eaao3048 (2018).
6. D. Matallanas *et al.*, *Genes Cancer* **2**, 232 (2011).
7. E. B. Ünal *et al.*, *FEBS Lett.* **591**, 2607 (2017).
8. T. Nakakuki *et al.*, *Cell* **141**, 884 (2010).
9. A. von Kriegsheim *et al.*, *Nat. Cell Biol.* **11**, 1458 (2009).
10. M. Holderfield *et al.*, *Nat. Rev. Cancer* **14**, 455 (2014).
11. S. Lu *et al.*, *Chem. Rev.* **116**, 6607 (2016).
12. Z. Yao *et al.*, *Cancer Cell* **28**, 370 (2015).
13. J. R. Haling *et al.*, *Cancer Cell* **26**, 402 (2014).
14. M. Holderfield *et al.*, *Cancer Cell* **23**, 594 (2013).
15. B. N. Kholodenko, *Cell Rep.* **12**, 1939 (2015).

10.1126/science.aau8059

AGRICULTURE

Insect threats to food security

Pest damage to crops will increase substantially in many regions as the planet continues to warm

By **Markus Riegler**

Globally, one out of nine people suffers from chronic hunger, and undernourishment is growing (*1*). Global average surface temperatures are also rising and are projected to increase by 2° to 5°C this century, with negative impacts on agricultural production. Even today, despite substantial plant protection efforts, about one-third of crops are lost to insect pests, pathogens, and weeds. How will climate warming affect these crop losses on a global scale? On page 916 of this issue, Deutsch *et al.* (*2*) evaluate the impact of rising average surface temperatures on yield losses due to insects in wheat, maize, and rice, which are staple foods for billions of people. The results show that insects will cause significantly increased grain loss across many regions of a warmer world.

Consumption by insects—and thus damage to crops—is directly linked to insect metabolism and population size, both of which generally increase with temperature. However, insect metabolism and demographics vary seasonally, latitudinally, and across insect species. Deutsch *et al.* capture some of this complexity with a spatially explicit insect population metabolism model, which integrates data on temperature-driven changes in metabolic and growth rates that were previously collated for 38 insect species from different latitudes. The authors estimate additional grain yield losses of 10 to 25% per 1°C of global warming. Projected losses are highest in the temperate regions of the Northern Hemisphere that produce most grain, as a result of increases in insect population growth and metabolic rates. In the tropics, projected yield loss due to warming is less acute because elevated insect metabolic rates will be offset by reduced insect population growth rates. However, tropical regions already experience the highest levels of yield loss.

Many previous studies have investigated temperature responses of pest insects, and some have predicted how pest populations affect yields in the context of climate change (*3*). However, these studies were mostly on individual crop types at local or regional scales. Few have attempted to link temperature responses of pest insects and the damage they cause under warming more widely, and never on a global scale.

Insects are highly sensitive to environmental temperatures; they can display a variety of responses to climate warming (*3*). The input data for the grain loss model derive mostly from laboratory studies undertaken at different constant-temperature regimes (*4*), yet thermal fitness can vary across life history stages and be modulated by insect thermoregulatory behavior. According to sensitivity tests undertaken by Deutsch *et al.*, the resulting uncertainty does not affect the overall estimates of global crop loss increases. Further research is required to verify the results for the many pest species of wheat, maize, and rice (see the photo) that Deutsch *et al.* could not include because of insufficient data.

A caterpillar of the rice leaf folder (*Cnaphalocrocis medinalis*) infects a rice plant. Because of limited data, Deutsch *et al.* could not include this and many other pest species of wheat, maize, and rice in their model.

Insect responses to climate change are further complicated by factors not included in the study. For example, many insect pests are vectors of plant pathogens that also cause crop losses. Predictions based on population growth and metabolic rates may thus underestimate crop damage due to insect vectors under global warming. Scientists should also aim to better understand the impact of geographic range expansions, human-assisted introductions, and biological invasions of pest species in a warming world (*5*).

Additional intricacies arise from the fact that climate warming also affects natural enemies of pests and that pests are exposed to indirect plant-mediated effects involving changed plant nutrient content and defense (*6*). Some responses may be compounded by interactive effects of warming with other factors, such as changing rainfall patterns and increasing atmospheric CO_2 concentrations. Furthermore, the evolutionary adaptation potential and the role of plasticity in insect responses to climate change remain poorly understood for field insect populations (*7*). Global warming also involves more frequent and intense temperature extremes. More common heatwaves may have larger impacts than average temperature increases (*8*).

The substantial increases in pest damage forecast by Deutsch *et al.* call for action on climate change mitigation and adaptation. Everyone must be involved in change: farmers, industries, policy-makers, and the wider society. Farming communities are already adapting, for example, in choosing which crop varieties to plant when and where. There is also an increased need to focus on plant protection, particularly given that many insecticides are being banned over human and environmental health concerns. Reinvigoration of debates about sustainable farming systems (*9*), including the role of genetically modified crop plants (*10*), is essential.

Humanity faces this food security challenge at a time when training and job opportunities for expert entomologists are shrinking (*11*). These experts are urgently needed to deal with insect pest problems and, ironically, with concurrent threats to insect biodiversity. Research will also require the interplay of empirical and theoretical approaches because model-based predictions can only be as good as the data that feed into them. Ecoinformatics (*12*) and the use of big data to answer ecological questions such as those posed by Deutsch *et al.* hold great promise to unravel large-scale system responses, and this improved understanding is necessary to develop and implement adaptation strategies. ■

Hawkesbury Institute for the Environment, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia. Email: m.riegler@westernsydney.edu.au

REFERENCES AND NOTES

1. FAO *et al.*, *The State of Food Security and Nutrition in the World* (FAO, Rome, 2017).
2. C. A. Deutsch *et al.*, *Science* **361**, 916 (2018).
3. J. S. Terblanche *et al.*, in *Climate Change and Insect Pests*, C. Björkman, P. Niemelä, Eds. (CABI, 2015), pp. 92–118.
4. M. R. Frazier *et al.*, *Am. Nat.* **168**, 512 (2006).
5. D. P. Bebber *et al.*, *Nat. Clim. Change* **3**, 985 (2013).
6. E. H. DeLucia *et al.*, *Plant Physiol.* **160**, 1677 (2012).
7. C. M. Sgrò *et al.*, *Annu. Rev. Entomol.* **61**, 433 (2016).
8. R. M. B. Harris *et al.*, *Nat. Clim. Change* **8**, 579 (2018).
9. E. M. Lichtenberg *et al.*, *Global Change Biol.* **23**, 4946 (2017).
10. R. J. Roberts, *J. Innov. Knowl.* **3**, 61 (2018).
11. H. D. Loxdale, *Entomol. Exp. Appl.* **159**, 121 (2016).
12. J. A. Rosenheim, C. Gratton, *Annu. Rev. Entomol.* **62**, 399 (2017).

ACKNOWLEDGMENTS

I thank the Australian Research Council for funding via the Industrial Transformation Training Centre scheme.

10.1126/science.aau7311

VIROLOGY

The reemergence of yellow fever

Since 2016, yellow fever outbreaks have become a major public health concern

By **Alan D. T. Barrett**

Yellow fever is a viral hemorrhagic fever with a case fatality rate up to 50%. It is caused by yellow fever virus (YFV), a mosquito-borne flavivirus that is related to dengue and Zika viruses. Despite an effective vaccine (17D), the virus still causes major outbreaks, as occurred in Brazil between December 2016 and March 2018 where there were >2000 confirmed cases, including >500 deaths, as well as >4000 epizootics (yellow fever in nonhuman primates) (*1*). On page 894 of this issue, Faria *et al.* (*2*) provide a genetic investigation of the outbreak in Brazil from December 2016 to October 2017, demonstrating the origins and movement of YFV during the outbreak. They determined that the outbreak originated in northeastern Brazil and moved southward to areas where the virus had not been found previously. Surprisingly, YFV moved at a rate of 4.25 km/day, which probably explains the magnitude of the outbreak. Modeling infectious disease outbreaks with phylogeographic tools (based on the geographic distribution of viruses according to viral genome sequence) as well as phylodynamic tools (based on the interaction of epidemiologic, immunologic, and evolutionary factors in viral genetics) has played a critical role in understanding outbreaks and developing public health countermeasures.

To date, there has been little modeling of YFV outbreaks because few isolates have been available to study, and we have instead relied on vaccination strategies. The study of Faria *et al.* demonstrates the potential of mapping viral incidence and spread almost in real time and their potential to contribute to control strategies, such as the current World Health Organization (WHO) program called Eliminate Yellow Fever Epidemics (EYE) that aims to eliminate the disease by 2026 (*3*).

However, the investigation of yellow fever outbreaks in real time is not straightforward. YFV has a sylvatic, or forest, transmission cycle involving tree hole–breeding mosquitoes and nonhuman primates. Human cases usually occur in forested areas; hence, other than during large outbreaks, it is difficult to obtain virus samples for analysis (*4*). For these reasons, until June 2016 there were only 42 YFV genomic sequences available for study. Today this has increased to ~135 genomes, with most of the additional sequences coming from the outbreak in Brazil. YFV is found in 44 countries in sub-Saharan Africa and tropical South America, and seven virus genotypes have been identified (*5*, *6*); many additional genomic sequences are needed to understand virus activity within the geographic range of the virus, especially because RNA viruses such as YFV continually evolve. Nonetheless, it is clear that the capabilities are now becoming available as the genome database expands over time and space.

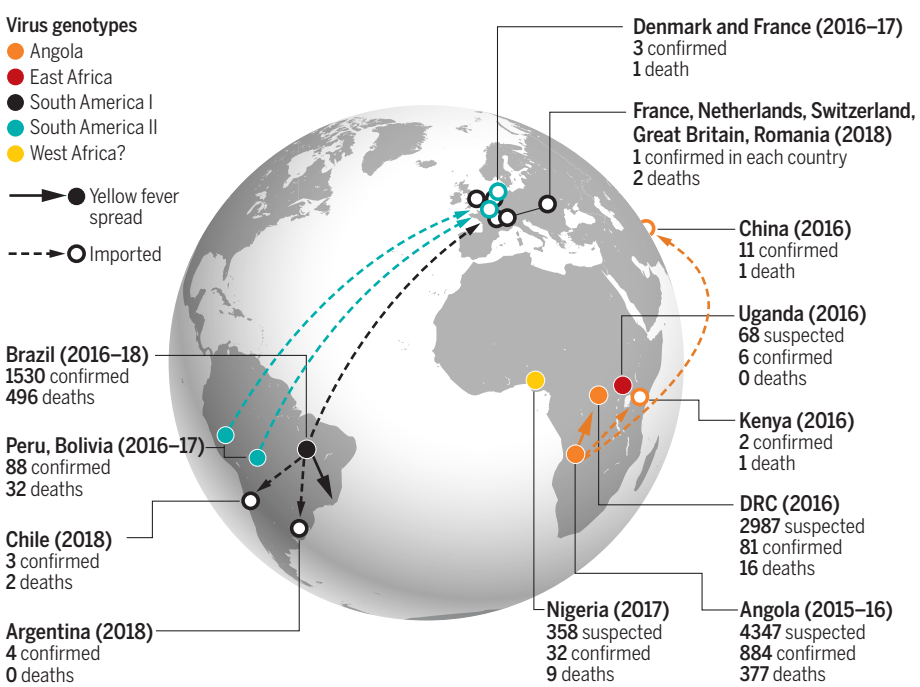
WHO, the Pan American Health Organization, and related sponsors have been successful at controlling yellow fever through mass vaccination campaigns such that there were no outbreaks in West Africa in 2015, a region where historically the most cases of disease have been recorded. However, dramatic increases in yellow fever incidence have occurred recently, and they are occurring in areas that had been considered free of yellow fever (see the figure).

The most dangerous form is urban yellow fever, where the transmission cycle involves domestic *Aedes aegypti* mosquitoes and humans. Urban yellow fever occurred in 2016 in both Angola and the Democratic Republic of Congo (DRC), and 11 Chinese workers infected in Angola returned to China, where they developed yellow fever (*7*). This was the first time that any cases of yellow fever had been reported in Asia, but these were imported cases and there were no secondary cases in China. Furthermore, there was a concurrent outbreak in Uganda caused by a virus genotype different from that in Angola (*8*). World vaccine supplies were exhausted twice during early 2016, and a dose-sparing regimen of 17D had to be used whereby one-fifth of a full dose was administered (*9*, *10*). The concurrent clinical trial of dose-sparing vaccination has shown that the immune response is not inferior to a full dose 4 to 5 weeks after immunization (*10*). The epidemic in Africa was controlled by September 2016. Unfortunately, yellow fever was then reported in Brazil in December 2016 (*1*). Again, 17D supplies were exhausted and dose sparing was used to immunize 24 million individuals in Brazil.

Whereas urban yellow fever outbreaks periodically occur in Africa, such as in Angola and the DRC in 2016, they are rare in South America. Faria *et al.* provide persuasive evidence that the recent Brazilian

Yellow fever outbreaks

In 2016–2018, there have been numerous outbreaks of suspected and confirmed cases of yellow fever, resulting in imported cases to other countries (dashed arrows), but these did not result in secondary cases.



GRAPHIC: N. CARY/SCIENCE

Sealy Institute for Vaccine Sciences and Department of Pathology, University of Texas Medical Branch, Galveston, TX 77555, USA. Email: abarrett@utmb.edu

outbreak was due to the forest yellow fever transmission cycle and not the urban cycle. The last documented large urban yellow fever outbreak in the Americas was likely in Brazil in 1928; subsequent urban yellow fever outbreaks in South America have been very small, each involving no more than nine cases. This suggests that the epidemiologies of urban yellow fever in Africa and South America are different and require further investigation.

The exhaustion of vaccine supplies on multiple occasions, together with the resurgence of YFV activity in the past 3 years, is a cause for concern. There were many outbreaks between 2016 and 2018, and in addition, international travelers caused multiple importations into countries outside these outbreaks (see the figure). Notably, the Brazil outbreak resulted in travelers transporting YFV to seven countries during 2018, including five in Europe (*11*, *12*). By comparison, only three cases were imported into Europe during the previous 16 years. Clearly, we cannot rely on YFV control by vaccination alone, and modeling is a critical component.

There are reasons to be optimistic. Advances in genomic sequencing technology of YFV isolates have enabled modeling of YFV activity, which has been routinely undertaken for other pathogens such as influenza virus. As the database improves, so will our understanding of YFV movement and our ability to identify areas where the virus has potential to cause outbreaks. Concurrent activities of the EYE strategy will result in production of vaccine that can be appropriately distributed with input from YFV modeling. Therefore, it may be possible to eliminate yellow fever epidemics by 2026, as planned. ■

REFERENCES AND NOTES

1. Pan American Health Organization, *Yellow Fever – Epidemiological Update 20 March 2018*; www.paho.org/hq/index.php?option=com_docman&task=doc_view&Itemid=270&gid=44111&lang=en.
2. N. R. Faria *et al.*, *Science* **361**, 894 (2018).
3. World Health Organization, *Eliminating Yellow Fever Epidemics (EYE) Strategy*; www.who.int/csr/disease/yellowfev/meeting-demand-for-vaccines/en/.
4. T. P. Monath, P. F. Vasconcelos, *J. Clin. Virol.* **64**, 160 (2015).
5. J. P. Mutebi *et al.*, *J. Virol.* **75**, 6999 (2001).
6. J. E. Bryant *et al.*, *PLOS Pathog.* **3**, e75 (2007).
7. R. Song *et al.*, *Emerg. Infect. Dis.* **24**, 1383 (2018).
8. H. R. Hughes *et al.*, *Emerg. Infect. Dis.* **24**, 1598 (2018).
9. A. D. T. Barrett, *N. Engl. J. Med.* **375**, 301 (2016).
10. S. Ahuka-Mundeke *et al.*, *N. Engl. J. Med.* 10.1056/NEJMoa1710430 (2018).
11. D. H. Hamer *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **67**, 340 (2018).
12. European Centre for Disease Prevention and Control, *Rapid Risk Assessment: Outbreak of Yellow Fever in Brazil, Third Update*; https://ecdc.europa.eu/en/publications-data/rapid-risk-assessment-outbreak-yellow-fever-brazil-third-update.

ACKNOWLEDGMENTS

Supported by grants from NIH and the Gillson Longenbaugh Foundation.

10.1126/science.aau8225

CANCER

Fusion oncogenes—genetic musical chairs

Ewing sarcoma–driver fusion genes can result from complex genomic rearrangements

By **Marcin Imielinski**[1,2] *and* **Marc Ladanyi**[3]

The cytogenetic definitions of many cancers predate the genome sequencing era. Indeed, some classes of cancers (largely subtypes of sarcomas, lymphomas, and leukemias) have long been defined by simple and distinct patterns of chromosomal changes, or karyotypes, that, in many cases, feature a single pathognomonic somatic translocation of two genomic regions that creates a fusion oncogene (for example, the Philadelphia chromosome translocation in chronic myelogenous leukemia results in the *BCR-ABL1* fusion oncogene) (*1*). Whereas many common cancers display genomic complexity consistent with multistep oncogenesis, such as carcinomas of breast and lung, cancers that are defined by translocations typically display simple karyotypes, suggesting that they were shaped by a single translocation. However, the cytogenetic simplicity of these cancers may mask more complex genomic events. On page 891 of this issue, Anderson *et al.* (*2*) report whole-genome sequencing (WGS) of 50 Ewing sarcomas (EWSs), an aggressive sarcoma that is defined by fusion between the EWS RNA binding protein 1 (*EWSR1*) gene on chromosome 22 and an E26 transformation-specific (ETS) family transcription factor gene, either *FLI1* at 11q24 or *ERG* at 21q11 (*3*). Anderson *et al.* show that ~40% of *EWSR1-FLI1* fusions and all *EWSR1-ERG* fusions arise via a complex rearrangement pattern called chromoplexy, which was first identified in prostate cancer (*4*). They suggest that chromoplexy "bursts" may be early initiating events in Ewing sarcomagenesis and mark a more aggressive form of the disease.

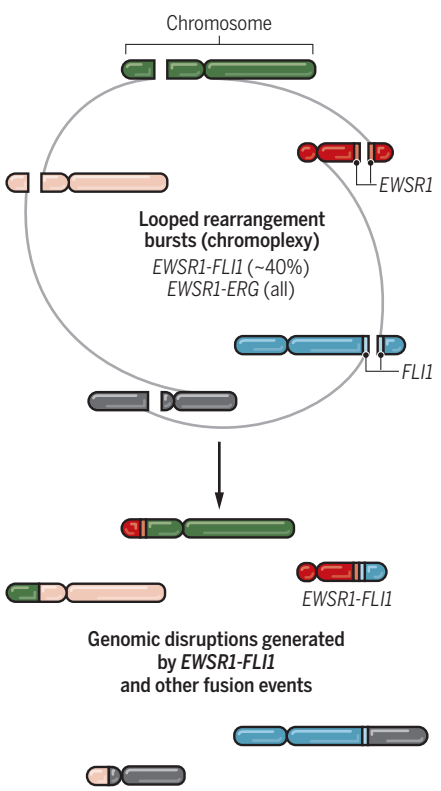
Whereas standard reciprocal translocations involve DNA breaks in two fusion partners, chromoplexy involves three or more breakpoints in the genome. Like the children's game of musical chairs, in which players are forced to stand up and find a new seat, three or more broken chromosome ends are forced to find a new partner. Unlike musical chairs, during which one of the chairs is removed at each round of play, every broken end finds a new partner, resulting in a loop pattern (see the figure). Chromoplexy is thus a complex means to an end: the formation of functional *EWSR1-FLI1* or *EWSR1-ERG* fusions that, upon expression, provide a selective growth or survival advantage.

Although *EWSR1-ERG* fusions were known to require more complex rearrangements because their opposing chromosome orien-

[1]*Department of Pathology and Laboratory Medicine, Englander Institute for Precision Medicine, Institute for Computational Biomedicine, Meyer Cancer Center, Weill Cornell Medicine, New York, NY, USA.* [2]*New York Genome Center, New York, NY, USA.* [3]*Department of Pathology and Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. Email: mai9037@med.cornell.edu; ladanyim@mskcc.org*

Fusions via chromoplexy

Bursts of complex rearrangements can generate Ewing sarcoma oncogenic fusion genes and are associated with genomic complexity, more frequent *TP53* mutations, and increased risk of relapse.



GRAPHIC: ADAPTED BY A. KITTERMAN/*SCIENCE*

tations preclude fusion-gene formation via a simple reciprocal translocation, the discovery by Anderson *et al.* of chromoplexy as the underlying genomic process helps to clarify this. However, that such a high proportion of *EWSR1-FLI1* fusions also arise by chromoplexy was not fully appreciated.

Notably, Anderson *et al.* provide evidence to suggest that chromoplexy-associated rearrangements occur simultaneously, presumably in the EWS cell of origin, the nature of which remains unclear. A fascinating question is whether these chromoplexy events can be linked to specific features of three-dimensional chromatin organization and/or of the transcriptional state of chromatin regions involved in chromoplexy rearrangements in this hypothetical cell type, given that they appear to be enriched in early replicating and transcriptionally active genomic regions, which may be prone to breakage as they are exposed during transcription. Gene pairs involved in fusions are often in close proximity in interphase nuclei, regardless of their chromosomal location (*5*), and this higher-order contiguity can be induced by specific transcription factors (*6*). A better understanding of the transcriptional regulation of the genes recurrently involved in chromoplexy-derived *EWSR1-FLI1* and *EWSR1-ERG* fusions might reveal the cell state or lineage in which EWS arises.

Given the young age of many EWS patients, one may speculate what exogenous or endogenous mutagen could be responsible for such a mutational "burst." Although radiation is a likely suspect in any disorder involving multiple chromosomal breaks, endogenous mutagens such as transposases and cytidine deaminases have also been linked to complex somatic rearrangements. Could EWS chromoplexy events be linked, for example, to the activity of an aberrantly expressed endogenous transposase such as PiggyBac transposase 5 (PGBD5), which was recently implicated in the genesis of the pathogenic gene rearrangements in childhood malignant rhabdoid tumors (*7*)? An alternative possibility is a constitutional or acquired DNA repair defect (*8*). Analysis of the sequence context surrounding chromoplexy breaks may provide clues and potentially point to a therapeutic vulnerability that could be used to treat EWS. Furthermore, perhaps EWS arising from chromoplexy may be responsive to immune checkpoint inhibition, given the preference of chromoplexy events for transcriptionally active regions that should result in multiple fusion transcripts, most of which are likely to be out of frame (except the driver fusion gene). Frameshift alterations represent an especially rich source of neoantigens (*9*), which can predict response to immune checkpoint inhibition.

Although Anderson *et al.* qualitatively validate a subset of chromoplexy events observed through WGS with spectral karyotype data, a comprehensive bridging of the divide between such large-scale views of chromosome structure and the detailed views of WGS will require the application of long-range WGS approaches using linked-read or proximity-ligation short-read sequencing and long-read sequencing, which uses more expensive and lower-throughput technologies to achieve read lengths that exceed 10 kilobase pairs. The long-range reconstruction of highly rearranged loci can yield insight into both the mutational processes generating complex structural variants and the consequences of these variants on DNA sequences (*10*).

The findings of Anderson *et al.* raise important clinical questions. The contribution of genetic analysis to the current standard of care for EWS is limited to confirmation of the diagnostic *EWSR1-FLI1* or *EWSR1-ERG* fusions. The discovery of genomic patterns associated with subsets of EWSs raises the question of whether additional molecular diagnostic modalities are warranted. If chromoplexy events are important clinical biomarkers for EWS disease aggressiveness, as the authors suggest, their findings may support a new indication for clinical WGS. However, additional analysis of more patient samples will be needed to confirm that the presence of chromoplexy is an independent prognostic predictor in EWS. This is because Anderson *et al.* find that chromoplexy-driven EWS more likely contains tumor protein 53 (*TP53*) mutations. Because *TP53* and stromal antigen 2 (*STAG2*) mutations and genomic complexity have each been associated with more aggressive EWS (*11*–*13*), dissecting the contribution of these factors to poor clinical outcomes in chromoplexy-derived EWS will be an important area of future work. More generally, the work of Anderson *et al.* has important clinical implications for the genomic diagnosis of these and other cancers, as well as the expanding biological role of complex rearrangements in cancer evolution. ■

REFERENCES

1. F. Mertens *et al.*, *Nat. Rev. Cancer* **15**, 371 (2015).
2. N. D. Anderson *et al.*, *Science* **361**, eaam8419 (2018).
3. T. G. P. Grünewald *et al.*, *Nat. Rev. Dis. Primers* **4**, 5 (2018).
4. S. C. Baca *et al.*, *Cell* **153**, 666 (2013).
5. M. N. Nikiforova *et al.*, *Science* **290**, 138 (2000).
6. R.-S. Mani *et al.*, *Science* **326**, 1230 (2009).
7. A. G. Henssen *et al.*, *Sci. Transl. Med.* **9**, eaam9078 (2017).
8. A. S. Brohl *et al.*, *Genet. Med.* **19**, 955 (2017).
9. S. Turajlic *et al.*, *Lancet Oncol.* **18**, 1009 (2017).
10. S. R. Viswanathan *et al.*, *Cell* **174**, 433 (2018).
11. H.-Y. Huang *et al.*, *J. Clin. Oncol.* **23**, 548 (2005).
12. P. Roberts *et al.*, *Genes Chromosomes Cancer* **47**, 207 (2008).
13. F. Tirode *et al.*, *Cancer Discov.* **4**, 1342 (2014).

10.1126/science.aau8231

ORGANOMETALLICS

18 electrons and counting

The bonding rule for transition metal complexes now extends to alkaline earth octacarbonyls

By **P. B. Armentrout**

The "octet rule" is based on the stability afforded to species with closed-shell electron configurations like the noble gases. Simple "second-row" compounds like methane, ammonia, and water have eight electrons surrounding the central atom, as do their third-row analogs (silane, phosphine, and hydrogen sulfide). For atoms in the fourth row of the periodic table and beyond (principal quantum number designated by n), the single ns, three np, and five (n – 1)d orbitals must be filled with two electrons per orbital, resulting in an analogous "18-electron" rule for a closed shell. This simple electron counting guides inorganic chemists working with transition metals in predicting stable compounds, just as the octet rule guides organic chemists working with carbon. For example, the stable transition metal carbonyls $Cr(CO)_6$, $Fe(CO)_5$, and $Ni(CO)_4$, as well as heavier homologs, can be formed, indicating that CO is a two-electron donor. Metals with an odd number of valence electrons must double up with a metal-metal bond, so $Mn_2(CO)_{10}$ and $Co_2(CO)_8$ form in order to satisfy the 18-electron rule. On page 912 of this issue, Wu *et al.* (*1*) demonstrate that the 18-electron guiding principle is not only limited to transition metals but can also be extended to nearby elements, the alkaline earths.

"Intriguingly...the $M(CO)_8$ species do not actually have a singlet spin state... [r]ather...a triplet spin state..."

Alkaline earth metals (Ms; Be, Mg, Ca, Sr, and Ba) have a valence electron configuration of ns^2 and generally form two covalent bonds

Department of Chemistry, University of Utah, Salt Lake City, UT 84112, USA. Email: armentrout@chem.utah.edu

with other elements, for example, MgO and $BaBr_2$. Grignard reagents RMgX, where R is generally an alkyl group and X is a halide, are common reagents in organic and organometallic synthesis. In solution, alkaline earths are readily oxidized and normally found in their +2 oxidation state. Surprisingly, Wu *et al.* show that the octacarbonyls of the heavier alkaline earths—Ca, Sr, and Ba—can be formed in the neutral state as a consequence of surrounding the two-electron M atoms with another 16 electrons.

Metal carbonyls have been known since the late 19th century (*2*) and are used as starting materials for transition metal compounds in organic synthesis and as catalysts in hydroformylation. At a fundamental level, metal carbonyls can be used to assess the relative strength of binding interactions in coordination complexes. The highest occupied molecular orbital (HOMO) in CO corresponds to a lone pair of electrons on the C atom (see the figure, left), so CO binds to metals at the C atom, except in extraordinary circumstances. Furthermore, this orbital can donate two electrons to the metal center. In the CO triple bond (one σ and two π bonds), four of the six electrons come from the more electronegative O atom. The dipole moment of CO is relatively small (0.122 debye), and counterintuitively, the C atom has the negative charge. However, as the C≡O bond stretches, the electrons follow the more electronegative O atom, and the dipole moment of CO increases dramatically. Thus, CO molecules "light up" in infrared (IR) spectroscopy because they have a large change in their dipole moment upon stretching.

This property of the CO ligand can then be used to assess how it binds to metals. The electron-counting procedure noted above, in which the CO ligand donates its pair of electrons into an empty orbital on the metal, provides one binding motif, called σ donation (after the symmetry of the bond being formed; see the figure, right). Because the HOMO is largely a nonbonding orbital on CO, σ-bond donation does not greatly affect the C≡O stretching frequency (*3*).

However, CO ligands are actually more promiscuous in their binding. The lowest unoccupied molecular orbitals (LUMOs) on CO are the two antibonding π* orbitals (see the figure, left). Transition metals can utilize electrons in d orbitals with the same π symmetry to augment the binding to the CO ligand, in essence, forming a second or even a third bond between the metal and the carbon (see the figure, right). This type of electron-exchange bonding is often referred to as the Dewar-Chatt-Duncanson model (*4*, *5*), although these authors actually addressed similar σ and π interactions between metals and olefins. The strength of these "backbonding" interactions can be assessed by measuring the C≡O stretching frequency. As electrons are donated into the antibonding LUMOs, the C≡O bond becomes weaker and its vibrational frequency is lowered. For example, in $Ni(CO)_4$, the stretch shifts down to ~2060 cm^{-1} from the free CO at 2143 cm^{-1}. The shift increases for isoelectronic anionic analogs that backbond more effectively, as in $Co(CO)_4^-$ at ~1890 cm^{-1} and $Fe(CO)_4^{2-}$ at ~1790 cm^{-1} (*6*).

Wu *et al.* generated complexes of Ca, Sr, and Ba with CO in a cold (4 K) neon matrix that allowed weakly bound species to form (*7*). They interrogated the matrix using IR irradiation, finding single intense absorptions at 1987, 1995, and 2014 cm^{-1} for saturated carbonyl complexes of Ca, Sr, and Ba, respectively, as well as other absorptions for smaller complexes. Because only a single C≡O stretching frequency was observed, the absorbing species must have high symmetry, assigned as cubic O_h for $M(CO)_8$.

These frequencies indicate that substantial backbonding occurs that gradually decreases with increasing metal size. Backbonding is critical in the formation of these complexes. The approach of CO to one of the Ms should entail a repulsive interaction between the occupied ns^2 orbital on the M and the HOMO of CO. To avoid this difficulty, the Ms must empty seven valence orbitals [one *n*s, three *n*p, and three of the (*n* – 1)d orbitals], which permits strong σ donor bonds with the CO ligands. The eighth "bonding" orbital needed for eight ligands is purely ligand based and has a_{2u} symmetry (a type for which there is no atomic orbital on the M). The two M valence electrons then occupy the remaining two (*n* – 1)d orbitals (having e_g symmetry), which augment the bonding by backbonding interactions, as demonstrated by the C≡O stretching frequencies measured.

Intriguingly, because these two orbitals are isoenergetic, the $M(CO)_8$ species do not actually have a singlet spin state in which all the electrons are paired (which one ordinarily associates with stable 18-electron complexes). Rather, each of the two (*n* – 1)d orbitals contains a single electron, yielding a triplet spin state, as confirmed by quantum chemical calculations. Wu *et al.* also examined the cationic analogs of $M(CO)_8$ complexes in the gas phase. Evidence for a saturated $M(CO)_8^+$ complex (now a 17-electron species) was obtained by observation of a single C≡O stretch, whereas $M(CO)_9^+$ exhibited a band characteristic of a CO ligand in a weakly bound second ligand shell.

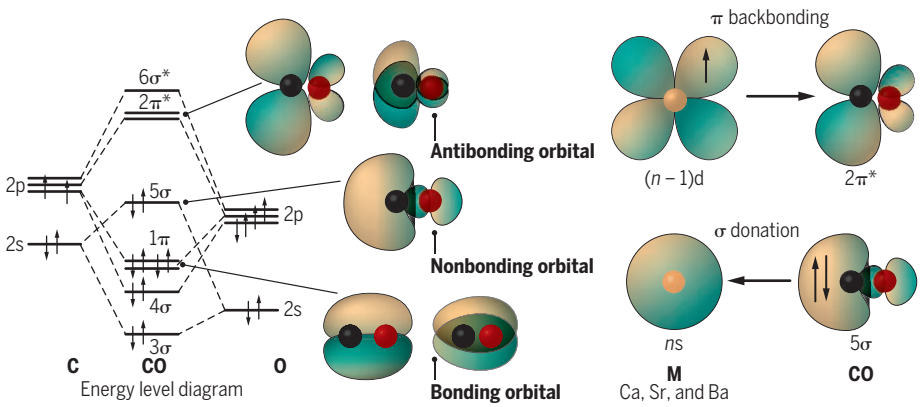
The study of Wu *et al.* challenges previous notions of limitations in the 18-electron rule and provides complexes exhibiting very interesting bonding motifs in the process. Whether the distinctive properties of such compounds can be exploited remains to be seen, but the study foreshadows additional complexes that might be generated and test the limits of the 18-electron rule. ■

Building an unexpected bond

Alkaline earth carbonyls reported by Wu *et al.* can be understood from simple bonding concepts.

Bonding in carbon monoxide
The molecular orbital diagram for the formation of carbon monoxide from carbon and oxygen atoms is shown. Several of the molecular orbitals are shown to the right.

Creating alkaline earth carbonyls
The bonding interactions between an alkaline earth metal (M) and carbon monoxide are shown. Vertical arrows indicate electrons in both parts.



REFERENCES

1. X. Wu *et al.*, *Science* **361**, 912 (2018).
2. L. Mond *et al.*, *J. Chem. Soc. Trans.* **57**, 749 (1890).
3. G. Bistoni *et al.*, *Chem. Sci.* **7**, 1174 (2016).
4. M. J. S. Dewar, *Bull. Soc. Chim. Fr.* **18**, C71 (1951).
5. J. Chatt, L. A. Duncanson, *J. Chem. Soc.* **1953**, 2939 (1953).
6. F. A. Cotton, G. Wilkinson, *Advanced Inorganic Chemistry* (Wiley, 1988).
7. L. Andrews, in *Encyclopedia of Spectroscopy and Spectrometry*, J. C. Lindon, Ed. (Academic Press, ed. 2, 1999), pp. 1479–1483.

10.1126/science.aau6622

POLICY FORUM

ENERGY AND CLIMATE

Global carbon intensity of crude oil production

New data enable targeted policy to lessen GHG emissions

By **Mohammad S. Masnadi, Hassan M. El-Houjeiri, Dominik Schunack, Yunpo Li, Jacob G. Englander, Alhassan Badahdah, Jean-Christophe Monfort, James E. Anderson, Timothy J. Wallington, Joule A. Bergerson, Deborah Gordon, Jonathan Koomey, Steven Przesmitzki, Inês L. Azevedo, Xiaotao T. Bi, James E. Duffy, Garvin A. Heath, Gregory A. Keoleian, Christophe McGlade, D. Nathan Meehan, Sonia Yeh, Fengqi You, Michael Wang, Adam R. Brandt**

Producing, transporting, and refining crude oil into fuels such as gasoline and diesel accounts for ~15 to 40% of the "well-to-wheels" life-cycle greenhouse gas (GHG) emissions of transport fuels (*1*). Reducing emissions from petroleum production is of particular importance, as current transport fleets are almost entirely dependent on liquid petroleum products, and many uses of petroleum have limited prospects for near-term substitution (e.g., air travel). Better understanding of crude oil GHG emissions can help to quantify the benefits of alternative fuels and identify the most cost-effective opportunities for oil-sector emissions reductions (*2*). Yet, while regulations are beginning to address petroleum sector GHG emissions (*3*–*5*), and private investors are beginning to consider climate-related risk in oil investments (*6*), such efforts have generally struggled with methodological and data challenges. First, no single method exists for measuring the carbon intensity (CI) of oils. Second, there is a lack of comprehensive geographically rich datasets that would allow evaluation and monitoring of life-cycle emissions from oils. We have previously worked to address the first challenge by developing open-source oil-sector CI modeling tools [OPGEE (*7*, *8*), supplementary materials (SM) 1.1]. Here, we address the second challenge by using these tools to model well-to-refinery CI of all major active oil fields globally—and to identify major drivers of these emissions.

We estimate emissions in 2015 from 8966 on-stream oil fields in 90 countries (SM 1.4.4). These oil fields represent ~98% of 2015 global crude oil and condensate production. This analysis includes all major resource classes (e.g., onshore/offshore and conventional/unconventional) and accounts for GHG emissions from exploration, drilling and development, production and extraction, surface processing, and transport to the refinery inlet (collectively called "upstream" hereafter). These results are based on data from nearly 800 references, including government sources, scientific literature, and public technical reports (SM 1.4.1, 1.4.4, and table S17). Proprietary databases are used to supplement these data when information is unavailable in the public domain (generally for small oil fields). The latest Intergovernmental Panel on Climate Change (IPCC) 100-year global warming potential (AR5/GWP100) factors are used in this work (SM 1.2.1).

See supplementary materials for author affiliations.
Email: abrandt@stanford.edu; masnadi@stanford.edu

COUNTRY-LEVEL UPSTREAM CARBON INTENSITY

The first figure presents the first upstream country-level volume-weighted-average CI estimates and their corresponding error bars (see fig. S22 for the global upstream CI map). Error bars are computed by using probabilistic uncertainty analysis solely associated with missing input data (SM 1.7 and 2.4). The CI estimates of some countries with poor data quality (e.g., Russia) are more uncertain (SM 1.4.6 and 2.3).

The global volume-weighted-average upstream CI estimate—shown by the vertical red line in the first figure—is 10.3 g CO_2equivalents (CO_2eq.)/megajoule (MJ) crude oil (+6.7 and –1.7), with country-level intensities ranging from 3.3 (Denmark) to 20.3 (Algeria) g CO_2eq./MJ. Carbon dioxide and methane contribute on average 65% and 34% of total CO_2eq. emissions, respectively (SM 2.2). The total petroleum well-to-refinery GHG emissions in 2015 are estimated to be ~1.7 Gt CO_2eq., ~5% of total 2015 global fuel combustion GHG emissions. This estimate of total emissions is ~42% higher than an industry-wide scaling of an estimate for 2015 from the International Association of Oil and Gas Producers (based on datasets comprising 28% of global production with uneven geographical coverage). See SM 3 for exploration of the differences between our analyses.

Emissions shown in the first figure can vary substantially over time (*9*), but time-series data are generally missing on a global basis and so are not explored here. In general, oil production declines with oil field depletion but is also accompanied by a substantial increase in per-MJ GHG emissions due to use of enhanced recovery practices. Other factors (e.g., oil price, geopolitics) could also affect oil production and thus the temporal CI (*9*).

Gas flaring (burning) practices have a considerable influence on the CI. If not economically salable, this gas is either flared, reinjected, or vented (directly emitting methane). The estimated share of flaring emissions in the global volume-weighted-average upstream CI is 22% (i.e., 2.3 g CO_2eq./MJ). Flaring data are not widely reported by governments or companies, so for most regions, our analysis relies on satellite-estimated volumes computed using nighttime radiometry (SM 1.2.4 and 1.4.3.18). Some important conventional crude oil producers with above-average global CI, such as Algeria, Iraq, Nigeria, Iran, and the United States, are also among the top 10 countries in flaring observed via satellite. The contributions of routine flaring to the total volume-weighted-average CI of these countries are estimated herein to be ~41, 40, 36, 21, and 18%, respectively. Variability between flaring data sources results in greater uncertainty for countries with high contribution of flaring to their CI. Figure S27 shows that gas venting instead of flaring increases the estimated GHG emissions substantially (SM 1.2.4 and 2.6). However, currently there is no reliable remote-sensing technology for measuring gas venting.

As the major global producers of unconventional heavy oils, Venezuela and Canada have high country-level CI. This is due to energy- and CO_2-intensive heavy oil extraction and upgrading. Enhanced oil recovery by steam flooding contributes to high CI in other locations, such as Indonesia, Oman, and California (USA).

Although some giant North Sea offshore fields have shown rapidly increasing per-bbl (barrel) emissions due to depletion (*9*), they have low upstream GHG intensities when compared to many other global oil fields. This is in part due to stringent regulations on gas processing and handling systems and renewable electric-power-from-shore initiatives. Saudi Arabia is the largest global oil producer but has a small number of extremely large and productive reservoirs. The country has low per-barrel gas flaring rates and low wa-

ter production—resulting in less mass lifted per unit of oil produced and less energy used for fluid separation, handling, treatment, and reinjection—and thus contributing to low CI.

FIELD-LEVEL UPSTREAM CARBON INTENSITY

The second figure shows a global field-level CI curve for our 8966 fields (sorted cumulatively) and illustrates CI heterogeneity of global crudes (SM fig. S19; Results Data Excel file). Fields in the highest 5th percentile emit more than twice as much as the median field. Upstream mitigation measures should focus on fields in the upper end of the CI curve.

Although crude density (requiring thermal extraction methods) and flaring are key determinants of a high CI (SM 1.5), the second figure shows that flaring is the more prevalent driver: For the highest CI quartile (i.e., >1.2 g CO_2eq./MJ) in this figure, 51% of crude volume comes from high flare fields (yellow, red), while 18% comes from heavy oil fields (yellow, blue). Only 4 and 9% of crude volumes from the rest of the sample (i.e., ≤11.2 g CO_2eq./MJ) come from high flare and heavy oil fields, respectively.

The cumulative CI curve uncertainty due to missing input data is computed via a Monte Carlo simulation and presented in fig. S25 (SM 1.7 and 2.4).

POLICY IMPLICATIONS

Although oil alternatives like electric vehicles are rapidly growing, society is likely to use large volumes of oil in the coming decades (*10*); thus, mitigation of crude oil CI is key. Our tools and dataset allow for improved analysis of the benefits of emissions mitigation policies. We highlight three broad strategies to reduce GHG impacts: (i) resource management, (ii) resource prioritization, and (iii) innovative technologies.

Performance-oriented fuel quality standard programs based on life-cycle analysis models have been implemented successfully and have created new regional market drivers (e.g. in California, British Columbia, the European Union). Relying on market forces and credit/debit mechanisms, these fuel-agnostic policies do not dictate specific technologies to reduce the emissions but rather encourage innovation to comply with the quality mandates. To achieve greater impacts, such regional fuel standard policies are emerging nationally (e.g., Canada's Clean Fuel Standard) and, subsequently, worldwide. These regulations should recognize the climate impact heterogeneity of different crude oils (see the second figure) to reward improved production practices with clear per-barrel incentives for the lowest CI producers (*10*).

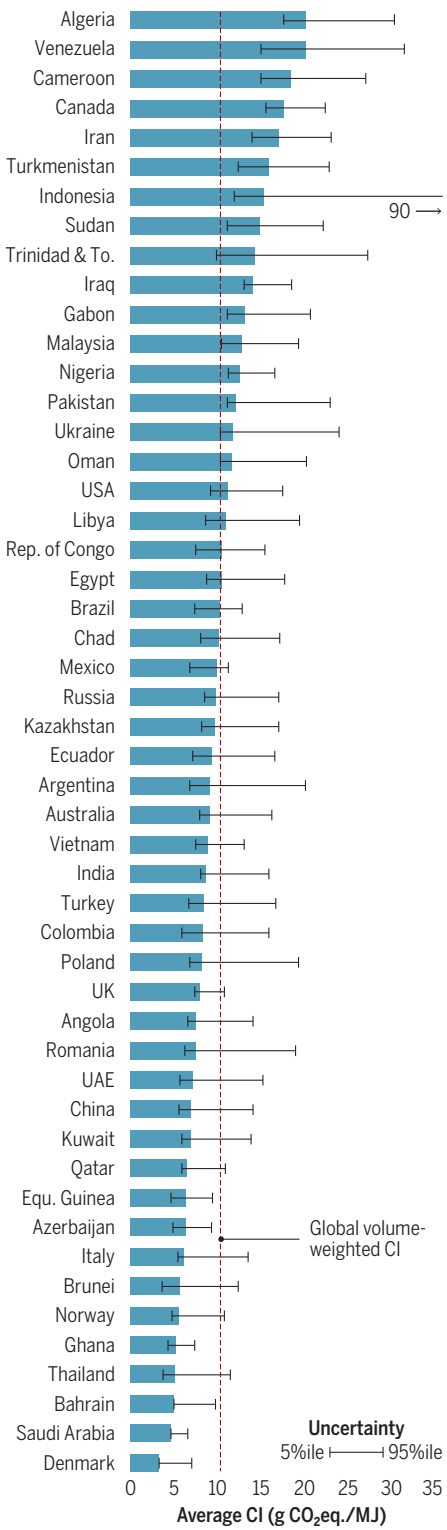
The current lack of transparency about global oil operations makes this type of analysis particularly challenging. Labor-intensive data gathering (as undertaken here) still results in large uncertainty in emissions estimates (SM 2.3 and 2.4). Thus, it is important to adopt policies to make data from oil and gas operations publicly available. If done correctly, these data can be released without affecting competitiveness of enterprises. Countries including Norway, Canada, the United Kingdom, Denmark, and Nigeria have led in this respect. As countries pledge their commitments to reduce country-level GHG emissions and transparent reporting under the Paris Agreement, it is essential for energy-intensive industries (such as the oil and gas sector) to regularly report their annual carbon footprints. New industry efforts such as the Oil and Gas Climate Initiative are beginning to tackle this challenge.

CI curves for four hypothetical GHG mitigation case studies are shown in fig. S26 (SM 1.2.2 and 2.5). Two "no routine flaring" case studies restrict the flare-to-oil ratio (FOR) to be no higher than the global 5th and 25th percentiles. A fugitive emissions reduction scenario sets fugitive and venting emissions to be 0.2 g CO_2eq./MJ, approximately the volume-weighted average from Norwegian oil fields in 2015 (SM 1.2.2). Cases with no routine flaring (moderate and extreme) have global volume-weighted-average CI reduced from 10.3 (current world) to 8.7 and 8.3 g CO_2eq./MJ. Achieving the fugitive and venting reduction scenario results in 7.9 g CO_2eq./MJ. These case studies mitigate 15% [262 megatons (Mt) CO_2eq.], 19% (332 Mt CO_2eq.), and 23% (397 Mt CO_2eq.) of the current annual global upstream estimate, respectively. A fourth case study, including both stringent flaring reduction and minimal fugitive and venting emissions, reduces the global average to 5.8 g CO_2eq./MJ and results in ~43% (~743 Mt CO_2eq.) annual CI reduction.

A simple calculation suggests that upstream emissions from oil extraction can materially affect cumulative emissions caps. Assume a reduction in the current global volume-weighted-average CI to the current 25th percentile (reducing emissions by ~3 g CO_2eq./MJ). Such reductions would be possible using the mitigation case studies from fig. S26. Given that a typical barrel of crude oil yields ~6000 MJ, this would result in ~18 kg CO_2eq./bbl emissions reduction. Also note that IPCC scenarios—even with aggressive adoption of alternative fuels used for transport—still result in projected cumulative oil consumption of >1 trillion barrels in the 21st century. Thus, at least 18 metric gigatons (Gt) CO_2eq. (~12 Gt as CO_2 and ~6 Gt as CH_4) could be saved over the century by mitigating oil-sector emissions through wise resource choices and improved gas management practices. Considering additional mitigation op-

National volume-weighted-average crude oil upstream GHG intensities (2015)

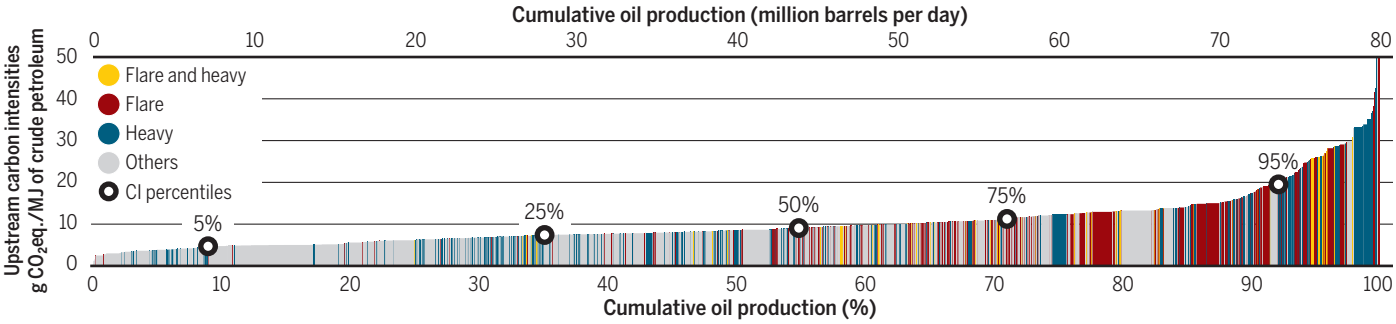
The global volume-weighted carbon CI estimate is shown (red line, ~10.3 g CO_2eq./MJ). Error bars reflect 5th to 95th percentiles of Monte Carlo simulation to explore the uncertainty associated with missing input data (see SM 1.7 and 2.4).



GRAPHIC: ADAPTED BY J. YOU/*SCIENCE*

Global field-level upstream carbon intensity supply curve (2015)

Contribution of high flaring ("Flare" with FOR >75th percentile of all fields) and oil density ("Heavy" with API gravity ≤22°) are shown. Bar width reflects the oil production of a particular field in 2015. Global GHG intensity percentiles (5%, 25%, 50%, 75%, 95%) are 4.7, 7.3, 9.1, 11.2, and 19.5 g CO_2eq./MJ crude oil, respectively.



portunities across the crude oil supply chain (e.g., improved refining), 18 Gt is likely an underestimate; other studies have estimated up to 50 Gt CO_2eq. reduction potential (*10*). For a >66% chance to keep global average temperature increases below 2°C, a total of ~800 Gt CO_2 can be emitted from 2017 forward (*11*). The petroleum sector reduction potentials outlined above are material on this scale.

Extraction and processing of heavy oils and oil sands with current technologies is very energy- and carbon-intensive, and the ability to reduce the intensities is challenging. Although market forces have recently led to investment shifts based on economics alone (*12*), other mechanisms exist to reduce emissions. Solar-powered steam generators developed for heavy oil fields in Oman and California can provide substantial mitigation benefit. More broadly, use of solar energy could result in sectorwide emissions reductions on the order of 5 kg CO_2eq./bbl (~1.7 g CO_2eq./MJ) (*13*). For some key regions with high seasonality and poor economics of solar technology (like Canada), using energy inputs with low carbon intensity (e.g., hydrogen sourced from wind and biomass), capturing CO_2 from oil sands extraction and upgrading facilities, and investing in new low-carbon technologies (e.g., nanoparticle-assisted in-situ recovery, or CO_2-free production of H_2 from CH_4 via catalytic molten metals) would be beneficial. In addition, low-value but high-carbon products such as petroleum coke from upgraded oil sands could be sequestrated in lieu of combustion (*10*). Countries with diverse resources could reduce their national CI by prioritizing less carbon-intensive assets (e.g., tight oil), accompanied by stringent flaring and venting management.

Flaring rates can also be reduced. The Global Gas Flaring Reduction Partnership (GGFR) reported a nearly continuous increase in global flared gas from 2010 to 2016. Flaring is a management and infrastructure problem and is not an unavoidable outcome of crude oil production. Plans for new oil field development should incorporate conservation methods (i.e., capture, utilization, and/or reinjection) to eliminate routine flaring. Canadian regulations point to a method for enforcement: For offshore fields where flaring is excessive, production rate restrictions are imposed until flaring reductions are made (*14*). Initiatives like the World Bank GGFR Zero Routine Flaring by 2030 are a start, though these could be strengthened with international advisory, financial, and technical aid to help countries implement flaring reduction policies. Moreover, continuous monitoring and verification are essential not only for flare management but also for eliminating venting and fugitive methane emissions in the oil and gas sector. Modern surveillance using remote-sensing technologies (e.g., flare- and methane-sensing satellites) could be supported and expanded (*10*).

Methane fugitive emissions and venting from oil and gas facilities are poorly detected, measured, and monitored, and thus, can increase the uncertainty associated with CI estimates. Recently, the International Energy Agency (IEA) estimated 76 Mt methane emissions from global oil and gas operations in 2015, with ~34 Mt due to oil production (*15*). This prorates to ~4.6 g CO_2eq./MJ crude oil, higher than this study's estimate of methane contribution (~2.6 g CO_2eq./MJ averaged from all global fields, from all fugitive emissions and venting). In many cases, reducing methane emissions can result in additional revenues from the captured methane. IEA estimates that around 40 to 50% of current methane emissions could be avoided at no net cost. The cost of mitigation is generally lowest in developing countries in Asia, Africa, and the Middle East, but in all regions, reducing methane emissions remains a cost-efficient way of reducing GHG emissions (*15*).

Important questions remain with regard to the interactions of economics and emissions. The CI curve in the second figure reflects differences in CI, but crude oil production choices are obviously influenced by the interaction of local production costs and the global price of oil. A market structure without carbon prices neglects differences in CI shown in the second figure. Future work needs to examine the interaction of supply economics and CI for different resource classes.

Data-driven CI estimates such as this work can encourage prioritizing low-CI crude oil sourcing, point to methods to manage crude oil CI, and enable governments and investors to avoid "locking in" development of high-CI oil resources. However, future progress in this direction will rely fundamentally on improved reporting and increased transparency about oil-sector emissions. ■

REFERENCES AND NOTES

1. IHS CERA, *Oil Sands, Greenhouse Gases, and US Oil Supply* (IHS, 2012).
2. M. S. Masnadi *et al.*, *Nat. Energy* **3**, 220 (2018).
3. California Environmental Protection Agency/Air Resources Board, Low Carbon Fuel Standard (LCFS).
4. European Commission, Fuel Quality Directive (FQD).
5. Environment and Climate Change Canada, Clean Fuel Standard (2017).
6. R. Baron, D. Fischer, *Divestment and Stranded Assets in the Low-carbon Transition* [Organization for Economic Cooperation and Development (OECD), 2015].
7. H. M. El-Houjeiri, A. R. Brandt, J. E. Duffy, *Environ. Sci.* **47**, 5998 (2013).
8. H. M. El-Houjeiri *et al.*, "Oil Production Greenhouse Gas Emissions Estimator OPGEE v2.0a, User guide & technical documentation" (2017).
9. M. S. Masnadi, A. R. Brandt, *Nat. Clim. Change* **7**, 551 (2017).
10. A. R. Brandt, M. S. Masnadi, J. G. Englander, J. Koomey, D. Gordon, *Environ. Res. Lett.* **13**, 044027 (2018).
11. C. Le Quéré *et al.*, Global carbon budget 2016. *Earth Syst. Sci. Data.* **8**, 605 (2016).
12. K. Gilblom, "Shell to Exit Canadian Natural Resources for $3.3 Billion," *Bloomberg*, 7 May 2018.
13. J. Wang, J. O'Donnell, A. R. Brandt, *Energy* **118**, 884 (2017).
14. C-NLOPB, Newfoundland and Labrador Offshore Area Gas Flaring Reduction Implementation Plan. Canada-Newfoundland Labrador Offshore Petroleum Board (2017).
15. T. Gould, C. McGlade, The environmental case for natural gas (International Energy Agency, 2017).

ACKNOWLEDGMENTS

The Natural Sciences and Engineering Research Council of Canada (NSERC) provided financial support to M.S.M. Aramco Services Company (MI, USA) and Ford Motor Company provided funding for D.S. and A.R.B. A.R.B. also received funding from the Carnegie Endowment for prior research and data collection. D.G. and J.K. were funded through the Carnegie Endowment for International Peace, with primary funding from Carnegie's endowment and additional funding from the Hewlett Foundation, ClimateWorks Foundation, and Alfred P. Sloan Foundation. The authors are grateful to G. Cooney from National Energy Technology Laboratory (PA, USA) for constructive comments.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/851/suppl/DC1

10.1126/science.aar6859

GRAPHIC: ADAPTED BY J. YOU/*SCIENCE*

Mary Robinson meets with local farmers in Tigray, Ethiopia, during a period of severe drought in 2016.

PHOTO: PA IMAGES/ALAMY STOCK PHOTO

BOOKS *et al.*

CLIMATE POLICY

A president amplifies unlikely activists

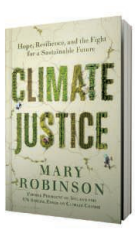
In dispatches from the front lines of global warming, Mary Robinson pushes for humane climate policies

By **Louise Fabiani**

One of the cruelest ironies of today's world is that many of the nations least responsible for causing climate change find themselves most affected. Widely fluctuating weather patterns and rising seas caused by melting glaciers are already displacing millions of people, endangering food security, increasing epidemics, and destabilizing political systems in these areas, most of which tend to hug the tropics. But even the Arctic is not safe: Temperatures have been rising in circumpolar regions more rapidly than anywhere, reducing sea ice and turning the polar bear into a poster child for the perils of global warming. Meanwhile, climate summits continue, scientists publish studies, and our collective addiction to fossil fuels goes on.

As the former president of Ireland, Mary Robinson knows about the doublespeak of global diplomacy all too well. Her new book, *Climate Justice*, is less about grim statistics and stalemates than how we might really address the problems facing our planet.

In her capacity as United Nations special envoy on climate change, a position she held from 2014 to 2015, Robinson attended countless meetings with world leaders. It is another kind of person she encountered, however—including, notably, many women who had never dreamed of becoming activists—who enlivens this book. Through the testimonies of ordinary citizens, we learn about the horrendous toll drought is taking in Chad, the way warmer winters threaten the continued existence of Saami reindeer culture in northern Europe, and the high cultural value the Vietnamese place on local forest products.



Climate Justice
Mary Robinson
Bloomsbury, 2018.
176 pp.

Island nations such as Kiribati may have to move their entire populations before rising sea levels cover the land masses they call home. "[B]ecause of its position on the international date line, Kiribati was the first country in the world to welcome in the new millennium," Robinson writes. "Now, in a tragic twist of fate, it may become the first one lost to the effects of climate change before the dawn of the next century." Yet in the 2009 Copenhagen climate talks, delegates ignored Kiribati's pleas to substantially reduce CO_2 levels. Mitigation, including a proposal to design floating islands, seems to appeal more to decision-makers than does prevention.

In this, as in nearly every other case of crisis management, the cost of "fixing" the problems that will arise as a result of global warming dwarfs that of almost every imaginable measure we could take to stop it from occurring in the first place. But for whatever reason—Robinson does not speculate as to why we prefer to pay for a pound of cure over an ounce of prevention—our climate strategies are stuck in a wait-and-see mindset.

Approaching climate change as an isolated issue would be wrong, no matter how much foresight we use. Robinson reminds the reader that we must also tackle the "poverty, inequality, and exclusion" that lock much of the world into patterns of destructive behavior. This, along with the recalcitrance of businesses and politicians to break out of age-old modes of thinking, is one reason Robinson looks for inspiration in more traditional forms of knowledge and the reason she advocates for increasing the participation of women and elders in indigenous communities in climate discussions. These individuals tend to be the ones with hands-on experience of natural systems, she argues.

Achieving "climate justice" also entails acknowledging the inevitable human cost of shifting to more ecofriendly systems. The "Just Transition" movement, for example, seeks to compensate workers who face reduced or lost employment as renewable energy sources replace coal, oil, and natural gas.

To Robinson, the 2015 Paris Agreement felt like a step in the right direction. However, in 2017 her worst fears were realized: President Donald Trump pulled the United States—the world's second-biggest polluter—out of the contract. "It is unconscionable that the United States has simply walked away from its responsibility to people both at home and abroad, in the interest of short-term fossil fuel profits, and abandoned an agreement that was negotiated by more than 190 world leaders, over decades," she writes.

In the end, however, Robinson is optimistic about our ability to meet the challenges that lie ahead, if pragmatic about what will compel us to do so. "Enlightened self-interest," she assures readers—in the form of an Alaskan compelled to call her congressperson over fears her house will fall into the sea, or an insurance company pressing for new tailpipe-emission regulations because of skyrocketing claims—can turn the drive for personal survival into help for many others. ■

The reviewer is a freelance science writer and culture critic based in Montreal, Québec, Canada. Email: m.l.fabiani47@gmail.com

10.1126/science.aau3320

QUANTUM PHYSICS

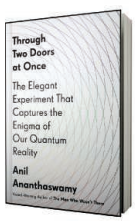
Understanding the double slit

Glimpses of quantum truth appear in diverse interpretations of the famous physics experiment

By **Mélanie Frappier**

In his famous *Lectures on Physics*, Richard Feynman argued that nothing more is needed to get a solid grasp of the behavior of quantum objects than the simple double-slit experiment, in which electrons or photons are fired toward two thin openings cut in a screen. To Feynman, the double-slit experiment encapsulated quantum physics's one and only mystery. Its results could be described but, he cautioned, could not be explained.

Despite Feynman's warning, the past 60 years have seen an explosion of interpretations appealing to devices as diverse as pilot waves and parallel universes in the hope of elucidating quantum behavior. So far, no interpretation has proven fully convincing, leading many physicists to conclude that the theory's mathematical formalism should be left uninterpreted and to demand a return to the "shut up and calculate" attitude that was prevalent among the physicists of Feynman's generation.



Through Two Doors at Once
Anil Ananthaswamy
Dutton, 2018. 301 pp.

Veteran science journalist Anil Ananthaswamy rejects this fatalistic perspective. On the contrary, he argues, a deeper understanding of the quantum world can only be achieved by embracing the diversity of interpretations available to us, a claim he persuasively defends in *Through Two Doors at Once*. In the book, he takes Feynman to task by offering a spirited introduction to the various quantum interpretations, examining their respective explanations of the supposedly inscrutable experiment.

Ananthaswamy starts his investigation with a description of the double-slit experiment that is so natural and elegant that one may forget that it took physicists close to 30 years to develop the mathematical framework needed to describe it adequately. He rapidly recaps this struggle, from Planck's original suggestion that energy might sometimes be quantized to the Bohr-Einstein debates of the 1930s. Here, *Through Two Doors at Once* offers little more than the usual narrative, apart from an amusing detour through G. I. Taylor's nonchalant approach to an early one-slit version of the experiment.

It is surprising to learn that the double-slit experiment played a minor role in the early development of quantum theory—that is, until Ananthaswamy explains that it was not performed in the laboratory until the 1960s. Until then, it was only a thought experiment.

Feynman might have obtained ideal data by imagining himself firing at the two fictive slits with a futuristic tungsten electron gun, but in real life, physicists wrestled for decades—relying on everything from spider silk to ingenious beamsplitters—to bring the experiment to life. But this is where the many iterations of the double-slit experiment really take center stage in the development of our understanding of quantum reality. Technological advances, we learn, prompted physicists to conduct ever-more-sophisticated versions of the experiment, which in turn fueled a greater variety of interpretations. This increased the need for ever-more-sophisticated experiments.

Step by step, *Through Two Doors at Once* reveals how physicists transformed, retooled, and repurposed the original double-slit setup to throw light on the fundamental principles of quantum physics. Each iteration of the experiment is, for Ananthaswamy, an opportunity to introduce readers to fundamental concepts, such as entanglement; to analyze iconic experiments, such as Aspect's test of Bell's inequalities; or, more importantly, to examine the most prominent interpretations of quantum physics, from the Copenhagen interpretation of the 1920s to the more recent "many interacting worlds" hypothesis.

Ananthaswamy's introduction of increasingly complex versions of the slit experiment proves extremely effective. Halfway through the book, even neophytes will likely find predicting the outcome of the delayed-choice quantum eraser experiment barely harder than figuring out the motions of a gear train. This approach also brings to the forefront the strengths and weaknesses of various interpretations, offering a perfectly balanced overview of each.

But Ananthaswamy carefully guards himself from offering any guiding principle that might help us decide which explanation is the best one. There is, he explains, no such thing as the "right" interpretation in good science. This does not mean, however, that we have to be mere instrumentalists and reject interpretations as misguiding fantasies. We have another, better option: We can decide to embrace the diversity of interpretations at our disposal because despite their respective flaws, each likely holds the key to at least one essential aspect of quantum behavior.

Through Two Doors at Once offers beginners the tools they need to seriously engage with the philosophical questions that likely drew them to quantum mechanics. But readers will also receive a more important lesson, one that Feynman would have approved: In science, a deep understanding is not achieved by limiting ourselves to a single perspective but by simultaneously exploring competing conceptions of reality. ■

10.1126/science.aav0128

The reviewer is at the History of Science and Technology Program, University of King's College, Halifax, NS B3H 2A1, Canada. Email: melanie.frappier@ukings.ca

PHOTO: ISTOCK.COM/MIDDELVELD

Does the rising sun cause the rooster to crow? Questions of causality can finally be answered, write Pearl and Mackenzie.

PODCAST

The Book of Why
The New Science of Cause and Effect
Judea Pearl and *Dana Mackenzie*
Basic Books, 2018. 429 pp.

"Correlation does not imply causation," we have all been cautioned. How, then, are we to determine whether one thing has caused another? This week on the *Science* podcast, Judea Pearl and Dana Mackenzie discuss strategies for causal thinking and describe its implications for artificial intelligence. Listen to the interview at: sciencemag.org/podcasts.

10.1126/science.aau9731

LETTERS

This intermittent stream in Colorado would benefit from EPA protection for temporary waterways.

PHOTO: DANIEL ALLEN

Edited by **Jennifer Sills**

Protecting U.S. temporary waterways

Protecting the ecological health of rivers relies on maintaining intact flows from source areas to downstream navigable waters (*1*). Yet the U.S. Environmental Protection Agency (EPA) intends to rescind legal protection of tributary rivers, streams, and wetlands that do not have year-round flows (temporary waterways) and whose surface waters contribute flow to permanent navigable waters (*2*). This decision would severely damage the condition and uses of many U.S. waters, both temporary and navigable.

Temporary waterways provide many ecosystem services, including water provision and purification, that contribute substantially to securing water quantity and quality (*3*–*5*). Fifty-eight percent of all waterways that provide drinking water to the continental United States are temporary or headwater streams, which support more than one-third of the United States's population (*6*). Furthermore, temporary waterways harbor important biodiversity (*5*) and imperiled species (*7*) and underpin global carbon and nutrient cycles (*8*). Even when dry, they provide ecosystem services such as providing groundwater, attenuating toxicants, buffering floods, and providing habitat for unique biodiversity (*5, 9*).

A comprehensive scientific review (*10*) of all the services provided by temporary waterways led to the decision in 2015 to recodify the definition of "waters of the United States" to include temporary waters hydrologically connected to navigable waters. This provided protection to many temporary waterways under the U.S. Clean Water Act and was hailed as a wise, well-informed decision (*4*). However, the recodification has not yet been implemented because the legal process is incomplete, and now reversal of the decision is expected (*2*).

We urge the EPA to uphold its 2015 decision and to ratify the legal status and protection of temporary waterways. This would provide U.S. temporary waterways with a level of protection similar to that in other countries, such as Australia (*5*). Failure to do so sets a poor global precedent and, more importantly, risks costly (*11*) and potentially irreversible harm to the ecosystem services supported by temporary waterways in the United States, including the provision of secure potable water.

Jonathan C. Marshall,[1,2*] Vicenç Acuña,[3] Daniel C. Allen,[4] Núria Bonada,[5] Andrew J. Boulton,[6] Stephanie M. Carlson,[7] Clifford N. Dahm,[8] Thibault Datry,[9] Catherine Leigh,[10,11,2] Peter Negus,[1] John S. Richardson,[12] Sergi Sabater,[3,13] R. Jan Stevenson,[14] Alisha L. Steward,[1,2] Rachel Stubbington,[15] Klement Tockner,[16,17] Ross Vander Vorste[7]

[1]Water Planning Ecology, Queensland Department of Environment and Science, Brisbane, QLD 4001, Australia. [2]Australian Rivers Institute, Griffith University, Nathan, QLD 4111, Australia. [3]Catalan Institute for Water Research, Girona, Spain. [4]Department of Biology, University of Oklahoma, Norman, OK 73019, USA. [5]Grup de Recerca Freshwater Ecology, Hydrology and Management (FEHM-Lab), Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona (UB), Barcelona, Catalonia, Spain. [6]School of Environmental and Rural Science, University of New England, Armidale, NSW 2350, Australia. [7]Department of Environmental Science, Policy, and Management, University of California Berkeley, Berkeley, CA 94720, USA. [8]Department of Biology, University of New Mexico, Albuquerque, NM 87131, USA. [9]IRSTEA, UR RiverLy, Centre de Lyon-Villeurbanne, Villeurbanne, France. [10]Institute for Future Environments and School of Mathematical Sciences, Queensland University of Technology, Brisbane, QLD 4001, Australia. [11]ARC Centre of Excellence for Mathematical and Statistical Frontiers, Queensland University of Technology, Brisbane, QLD 4001, Australia. [12]Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC, Canada. [13]Institute of Aquatic Ecology, University of Girona, Girona, Spain. [14]Center for Water Science, Michigan State University, East Lansing, MI 48824, USA. [15]School of Science and Technology, Nottingham Trent University, Nottingham, UK. [16]Leibniz-Institute of Freshwater Ecology and Inland Fisheries, IGB, Berlin, Germany. [17]Austrian Science Fund, FWF, Vienna, Austria.
*Corresponding author. Email: jonathan.marshall@des.qld.gov.au

REFERENCES AND NOTES

1. D. C. Goodrich, W. G. Kepner, L. R. Levick, P. J. Wigington Jr., *J. Am. Water Resour. Assoc.* **54**, 400 (2018).
2. C. Whyte, "Six pollution policies gutted by Scott Pruitt – and what happens next," *New Scientist Daily News* (2018); www.newscientist.com/article/2173680-six-pollution-policies-gutted-by-scott-pruitt-and-what-happens-next.
3. T. L. Nadeau, M. C. Rains, *J. Am. Water Resour. Assoc.* **43**, 118 (2007).
4. V. Acuña *et al.*, *Science* **343**, 1080 (2014).
5. T. Datry, N. Bonada, A. Boulton, Eds., *Intermittent Rivers and Ephemeral Streams: Ecology and Management* (Academic Press, London, 2017).
6. U.S. EPA, "Geographic information systems analysis of the surface drinking water provided by intermittent, ephemeral, and headwater streams in the U.S." (EPA, 2017); www.epa.gov/cwa-404/geographic-information-systems-analysis-surface-drinking-water-provided-intermittent.
7. P. J. Wigington Jr. *et al.*, *Front. Ecol. Environ.* **4**, 513 (2006).
8. T. Datry *et al.*, *Nat. Geosci.* **11**, 497 (2018).
9. A. L. Steward, D. von Schiller, K. Tockner, J. C. Marshall, S. E. Bunn, *Front. Ecol. Environ.* **10**, 202 (2012).

10. U.S. EPA, "Connectivity of streams and wetlands to downstream waters: A review and synthesis of the scientific evidence" (Publication EPA/600/R-14/475F, EPA, 2015).
11. U.S. EPA, "Economic analysis of proposed revised definition of waters of the United States" (EPA, 2014); https://archive.epa.gov/epa/sites/production/files/2014-03/documents/wus_proposed_rule_economic_analysis.pdf.
12. To view or to add to a list of signatories in support of this letter see: https://docs.google.com/document/d/1AS5ew4S3CI3C-5TECAmNgi4cu8O48BY4WI1AqqIKpog/edit?usp=sharing.

10.1126/science.aav0839

Privacy and genetic genealogy data

Law enforcement use of genetic genealogy has recently led to identifications of missing persons and suspected criminals (*1*). These successes have prompted discussions about the genetic privacy of individuals whose DNA data are used in these investigations, particularly with regard to control over the usage of one's data and the sensitivity of the information that is obtained ["Genealogy databases and the future of criminal investigation," N. Ram *et al.*, Policy Forum, 8 June, p. 1078, and (*2*, *3*)]. As policies for law enforcement use of genetic genealogy are contemplated, several factors that mitigate the threat to privacy should be considered.

First, only data voluntarily uploaded and explicitly made public are searched. Direct-to-consumer DNA testing companies, from which most data arise, do not voluntarily participate in law enforcement investigations. Although it is possible for investigators to compel such companies to disclose user information, neither Ancestry.com nor 23andMe has turned over genetic information to law enforcement (*4*, *5*), and forensic genetic genealogy has not involved acquiring data in this way. Instead, investigations have relied on data that individuals have chosen to download from a testing company's database and upload to GEDmatch, a public genealogy database. GEDmatch is open to anyone, including law enforcement, who wants to check the database for indications of kinship with DNA data in their possession (including data from crime-scene or victim samples). Before allowing a new or existing user to access the site, GEDmatch prominently displays the full text of the Terms of Service and Privacy Policy, which advises individuals that GEDmatch can be used "by third parties such as law enforcement agencies to identify the perpetrator of a crime, or to identify remains" (*6*).

Because no one is legally required to contribute to a genetic genealogy database, and because the samples are not in the possession of government agencies, these searches are substantially different from familial searching of law enforcement databases (*7*). Jurisdictions that prevent or limit familial searching of those databases (*8*–*12*) should not automatically adopt identical restrictions on genetic genealogy investigations of publicly available databases.

Another factor that lessens privacy concerns is that raw genetic data are not disclosed to law enforcement. Raw data contain highly personal and health-related information. Search results display only the length and chromosomal location of shared DNA blocks, which are used to determine approximate kinship relationships between individuals. The raw data cannot be accessed; only the possible genetic kinship among individuals is shown. Customer relations create an incentive for testing companies and GEDmatch to maintain current policies of not releasing raw data without consent.

Finally, genetic genealogy is for lead generation, not conviction. Genetic genealogy leads are tested by direct DNA matching to samples from persons of interest using standard forensic identification loci; only matches obtained with these well-established methods will result in continued investigation.

Ellen M. Greytak,[1]* David H. Kaye,[2] Bruce Budowle,[3] CeCe Moore,[1] Steven L. Armentrout[1]
[1]Parabon NanoLabs, Inc., Reston, VA 20190, USA. [2]Penn State Law, Pennsylvania State University, University Park, PA 16802, USA. [3]Center for Human Identification, University of North Texas Health Sciences Center, Fort Worth, TX 76107, USA.
*Corresponding author. Email: ellen@parabon.com

REFERENCES

1. H. Murphy, "Genealogists turn to cousins' DNA and family trees to crack five more cold cases," *The New York Times* (2018); www.nytimes.com/2018/06/27/science/dna-family-trees-cold-cases.html.
2. B. E. Berkman, W. K. Miller, C. Grady, *Ann. Intern. Med.* 10.7326/M18-1348 (2018).
3. D. S. Court, *Forensic Sci. Int. Genet.* **36**, 203 (2018).
4. 23andMe Transparency Report (15 May 2018); www.23andme.com/transparency-report.
5. Ancestry.com 2017 Transparency Report (2017); www.ancestry.com/cs/transparency.
6. GEDmatch.Com, Terms of Service and Privacy Policy (www.gedmatch.com/tos.htm).
7. B. Budowle, R. P. Harmon, P. Melton, S. Hamstra, A. Ambers, "Familial DNA Searching: Current Approaches" (2015).
8. J. Kim, D. Mammo, M. B. Siegel, S. H. Katsanis, *Investig. Genet.* **2**, 22 (2011).
9. E. Murphy, *Mich. Law Rev.* **109**, 291 (2010).
10. S. M. Suter, *Harv. J. Law Technol.* **23**, 309 (2010).
11. N. Ram, *Stanford Law Rev.* **63**, 751 (2011).
12. D. H. Kaye, *Am. Crim. Law Rev.* **50**, 109 (2013).

COMPETING INTERESTS

E.M.G., C.M., and S.L.A. are employees of Parabon NanoLabs, Inc., which provides genetic genealogy services to law enforcement.

10.1126/science.aav0330

Chile's salmon escape demands action

After heavy winds and stormy conditions, about 650,000 salmon recently escaped from a net-pen aquaculture facility in southern Chile (*1*). This unintentional influx of salmon, a potentially invasive species in Chile (*2*), is just the latest of many escapes of farmed salmon and trout (*3*). The escape of these non-native species highlights the risk that such aquaculture facilities pose to native ecosystems (*3*, *4*).

Chile's US$4.6 billion aquaculture industry has positioned the country as a global provider of salmon and trout food products (*5*), and government institutions have contributed to this success (*6*). However, the Ministry of Environment, Superintendency of Environment, and National Fisheries Service, which are responsible for safeguarding biodiversity and fishery resources against potentially invasive species, have failed to coordinate with the aquaculture industry to provide the necessary short- and long-term monitoring of the escaped salmon (*7*).

To balance industry development with protection of native ecosystems and species, Chile must initiate new measures and invest critical institutional funding. The government should begin by passing legislation requiring industry to limit escapes and develop enhanced biosecurity technology to prevent them [e.g., (*8*)]. Chile should also allow catch and commercialization of escapees by local fishers and provide testing for antibiotics to determine whether the catch is suitable for human consumption. Finally, to mitigate the impacts of salmon aquaculture, Chile should implement ecosystem approaches recognizing that aquaculture affects other stakeholders and multiple ecological goods and services (*9*). Chile's public is largely resistant to salmon farming because of the environmental risks (*10*), whereas the government and industry focus instead on the opportunities aquaculture provides for jobs and development. All sides need to come together if there is to be a future for salmon aquaculture in the region.

Daniel Gomez-Uchida,[1,2]* Maritza Sepúlveda,[1,3] Billy Ernst,[1,4] Tamara A. Contador,[1,5] Sergio Neira,[1,4] Chris Harrod[1,6]
[1]Núcleo Milenio INVASAL, Concepcion, Chile. [2]Departamento de Zoologia, Facultad de Ciencias Naturales y Oceanograficas, Universidad de Concepcion, Concepcion, Chile. [3]Instituto de Biologia, Facultad de Ciencias, Universidad de Valparaiso, Valparaiso, Chile. [4]Departamento de Oceanografia, Facultad de Ciencias Naturales y Oceanograficas, Universidad de Concepcion, Concepcion, Chile.

[5]Departamento de Ciencias & Parque Etnobotanico Omora, Universidad de Magallanes, Punta Arenas, Chile. [6]Instituto de Ciencias Naturales Alexander von Humboldt, Universidad de Antofagasta, Antofagasta, Chile.
*Corresponding author. Email: dgomezu@udec.cl

REFERENCES

1. D. Guzman, "Great salmon escape threatens to taint Chile's fish farms," *Bloomberg* (2018); www.bloomberg.com/news/articles/2018-07-09/great-salmon-escape-threatens-to-taint-chile-fish-farm-industry.
2. I. Arismendi *et al*., *Rev. Fish Biol. Fish*. **24**, 919 (2014).
3. M. Sepulveda, I. Arismendi, D. Soto, F. Jara, F. Farias, *Aquacult. Environ. Interact*. **4**, 273 (2013).
4. D. Soto, F. Jara, C. Moreno, *Ecol. Appl*. **11**, 1750 (2001).
5. "Value of Chilean salmon exports up 20% in 2017," *Fish Farming Expert* (2018); www.fishfarmingexpert.com/article/value-of-chilean-salmon-exports-up-20-in-2017.
6. C. Montero, V. Konde, F. Farinelli, "A case study of the salmon industry in Chile" (United Nations, 2006).
7. AQUA, "Autoridades de Los Lagos se refieren a escape de salmones," *AQUA—Acuicultura + Pesca* (2018); www.aqua.cl/2018/07/10/autoridades-de-los-lagos-se-refieren-a-escape-de-salmones [in Spanish].
8. Ø. Jensen, T. Dempster, E. B. Thorstad, I. Uglem, A. Fredheim, *Aquacult. Environ. Interact*. **1**, 71 (2010).
9. D. Soto, J. Aguilar-Manjarrez, N. Hishamunda, "Building an ecosystem approach to aquaculture" (FAO Fisheries and Aquaculture Proceedings, 2008); www.fao.org/3/a-i0339e.pdf.
10. H. Salgado, J. Bailey, R. Tiller, J. Ellis, *Ocean Coast. Manage*. **118**, 189 (2015).

ACKNOWLEDGMENTS

Funding to Núcleo Milenio INVASAL was provided by Iniciativa Científica Milenio from Chile's Ministerio de Economía, Fomento & Turismo.

10.1126/science.aau7973

TECHNICAL COMMENT ABSTRACTS

Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Robert B. Best, Wenwei Zheng, Alessandro Borgia, Karin Buholzer, Madeleine B. Borgia, Hagen Hofmann, Andrea Soranno, Daniel Nettels, Klaus Gast, Alexander Grishaev, Benjamin Schuler

Riback *et al.* (Reports, 13 October 2017, p. 238) used small-angle x-ray scattering (SAXS) experiments to infer a degree of compaction for unfolded proteins in water versus chemical denaturant that is highly consistent with the results from Förster resonance energy transfer (FRET) experiments. There is thus no "contradiction" between the two methods, nor evidence to support their claim that commonly used FRET fluorophores cause protein compaction.

Full text: dx.doi.org/10.1126/science.aar7101

Response to Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Joshua A. Riback, Micayla A. Bowman, Adam Zmyslowski, Catherine R. Knoverek, John Jumper, Emily B. Kaye, Karl F. Freed, Patricia L. Clark, Tobin R. Sosnick

Best *et al.* claim that we provide no convincing basis to assert that a discrepancy remains between FRET and SAXS results on the dimensions of disordered proteins under physiological conditions. We maintain that a clear discrepancy is apparent in our and other recent publications, including results shown in the Best *et al*. comment. A plausible origin is fluorophore interactions in FRET experiments.

Full text: dx.doi.org/10.1126/science.aar7949

Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Gustavo Fuertes, Niccolo Banterle, Kiersten M. Ruff, Aritra Chowdhury, Rohit V. Pappu, Dmitri I. Svergun, Edward A. Lemke

Editors at *Science* requested our input on the above discussion (comment by Best *et al.* and response by Riback *et al*.) because both sets of authors use our data from Fuertes *et al.* (2017) to support their arguments. The topic of discussion pertains to the discrepant inferences drawn from SAXS versus FRET measurements regarding the dimensions of intrinsically disordered proteins (IDPs) in aqueous solvents. Using SAXS measurements on labeled and unlabeled proteins, we ruled out the labels used for FRET measurements as the cause of discrepant inferences between the two methods. Instead, we propose that FRET and SAXS provide complementary readouts because of a decoupling of size and shape fluctuations that is intrinsic to finite-sized, heteropolymeric IDPs. Accounting for this decoupling resolves the discrepant inferences between the two methods, thus making a case for the utility of both methods.

Full text: dx.doi.org/10.1126/science.aau8230

Editor's Note: To expedite publication, we have decided to post some Technical Comments before their responses, which will run in a later issue.

Comment on "Unexpected reversal of C_3 versus C_4 grass response to elevated CO_2 during a 20-year field experiment"

Ming Nie, Junyu Zou, Xiao Xu, Chao Liang, Changming Fang, Bo Li

Reich *et al.* (Reports, 20 April 2018, p. 317) reported that elevated carbon dioxide (eCO_2) switched its effect from promoting C_3 grasses to favoring C_4 grasses in a long-term experiment. We argue that the authors did not appropriately elucidate the interannual climate variation as a potential mechanism for the reversal of C_4-C_3 biomass in response to eCO_2.

Full text: dx.doi.org/10.1126/science.aau3016

TECHNICAL COMMENT

DISORDERED PROTEINS

Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Robert B. Best[1]*, Wenwei Zheng[2], Alessandro Borgia[3], Karin Buholzer[3], Madeleine B. Borgia[3], Hagen Hofmann[4], Andrea Soranno[5], Daniel Nettels[3], Klaus Gast[6], Alexander Grishaev[7], Benjamin Schuler[3,8]*

Riback *et al*. (Reports, 13 October 2017, p. 238) used small-angle x-ray scattering (SAXS) experiments to infer a degree of compaction for unfolded proteins in water versus chemical denaturant that is highly consistent with the results from Förster resonance energy transfer (FRET) experiments. There is thus no "contradiction" between the two methods, nor evidence to support their claim that commonly used FRET fluorophores cause protein compaction.

Riback *et al*. (*1*) recently presented a "molecular form factor" (MFF) method addressing the well-known challenges (*2*) of analyzing small-angle x-ray scattering (SAXS) data for unfolded or intrinsically disordered proteins (IDPs) (*3*, *4*). Combined with the precision of SAXS measurements coupled to size exclusion chromatography, their method yielded the following results: (i) Unfolded proteins in water have a polymer scaling exponent $\nu \approx 1/2$, near the theta-solvent condition where protein-protein and protein-solvent interactions are balanced; in denaturant, this increases to $\nu \approx 3/5$, the limit where the protein-solvent interactions dominate. (ii) This change of scaling exponent is accompanied by an increase in radius of gyration (R_g) of 10% to 20%, depending on the sequence. We are pleased that these findings are in overall agreement with SAXS and Förster resonance energy transfer (FRET) studies from our laboratories (*3*, *5*, *6*) and others (*4*).

The chain expansion observed by Riback *et al*. helps to resolve a long-standing controversy between SAXS and FRET experiments (*7*): With increasing denaturant concentration, FRET experiments had generally shown chain expansion, while until recently (*3*, *4*) most SAXS studies observed no statistically significant change of R_g [(*8*) and references therein]. Their results are consistent with our recent collaborative study in which we compared SAXS and FRET estimates of R_g for each of two proteins (necessary because chain dimensions can be sequence-dependent). We found that the results are mutually consistent if both data types are analyzed with state-of-the-art methods (*3*, *6*) (Fig. 1). A second study of a large group of IDPs reached a similar conclusion (*4*).

The main reasons for the discrepancy were deficiencies in the analysis of both SAXS and FRET data. Earlier SAXS experiments underestimated expansion because the unfolded state of the foldable sequences studied could only be accessed above a certain denaturant concentration (*3*), as also pointed out by Riback *et al*., and because obtaining precise and accurate R_g values from SAXS data of IDPs using the Guinier approximation is challenging (*1*, *3*, *4*). The former difficulty has been overcome by studying destabilized or intrinsically disordered proteins (*1*, *3*, *4*), the latter by improved analysis such as Bayesian ensemble refinement (*3*, *4*, *6*, *9*) or the closely related MFF method (*1*). On the FRET side, the use of polymer models, such as a Gaussian chain or self-avoiding random walk (SARW), to interpret experimental results can overestimate the change in R_g (*3*, *10*, *11*) (Fig. 1), largely because the relative change of R_g upon chain expansion is intrinsically less than that of the end-to-end distance most commonly measured by FRET (*3*, *4*, *10*, *11*). With ensemble refinement applied to either SAXS or FRET, or both combined, the data from both experiments yield consistent distributions of conformations, considering statistical error (*3*, *4*, *6*), as shown in Fig. 1 (*3*).

We therefore dispute the authors' claim that their results are "in apparent contradiction to a variety of FRET measurements," given that FRET experiments have not been reported on the sequences they studied. Their results are consistent with the magnitude of the change in both R_g and ν with denaturant inferred from recent FRET studies (*3*, *4*), including the larger change in R_g and ν below 2 M GdmCl (Fig. 2A). Their ν values of 0.48 to 0.54 in water or low denaturant concentration are within the range (reflecting sequence-dependent variation of ν) obtained on the basis of ensemble refinement of data from previous FRET studies (*3*–*6*) (Fig. 2B).

Despite this consistency, the authors suggest that "addition of fluorophores with hydrophobic character may lead to chain compaction and may

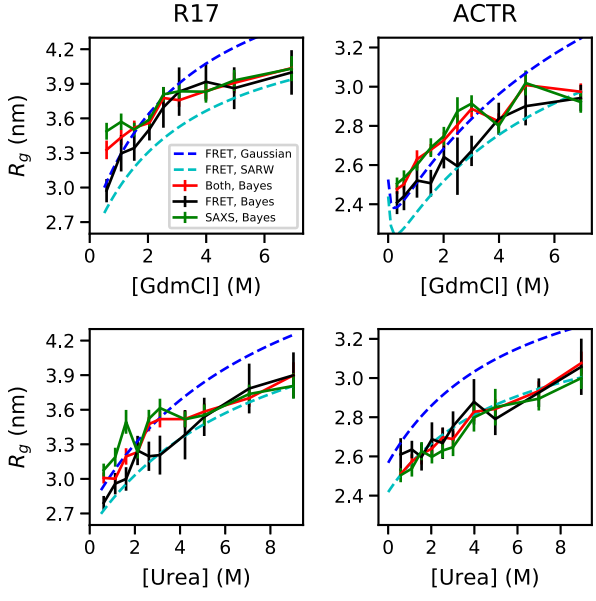


Fig. 1. R_g from Bayesian ensemble refinement against experimental data for unfolded proteins in denaturant using FRET, SAXS, or both experiments. Results are shown for two proteins (ACTR, R17) in urea and guanidinium chloride (GdmCl) (*3*). R_g values from FRET using Gaussian chain or SARW models (*3*) are shown for reference.

[1]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, MD 20892, USA. [2]College of Integrative Sciences and Arts, Arizona State University, Mesa, AZ 85212, USA. [3]Department of Biochemistry, University of Zurich, 8057 Zurich, Switzerland. [4]Department of Structural Biology, Weizmann Institute of Science, Rehovot 76100, Israel. [5]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA. [6]Physical Biochemistry, University of Potsdam, 14476 Potsdam, Germany. [7]National Institute of Standards and Technology and Institute for Bioscience and Biotechnology Research, Rockville, MD 20850, USA. [8]Department of Physics, University of Zurich, 8057 Zurich, Switzerland.
***Corresponding author. Email: robert.best2@nih.gov (R.B.B.); schuler@bioc.uzh.ch (B.S.)**

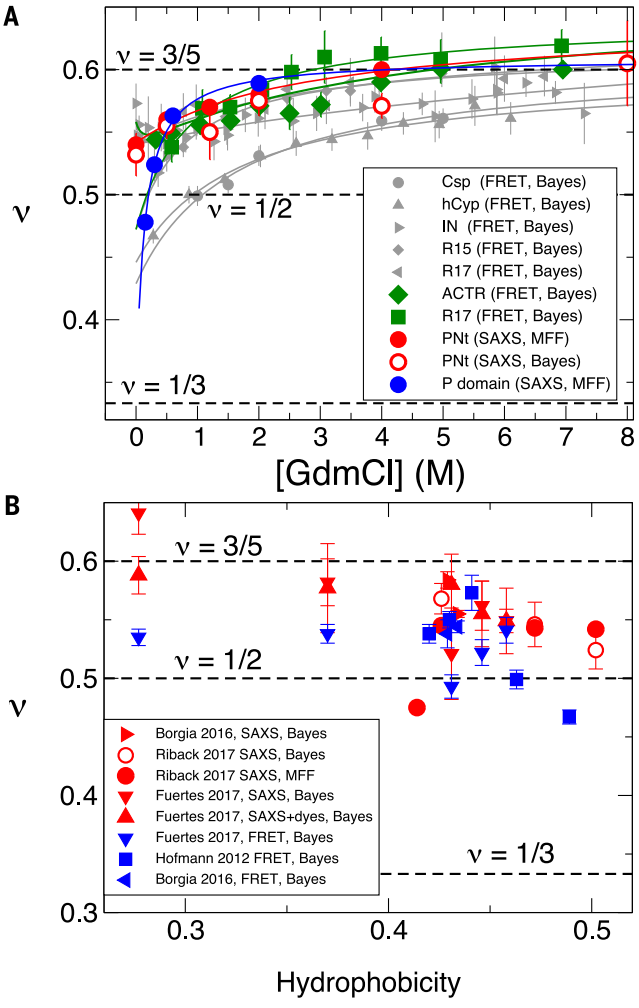


Fig. 2. Polymer scaling exponents, ν, for unfolded or disordered proteins. (**A**) Denaturant-dependent ν from SAXS data of Riback *et al*. for PNt (red) and P domain (blue) (*1*) compared with those from FRET data for a variety of unfolded and disordered proteins (gray) (*5*) and for the IDP ACTR and a destabilized spectrin R17 domain (green) (*3*). Exponents were obtained from Bayesian ensemble refinement (*3*) of primary FRET or SAXS data ("Bayes") or MFF analysis where indicated. Curves are fits to a binding model (*5*) or to a polyelectrolyte model for IN and ACTR (*3*). (**B**) Scaling exponents versus Kyte-Doolittle hydrophobicity (*15*) (rescaled between 0 and 1) for the same proteins in water or low denaturant concentration, as well as additional data for a set of IDPs in water from Fuertes *et al*. (*4*). Results for MFF (*1*) and Bayesian ensemble refinement (*3*) are highly consistent. Error bars indicate statistical error.

contribute to FRET signal changes" (*1*). Although some extremely hydrophobic FRET fluorophores can indeed cause additional compaction under native conditions, ensemble refinement identified the inconsistency between the resulting FRET and SAXS data (*3*). However, results for the more hydrophilic fluorophores most commonly used were in good agreement with SAXS data (*3*) (Fig. 1). Furthermore, a recent tour-de-force SAXS study of proteins with and without fluorophores showed only small perturbations and no systematic changes of R_g and ν upon labeling (*4*) (Fig. 2B). The evidence presented by Riback *et al*. to support their claim comes not from a protein but from earlier small-angle neutron scattering (SANS) and FRET measurements on polyethylene glycol (PEG) (*12*). PEG lacks complications from a folded state, such as those that previously (*13*) caused these authors to overlook ubiquitin expansion (*1*). The PEG study, however, used old protocols to analyze the data. Applying such earlier methods to a protein lacking a folded state, the authors had determined that "fully reduced ribonuclease A does not expand at high denaturant concentration" (*14*), but they now find an expansion for the same protein [figure 3C in (*1*)].

Riback *et al*. thus do not provide a convincing basis for their assertion that the conclusions of FRET and SAXS experiments are contradictory. Rather, their results add to the increasingly consistent picture of unfolded and intrinsically disordered proteins that has been emerging in recent years.

REFERENCES AND NOTES

1. J. A. Riback *et al*., *Science* **358**, 238–241 (2017).
2. P. Calmettes *et al*., *J. Phys. IV* **03**, C8-253–C8-256 (1993).
3. A. Borgia *et al*., *J. Am. Chem. Soc.* **138**, 11714–11726 (2016).
4. G. Fuertes *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6342–E6351 (2017).
5. H. Hofmann *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16155–16160 (2012).
6. W. Zheng *et al*., *J. Am. Chem. Soc.* **138**, 11702–11713 (2016).
7. K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1528–1533 (2007).
8. T. Y. Yoo *et al*., *J. Mol. Biol.* **418**, 226–236 (2012).
9. G. Hummer, J. Köfinger, *J. Chem. Phys.* **143**, 243150 (2015).
10. E. P. O'Brien, G. Morrison, B. R. Brooks, D. Thirumalai, *J. Chem. Phys.* **130**, 124903 (2009).
11. J. Song, G.-N. Gomes, C. C. Gradinaru, H.-S. Chan, *J. Phys. Chem. B* **119**, 15191–15202 (2015).
12. H. M. Watkins *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6631–6636 (2015).
13. J. Jacob, B. Krantz, R. S. Dothager, P. Thiyagarajan, T. R. Sosnick, *J. Mol. Biol.* **338**, 369–382 (2004).
14. J. Jacob, R. S. Dothager, P. Thiyagarajan, T. R. Sosnick, *J. Mol. Biol.* **367**, 609–615 (2007).
15. J. Kyte, R. F. Doolittle, *J. Mol. Biol.* **157**, 105–132 (1982).

ACKNOWLEDGMENTS

We thank D. Thirumalai, H. S. Chan, R. Pappu, W. Eaton, and A. Bax for helpful comments, and P. Clark, T. Sosnick, D. Svergun, E. Lemke, and R. Pappu for sharing their SAXS data. Our analysis used the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov). **Funding:** Supported by the NIDDK Intramural Research Program of NIH (R.B.B.) and Arizona State University (W.Z.). Work at the University of Zurich was supported by funding from the Swiss National Science Foundation. **Author contributions:** W.Z. and R.B.B. analyzed data. R.B.B. and B.S. wrote the paper with the help of all authors. **Competing interests:** None. **Data availability:** All analysis uses primary data from published work as cited.

8 December 2017; accepted 30 July 2018
10.1126/science.aar7101

TECHNICAL RESPONSE

DISORDERED PROTEINS

Response to Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Joshua A. Riback[1], Micayla A. Bowman[2], Adam Zmyslowski[3], Catherine R. Knoverek[2], John Jumper[3,4], Emily B. Kaye[2], Karl F. Freed[4], Patricia L. Clark[2]*, Tobin R. Sosnick[3,5]*

Best *et al*. claim that we provide no convincing basis to assert that a discrepancy remains between FRET and SAXS results on the dimensions of disordered proteins under physiological conditions. We maintain that a clear discrepancy is apparent in our and other recent publications, including results shown in the Best *et al*. comment. A plausible origin is fluorophore interactions in FRET experiments.

Using our new molecular form factor (MFF), we analyzed small-angle x-ray scattering (SAXS) data for three intrinsically disordered proteins (IDPs) and found that upon a shift from 6 to 0 M guanidinium chloride (Gdn), there was a mild decrease in radius of gyration (R_g) (*1*). For these 124- to 334-residue IDPs with amino acid sequences typical of foldable proteins, the value of R_g decreased by 20 to 28% in water. As predicted from scaling laws, the denatured states of smaller proteins such as ubiquitin and protein L (76 and 64 amino acids, respectively) will contract less (17 and 15%, respectively). Notably, approximately half of this contraction occurs below 1 M Gdn, beyond the limit of many prior kinetic studies. For example, for ubiquitin, only an 8% decrease (~2 Å) in R_g is expected from 6 to 1 M Gdn. This mild and nonlinear decrease in R_g explains why many previous SAXS studies of small proteins did not observe a statistically significant decrease in R_g (*2*–*6*)—a conclusion we highlighted in figure 4A of our publication (*1*).

In contrast, Förster resonance energy transfer (FRET) studies have observed much more contraction, up to a 50% decrease in R_g in the absence of denaturant, even for small proteins (*3*, *7*, *8*). Furthermore, the majority of this contraction occurs above 1 M Gdn (Fig. 1). These findings are different from those observed using SAXS and are noticeable in figure 1 of Best *et al*. (*9*) (compare green and black points at lowest denaturant concentration). Hence, our methods and those of Best *et al*. concur that in the absence of denaturant, SAXS analysis returns R_g values ~15% larger than FRET studies when each data set is analyzed independently. In our publication, we hypothesized that this discrepancy could be due to the addition of dyes necessary for FRET studies (*1*). In contrast, Best *et al*. provide no physical explanation and instead rely on a joint analysis of both types of data that only reduces the appearance of the discrepancy.

An alternative, chain length–independent method to compare results from SAXS and FRET studies is to compare the Flory exponent (ν in the relationship $R_g \propto N^{\nu}$, where N = protein length). Values of ν equal to 0.33 or 0.60 correspond to a globule or a self-avoiding random walk (SARW), respectively. In a theta solvent, ν = 0.50 and intrachain interactions are equally as favorable as solvent-chain interactions; this value defines the boundary between good and poor solvents. For the three foldable sequences we examined using SAXS, ν in water is 0.54. In contrast, using FRET, Hofmann *et al*. (*10*) found that ν for four foldable sequences in water ranged from 0.4 to 0.51. Because the range of ν between a globule and a SARW is only 0.27, a difference of 0.1 units represents 38% of the entire range. Hence, to argue that both methods are in agreement if the values of ν are near ½ is imprecise.

This discrepancy is not limited to the specific proteins selected by us or by Hofmann *et al*. (*10*). We reanalyzed available SAXS data for other IDPs with sequences typical of folded proteins (*11*–*17*) using our new MFF procedure and compared these results to ν extracted from available FRET studies of disordered proteins with foldable sequences (Fig. 2A). When measured by SAXS, ν is typically above 0.54, whereas ν from FRET studies is typically below the theta condition of 0.50. Note that our Fig. 2A should, in principle, match figure 2B of Best *et al*. However, figure 2B of Best *et al*. does not include results from two of the most collapsed proteins that Fuertes *et al*. (*11*) studied by FRET: R98 and NSP. The FRET and SAXS values for these proteins differ significantly (ν = 0.44 versus 0.56 for R98 and 0.49 versus 0.60 for NSP, respectively). Moreover, Best *et al*. plot results for other proteins (e.g., Csp, hCyp, and R15) in ~1 M denaturant rather than the value extrapolated to 0 M denaturant, as they calculated in figure 2A and reported in their original study

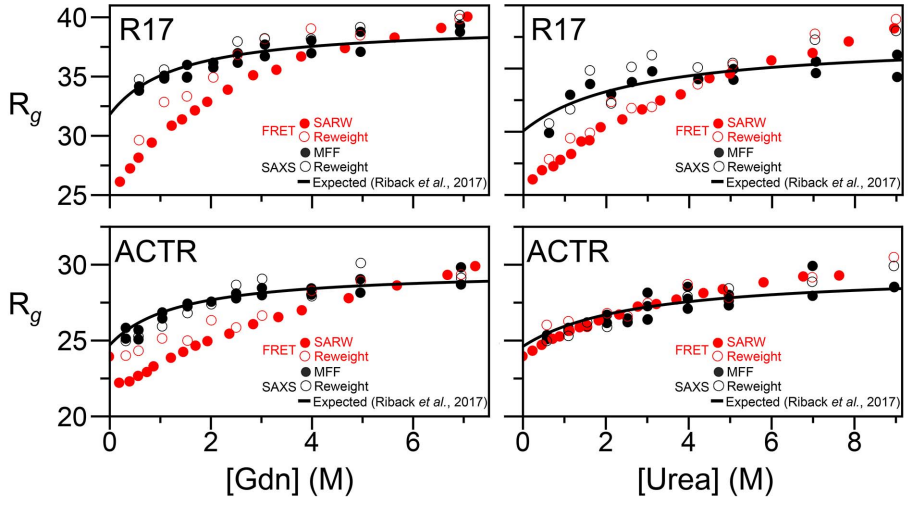


Fig. 1. SAXS and FRET exhibit different denaturant dependence for two proteins, R17 and ACTR, studied by Best and co-workers. All primary data are from (*22*). Open and solid red circles are two models applied to FRET data (reweighted simulations and a SARW, respectively), as presented in (*22*). Open and solid black circles are SAXS data fit using reweighted simulations [as presented in (*22*)] and our MFF (*1*), respectively. The black line is the expected SAXS denaturant dependence of R_g taken from (*1*), assuming that 2 M urea is equivalent to 1 M Gdn.

[1]Graduate Program in Biophysical Sciences, University of Chicago, Chicago, IL 60637, USA. [2]Department of Chemistry and Biochemistry, University of Notre Dame, Notre Dame, IN 46556, USA. [3]Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, IL 60637, USA. [4]Department of Chemistry and James Franck Institute, University of Chicago, Chicago, IL 60637, USA. [5]Institute for Biophysical Dynamics, University of Chicago, Chicago, IL 60637, USA.
***Corresponding author. Email: pclark1@nd.edu (P.L.C.); trsosnic@uchicago.edu (T.R.S.)**

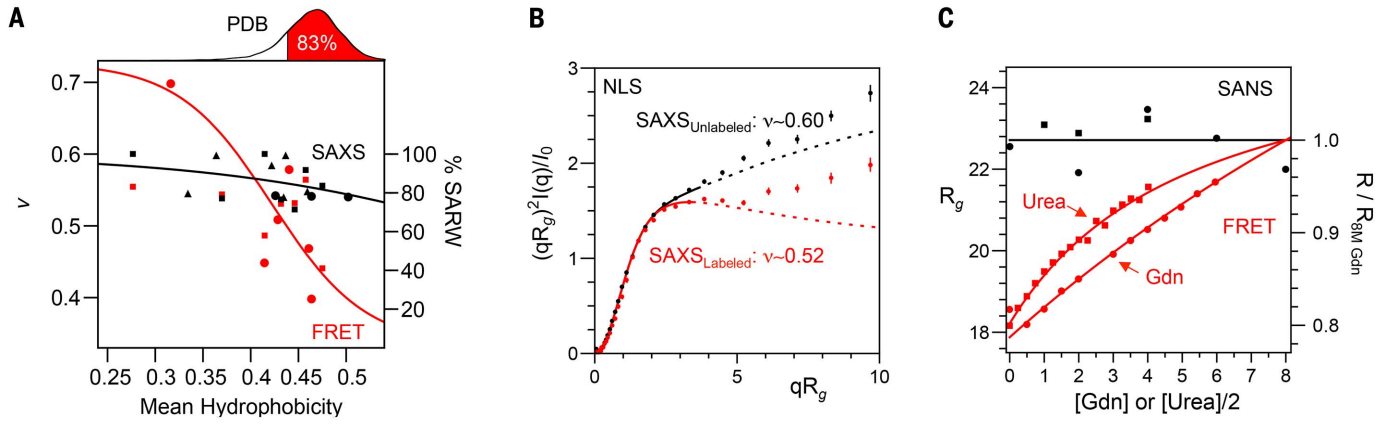


Fig. 2. SAXS and FRET show inconsistent solvent qualities for IDPs in water. (**A**) Trends of hydrophobicity (Kyte-Doolittle scale) versus ν in water for SAXS data of foldable protein sequences fit to our MFF for data from our recent study (*1*) (black circles), Fuertes *et al.* (*11*) (black squares), and other studies (*12*–*17*) (black triangles). Also shown are results from FRET studies calculated as in (*10*) for data from (*10*) (red circles) and (*11*) (red squares). The red trend line for FRET results is from (*10*). The black trend line is the best fit to the SAXS data shown. At the top is a histogram of hydrophobicity of representative proteins in the PDB [dataset from (*1*)] with sequences in the FRET-inferred poor solvent region highlighted in red. (**B**) SAXS profiles from (*11*) for unlabeled and labeled versions of the NLS protein fit using our MFF. Solid lines span data points used for fitting. Error bars are the SD appropriate for Poisson counting statistics. (**C**) PEG results from (*8*) remain unchanged with improved analysis. Small-angle neutron scattering (SANS) data (black) are fit to our MFF; FRET data (red) are fit assuming a SARW distribution and normalized to high denaturant.

[Hofmann *et al.* (*10*)]. These differences misleadingly minimize the discrepancy between the SAXS and FRET results. Plotting instead results for all proteins with sequences typical of folded proteins in the absence of denaturant, as in our Fig. 2A, reveals clear discrepancies between results from these methods.

Figure 2 of Best *et al.* emphasizes the sole exception to ν > 0.5 in our SAXS data: the P domain, a hydrophobic IDP region of low sequence complexity whose collapse correlates with stress granule formation (*18*). Although this finding indicates that some hydrophobic, nonfoldable sequences can collapse, it does not speak directly to the discrepancy between SAXS and FRET results for protein-like sequences.

Having established that discrepancies remain between results of SAXS and FRET studies, we considered possible origins. As we discussed (*1*), Chan and others have proposed that complications in converting FRET efficiency to R_g could account for some of this difference (*11*, *19*–*21*). The other obvious issue is the presence of fluorophores.

Schuler and co-workers have found that some fluorophore pairs influence FRET signals but maintain that the Alexa 488/594 pair is suitable (*22*). Their all-atom molecular dynamics simulations, however, reported a 10% contraction for an IDP with versus without fluorophores in 1 M urea (*23*). In the absence of denaturant, this dye effect would presumably be even larger. Recently, Fuertes *et al.* conducted SAXS measurements on five proteins with and without Alexa 488/594 (*11*). In all cases, addition of fluorophores changed the SAXS profile. From the change in scattering between the labeled and unlabeled proteins, the authors calculated meaningful differences ($\nu_{unlabeled} - \nu_{labeled} = \Delta\nu = 0.08, 0.03, 0.03, -0.02, -0.04$). Reanalysis of these data with our MFF (*1*) finds $\Delta\nu = 0.09, 0.06, 0.03, -0.02, -0.08$ (Fig. 2B). To address an additional concern raised by Best *et al.*, we reanalyzed published polyethylene glycol (PEG) scattering data (*8*) with our MFF and confirmed the published observation that whereas FRET-measured PEG labeled with Alexa 488/594 contracts in lower denaturant, PEG without labels does not contract (Fig. 2C). Hence, significant discrepancies exist between FRET and scattering techniques even when using fluorophores considered suitable.

We are gratified that Best *et al.* agree that water should be considered to be a good solvent for the denatured state of most foldable sequences (ν > 0.5). However, an analysis of the FRET results alone would lead one to believe the opposite (ν < 0.5). The presence of fluorophores in FRET studies is a likely origin of this difference.

REFERENCES AND NOTES

1. J. A. Riback *et al.*, *Science* **358**, 238–241 (2017).
2. K. W. Plaxco, I. S. Millett, D. J. Segel, S. Doniach, D. Baker, *Nat. Struct. Biol.* **6**, 554–556 (1999).
3. T. Y. Yoo *et al.*, *J. Mol. Biol.* **418**, 226–236 (2012).
4. J. Jacob, B. Krantz, R. S. Dothager, P. Thiyagarajan, T. R. Sosnick, *J. Mol. Biol.* **338**, 369–382 (2004).
5. J. J. Skinner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15975–15980 (2014).
6. S. V. Kathuria *et al.*, *J. Mol. Biol.* **426**, 1980–1994 (2014).
7. K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, W. A. Eaton, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 1528–1533 (2007).
8. H. M. Watkins *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6631–6636 (2015).
9. R. B. Best *et al.*, *Science* **361**, eaar7101 (2018).
10. H. Hofmann *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 16155–16160 (2012).
11. G. Fuertes *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6342–E6351 (2017).
12. Z. P. Gates *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2241–2246 (2017).
13. S. Contreras-Martos *et al.*, *Sci. Rep.* **7**, 4676 (2017).
14. E. W. Martin *et al.*, *J. Am. Chem. Soc.* **138**, 15323–15335 (2016).
15. M. Kolonko *et al.*, *PLOS ONE* **11**, e0162950 (2016).
16. S. S. Shell, C. D. Putnam, R. D. Kolodner, *Mol. Cell* **26**, 565–578 (2007).
17. M. Renner *et al.*, *Sci. Rep.* **7**, 14865 (2017).
18. J. A. Riback *et al.*, *Cell* **168**, 1028–1040.e19 (2017).
19. J. Song, G. N. Gomes, T. Shi, C. C. Gradinaru, H. S. Chan, *Biophys. J.* **113**, 1012–1024 (2017).
20. J. Song, G. N. Gomes, C. C. Gradinaru, H. S. Chan, *J. Phys. Chem. B* **119**, 15191–15202 (2015).
21. H. Maity, G. Reddy, *J. Am. Chem. Soc.* **138**, 2609–2616 (2016).
22. A. Borgia *et al.*, *J. Am. Chem. Soc.* **138**, 11714–11726 (2016).
23. W. Zheng *et al.*, *J. Am. Chem. Soc.* **138**, 11702–11713 (2016).

ACKNOWLEDGMENTS

We thank H. S. Chan, R. Best, W. Zhang, L. Lapidus, K. Plaxco, E. Martin, and A. Holehouse for useful discussions. The analysis used in this and our original publication are available from https://github.com/sosnicklab/SAXSonIDPs. A web server to fit SAXS data using our MFF is available at http://sosnick.uchicago.edu/SAXSonIDPs. **Funding:** Supported by NIH grants GM055694 (T.R.S., K.F.F.), GM097573 (P.L.C.), GM103622 and 1S10OD018090-01 (T. C. Irving), T32 EB009412 (T.R.S.), T32 GM007183 (B. Glick), and T32 GM008720 (J. Picirrilli) and by NSF grants GRF DGE-1144082 (J.A.R.) and MCB 1516959 (C. R. Matthews). Use of the Advanced Photon Source was supported by the U.S. Department of Energy under contract DE-AC02-06CH11357. **Author contributions:** All authors contributed to and approved the manuscript. J.A.R. performed the data analysis. **Competing interests:** None.

3 January 2018; accepted 30 July 2018
10.1126/science.aar7949

TECHNICAL COMMENT

DISORDERED PROTEINS

Comment on "Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water"

Gustavo Fuertes[1,2,3]*, Niccolo Banterle[1]*, Kiersten M. Ruff[4]*, Aritra Chowdhury[1]*, Rohit V. Pappu[4]†, Dmitri I. Svergun[2]†, Edward A. Lemke[1,5,6]†

Editors at *Science* requested our input on the above discussion (comment by Best *et al*. and response by Riback *et al*.) because both sets of authors use our data from Fuertes *et al*. (2017) to support their arguments. The topic of discussion pertains to the discrepant inferences drawn from SAXS versus FRET measurements regarding the dimensions of intrinsically disordered proteins (IDPs) in aqueous solvents. Using SAXS measurements on labeled and unlabeled proteins, we ruled out the labels used for FRET measurements as the cause of discrepant inferences between the two methods. Instead, we propose that FRET and SAXS provide complementary readouts because of a decoupling of size and shape fluctuations that is intrinsic to finite-sized, heteropolymeric IDPs. Accounting for this decoupling resolves the discrepant inferences between the two methods, thus making a case for the utility of both methods.

Quantitative descriptions of conformational ensembles of intrinsically disordered proteins (IDPs) are directly relevant for understanding the functions and cellular processes controlled by IDPs. Small-angle x-ray scattering (SAXS) and Förster resonance energy transfer (FRET) are two experimental techniques that have been widely used to quantify the overall sizes and shapes of IDPs in different milieus. The two techniques yield convergent descriptions regarding the dimensions of IDPs in high concentrations of denaturants, but they can yield discrepant inferences in the absence of denaturants (*1*–*3*). What is the source of these discrepant inferences? Are they caused by the dyes used in FRET measurements, as proposed by Riback *et al*. (*4*)? Or does the discrepancy come from the method of analysis of SAXS and FRET data, as proposed by Best *et al*. (*5*)?

On the basis of direct SAXS measurements of IDPs with and without dyes, we have experimentally demonstrated that the dyes are not the source of systematic biases and the discrepant inferences (*1*). However, as a general resolution of the conflict between SAXS and FRET, our results point to sequence-specific decoupling between end-to-end distances (R_e) and radii of gyration (R_g). It is important to note that R_e is directly derivable from FRET data, whereas R_g is directly derivable from SAXS data. FRET data cannot be used to extract R_g without making a series of simplifying assumptions about the connections between R_g and R_e, and the converse is true of SAXS data. Indeed, much of the conflict in inferences drawn from SAXS versus FRET originates from the fact that the two methods provide access to two different quantities. When we allow for the possibility that R_e and R_g can be decoupled from one another because of shape fluctuations that are consequences of the finite size and the heteropolymeric nature of IDPs, we arrive at a reconciled view, which suggests that SAXS and FRET yield complementary rather than contradictory insights. Therefore, we propose that the debate should not be about the merits or demerits of the two methods. Both methods have their strengths and weaknesses; therefore, the focus should be on the growing consensus regarding methodological advances that rely on improved numerical/theoretical analysis and the use of sophisticated atomistic simulations to analyze data from SAXS and FRET measurements.

To directly test the effects of the dyes, we performed SAXS measurements with labeled and unlabeled IDPs on a series of molecules with different sequence attributes (*1*). We pursued global fits to analyze the relationship of measured R_g and R_e from several proteins giving a broader coverage of sequence space. Scaling theories suggest that $R_g \propto N^{\nu}$, where the critical exponent ν describes how the protein dimensions scale with the number of residues (N). We agree that comparing $\Delta\nu = \nu_{labeled} - \nu_{unlabeled}$ is one way to detect potential dye interferences. Indeed, it is reassuring that estimates of $\Delta\nu$ in the absence of denaturant show the lack of systematic trends in $\Delta\nu$. The values for $\Delta\nu$ range from being negative to positive, and importantly, $\Delta\nu$ is zero on average ($\Delta\nu$ = 0.03, 0.08, 0.03, −0.02, −0.04; average 0.02 ± 0.04). To test whether this level of deviation is meaningful or lies within the experimental uncertainty, we conducted SAXS and FRET under highly denaturing conditions, where all independent assessments converge. These experiments yield similarly small $\Delta\nu$ values ranging from positive to negative ($\Delta\nu$ = −0.07, −0.09, −0.01, 0.04, 0.00; average −0.03 ± 0.05). Our direct measurements negate the hypothesis that the selected FRET dyes used by many groups including Fuertes *et al*. and Borgia *et al*. (*1*, *3*) are systematic modifiers of conformational ensembles, although sequence-specific interactions involving specific dyes may prevail in some cases.

In their response, Riback *et al*. picked out the sequence designated as NLS from our set of IDPs for their analysis [figure 2B of (*4*)]. For this sequence, there is indeed a small but measurable difference between the SAXS data for unlabeled and labeled proteins. This observation is consistent with our analysis (*1*) but is limited to NLS, whereas other IDPs in our study showed virtually identical normalized scattering behavior for labeled versus unlabeled IDPs. Of course the label contributes a detectable increase in the molecular weight of the protein, and this is accounted for in the analysis and interpretation of the SAXS data (*1*). Focusing on the NLS sequence reveals a pitfall of isolating a single measurement from an entire dataset to draw general conclusions. Of all the proteins in our dataset, NLS was the only protein for which we could not perform measurements in phosphate-buffered saline (PBS) because of confounding effects of the milieu. Instead, we had to perform measurements in a zwitterionic HEPES buffer. This "complication" is in line with this particular protein being highly charged, and thus sensitivity toward certain buffers as well as charged dyes is to be expected. If one were to extrapolate from the NLS data and assert that dyes have an unambiguous perturbing effect on FRET measurements, then one would need to conclude that in general no measurements can be made for proteins in physiologically relevant PBS buffers, thus highlighting the issues one confronts with making extrapolations from a specific data point.

Having largely exonerated the dyes as the source of any form of systematic bias, we are left with the question of why the extent of contraction previously inferred from FRET is higher than from SAXS and why this difference is manifested for IDPs in the absence of denaturant. We propose that the discrepant degrees of contraction inferred from FRET versus SAXS originate

[1]European Molecular Biology Laboratory, 69117 Heidelberg, Germany. [2]European Molecular Biology Laboratory, 22607 Hamburg, Germany. [3]Institute of Biotechnology CAS, 252 50 Vestec, Czech Republic. [4]Department of Biomedical Engineering and Center for Biological Systems Engineering, Washington University, St. Louis, MO 63130, USA. [5]Biocenter, Departments of Biology and Chemistry, Pharmacy and Geosciences, Johannes Gutenberg University, 55128 Mainz, Germany. [6]Institute of Molecular Biology, 55128 Mainz, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: edlemke@uni-mainz.de (E.A.L.); svergun@embl-hamburg.de (D.I.S.); pappu@wustl.edu (R.V.P.)

in the distinct types of quantities that the two techniques measure and the assumptions used to convert between them. This proposal is consistent with issues raised by Borgia *et al.* (*3*). Converting FRET efficiency (E_{FRET}) into an average R_e value requires the use of theoretical polymer models or of ensembles derived from computer simulations. Additionally, SAXS provides inferences regarding R_g, whereas FRET provides inferences regarding R_e. To compare the two inferences, one simple approach is to convert between R_e and R_g using a multiplicative factor—for example, $R_g = R_e/\sqrt{6}$ for a Gaussian chain. However, all IDPs cannot be taken for granted to be Gaussian chains in different milieus.

Motivated by the dependence on polymer models for the values of R_g extracted from FRET efficiencies (*6*, *7*), we abandoned the ansatz that R_g and R_e must be coupled by a unique constant. This is justified because (i) proteins are not homopolymers [e.g., PEG (*8*)], and (ii) different geometrical objects have very different R_g/R_e relationships (defined as $G = R_g^2/R_e^2$) such that $G = 1.31$ for a sphere, $G = 12$ for a rod, $G = 6$ for an ideal Gaussian chain with a scaling exponent $\nu = 0.5$, and $G = 7.04$ for a self-avoiding random walk, where $\nu = 0.6$. Consequently, when measuring only R_g, there exist multiple solutions for R_e depending on the shape of the IDP ensemble and vice versa. We showed that R_g and R_e are two related but genuinely distinct measures of IDP conformations. Away from the Gaussian chain limit, the two quantities can be readily decoupled from one another. Accordingly, the two measures together provide complementary albeit distinct insights regarding the dimensions of IDPs in the absence of denaturant. Using only one measure as opposed to both R_e and R_g leads to discrepant inferences because the models used to interpret either dataset end up imposing a homogeneity that appears to be unwarranted for IDPs in the absence of denaturant (*1*). We propose that the extent of decoupling between R_g and R_e depends on the amino acid sequence and solution conditions.

Overall, using ensemble-based numerical analysis methods that avoid simplifications for the conversion between R_g and R_e can help to minimize apparent discrepancies between SAXS and FRET (*1*, *3*, *7*). Along these lines, Best *et al.* summarize a method that was used by Borgia *et al.* to integrate SAXS and FRET data (*3*). The approaches used by Borgia *et al.* (*3*) and Fuertes *et al.* (*1*) are complementary and the findings that emerge are in general agreement with one another. Both groups recognize the weaknesses of simplified homopolymer models for analyzing SAXS and FRET data. Neither group finds detectable evidence for systematic compacting effects of dyes. Minor positive and negative deviations in the inferred scaling exponents from FRET vis-à-vis SAXS are to be expected for finite-sized heteropolymers, given the errors and features that are associated with both experimental methodologies.

To conclude, it is worth reiterating that our statement that "dyes are not the source of the discrepancy" results from direct measurements rather than conjecture. Further, our explanations for resolving the discrepant inferences between SAXS and FRET invoke a decoupling of R_g and R_e. This hypothesis is substantiated by numerical evidence for the decoupling that explains all the available data. Nonetheless, it remains a phenomenological model and demands careful theoretical and experimental scrutiny.

REFERENCES AND NOTES

1. G. Fuertes *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6342–E6351 (2017).
2. J. A. Riback *et al.*, *Science* **358**, 238–241 (2017).
3. A. Borgia *et al.*, *J. Am. Chem. Soc.* **138**, 11714–11726 (2016).
4. J. A. Riback *et al.*, *Science* **361**, eaar7949 (2018).
5. R. B. Best *et al.*, *Science* **361**, eaar7101 (2018).
6. E. P. O'Brien, G. Morrison, B. R. Brooks, D. Thirumalai, *J. Chem. Phys.* **130**, 124903 (2009).
7. J. Song, G. N. Gomes, T. Shi, C. C. Gradinaru, H. S. Chan, *Biophys. J.* **113**, 1012–1024 (2017).
8. H. M. Watkins *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6631–6636 (2015).
9. E. Valentini, A. G. Kikhney, G. Previtali, C. M. Jeffries, D. I. Svergun, *Nucleic Acids Res.* **43**, D357–D363 (2015).

ACKNOWLEDGMENTS

Funding: Supported by funding from the Gutenberg Research College, SFB1129, CellNetworks, and the European Research Council (SMPFv2.0) (E.A.L.); the European Union (H2020 grant iNEXT # 653706) and the Joachim Herz Foundation (D.I.S.); NIH grant (5R01NS056114 (R.V.P.); and the project "Structural dynamics of biomolecular systems" (CZ.02.1.01/0.0/0.0/15_003/0000447) from European Regional Development Fund (ELIBIO) (G.F.). **Author contributions:** All authors contributed to data analysis and reanalysis. All authors co-wrote the manuscript. **Competing interests:** None. **Data and materials availability:** The complete SAXS dataset discussed in (*1*) has been deposited to the Small-Angle Scattering Biological Database (SASBDB) (*9*) under the following accession codes: ASDEH2; SASDEJ2; SASDEK2; SASDEL2; SASDEM2; SASDEN2; SASDEP2; SASDEQ2; SASDER2; SASDES2; SASDET2; SASDEU2; SASDEV2; SASDEW2; SASDEX2; SASDEY2; SASDEZ2; SASDE23; SASDE33; SASDE43; SASDE53; SASDE63; SASDE73; SASDE83; SASDE93; SASDEA3; SASDEB3; SASDEC3.

18 July 2018; accepted 30 July 2018
10.1126/science.aau8230

Cite as: M. Nie *et al.*, *Science* 10.1126/science.aau3016 (2018).

Comment on "Unexpected reversal of C_3 versus C_4 grass response to elevated CO_2 during a 20-year field experiment"

Ming Nie,[1]*† Junyu Zou,[1]† Xiao Xu,[1]† Chao Liang,[2] Changming Fang,[1] Bo Li[1]

[1]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Coastal Ecosystems Research Station of the Yangtze River Estuary, Fudan University, 2005 Songhu Road, Shanghai 200438, China. [2]Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, China.

*Corresponding author. Email: mnie@fudan.edu.cn

†These authors contributed equally to this work.

Reich *et al.* (Reports, 20 April 2018, p. 317) reported that elevated carbon dioxide (eCO_2) switched its effect from promoting C_3 grasses to favoring C_4 grasses in a long-term experiment. We argue that the authors did not appropriately elucidate the interannual climate variation as a potential mechanism for the reversal of C_4-C_3 biomass in response to eCO_2.

Reich *et al.* (*1*) presented results of a long-term free-air CO_2 enrichment experiment. The results showed that elevated CO_2 (eCO_2) favored C_3 grasses rather than C_4 grasses during the first 12 years; however, the pattern reversed during the subsequent eight years. It appears that their observations regarding the changes in C_4-C_3 grasses under eCO_2 condition did not reflect the effects of inter-annual variations in the ambient rainfall and temperature during the 20-yr experimental period, leading to uncertainties in their results.

The effect of eCO_2 on plant biomass largely depends on the ambient rainfall and temperature (*2*, *3*). However, Reich *et al.* (*1*) found that the responses of C_4 and C_3 grasses to eCO_2 had negligible dependence on these important climatic factors, determined by estimating the effects of multiple variables on C_4-C_3 biomass with repeated-measures analysis. According to the Cedar Creek weather data from Fort Snelling near the Saint Paul International Airport, the ambient total rainfall (316–722 mm) and average temperature (18.6–21.4°C) during the growing season had considerable interannual variations during the 20-yr experimental period. Using generalized linear models, we found that both the growing-season rainfall and average temperature positively correlated with the effect of CO_2 on C_4 biomass and the growing-season average temperature negatively correlated with the effect of CO_2 on C_3 biomass (Fig. 1). Without potential collinearity among the explanatory variables and order effects of repeated-measures analysis, our analysis is more appropriate to estimate the effect of individual variable on the response of C_4 or C_3 biomass to eCO_2 with an accurate and interpretable predictor.

The change in C_4 biomass showed a sharp decrease from 2005 to 2008 (Fig. 2), and the C_3 biomass also reached the lowest level during this period (*1*). Water stress during summer might have led to the decrease in biomass because summer rainfall during these dry years was about 53% less than the average of other years (Fig. 2). After these dry years, eCO_2 favored C_4 but not C_3 grasses. Besides the asymmetric changes in net nitrogen mineralization rates between C_4 and C_3 soils as suggested by Reich *et al.* (*1*), we offer two other possible mechanisms for the "winner"—C_4 grasses. First, increased growing-season average temperature might favor C_4 than C_3 grasses under eCO_2 condition. The growing-season average temperature significantly increased by approximately 0.98°C before and after the dry years (Fig. 2; $t = -3.6$; $P < 0.01$). By the two-way ANOVA with CO_2 (ambient CO_2 versus eCO_2) and average growing-season temperature (before versus after the dry years) as fixed factors to determine the effects of eCO_2 and temperature on the 3-yr moving averaged C_4 biomass, we found that increased growing-season temperature might interact with eCO_2 to affect $\log_{10}$-transformed C_4 biomass ($F = 4.4$; $P < 0.05$). As suggested by other studies and as shown in Fig. 1, the warm-season C_4 grasses can grow better than C_3 grasses under higher temperature conditions (*4*), and can enhance their sensitivity to eCO_2 with increasing temperature when soil moisture content is not limited (*4*–*6*). Second, C_3 grasses as cool season species lose their positive responses to eCO_2 with increase in the ambient temperature as shown in Fig. 1.

Understanding the directions and magnitudes of responses of C_4 and C_3 grasses to eCO_2 is crucial in modeling carbon-climate feedbacks. It is difficult to predict the changes in plant biomass dynamics in an intricate ecosys-

tem based only on the photosynthetic pathways. Several studies have shown that the relative effects of eCO_2 on the biomass of C_4 and C_3 grasses are highly influenced by soil water availability and temperature (*2*–*6*). We argue that the interpretation of the biomass data would be more meaningful by appropriately considering the effects of inter-annual variations in the ambient rainfall and temperature. Our analysis and interpretation of the biomass data provides insights different from those of Reich *et al.* (*1*), but we fully support their call for long-term experiments.

REFERENCES AND NOTES

1. P. B. Reich, S. E. Hobbie, T. D. Lee, M. A. Pastore, Unexpected reversal of C_3 versus C_4 grass response to elevated CO_2 during a 20-year field experiment. *Science* **360**, 317–320 (2018). doi:10.1126/science.aas9313 Medline
2. J. A. Morgan, D. R. LeCain, E. Pendall, D. M. Blumenthal, B. A. Kimball, Y. Carrillo, D. G. Williams, J. Heisler-White, F. A. Dijkstra, M. West, C4 grasses prosper as carbon dioxide eliminates desiccation in warmed semi-arid grassland. *Nature* **476**, 202–205 (2011). doi:10.1038/nature10274 Medline
3. K. E. Mueller, D. M. Blumenthal, E. Pendall, Y. Carrillo, F. A. Dijkstra, D. G. Williams, R. F. Follett, J. A. Morgan, Impacts of warming and elevated CO_2 on a semi-arid grassland are non-additive, shift with precipitation, and reverse over time. *Ecol. Lett.* **19**, 956–966 (2016). doi:10.1111/ele.12634 Medline
4. H. Poorter, M. L. Navas, Plant growth and competition at elevated CO2: On winners, losers and functional groups. *New Phytol.* **157**, 175–198 (2003). doi:10.1046/j.1469-8137.2003.00680.x
5. O. Ghannoum, C4 photosynthesis and water stress. *Ann. Bot.* **103**, 635–644 (2009). doi:10.1093/aob/mcn093 Medline
6. O. Ghannoum, S. Von Caemmerer, L. H. Ziska, J. P. Conroy, The growth response of C4 plants to rising atmospheric CO2 partial pressure: A reassessment. *Plant Cell Environ.* **23**, 931–942 (2000). doi:10.1046/j.1365-3040.2000.00609.x

ACKNOWLEDGMENTS

We thank Prof. Reich for discussing with us. The work was supported by the National Key Research and Development Program of China (2017YFC1200100), National Science Foundation of China (41630528 and 31670491), and Young Thousand Talents Program Scholar.

14 June 2018; accepted 15 August 2018
Published online 31 August 2018
10.1126/science.aau3016

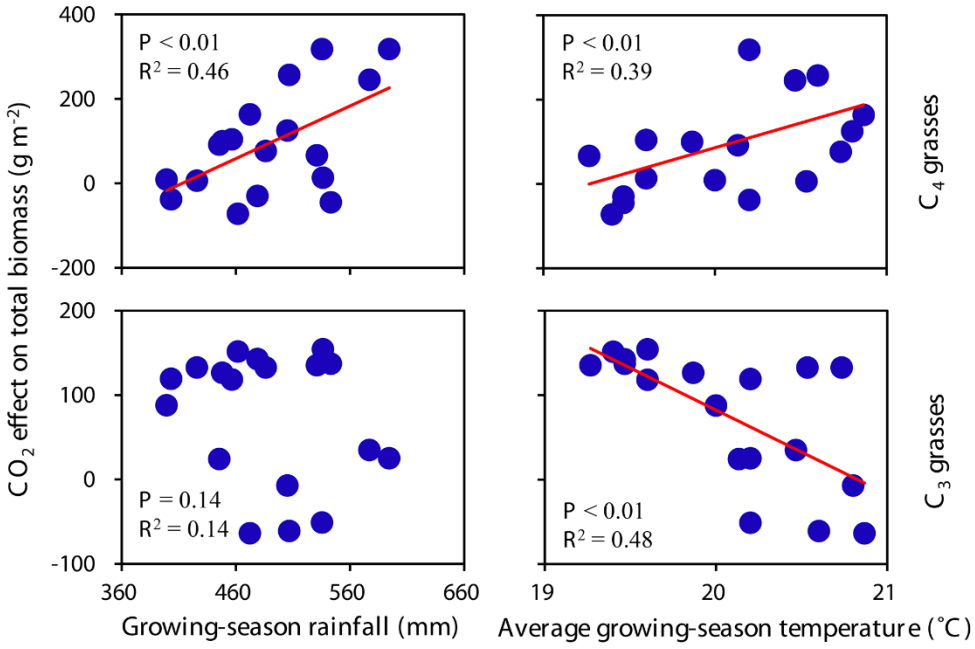


Fig. 1. Relationships between the CO_2 effect on total C_4-C_3 biomass and growing-season climate. The biomass data are from the measurements of Reich *et al*. (*1*). CO_2 effect size = biomass under eCO_2 condition – biomass under ambient CO_2 condition. The Cedar Creek weather data are from Fort Snelling near the Saint Paul International Airport (www.wunderground.com/history/airport/KMSP). The biomass and weather data are shown as 3-yr moving averages centered over the middle of each 3-yr group. The relationships between CO_2 effect size and climatic factors were analyzed using generalized linear models (CO_2 effect size – temperature + rainfall; family = Gaussian; link = identity). The partial R^2 of each climatic factor was obtained using the rsq. partial function with the rsq package in the R version 3.2.2.

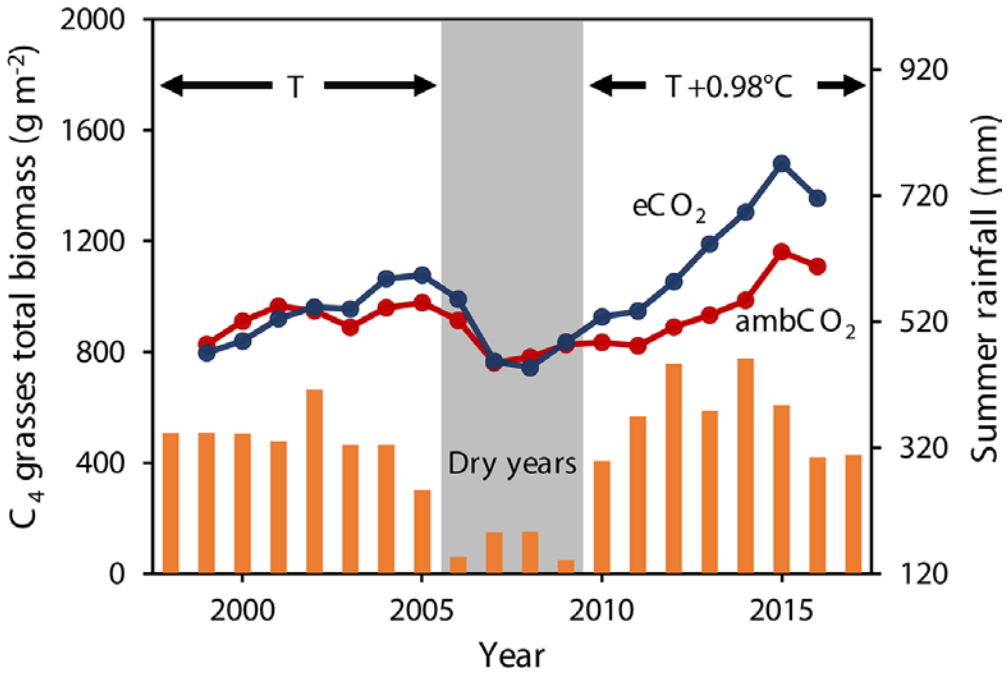


Fig. 2. Inter-annual trajectories of C_4 total biomass at the ambient (red) and elevated CO_2 (blue) levels and summer rainfall (orange). The Biomass data from the measurements of Reich *et al*. (*1*) are shown as 3-yr moving averages centered over the middle of each 3-yr group. The Cedar Creek weather data are from Fort Snelling near the Saint Paul International Airport.

AAAS analyst assists a human rights organization in gathering data during an exhumation.

PHOTO: JONATHAN DRAKE/AAAS

Emerging scientific technologies help defend human rights

Scientists leverage advancing tools to gather evidence and expand the capacity of human rights groups

By **Anne Q. Hoy**

Against a backdrop of summer heat and a constant roar of distant howler monkeys, a scientific analyst piloted a drone to collect data from a hillside in northern Guatemala. At his side, anthropologists affiliated with a regional human rights group painstakingly cleared soil and roots from human remains in a mass grave.

"Remains contorted, overlapping, interlaced, a cruel, tragic mashup of Hieronymus Bosch and H.R. Giger," noted Jonathan Drake, senior program associate of the American Association for the Advancement of Science's Geospatial Technologies Project, summoning images from 15th- and 20th-century artists to describe the nightmarish remnants of an atrocity estimated to have occurred sometime after 1980, during Guatemala's lengthy civil war.

Multiple skeletons were exhumed. Clothing with burnt edges stuck to the bones of some. A blindfold encircled a skull. Leg bones bore evidence of a child. Those were among the observations Drake shared after maneuvering a commercial-grade drone at specific angles for optimal data collection and documentation.

Drake and the representatives of the Guatemalan Forensic Anthropology Foundation (FAFG) each used drones to collect several hundred overlapping photographs documenting the location of the mass grave and to record each step of the exhumation process.

AAAS and FAFG have worked together on six projects since AAAS first offered its scientific expertise and technical and analytical skills to assist in the search and exhumation of mass graves in Guatemala. The collaborations went from "proof of concept, to training on data processing, to capacity building, to implementation in a real-world context," said Drake.

> ***"All of these technologies are sort of converging in a way that is really very effective…"***
>
> **Jonathan Drake,** AAAS

AAAS's use of drones for geospatial documentation in Guatemala grew out of an earlier alliance with EQUITAS, an independent, nonprofit organization of scientific investigators of human rights violations in Bogotá, Colombia. The organization sought help locating remains of victims of forced disappearances during Colombia's decades-long war.

EQUITAS asked for help with a suspected mass grave in a cloud-covered, mountainous region likely inaccessible to remote sensing satellites. While remote sensing collects data without requiring a visit and some types can penetrate cloud cover, high-resolution optical imagery offers better chances of capturing data from locations blanketed by clouds.

It was in Colombia where AAAS first used and tested a drone to collect data from the suspected mass grave site. The site turned out to have no graves, but the data gathered allowed AAAS to produce a comprehensive map of the canopied jungle, providing EQUITAS with its first detailed survey of the region and a tool for planning future digs, Drake said.

The AAAS geospatial project has evolved since its establishment in 2005 by putting technologies to work as they emerge. Drake and

earlier project participants have incorporated emerging technological advances into scientific collaborations with human rights practitioners around the globe.

Lars Bromley, a former AAAS staff member and participant in the geospatial project's collaborations with human rights organizations, took on one of AAAS's initial initiatives. In an alliance with Amnesty International, Bromley used available technology and applied his background in geospatial analysis to scrutinize satellite imagery of four communities in Zimbabwe.

The government asserted that homes in the communities were illegal and launched a demolition campaign that organizations considered to be human rights violations. AAAS analyzed the destruction and removal of more than 5000 homes captured by data that were provided to human rights lawyers in Zimbabwe.

Leveraging drones to gather evidence that can later be used in courts is just one of the emerging technologies AAAS is now testing and using in its alliances with human rights organizations, particularly those that lack significant resources.

Scientific advances in machine learning and artificial intelligence are being studied and, in some cases, tested for their ability to capture and analyze copious amounts of data. More recently, blockchain, microdrones, and nanodrones are being recognized as technologies that could assist in human rights investigations.

Blockchain technology, Drake noted, could be useful in establishing a chain of custody for scientific evidence, as well as providing verifiable provenance of digital data necessary in human rights cases.

"They all fit together. Artificial intelligence fits with satellite imagery in terms of classifying images, and it is also critical in linking hundreds and hundreds of images collected by a drone into a 3-D model," Drake explained. "You can't do that sort of thing without machine learning."

Insights gleaned from the drone test in Colombia were put to work in Guatemala. The drone's effectiveness in data gathering opened the door to AAAS's eventual ability to produce 3-D models from high-resolution images, assign global coordinates, map large areas, and conduct such analyses relatively quickly.

During AAAS's first collaboration with FAFG in Guatemala, Drake used a drone to collect extensive photographic data and later transformed the photographs into a 3-D representation using a computer-assisted process known as photogrammetry. The process permits photographs of a single skeleton to be merged with images of subsequent exhumations, producing a dynamic model of a mass grave's human remains and contents. The model offers an integrated 3-D image of all or any part of the scene.

Data collection by drones can show the precise orientation of one skeleton in relation to another—such as one facing another—a heartbreaking and rarely noticed view at a site during an exhumation since skeletons are most often removed separately.

Removal is a laborious process that can consume a day for a single extraction. Precise documentation of each exhumation is required. The scientific methods employed to collect data also must be described.

"All of these technologies are sort of converging in a way that is really very effective and has the potential of being really effective in promoting human rights around the world," noted Drake.

The AAAS Scientific Responsibility, Human Rights and Law Program released a report in July examining the lessons learned by the organization in providing "geospatial analysis in a human rights context."

The report includes reviews of dozens of legal cases in which geospatial technology provided evidence used in international criminal prosecutions brought before the International Criminal Court; conflict-specific courts in Yugoslavia, Sierra Leone, and Cambodia; and human rights courts in Europe and Central and South America.

Among its recommendations, the report calls on the judicial branch to appoint independent scientific advisers for cases that involve highly technical or specialized research; geospatial analysts to document where and how their findings were collected and explain the scientific methods used; and human rights groups to use independent, experienced analysts. It urges government agencies to safeguard metadata and preserve the chain of custody of data, which are vital to the legal community. It also advises private-sector satellite data providers to protect their images from manipulation.

Bromley, who now serves as a principal analyst and research adviser for a United Nations institute and a satellite operations program, expressed surprise about how often fundamental education about geospatial technologies continues to be required by human rights organizations. Over time, Bromley said, he has "come to accept that capacity building is a slow and never-ending process." AAAS integrates extensive training in the collection of evidence from emerging technologies in its work with human rights groups.

"Now, 10 years later, these technologies are firmly embedded in the human rights landscape and are in relatively common use," said Bromley, referring to the geospatial technologies that were actively used when the program began and the pathway AAAS continues to pursue with emerging technologies.

"We really did manage to take these from an infant technology to something useful and valued," he added.

AAAS annual election: Preliminary announcement

The 2018 AAAS election of general and section officers is scheduled to begin in October. All members will receive a ballot for election of the president-elect, members of the Board of Directors, and members of the Committee on Nominations. Additionally, members registered in sections (up to three) will receive ballots for the specified section elections. Biographical information for the candidates will be provided along with ballots. The general election slate is listed below. The list of section candidates can be viewed at www.aaas.org/annual-election.

Notice to our international members:
In an effort to conserve resources, AAAS will be sending electronic election ballots to our non-U.S.-based members. In order to ensure you receive your ballot, please make sure your email is up-to-date with AAAS by logging on to www.aaas.org. 1) Click on "Member Login" (if you have not yet created an account, you will be prompted to do so); 2) After you log in, click on the red "My Profile" button in the upper right-hand corner of the page; 3) Click on "Edit My Contact Information" in the left-hand side bar; 4) Update your email and click on the "Save" button.

If you would like to request a special paper ballot, please send an email with your name and address with your request to elections@aaas.org.

General Election Slate

President-Elect:
Jared D. Cohon, Carnegie Mellon University; Claire M. Fraser, University of Maryland School of Medicine

Board of Directors:
Ann Bostrom, University of Washington; Maria Klawe, Harvey Mudd College; Peter R. MacLeish, Morehouse School of Medicine; Griffin P. Rodgers, National Institute of Diabetes and Digestive and Kidney Diseases

Federal research funding aims to ease societal challenges

A key goal is to spread access to scientific excellence and federal research funding nationwide

By **Anne Q. Hoy**

Half of all federal research funding in the United States goes to recipients based in six states and the District of Columbia, leaving the other 50% of funding split among those in the remaining 44 states, the National Science Foundation's annual Survey of Federal Funds for Research and Development shows.

The concentration of geographic funding to primary recipients in California, Maryland, Massachusetts, New York, Texas, Virginia, and the District of Columbia was cited by Kei Koizumi, a senior science policy adviser for the American Association for the Advancement of Science, during a panel presentation exploring the evolution of federal research funding in the United States, France, Japan, and other Organization for Economic Co-operation and Development member countries.

"We have abundant research that shows the majority of U.S. students go to school in their home states," said Koizumi during the presentation on 13 July. "So, if research funding is not happening in their state, then they are missing out on an opportunity to participate in our science and technology enterprise." Concentrated research funding distribution levels, he added, deny states economic development, growth, and jobs that the system develops.

The presentation was held at the biennial EuroScience Open Forum 2018 in Toulouse, France, a gathering of more than 3000 scientists, innovators, policy-makers, and business representatives 9–14 July to discuss scientific research, innovation, and science policy issues. AAAS CEO Rush Holt moderated a session on science diplomacy and AAAS staff highlighted activities of AAAS's Cambridge, U.K., office and the online, global news service EurekAlert!

Kei Koizumi (right) fields questions about scientific research funding mechanisms during a break at biennial ESOF in Toulouse, France.

PHOTO: AAAS

In recent years, geographic funding concentration levels in the United States have remained fairly consistent. Yet the country's leading federal research funding institutions have been testing experimental programs to spread federal research funding more equally across the country to address economic and social inequities.

Science and engineering research funding programs are searching for ways to provide university students in every state an opportunity to search for knowledge, extend scientific excellence, and, in so doing, ensure that the system tackles larger societal issues, said Koizumi.

"It is important, both politically and socially, to address inequalities on multiple dimensions, and science funding is not exempt from that imperative," Koizumi said. "We have seen that competitive research funding mechanisms, left to their own devices, can result in inequalities. They can perpetuate other inequalities that exist in society."

Adjusting the funding system to support multiple societal objectives as it also seeks to produce scientific excellence is not easy, noted Koizumi in his session on "Supporting long-term research in a world of sudden change: The evolution of research and funding in current financial and political contexts."

"It is, of course, a common insight now to see that the U.S. scientific workforce does not look like the U.S. population, and so diversity and inclusion are important considerations for how we support the U.S. scientific enterprise," added Koizumi.

To overcome impediments, the National Science Foundation, the Departments of Energy and Agriculture, and NASA have established programs under the Established Program to Stimulate Competitive Research, or EPSCoR, which was established in 1978 to enable universities across the country to compete for federal research funding. The National Institutes of Health began a similar program 25 years ago known as Institutional Development Awards, or IDeA. Both programs continue to grow.

For two decades, AAAS has supported more than 30 states by providing 151 assessments of more than $1.2 billion in research projects funded by NSF's EPSCoR and NIH's biomedical research IDeA programs. AAAS's Research Competitiveness Program, or RCP, conducts the work and provides peer-to-peer insights from independent U.S. experts and, more recently, quantitative evaluations of projects. In addition, RCP is now working on a NSF study to devise a framework for measuring academic research excellence and competitiveness for EPSCoR and potentially other NSF programs.

"Our programs strengthening STEM ecosystems within the U.S. have parallels to national STEM initiatives in other countries," said Charles Dunlap, RCP's program director. "While we continue to support institutions in the U.S., institutions abroad are increasingly contacting AAAS for support as well."

A range of other collaboration models also have emerged. One is organized around national objectives such as improving health care, addressing climate change, or expanding manufacturing opportunities. Other competitive funding initiatives promote cross-sector collaborations that align private businesses with research universities and federal research laboratories and international collaborations that match scientific research groups with global partners to pursue shared scientific goals.

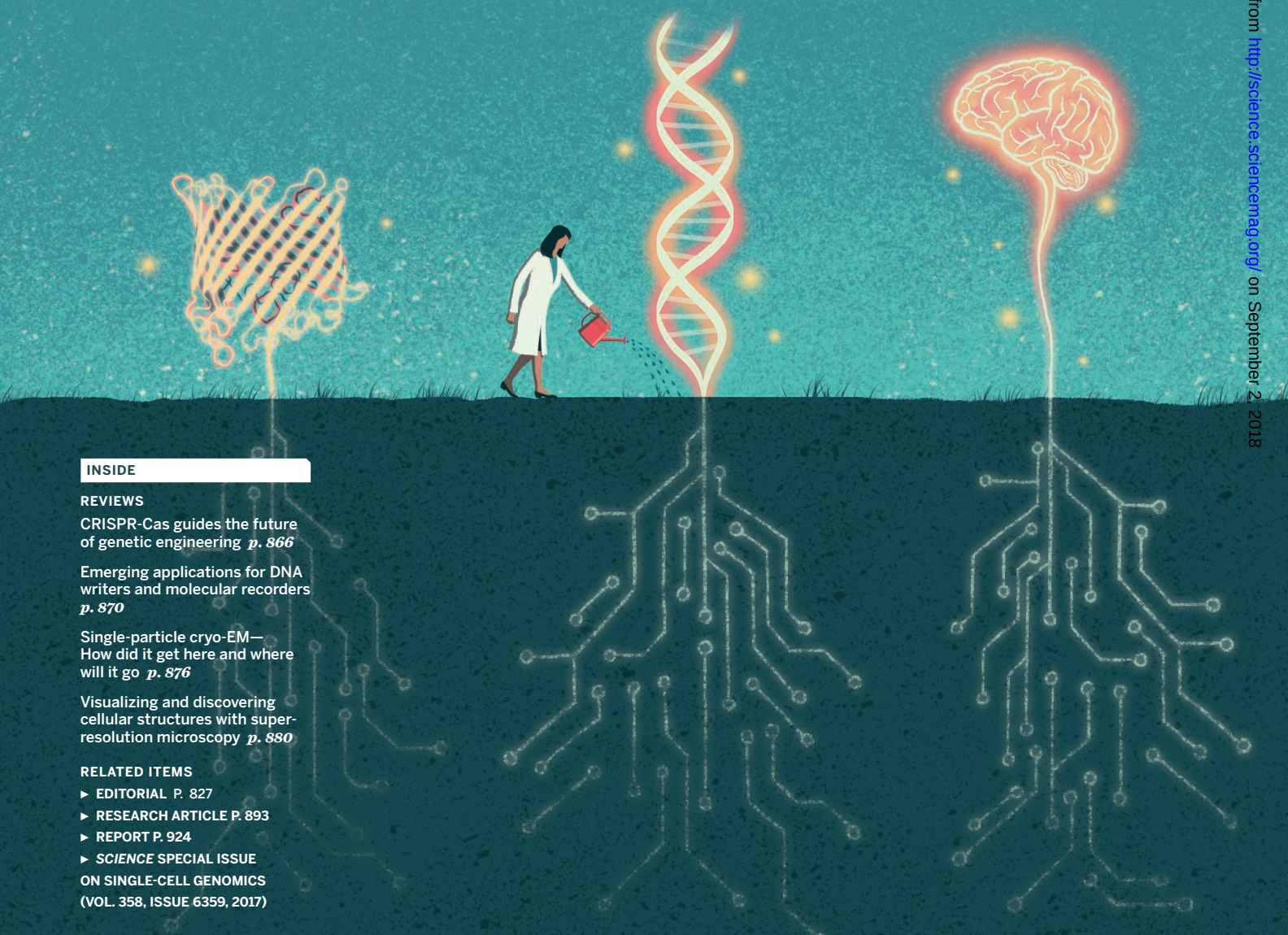
Competitive research funding endeavors also focus on "high-risk, high-reward research, or potentially transformative research," said Koizumi, in a drive to offset the tendency among experienced researchers in a highly competitive funding arena to pitch less risky and shorter-term proposals.

Since World War II, the federal research funding system has helped the United States become the world's leader in science and engineering innovation. With time, though, system flaws have emerged, including stress on success rates due to the growing number of research proposals that fall short of funding, a development that raises the cost of the scientific review process. The system also has failed to expand the ranks of underrepresented minorities and women in the scientific enterprise.

In addressing the state of today's competitive research funding system, Koizumi said, "With careful attention we can use competitive research funding to attempt to address the challenges of inequalities both inside the scientific enterprise but also with our society at large."

POWER COUPLE: SCIENCE AND TECHNOLOGY

By **Steve Mao** *and* **Valda Vinson**



INSIDE

REVIEWS

RELATED ITEMS

- EDITORIAL P. 827
- RESEARCH ARTICLE P. 893
- REPORT P. 924
- *SCIENCE* SPECIAL ISSUE ON SINGLE-CELL GENOMICS (VOL. 358, ISSUE 6359, 2017)

Every revolution in technology is followed by an explosion of new scientific knowledge. Biology is no exception. Van Leeuwenhoek's microscope jump-started microbiology, polymerase chain reaction revolutionized molecular biology (see Editorial in this issue), and single-cell imaging and sequencing approaches remarkably advanced immunology, cancer research, developmental biology, and beyond.

Recent years have witnessed disruptive innovations in biotechnology. Researchers have never been equipped with more powerful tools to probe biology. Breakthroughs in electron microscopy allow biomolecular complexes to be visualized with high resolution, offering insight into the workings of these molecular machines. Ingenious methods to break the diffraction limit in microscopy enable single molecules to be observed and tracked in single cells. Not only do we see cells in ever-greater detail, but with CRISPR-mediated gene editing techniques, biologists can precisely and easily manipulate cellular genomes of diverse organisms. As we gain understanding of biological networks, tools including those based on CRISPR give us the ability to record biological events, to detect and treat disease, and to engineer plants with new traits and greater productivity.

When biologists encounter problems that cannot be solved by currently available technologies, shared creativity between researchers drives the development of better and smarter tools. New technologies, in turn, push the frontier of biology. This synergy has been moving our society and humanity forward, and the advent of artificial intelligence is likely to speed up this cycle of discovery.

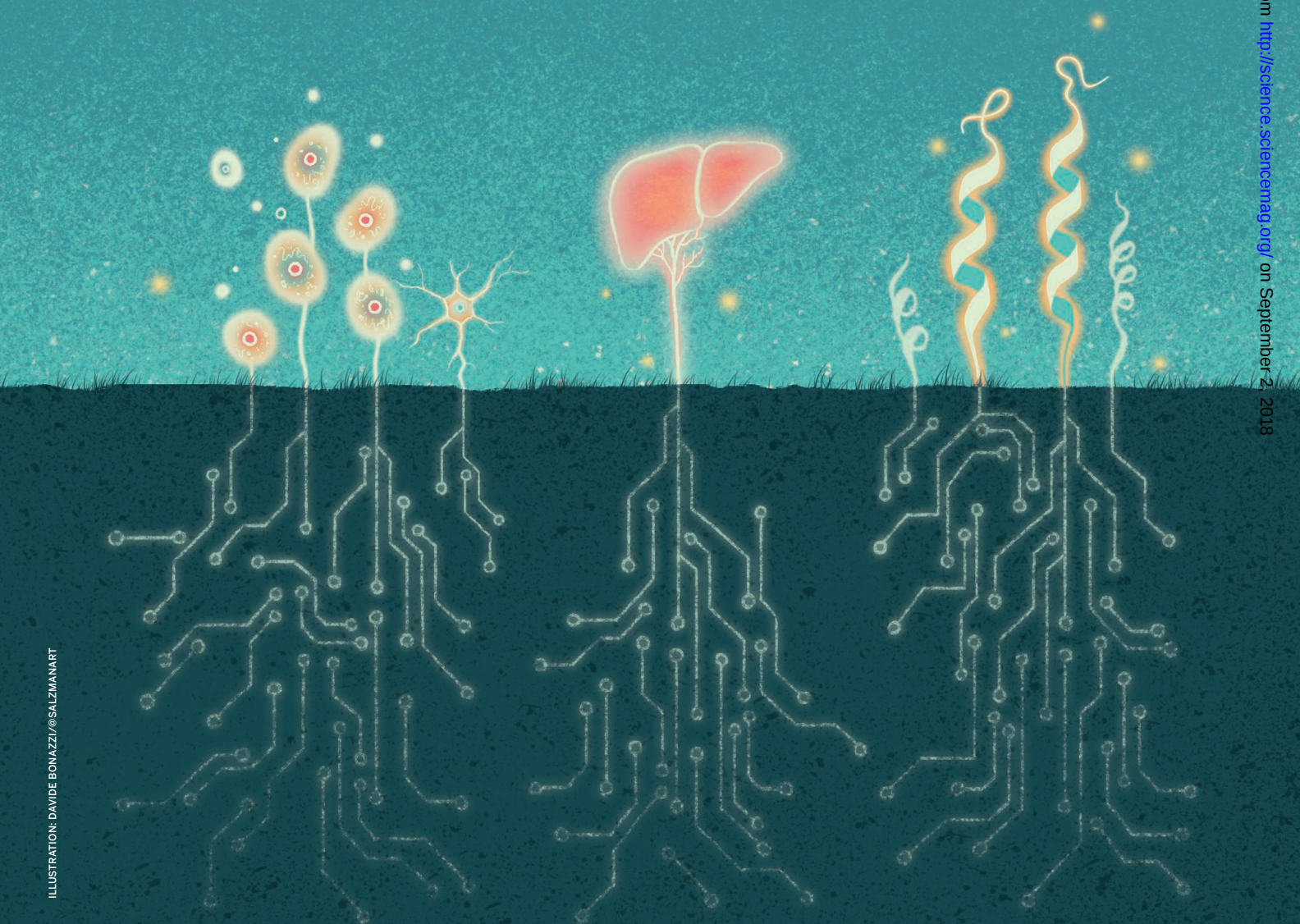


ILLUSTRATION: DAVIDE BONAZZI/@SALZMANART

REVIEW

CRISPR-Cas guides the future of genetic engineering

Gavin J. Knott[1] and Jennifer A. Doudna[1,2,3,4,5]*

The diversity, modularity, and efficacy of CRISPR-Cas systems are driving a biotechnological revolution. RNA-guided Cas enzymes have been adopted as tools to manipulate the genomes of cultured cells, animals, and plants, accelerating the pace of fundamental research and enabling clinical and agricultural breakthroughs. We describe the basic mechanisms that set the CRISPR-Cas toolkit apart from other programmable gene-editing technologies, highlighting the diverse and naturally evolved systems now functionalized as biotechnologies. We discuss the rapidly evolving landscape of CRISPR-Cas applications, from gene editing to transcriptional regulation, imaging, and diagnostics. Continuing functional dissection and an expanding landscape of applications position CRISPR-Cas tools at the cutting edge of nucleic acid manipulation that is rewriting biology.

Researchers have long pursued a means of efficiently manipulating DNA and RNA to tailor genes and their regulation. Genetic perturbation enables scientists to probe gene function or correct mutations but is often intractable due to a technical challenge: site-specific nucleic acid targeting. Targeted gene editing has been achieved by induced double-stranded DNA (dsDNA) breaks in eukaryotic chromosomes (*1*), but with challenging technologies based on engineering direct protein-DNA recognition. The history recounting the discovery, development, and application of such engineered nucleic acid binding proteins—including zinc fingers, TALENS, and meganucleases—is rich in noteworthy scientific feats (*2*). Over the past 6 years, however, transformative discoveries have shaped the CRISPR (clustered regularly interspaced short palindromic repeats) Cas (CRISPR-associated) toolbox for genetic manipulation on the basis of simpler RNA-guided DNA recognition. This toolbox now provides important scientific opportunities for curing genetic diseases and engineering desirable genetic traits, as well as new approaches to live-cell imaging, high-throughput functional genomic screens, and point-of-care diagnostics. In this Review, we summarize the basic mechanisms of RNA-guided single-component CRISPR-Cas systems and their general applications. The basis for the CRISPR revolution goes beyond inherent programmability, lending itself to the naturally evolved diversity of systems that extend CRISPR-based technology beyond precision gene editing. To capture the broadened landscape of Cas applications and their impact as a force for revolution in molecular biology, where appropriate, we refer readers to recent reviews for a more detailed discussion.

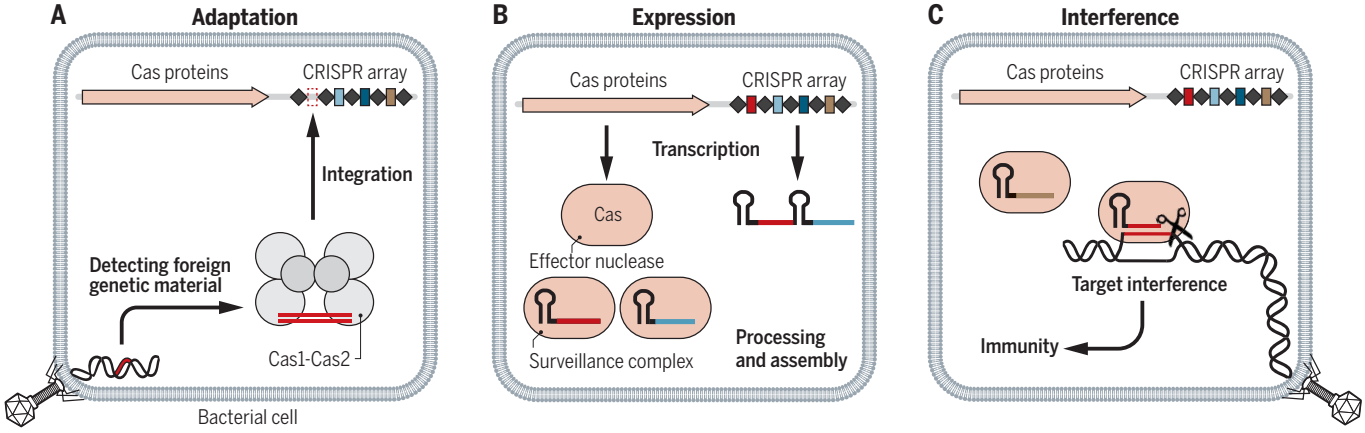
Diverse RNA-programmable CRISPR-Cas enzymes

CRISPR-Cas systems provide microbes with RNA-guided adaptive immunity to foreign genetic elements by directing nucleases to bind and cut specific nucleic acid sequences (*3*–*5*) (Fig. 1). Through a process termed adaptation, microbes capture snippets of foreign genetic elements and incorporate them into their genomic CRISPR array. Transcription of CRISPR arrays creates CRISPR RNAs (crRNAs) that bind to Cas nucleases and provide specificity by base-pairing with target nucleic acids (*4*, *5*). Among the diverse naturally evolved CRISPR-Cas systems, those designated class 2 constitute a single large RNA-guided Cas nuclease that mediates target interference or cleavage [reviewed in (*6*)].

The class 2 type II DNA-targeting endonuclease Cas9 (the first Cas effector to be harnessed for genome engineering) has several properties that ensure precise and efficient editing (Box 1A and Fig. 2A). Cas9 assembles with only the intended guide RNA through specific recognition of the crRNA and its interaction with a trans-activating crRNA (tracrRNA). In addition, the dual crRNA-tracrRNA can be fused into a chimeric single-guide RNA (sgRNA), thereby creating a two-component system composed of Cas9 and its sgRNA (*7*). Finally, stable binding to target DNA adjacent to a specific motif [protospacer adjacent motif (PAM) (*8*, *9*)] with the correct nucleotide sequence acts as a switch, triggering Cas9 to introduce a dsDNA break (*7*). Scientists worldwide have deployed Cas9 because of this switchable nuclease activity and the ease of redirecting the enzyme by altering the sgRNA-targeting region (or spacer sequence) (*10*–*12*).

Although *Streptococcus pyogenes* Cas 9 (SpCas9) remains the most commonly used Cas effector, it is not alone in the evolutionary arms race against mobile genetic elements. Bacteria and archaea have evolved numerous functionally distinct

[1]Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA. [2]Department of Chemistry, University of California, Berkeley, CA, USA. [3]Howard Hughes Medical Institute, University of California, Berkeley, CA, USA. [4]Innovative Genomics Institute, University of California, Berkeley, CA, USA. [5]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA.
*Corresponding author. Email: doudna@berkeley.edu



CREDIT: ADAPTED BY V. ALTOUNIAN/*SCIENCE* FROM THE DOUDNA LAB

Fig. 1. CRISPR-Cas adaptive immunity. (**A**) Foreign genetic elements are acquired by Cas1-Cas2 and integrated into the CRISPR array in a process broadly termed adaptation. (**B**) The CRISPR array and associated Cas proteins are expressed. The CRISPR array is processed and Cas effector nucleases associate with a crRNA to form a surveillance complex. (**C**) The Cas effector nucleases target foreign genetic elements complementary to their crRNA, leading to target interference and immunity.

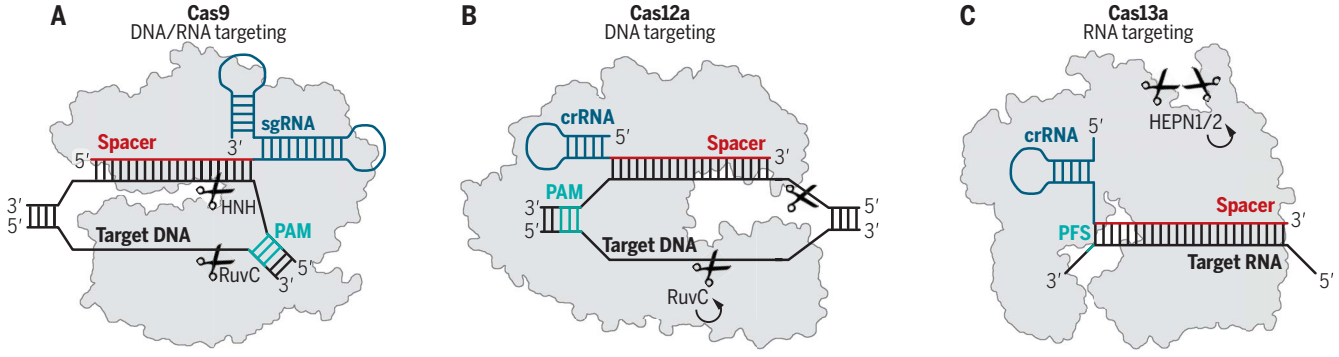


Fig. 2. Schematic of class 2 CRISPR-Cas systems. (**A**) Class 2 type II CRISPR-Cas9 shown schematically with an sgRNA (blue) encoding a spacer (red) bound to a target dsDNA (black) proximal to a PAM (teal). Correct base-pairing activates the HNH and RuvC nuclease domains, cleaving both strands (scissors). (**B**) Class 2 type V CRISPR-Cas12a shown schematically with a crRNA (blue) encoding a spacer (red) bound to a complementary dsDNA target (black) proximal to a PAM (teal). Correct base-pairing activates the RuvC nuclease, cleaving both strands (scissors) with multiple-turnover general ssDNase activity (arrow). (**C**) Class 2 type VI CRISPR-Cas13a shown schematically with a crRNA (blue) encoding a spacer (red) bound to a complementary RNA target (black). Correct base-pairing activates HEPN nuclease general ssRNase activity (arrow). See also Box 1.

CRISPR-Cas systems that maintain the programmable characteristics key to the success of SpCas9. Scientists have tapped the evolutionary diversity of type II systems, incorporating divergent homologs and engineered variants of SpCas9 into an arsenal of genome editors. At the tail end of 2015, class 2 systems expanded to include a number of candidate systems, which were later designated type V CRISPR-Cas12a (formerly Cpf1) and type VI CRISPR-Cas13a (formerly C2c2) (Box 1, B and C, and Fig. 2, B and C). Today, SpCas9 shares the spotlight with a diversity of Cas9 homologs, DNA-targeting Cas12, and RNA-targeting Cas13, all of which are programmable RNA-guided nucleases [reviewed in (*13*)]. It is this inherent programmability present in a diversity of naturally evolved systems that extends CRISPR-Cas applicability beyond precision genome editing.

Applications of Cas-mediated genome editing

Although the scope of Cas application has broadened, precision genome engineering remains at the forefront of the CRISPR revolution. Cas9 and Cas12a are RNA-guided nucleases that can induce genome editing by triggering dsDNA break repair at a specific site (Fig. 3). Editing occurs after cellular DNA repair pathways resolve the break by nonhomologous end joining (NHEJ), which can introduce small insertions or deletions, or by homology-directed repair (HDR) with a donor sequence at the site of the dsDNA break [reviewed in (*14*)].

As tools for precision genome engineering, Cas9 and Cas12a work in a wide range of cell types and organisms. Cas-mediated gene editing has prompted genome-wide screens to probe basic biological function, in addition to identification and validation of potential drug targets in complex heritable diseases [reviewed in (*15*)]. Agricultural applications of Cas-nucleases [reviewed in (*16*)] have produced modified crops that now have a streamlined path to the market (*17*). In the clinic, Cas-nucleases allow diseases with a known genetic basis to be treated and, in an era of high-throughput DNA sequencing, personalized to a patient's disease etiology. Examples include gene editing to correct mutations or induce skipping of defective exons in Duchenne muscular dystrophy (DMD), strategies that are already showing efficacy in animal models (*18*, *19*). Cas9 has also been used to inactivate defective genes that underlie neurological diseases, including amyotrophic lateral sclerosis (*20*) and Huntington's disease (*21*). Scientists have used Cas9 to eliminate an entire chromosome in aneuploid human pluripotent stem cells (*22*), to inactivate an endogenous retrovirus in pigs (*23*), and to engineer T cells as a prelude to developing advanced immunotherapies to target cancers (*24*). Furthermore, Cas9 has allowed targeting of the genetic basis for sickle cell disease (*25*) such that there are now established protocols for the correction of genetic defects in this cell type (*26*). Beyond such somatic cell editing, the potential to correct genetic mutations in human embryos is on the horizon, raising societal and ethical questions about creating heritable changes in the human germline (*27*).

However, it is important to note that precision editing remains challenging, particularly given competing repair outcomes (such as NHEJ) restraining the efficiency of more desirable HDR repair outcomes (*28*). An alternative approach utilizes Cas effectors fused to base editors, limiting unintended edits and eliminating the requirement for repair templates. Distinct from DNA cleavage and repair, nickase Cas9 (nCas9)–mediated base editing carries a single-base editor to a target locus facilitating base conversion without dsDNA cleavage (Fig. 3) [reviewed in (*29*)]. Recently, the toolbox of base editors expanded to include a laboratory-evolved deaminase enabling nCas9-mediated single-base editing to catalyze A-T to G-C transitions (*30*). The existing Cas9-mediated base editors now enable researchers to create any of the four possible transition mutations at a specific genomic locus (*30*–*33*). Although single-base editors provide the potential to correct disease-causing mutations without inducing a dsDNA break, the technology requires further development to limit off-target editing. Looking forward, the next generation of Cas-mediated genome editors will likely include base editors, ideally with base-editing activity conformationally coupled to Cas9 target DNA binding.

Box 1. Crash course in class 2 CRISPR-Cas systems.

(**A**) Class 2 type II CRISPR-Cas systems target dsDNA using the effector nuclease Cas9, a crRNA, and tracrRNA [crRNA-tracrRNA fusion creates the sgRNA (*7*)]. Cas9 binds to a DNA sequence complementary to the sgRNA spacer adjacent to a PAM. Cas9 senses correct base-pairing, thus activating its RuvC and HNH nucleases to cleave the nontarget and target DNA strands. (**B**) Class 2 type V CRISPR-Cas systems, specifically subtype Cas12a, target ssDNA and dsDNA using the effector nuclease Cas12a (formerly Cpf1) guided by a single crRNA (*75*). Cas12a binds to a DNA sequence complementary to the crRNA spacer, adjacent to a PAM for dsDNA targets. Cas12a senses correct base-pairing to activate its RuvC nuclease for general ssDNase activity, cleaving the nontarget and target DNA strands and trans-ssDNA substrates. (**C**) Class 2 Type VI CRISPR-Cas systems, specifically subtype Cas13a, target ssRNA using the effector nuclease Cas13a (formerly C2c2) guided by a single crRNA (*46*, *47*). Cas13a binds to a ssRNA sequence complementary to the crRNA spacer. Cas13a senses correct base-pairing to activate the HEPN nuclease for general ssRNase activity. See also Fig. 2.

Transcriptional regulation with dCas9

Cas9 has proven to be a modular platform with functionally distinct DNA binding and nuclease activities. Decoupling DNA binding from the enzymatic

activity of Cas9 by mutating the nuclease domains creates catalytically deficient Cas9 (dCas9), a functional scaffold for recruiting protein or RNA components to a specific locus to perturb transcription without permanently altering DNA [reviewed in (*34*, *35*)] (Fig. 3). The use of dCas9 has revolutionized functional genetic screening by enabling specific, rapid, and multiplexed genetic knockdowns in a range of cell types, including immune cells and neurons (*36*, *37*). These advances with dCas9 highlight the practicality of genomic perturbation without the risk of DNA damage, an attribute that has motivated studies in model systems to drive therapeutic development. For example, dCas9 fused to TET1, a demethylase, targeted to the dysregulated *FMR1* locus reversed the phenotype of fragile X syndrome in neurons and mouse models (*38*). Gain-of-function studies have successfully implemented a modified dCas9 target gene activation system to treat type 1 diabetes, acute kidney injury, and murine muscular dystrophy (*39*). The ability to conduct suppressor screens and synthetic lethal screens in virtually any cell type offers the potential to discover gene functions, effector interactions, and pathways at a pace never before possible. However, challenges remain: dCas9-effector fusions have complex off-target effects due to the fused catalytic domains targeting neighboring or even unrelated loci. Additionally, unpredictable locus-specific effects on chromatin, and in turn transcription, can confound analysis and obscure causality (*40*). Future work should appropriately control for unpredictable locus-specific effects with systematic validation and should aim to further improve specificity.

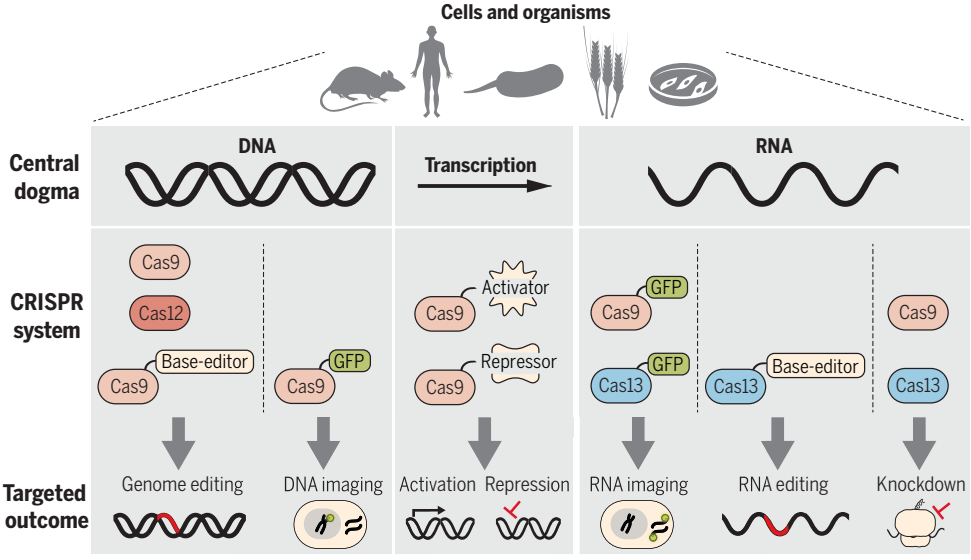


Fig. 3. CRISPR-Cas systems allow genetic manipulation across the central dogma. From left to right, Cas9 and Cas12a are used for inducing dsDNA breaks for genome editing. nCas9 can be fused to base editors to modify nucleotides in dsDNA for genome editing without introducing a dsDNA break. dCas9 can be fused to transcriptional activators, repressors, or epigenetic modifiers to regulate transcription. Cas9 and Cas13a can be used for targeted RNA interference. Cas13a fused to base editors can be used to modify nucleotides in RNA. dCas9 or dCas13a can be fused to green fluorescent protein (GFP) to visualize DNA or RNA.

Posttranscriptional engineering with RNA-targeting Cas

As an alternative to permanent genetic alteration, Cas effectors can be applied to transiently perturb the transcriptome through direct RNA targeting (Fig. 3). Engineering SpCas9 to create a programmable RNA-targeting system with the use of a PAM-presenting oligonucleotide (PAMmer) (*41*) ushered in applications for RNA-targeting with Cas9 (RCas9). Targeting RCas9 to RNA can eliminate pathogenic RNA foci, rescue mRNA splicing defects, or attenuate polyQ-containing protein production from RNAs with trinucleotide CAG repeats (*42*). To date, the arsenal of RNA-targeting Cas9s has expanded to include related homologs with programmable RNA-targeting activity that is PAMmer independent (*43*–*45*). Given its success, Cas9 lends itself to further development for posttranscriptional engineering, such as fusions to single-base RNA modifiers to achieve site-specific RNA modifications.

Cas13 has also emerged as a highly versatile tool for RNA targeting. Reconstituting Cas13a in *Escherichia coli* (*46*) and in vitro (*46*, *47*) established type VI systems as an RNA-guided general ribonuclease (RNase). Cas13a has been employed in vivo as a tool for specific knockdown in mammalian (*48*) and plant cells (*49*). Evolutionarily and functionally related to Cas13a, Cas13b enzymes have programmable RNase activity that has been functionalized for RNA interference and RNA editing in mammalian cells (*50*, *51*) (Fig. 3). More recently, CRISPR-Cas13d was identified (*52*, *53*) and reconstituted for modulating splicing in vivo (*52*). RNA-targeting systems such as Cas9 and Cas13 support targeted RNA-guided research in addition to clinical applications akin to antisense oligonucleotide therapies for the treatment of acute non-Mendelian pathologies, avoiding the risks associated with permanent genetic modification. However, future studies are needed to determine how RNA-targeting Cas-effectors interface with a structured or protein-occluded RNA landscape and how trans-RNA cleavage by Cas13 is attenuated in vivo.

Programmable nucleic acid imaging

Correct spatiotemporal localization is critical to the function of specific genomic loci, mRNAs, and noncoding RNAs, with dysregulated molecular localization strongly implicated in disease. Current technologies for live-cell imaging of genomic loci or nascent RNA are limited by the need for protein engineering or the introduction of targetable sequences into a transcript of interest. Leveraging dCas9, researchers have imaged repetitive genomic loci in live cells using dCas9 fused to fluorescent reporters [reviewed in (*54*)] (Fig. 3). Exploiting the stringency of dCas9 PAM recognition, a method was developed that allows high-resolution single-nucleotide polymorphism CRISPR live-cell imaging of DNA loci (*55*). However, widespread use of dCas9 to study localization of specific genomic loci has been limited by a low signal-to-noise ratio at nonrepetitive genomic sequences. One strategy to overcome inadequate signal-to-noise ratio involves appending multiple bacteriophage MS2 operator RNA hairpins (MS2 motifs) to the sgRNA (*56*). Tandem MS2 motifs act as high-affinity binding sites recruiting multiple MS2 motif binding proteins fused to a fluorescent reporter, effectively amplifying the signal to allow detection of a single dCas9-sgRNA bound to DNA in vivo (*56*). Leveraging RNA-targeting RCas9 has allowed researchers to track RNA in live cells (*57*), thus making it possible to visualize clinically relevant repeat expansion-containing transcripts (*42*) (Fig. 3). With the growth of the RNA-guided RNA-targeting toolbox, RNA imaging tools now also include catalytically deficient Cas13a (dCas13a) (*48*). Though both RCas9 and dCas13a show promise when targeted to repetitive elements, further development is required to realize either platform as a reliable tool for low-abundance transcripts lacking in repetitive sequences. Furthermore, it is unclear whether localizing large exogenous ribonucleoproteins (RNPs) to transcripts might perturb cellular processes.

Nucleic acid detection and diagnostics

The RNA-guided nuclease activities of Cas13a and Cas12a have driven development of innovative tools for nucleic acid detection. For both Cas13 and Cas12a, which are functionally distinct from Cas9, a target nucleic acid (or activator RNA or DNA) activates general multiple-turnover nuclease activity through correct base-pairing to the guide RNA (Box 1, B and C, and Fig. 2, B and C). Leveraging this switchable nuclease activity, Cas13a was first functionalized as a tool for detecting target RNA transcripts of interest in a

pool of RNA by detecting its RNase activity (*47*). Expanding on this work, SHERLOCK (Specific High-Sensitivity Enzymatic Reporter UnLOCKing) was developed as a platform incorporating pre-amplification of the input material to create a tractable paper-based assay with improved sensitivity (*58*). Biochemical dissection identified that divergent Cas13a homologs have discrete crRNA and substrate preferences enabling orthogonal use to simultaneously detect two different transcripts (*59*). Similar dissection of Cas13b homologs revealed substrate preferences that supported expansion of the SHERLOCK platform, now SHERLOCKv2, to simultaneously detect dengue and Zika virus single-stranded RNA (ssRNA) (*60*) in a readily deployable format (*61*).

Analogous to Cas13, Cas12a has evolved a functionally convergent switchable general nuclease that targets ssDNA (*62*). Exploiting this activity, DETECTR (DNA endonuclease-targeted CRISPR trans reporter) was developed as a CRISPR-based DNA detection and diagnostic platform (*62*). Coupled with isothermal pre-amplification, DETECTR was shown to rapidly and accurately detect clinically relevant types of human papillomavirus (*62*). SHERLOCKv2 integrated Cas12a-based DNA-targeting to detect either *Pseudomonas aeruginosa* or *Staphylococcus aureus* DNA targets in parallel to detection of RNA targets by Cas13a and Cas13b (*60*). Akin to Cas13, tapping the functional diversity of Cas12 systems may yield functional variants that enable further development of DNA-based diagnostics. Looking ahead, the detection of a specific transcript using CRISPR-Cas is rapid and readily adaptable in the clinic, setting the stage for inexpensive point-of-care diagnostics.

Specificity and delivery of CRISPR-Cas

Unintended binding, modification, and cleavage of nucleic acids pose a challenge to all technologies for genetic manipulation. Compared with the side effects caused by off-target interactions of small-molecule drugs or antibody therapeutics, off-target Cas nuclease activity is especially deleterious because of the permanence of genome editing. Indeed, this further reinforces the necessity for nuclease specificity and targeted delivery. Researchers have made considerable advances in evolving and engineering Cas enzymes (*63*, *64*) or sgRNAs (*65*) to improve nuclease specificity. In addition, robust methods for predicting targeting outcomes (*66*) and achieving spatiotemporal gene regulation (*67*) provide comprehensive strategies to reduce off-targets. Beyond engineering the Cas nuclease, researchers are also developing a deeper understanding of cellular DNA repair to improve the likelihood of achieving a desired editing outcome (*68*).

Optimizing vehicles for efficient and specific delivery of the Cas payload remains a major obstacle, particularly in light of immune responses to sgRNA and Cas9 in humans (*69*, *70*). Within the lab, researchers have a number of options (electroporation, transfection, direct injection, and viral vectors) for delivering the DNA encoding the Cas payload, the sgRNA and mRNA encoding the Cas proteins, or preformed RNPs to cells ex vivo (*71*) or within immune privileged environments (*21*). Unfortunately, many of these options cannot be broadly translated in clinical settings where the specific requirements for efficient in vivo delivery vary with disease etiology. Furthermore, the large size of Cas nucleases bound to their guide RNA presents a challenge for packaging within viral-based vectors. One strategy to solve this problem is leveraging smaller related Cas homologs, or minimized systems that support packaging into viral vectors (*72*, *73*). Alternatively, functionalized nanomaterials enable specific delivery to a cell type of interest. Recent studies have shown that directly injecting nanoparticles containing Cas9-sgRNA efficiently corrects the causative DMD mutation, leading to improved clinical phenotypes in mice (*74*). In all likelihood, the success of CRISPR-based therapeutics will depend on further development of suitable vehicles for delivering the Cas payload.

Conclusions and future directions

CRISPR-Cas based technologies provide an accessible and adaptable means to alter, regulate, and visualize genomes, enabling biological research and biotechnological applications in a wide range of fields. CRISPR-Cas tools have vastly accelerated the pace of research, from understanding the genetics of previously unstudied organisms to discovering genes that contribute directly to disease. The field of Cas-based biotechnology is developing at a rapid pace, with multiple Cas9-based clinical trials in progress or beginning soon, the results of which will likely guide future use for somatic cell editing both ex vivo and in patients. Outside of the clinic, agricultural applications of CRISPR-Cas9 are already creating products for various markets, leading to recent rulings by the U.S. Department of Agriculture about their regulation (*17*). This ever-expanding repertoire of applications firmly positions the CRISPR-Cas toolkit at the cutting edge of genome editing and, more broadly, genetic engineering.

REFERENCES AND NOTES

1. M. Jasin, R. Rothstein, *Cold Spring Harb. Perspect. Biol.* **5**, a012740 (2013).
2. F. D. Urnov, *CRISPR J.* **1**, 34–46 (2018).
3. R. Barrangou *et al.*, *Science* **315**, 1709–1712 (2007).
4. L. A. Marraffini, E. J. Sontheimer, *Science* **322**, 1843–1845 (2008).
5. S. J. J. Brouns *et al.*, *Science* **321**, 960–964 (2008).
6. E. V. Koonin, K. S. Makarova, F. Zhang, *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
7. M. Jinek *et al.*, *Science* **337**, 816–821 (2012).
8. L. A. Marraffini, E. J. Sontheimer, *Nature* **463**, 568–571 (2010).
9. F. J. Mojica, C. Díez-Villaseñor, J. García-Martínez, C. Almendros, *Microbiology* **155**, 733–740 (2009).
10. P. Mali *et al.*, *Science* **339**, 823–826 (2013).
11. L. Cong *et al.*, *Science* **339**, 819–823 (2013).
12. M. Jinek *et al.*, *eLife* **2**, e00471 (2013).
13. K. Murugan, K. Babu, R. Sundaresan, R. Rajan, D. G. Sashital, *Mol. Cell* **68**, 15–25 (2017).
14. K. S. Pawelczak, N. S. Gavande, P. S. VanderVere-Carozza, J. J. Turchi, *ACS Chem. Biol.* **13**, 389–396 (2018).
15. C. Fellmann, B. G. Gowen, P.-C. Lin, J. A. Doudna, J. E. Corn, *Nat. Rev. Drug Discov.* **16**, 89–100 (2017).
16. K. Yin, C. Gao, J.-L. Qiu, *Nat. Plants* **3**, 17107 (2017).
17. E. Waltz, *Nat. Biotechnol.* **36**, 6–7 (2018).
18. Y. Zhang *et al.*, *Sci. Adv.* **3**, e1602814 (2017).
19. C. Long *et al.*, *Sci. Adv.* **4**, eaap9004 (2018).
20. T. Gaj *et al.*, *Sci. Adv.* **3**, eaar3952 (2017).
21. B. T. Staahl *et al.*, *Nat. Biotechnol.* **35**, 431–434 (2017).
22. E. Zuo *et al.*, *Genome Biol.* **18**, 224 (2017).
23. D. Niu *et al.*, *Science* **357**, 1303–1307 (2017).
24. L. J. Rupp *et al.*, *Sci. Rep.* **7**, 737 (2017).
25. J. Eyquem *et al.*, *Nature* **543**, 113–117 (2017).
26. R. O. Bak, D. P. Dever, M. H. Porteus, *Nat. Protoc.* **13**, 358–376 (2018).
27. H. Ma *et al.*, *Nature* **548**, 413–419 (2017).
28. M. D. Canny *et al.*, *Nat. Biotechnol.* **36**, 95–102 (2018).
29. G. T. Hess, J. Tycko, D. Yao, M. C. Bassik, *Mol. Cell* **68**, 26–43 (2017).
30. N. M. Gaudelli *et al.*, *Nature* **551**, 464–471 (2017).
31. K. Nishida *et al.*, *Science* **353**, aaf8729 (2016).
32. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, *Nature* **533**, 420–424 (2016).
33. A. C. Komor *et al.*, *Sci. Adv.* **3**, eaao4774 (2017).
34. M. Kampmann, *ACS Chem. Biol.* **13**, 406–416 (2018).
35. J. Pulecio, N. Verma, E. Mejía-Ramírez, D. Huangfu, A. Raya, *Cell Stem Cell* **21**, 431–447 (2017).
36. R. O. Bak *et al.*, *eLife* **6**, e27873 (2017).
37. Y. Zheng *et al.*, *Nat. Neurosci.* **21**, 447–454 (2018).
38. X. S. Liu *et al.*, *Cell* **172**, 979–992.e6 (2018).
39. H.-K. Liao *et al.*, *Cell* **171**, 1495–1507.e15 (2017).
40. F. S. Howe *et al.*, *eLife* **6**, e29878 (2017).
41. M. R. O'Connell *et al.*, *Nature* **516**, 263–266 (2014).
42. R. Batra *et al.*, *Cell* **170**, 899–912.e10 (2017).
43. S. C. Strutt, R. M. Torrez, E. Kaya, O. A. Negrete, J. A. Doudna, *eLife* **7**, e32724 (2018).
44. B. A. Rousseau, Z. Hou, M. J. Gramelspacher, Y. Zhang, *Mol. Cell* **69**, 906–914.e4 (2018).
45. G. Dugar *et al.*, *Mol. Cell* **69**, 893–905.e7 (2018).
46. O. O. Abudayyeh *et al.*, *Science* **353**, aaf5573 (2016).
47. A. East-Seletsky *et al.*, *Nature* **538**, 270–273 (2016).
48. O. O. Abudayyeh *et al.*, *Nature* **550**, 280–284 (2017).
49. R. Aman *et al.*, *Genome Biol.* **19**, 1 (2018).
50. A. A. Smargon *et al.*, *Mol. Cell* **65**, 618–630.e7 (2017).
51. D. B. T. Cox *et al.*, *Science* **358**, 1019–1027 (2017).
52. S. Konermann *et al.*, *Cell* **173**, 665–676.e14 (2018).
53. W. X. Yan *et al.*, *Mol. Cell* **70**, 327–339.e5 (2018).
54. S. C. Knight, R. Tjian, J. A. Doudna, *Angew. Chem. Int. Ed.* **57**, 4329–4337 (2018).
55. P. G. Maass *et al.*, *Nat. Struct. Mol. Biol.* **25**, 176–184 (2018).
56. P. Qin *et al.*, *Nat. Commun.* **8**, 14725 (2017).
57. D. A. Nelles *et al.*, *Cell* **165**, 488–496 (2016).
58. J. S. Gootenberg *et al.*, *Science* **356**, 438–442 (2017).
59. A. East-Seletsky, M. R. O'Connell, D. Burstein, G. J. Knott, J. A. Doudna, *Mol. Cell* **66**, 373–383.e3 (2017).
60. J. S. Gootenberg *et al.*, *Science* **360**, 439–444 (2018).
61. C. Myhrvold *et al.*, *Science* **360**, 444–448 (2018).
62. J. S. Chen *et al.*, *Science* **360**, 436–439 (2018).
63. J. H. Hu *et al.*, *Nature* **556**, 57–63 (2018).
64. J. S. Chen *et al.*, *Nature* **550**, 407–410 (2017).
65. H. Yin *et al.*, *Nat. Chem. Biol.* **14**, 311–316 (2018).
66. H. K. Kim *et al.*, *Nat. Biotechnol.* **36**, 239–241 (2018).
67. F. Richter *et al.*, *Curr. Opin. Biotechnol.* **48**, 119–126 (2017).
68. Y. Kan, B. Ruis, T. Takasugi, E. A. Hendrickson, *Genome Res.* **27**, 1099–1111 (2017).
69. C. T. Charlesworth *et al.*, bioRxiv 243345 [Preprint]. 5 January 2018.
70. S. Kim *et al.*, *Genome Res.* **28**, 367–373 (2018).
71. Z. Glass, M. Lee, Y. Li, Q. Xu, *Trends Biotechnol.* **36**, 173–185 (2018).
72. F. A. Ran *et al.*, *Nature* **520**, 186–191 (2015).
73. E. Kim *et al.*, *Nat. Commun.* **8**, 14500 (2017).
74. K. Lee *et al.*, *Nat. Biomed. Eng.* **1**, 889–901 (2017).
75. B. Zetsche *et al.*, *Cell* **163**, 759–771 (2015).

ACKNOWLEDGMENTS

We apologize to those whose work was not cited due to space limitations. We thank members of the Doudna laboratory for discussion. Special thanks to S. C. Strutt, J. C. Cofsky, and C. Fellmann for editing. **Funding:** J.A.D. receives support for her research from the HHMI, the NIH, the NSF (grant 1244557), the Allen Distinguished Investigator Program through The Paul G. Allen Frontiers Group, the William M. Keck foundation, and the Defense Advanced Research Projects Agency (DARPA) (award HR0011-17-2-0043). The views, opinions, and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. government. G.J.K. receives support from the American Australian Association. **Competing interests:** J.A.D. is an investigator of the HHMI and executive director of the Innovative Genomics Institute at the University of California, Berkeley, and the University of California, San Francisco. J.A.D. is a cofounder of Editas Medicine, Intellia Therapeutics, Caribou Biosciences, Scribe Therapeutics, and Mammoth Biosciences. J.A.D. is a scientific adviser to Caribou Biosciences, Intellia Therapeutics, eFFECTOR Therapeutics, Scribe Therapeutics, Sythego, Metagenomi, and Inari and is a member of the Board of Directors of Driver and Johnson & Johnson. The Regents of the University of California have patents pending for CRISPR technologies on which the authors are inventors.

10.1126/science.aat5011

REVIEW

Emerging applications for DNA writers and molecular recorders

Fahim Farzadfard and Timothy K. Lu

Natural life is encoded by evolvable, DNA-based memory. Recent advances in dynamic genome-engineering technologies, which we collectively refer to as in vivo DNA writing, have opened new avenues for investigating and engineering biology. This Review surveys these technological advances, outlines their prospects and emerging applications, and discusses the features and current limitations of these technologies for building various genetic circuits for processing and recording information in living cells.

Genomic DNA is an ideal medium for artificial biological information storage because of its ubiquitous presence, durability, and compatibility with biological functions, especially as the throughput of DNA sequencing has substantially increased along with drops in cost (*1*). With the advent of genome-editing technologies, we can now dynamically change genetic information and harness the vast capacity of genomic DNA for information processing and storage in living cells. These dynamic in vivo DNA-writing technologies have opened new avenues for investigating and engineering biology, ranging from building molecular recorders and living biosensors for the longitudinal study of signaling dynamics in biological processes (*2*–*6*) to rationally designing genetic memory elements and computation operations in living cells (*6*–*10*) to tracing cellular lineages during development and differentiation (*11*–*13*). Here, we first review the applications, prospects, and potential uses of these technologies in various biological and biomedical settings. We then outline current in vivo DNA-writing technologies, summarize the memory architectures and features that each of these technologies offers, and discuss their current limitations.

DNA writers

DNA writers are genetically encoded devices that enable targeted, dynamic, and recurring modifications of DNA in living cells (*2*–*7*, *14*). These modifications can take the form of targeted insertions, deletions, inversions, or base substitution mutations and can serve as distinct DNA memory states (Fig. 1A). On the basis of the mutational outcomes, these devices can be broadly categorized into two classes: precise and pseudorandom writers. Precise DNA writers generate predetermined mutational outcomes, resulting in well-defined transitions between memory states in cell populations. Pseudorandom DNA writers generate targeted but stochastic mutational outcomes, resulting in unpredictable mutation signatures in cell populations. These two classes of DNA writers offer different levels of encoding capacity and control over memory states and operations, making them suitable for different sets of applications (summarized in Table 1).

Synthetic Biology Group, Research Laboratory of Electronics, Department of Electrical Engineering and Computer Science and Department of Biological Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA.
Email: ffard@mit.edu (F.F.); timlu@mit.edu (T.K.L.)

DNA-writing applications

Molecular recording

Many molecular events that occur in biological systems are transient and thus difficult to monitor and study within their native context. DNA writing can be used to create molecular recorders that capture these transient signals and stably encode them into the DNA of cell populations or individual cells in vivo and in situ (Fig. 1B, left). The accumulated mutations can then be retrieved by DNA sequencing or functional assays to infer information about the original signals, even after the original signals are gone. This principle converts living cells into recording devices that memorize the history of their own signaling dynamics into permanent DNA records, which in turn can provide longitudinal insights into biological processes in their native contexts, as opposed to snapshots in time obtained by current approaches.

"With the advent of genome-editing technologies, we can now dynamically change genetic information and harness the vast capacity of genomic DNA for information processing and storage in living cells."

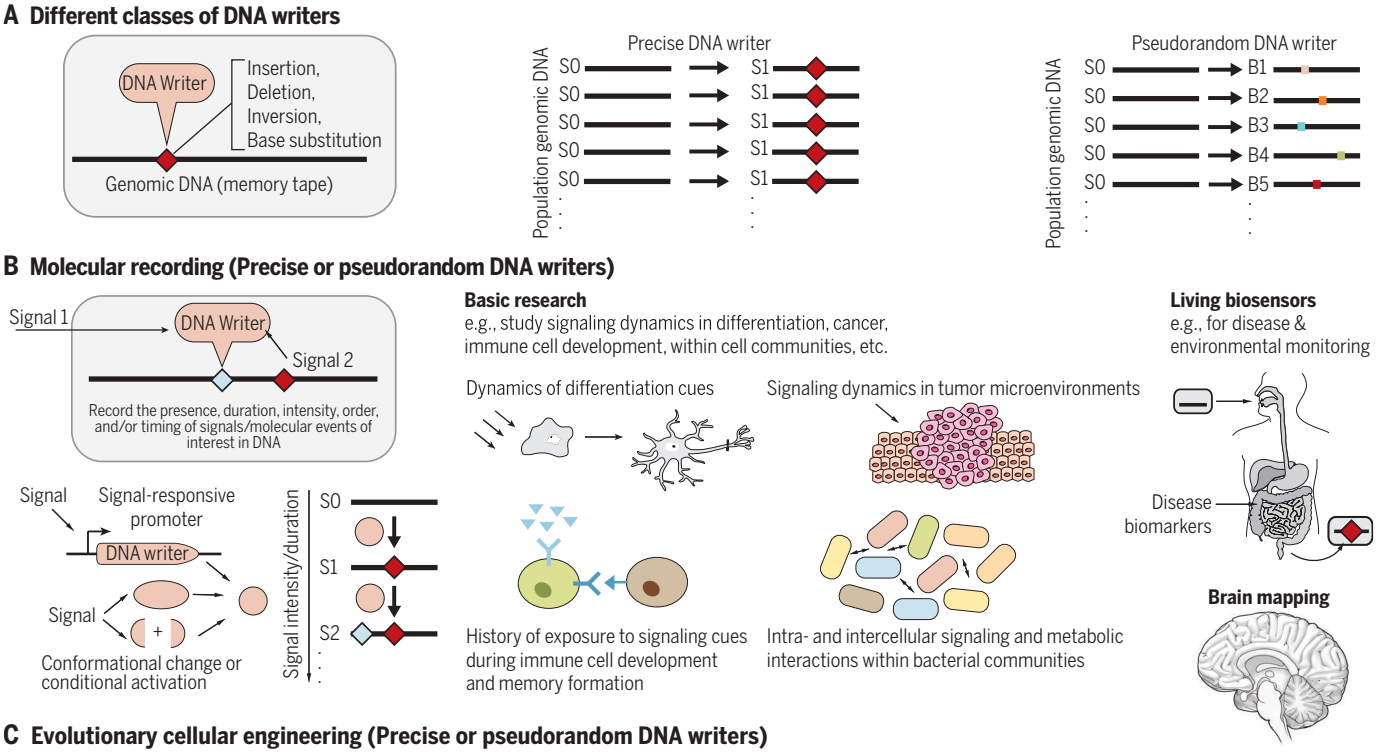
Several strategies, including conditional transcriptional or posttranscriptional activation of DNA writer components, can be used to couple the activity of a given DNA writer to signals of interest (Fig. 1B, bottom left). For example, by using signal-responsive promoters, information regarding the presence, duration, intensity, order, and timing of biological cues (such as metabolites and cytokines) or environmental cues (such as light, pollutants, exposure to phages, or changes in temperature) can be recorded in DNA (*2*–*6*). Naturally occurring signal-responsive promoters could be linked to DNA writer activity and used as a proxy to record and study the dynamics of the corresponding signaling pathways. If desired, rational design or directed evolution could be used to decouple natural promoters from unwanted overlapping pathways (e.g., by removing binding sites of corresponding transcription factors) or to engineer synthetic promoters with altered response dynamics (*15*–*17*). Alternatively, conditional activation of DNA writers in response to a desired signal can be achieved posttranscriptionally: for example, by implementing signal-dependent changes in conformation or interactions between DNA writer components.

Basic research

By offering an unprecedented ability to capture transient spatiotemporal molecular events in their native contexts, molecular-recording technologies could have broad utility across various disciplines (Fig. 1B, middle). For example, developmental biologists could use these DNA recorders to study the dynamics of differentiation cues and developmental pathways. Cancer biologists could use these recorders to study tumor development and to gain deeper insight into the cellular and environmental cues in tumor microenvironments that are involved in cancer heterogeneity. Immunologists could use these recorders to study signaling in immune cell maturation, memory formation, and immune responses. Microbiologists could use these recorders to study signaling dynamics and molecular interactions within bacterial communities and biofilms.

Various biological signals, ranging from small molecules to immunological cues to light, have been successfully recorded in both prokaryotic and eukaryotic cells (*2*–*6*). However, those recordings have been applied mainly to in vitro settings and relied on population-averaged readouts (see Box 1). Future work is needed to improve these technologies for single-cell recording or

Fig. 1. DNA-writing technologies and their emerging applications. (**A**) A schematic representation of a DNA writer (left) and mutation signatures generated by precise (middle) and pseudorandom (right) DNA writers. S0 and S1 indicate unmodified (memory state 0) and mutated (memory state 1) alleles, respectively. B1 to B5 indicate random memory states 1 to 5 that are generated by pseudorandom DNA writers and could serve as distinct barcodes. (**B**) Schematic representation of a molecular recorder and strategies that can be used to couple its activity to signals of interest (left), along with examples of applications in basic research (middle) and biotechnology (right). (**C**) Examples of evolutionary cellular engineering applications enabled by DNA writers. (**D**) Various forms of logic and computation can be achieved by layering multiple precise recorders. (**E**) Examples of strategies for high-throughput mapping of interactions or activities of variant libraries by DNA writers. (**F**) Pseudorandom DNA writers can be used to develop dynamic in vivo genetic barcoding schemes that distinctively and progressively mark cellular lineages over time.



to demonstrate the transformative use of molecular recorders in live animals, where the longitudinal study of in situ biology is currently limited. Memory architectures that impose minimal fitness effects will be important for realizing the use of molecular recorders in challenging in vivo conditions.

Living biosensors

Nonbiological sensors are not optimized to interact with biological systems. Living cells, on the other hand, are useful chassis for hosting sensors that can respond to various biological cues. DNA-writing technologies can be used to create living biosensors for longitudinal health and environmental monitoring. For example, bacterial cells endowed with disease biomarker sensors coupled with DNA recorders could be consumed orally, transit through the gastrointestinal tract to record disease biomarkers, and report this information later when they exit the body (Fig. 1B, top right). Engineered human cells harboring molecular recorders could be deployed into the body to report on early signs of disease, such as cancer or neurodegeneration. Finally, engineered cells and animals equipped with recording capacities could be used to continuously monitor and record the levels and activities of biological and environmental cues (such as toxins, heavy metals, metabolites, and light) without requiring artificial power supplies and in conditions and places that are not readily accessible to nonbiological sensors. Similar to basic research purposes, biosensing applications will need memory architectures with minimized fitness effects and extended recording capacities to achieve continuous and robust recording.

Brain mapping

Mapping the activities and connectome of neural circuits in the brain is one of the greatest challenges of our time (Fig. 1B, bottom right). As an alternative to current imaging-based techniques, which suffer from trade-offs between resolution and throughput, DNA-based ticker tape circuits that allow for the dynamic logging of signals have been proposed for recording spatiotemporal neural activities (*18*). Although existing molecular-recording technologies offer temporal resolutions that are orders of magnitude longer than neural pulses (Box 1), they could potentially be used to study time-averaged neural activities. For example, neural activities can be linked to molecular recorders via neural activity–responsive regulatory elements, such as immediate early gene promoters (*19*). Live animals harboring these genetic recorders could then be subjected to different neural stimuli, and the resulting mutational signatures could be used to infer time-averaged activities across the entire animal brain. Alternatively, DNA writers encoded on mobilizable genetic elements that can pass through synapses, such as rabies or pseudorabies viruses, could be used to distinctively mark neural connections by DNA barcodes that could then be used to map the connectome in a high-resolution and high-throughput fashion (*20*). Despite many technical challenges, we envision that applying molecular-recording technologies to decipher the functional architecture of the brain will be a strong driving force for the advancement of these technologies, especially in terms of scalability, recording capacity, and temporal and spatial resolution.

Evolutionary cellular engineering

Continuous in vivo evolution

In vivo DNA-writing technologies could be used to recurrently mutate desired genomic segments and achieve targeted genetic diversification within a short period. Once coupled with continuous selection, this strategy could enable continuous rounds of evolution to improve cellular traits of interest or to quickly evolve protein and RNA scaffolds for biotechnological and therapeutic applications (Fig. 1C, left). Unlike in molecular-recording applications, where it is desirable to minimize fitness effects to achieve robust recording, in evolutionary engineering applications, a selective pressure is applied to direct evolutionary trajectories toward desired outcomes. DNA-writing technologies with relaxed (*21*) or obviated (*2*) requirements for cis-encoded elements and extended mutational spectra (*22*) could be especially useful for these applications.

Synthetic Lamarckian evolution

Living cells have evolved mechanisms to elevate their local mutation rate under certain conditions and in response to specific signals. For example, during antibody maturation, CRISPR-Cas9 spacer acquisition, and mutagenesis processes mediated by diversity-generating retroelements in phages and bacteria, a series of actively regulated molecular events lead to targeted mutagenesis in certain genomic loci. These Lamarckian evolutionary strategies can increase the overall fitness of cell populations in uncertain environments and help them to adapt to environmental changes at greater rates than are possible by random Darwinian mutations. DNA-writing technologies could be used to emulate Lamarckian evolution by increasing the local mutation rates of desired genetic loci in response to signals of interest and in the presence of suitable selective pressures (Fig. 1C, right). Cells engineered with such a capacity could evolve faster than possible by natural evolution and enable adaptive cell-based therapeutics that tune their responses to the conditions

Box 1. DNA memory features.

Population-distributed versus single-cell recording

Because of the probabilistic nature of DNA writing at the single-molecule level, a statistically significant number of recording substrates (i.e., DNA molecules) are required to achieve robust recording. All the molecular recorders described so far have utilized the distributed genomic DNA of cell subpopulations to achieve robust recording. Developing efficient writers and/or using these together with high-copy-number recording substrates could pave the way toward single-cell recording.

Write cycles

We define write cycles as the number of iterations in which new information can be added to a memory register encoded on a single molecule of DNA by a single DNA writer or recorder complex before the memory register becomes nonresponsive to that complex. With the use of base editing (*6*), stgRNA (*4*, *14*), and Cas1-Cas2 (*3*, *28*) technologies, memory architectures with write cycles of >1 have been demonstrated.

Recording capacity

We define recording capacity as the number of distinct memory states that can be recorded (and practically retrieved) in the entire storage unit (the cell population for population-level recording or an individual cell for single-cell recording) by using a single DNA writer or recorder complex.

Digital versus analog recording

Depending on the writer efficiency and the potential memory states in the population (the number of memory states in each cell × the number of cells), two signal-recording regimes can be defined (Fig. 2D). Digital recording (a sharp, saturating increase in the mutation frequency in the population in response to an input) can be achieved when highly efficient DNA writers are used or when the number of potential memory states is limited. Analog recording (a gradual accumulation of mutations in the population in response to an input) can be implemented when moderately efficient writers are used or when there are many potential memory states. This extended dynamic range enables one to infer information regarding signal intensity and duration, which are analog properties, as opposed to the absence or presence of a signal, which is digital information.

Sequential and temporal resolution

Analog recorders integrate a signal over time but do not necessarily preserve information about the relative order or timing of multiple signals or the recurrence of a signal. Memory architectures with the capacity to record sequential and temporal information have been developed by using site-specific recombinases (*7*), base editing (*6*), and Cas1-Cas2 (*3*), although their resolution, write cycles, and recording capacity still need to be improved for demanding applications. Notably, ticker tape memory architectures (*18*) that record signaling dynamics in a temporally resolved fashion could enable one to infer signal intensity as a function of time.

they encounter. Alternatively, engineered bacteriophages endowed with the capacity to target and mutagenize their own host-range determinants could be useful for the streamlined development of phage-based antimicrobials that could adapt to infect new hosts faster than natural phages.

Applications specific to precise DNA writers

Layered molecular-recording, computation, and artificial-learning gene circuits

The precise and well-defined nature of the mutational outcomes generated by precise DNA writers allows them to be layered into more sophisticated genetic circuits in which the mutational outcome of one element can be used as inputs for other elements. By doing so, information regarding a series of input signals can be recorded in the form of well-defined transitions between multiple memory states. This strategy has been used to encode various forms of combinatorial, sequential, and temporal logic and other increasingly complex computing operations in living cells (Fig. 1D) (*6*–*10*). Additionally, because of the predictable nature of precise writers, their mutational output can be linked to functional genetic elements and used to control gene expression. These rationally designed genetic programs could be used, for example, to study or control the sequence and timing of developmental programs or to build gene circuits that classify disease conditions on the basis of multiple inputs. In addition, genetic programs could be created to endow cells with artificial learning capabilities such that specific circuit responses are gradually reinforced (or degraded) in response to signals (*6*), much like the reinforcement of synaptic interconnections in neurons.

High-throughput interaction and activity mapping

Transient cellular events, such as protein-protein interactions, can be converted into transcriptional outputs and therefore captured into DNA memory. For example, a split DNA-writing system, where the N- and C-terminal domains of a precise DNA writer are fused to barcoded bait and prey, respectively, could be used to record protein-protein interactions (Fig. 1E, left). In cells harboring interacting partners, a functional DNA writer could be reconstituted and write a prey-specific barcode next to a bait-specific barcode. The joined barcode could then be retrieved by sequencing to identify interacting partners in pooled libraries in a high-throughput fashion. Analogous strategies could be used to study the activities of RNA and protein variant libraries in a high-throughput fashion (Fig. 1E, right).

Application specific to pseudorandom DNA writers: lineage tracing

Capturing cellular ancestry relationships during development and creating corresponding lineage maps, especially in larger animals, have been a long-standing challenge in developmental biology. Traditionally, various static genetic and nongenetic barcoding approaches have been used for lineage tracing (*23*–*25*). In these methods, once a cell receives a barcode, it passes the barcode to its progenies with no change. Therefore, lineages that are generated in later stages are not differentially barcoded and, as a result, only a low-resolution lineage tree can be constructed. DNA writers can be used to devise dynamic genetic barcoding schemes that continuously and distinctively mark cell lineages as they progress in vivo, thus enabling higher-resolution lineage maps (Fig. 1F). Lineage tracing can be considered a specific application of molecular recording, where, instead of a transient signal, the chronicle of transient events (e.g., cell divisions) is recorded in DNA and later retrieved by sequencing. Pseudorandom writers are especially useful for lineage-tracing applications because they can generate many distinct mutational signatures in an initially clonal population.

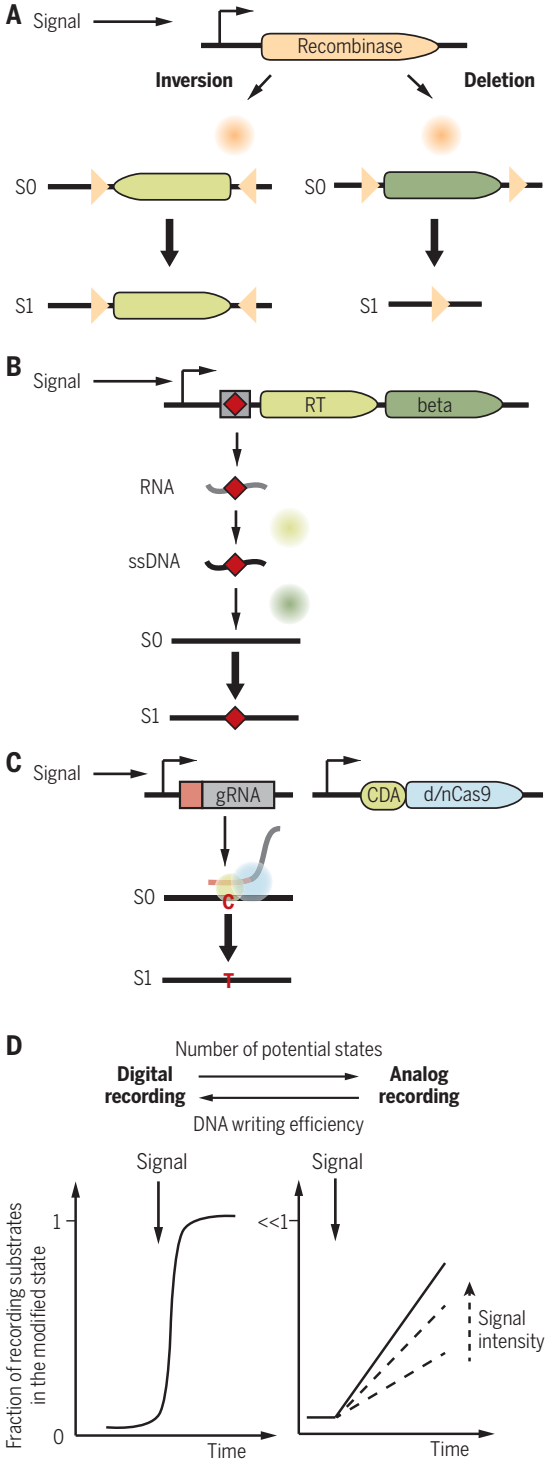


Fig. 2. Precise DNA-writing technologies. (**A**) Site-specific recombinases. (**B**) Recombineering. (**C**) Base editing. CDA, cytidine deaminase; d/nCas9, dCas9 or nickase Cas9. (**D**) Digital versus analog recording (see Box 1).

Precise DNA writers

Three classes of precise DNA writers have been described to date (Table 1), each featuring a different DNA-writing efficiency and thus a different recording regime (Box 1). Site-specific recombinases are the most efficient and well-established class of precise DNA writers. Depending on the orientation of their DNA recognition sites, these enzymes can either flip or excise a piece of DNA that lies between their cognate sites, thus memorizing the history of exposure to a signal in the form of defined and permanent DNA reconfiguration (Fig. 2A, transition from S0 to S1). Because of their relatively high efficiency, these DNA writers have been used mainly in digital recording (Box 1 and Fig. 2D) and building layered synthetic gene circuits for digital computation (*7*–*9*).

The second class of precise DNA writers relies on reverse transcriptase (RT)–mediated in vivo single-stranded DNA (ssDNA) expression followed by recombineering to achieve cis element–independent DNA writing in bacteria (Fig. 2B) (*2*). The moderate writing efficiency of this system offers wider-dynamic-range molecular recording in which the analog properties of biological signals, such as signal intensity and exposure duration, are recorded into the overall genomic DNA of cell populations (Box 1 and Fig. 2D). Because these DNA writers do not require cis-encoded elements on the target, they are desirable for evolutionary engineering applications.

The third class of precise DNA writers performs nucleotide-resolution manipulation of DNA via base editing (*26*). In this system, a base editor, such as a cytidine deaminase domain fused to dead Cas9 (dCas9), is addressed to a desired target site by expression of a complementary

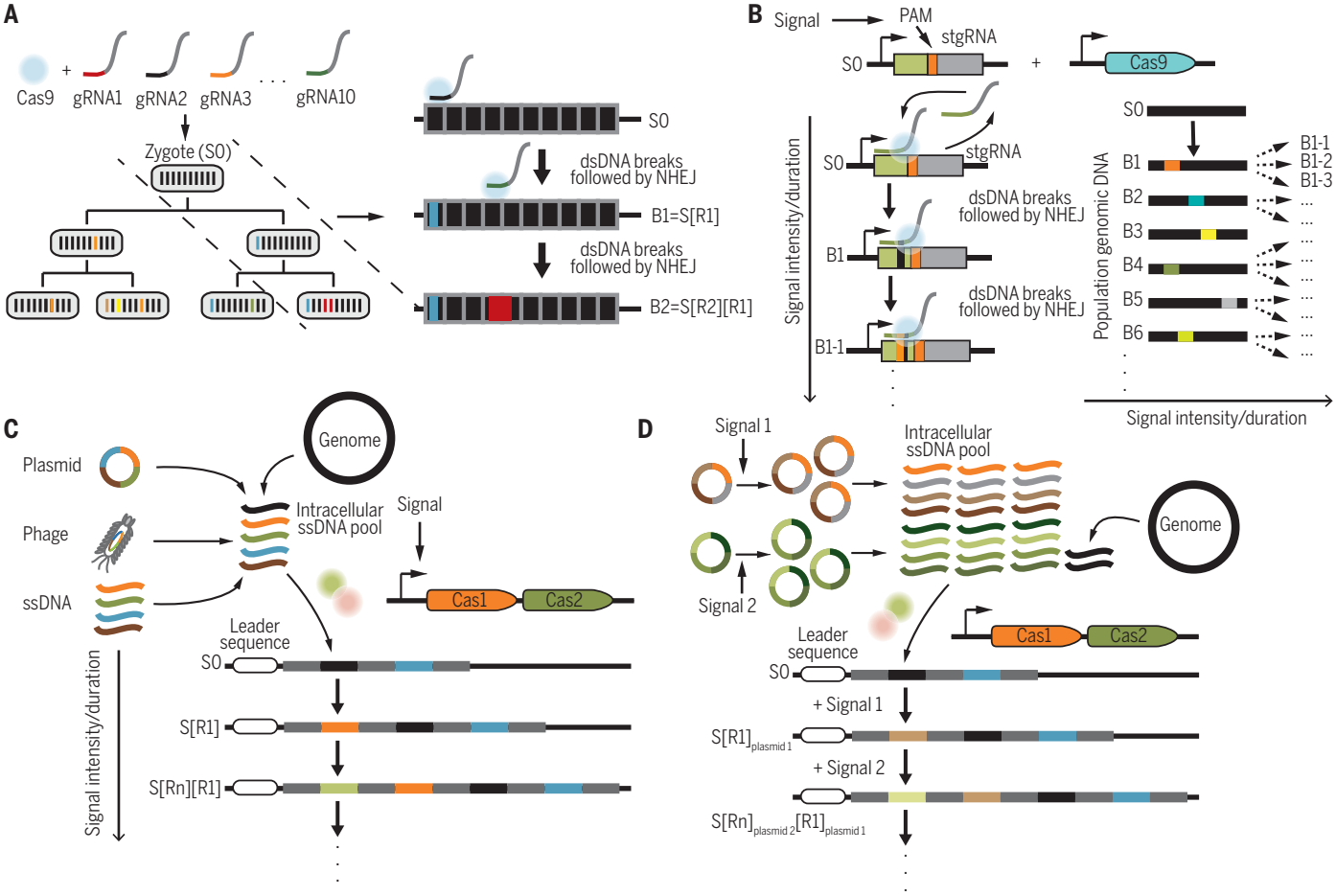


Fig. 3. Pseudorandom DNA-writing technologies. (**A**) A schematic representation of the Cas9 nuclease-based DNA-writing system used for lineage tracing in zebrafish in (*11*). S[R1], S[R2], and S[Rn] represent random memory states 1, 2, . . . , *n* generated by the pseudorandom DNA writers. These random memory states individually, or in combination with other random memory states, can serve as distinct barcodes (e.g., B1 = S[R1] and B2 = S[R2][R1]). (**B**) A schematic illustration of the Cas9 + stgRNA recording system. (**C**) DNA writing by the Cas1-Cas2 spacer acquisition system. (**D**) The strategy used in (*3*) to record temporal information into a CRISPR array.

Table 1. Features and demonstrated applications for the current DNA-writing technologies. TBD, to be determined; RSM, recombinase-based state machine; BLADE, Boolean logic and arithmetic through DNA excision; SCRIBE, synthetic cellular recorders integrating biological events; CAMERA, CRISPR-mediated analog multi-event recording apparatus; DOMINO, DNA-based ordered memory and iteration network operator; GESTALT, genome editing of synthetic target arrays for lineage tracing; MEMOIR, memory by engineered mutagenesis with optical in situ readout; mSCRIBE, mammalian synthetic cellular recorders integrating biological events; TRACE, temporal recording in arrays by CRISPR expansion.

DNA-writing technology	Mutation type	Used in	Host-specific requirement	Example platform(s)	Application(s)
			Precise writing mode		
Site-specific recombinases	Inversion and deletion	Bacteria and eukaryotes	–	RSM (*7*), BLADE (*10*)	Digital recording and computation
Recombineering	Small insertion, deletion, and base substitution	Bacteria	ssDNA-mediated recombineering	SCRIBE (*2*)	Analog recording
Base editing	dC to dT or dA to dG	Bacteria and eukaryotes	–	CAMERA (*5*), DOMINO (*6*)	Analog and digital recording and computation
			Pseudorandom writing mode		
Cas9 nuclease + gRNA	Indel	Eukaryotes	NHEJ	GESTALT (*11*), MEMOIR (*13*)	Lineage tracing
Cas9 nuclease + stgRNA	Indel	Eukaryotes	NHEJ	mSCRIBE (*4*, *14*)	Analog recording, lineage tracing
Cas1-Cas2 (spacer acquisition)	Insertion	Bacteria	TBD	TRACE (*3*, *27*, *28*)	Analog recording, lineage tracing, digital information storage

guide RNA (gRNA), generating deoxycytidine (dC)-to-deoxythymidine (dT) mutations within a narrow window in the target vicinity (Fig. 2C). As these memory operators are CRISPR-Cas9 based, they are more scalable than other precise writers. Additionally, they can be functionalized with regulatory modules (such as CRISPR interference and activation) to achieve complex recording and computation operations in living cells (*6*). Recently, an adenosine deaminase base editor, which writes deoxyadenosine (dA)-to-deoxyguanosine (dG) mutations, was developed (*22*), further expanding the mutation spectrum and utility of this class of DNA writers and paving the way toward bidirectional DNA-writing systems that could be used for advanced computation and evolutionary engineering applications.

"...there is plenty of room for improving existing memory architectures or developing new ones with desirable features, especially in terms of recording capacity, scalability, robustness, fitness effects, cellular resource consumption, write cycles, temporal resolution, and recording kinetics."

Pseudorandom DNA writers

Two main classes of pseudorandom DNA writers have been described to date. The first class relies on targeted double-stranded DNA (dsDNA) breaks generated by site-specific nucleases, such as CRISPR-Cas9, followed by error-prone repair of the breaks by the nonhomologous end–joining (NHEJ) pathway (*11*). During this process, each individual cell can acquire a pseudorandom mutational signature (i.e., indel mutations) in the target locus. Several studies have used these mutational signatures as barcodes to trace cellular lineages during embryo development in zebrafish and other small animals or in situ in cell cultures (Fig. 3A) (*11*–*13*). Efforts to extend the write cycles (see Box 1) of these molecular recorders led to the development of evolving barcodes (*4*, *14*). This memory architecture was built by engineering a protospacer adjacent motif (PAM) into the gRNA-encoding locus, resulting in self-targeting gRNAs (stgRNAs) that undergo iterative barcoding cycles during which the stgRNA locus is repeatedly diversified (Fig. 3B). This memory architecture was leveraged to build a population-level analog recorder (*4*) and to dynamically barcode cellular lineages in mammalian cells (*14*).

Despite these improvements, the reliance of these memory architectures on dsDNA breaks and NHEJ still limits their write cycles and makes them unsuitable for usage in organisms that lack an efficient NHEJ pathway, such as most prokaryotes. The prevalence of deletions in NHEJ can result in shortening of the stgRNA and loss of the PAM, thereby rendering the recorder nonfunctional over time. Moreover, the stochastic and deletion-based nature of the mutations generated by these strategies can result in nonpersistent encoding, where new memory states overwrite previous ones, thus making it difficult to infer ancestral relationships. Furthermore, encoding multiple stgRNAs in the same cell could result in unwanted chromosomal rearrangements and cellular toxicity. To extend the use of DNA writers for high-resolution lineage-tracing applications, particularly in larger animals, new memory architectures with improved efficiency, extended write cycles, reduced toxicity, and persistent (i.e., non–deletion-based) barcoding are desired. Alternative strategies, for example, using C- or A-rich stgRNAs in combination with base editors, could be devised to combine the high storage capacity of pseudorandom memory architectures with the well-defined and persistent memory states offered by precise DNA writers to address some of the above-mentioned limitations.

The second class of pseudorandom DNA writers was built upon Cas1 and Cas2 proteins, which naturally mediate spacer acquisition in the CRISPR bacterial immune system. In this system, the Cas1-Cas2 complex samples the intracellular ssDNA pool, which can originate from various intracellular or extracellular sources, and integrates short (~20 to 30 base pairs) ssDNA fragments from this pool into a preexisting CRISPR array, resulting in extension of the array over time (Fig. 3C). New spacers are added to the leader-proximal site, so the chronological order of spacer addition events is preserved within the array configuration. By placing the expression of the Cas1-Cas2 cassette under the control of signal-inducible promoters and introducing exogenous oligonucleotides into *Escherichia coli* cell populations, Shipman *et al.* (*27*) demonstrated that the signal intensity and duration can be inferred from array extensions and that the temporal order of the addition of the oligonucleotide pools can be inferred from the array composition. In a follow-up study, the authors demonstrated that artificial digital information, such as small pictures and movies, could be encoded into oligonucleotide pools and recorded into the distributed genomic DNA of a cell population (*28*). Building on these results, Sheth *et al.* (*3*) showed that instead of providing exogenous oligonucleotides, the intracellular ssDNA pool could be dynamically modulated by using template plasmids with tunable copy numbers. Using this strategy and multiple template plasmids, the authors demonstrated that the temporal order of multiple signals and lineage information of bacterial populations could be recorded in the CRISPR array composition (Fig. 3D). Though the Cas1-Cas2 writing system offers a relatively persistent memory architecture with desirable features, such as sequentially and temporally resolved recording and extended write cycles, it is currently limited to bacteria. The system could offer an attractive strategy for lineage tracing if it could be adapted to eukaryotes and function stably over multiple generations.

Conclusion and future prospects

In the past few years, we have witnessed the transition from the read-only genomic era to the read-and-write era. DNA-writing technologies have transformed genomic DNA into a dynamic medium for processing and storing biological and artificial information in living cells. These advances herald a new generation of powerful approaches for investigating and engineering in situ biology in basic research, biotechnology, and medicine. Although substantial progress has been made, there is plenty of room for improving existing memory architectures or developing new ones with desirable features, especially in terms of recording capacity, scalability, robustness, fitness effects, cellular resource consumption, write cycles, temporal resolution, and recording kinetics. These technologies promise to further advance our ability to manipulate life's natural memory storage media in a dynamic, longitudinal, and multiplexed fashion.

REFERENCES AND NOTES

1. G. M. Church, Y. Gao, S. Kosuri, *Science* **337**, 1628–1628 (2012).
2. F. Farzadfard, T. K. Lu, *Science* **346**, 1256272 (2014).
3. R. U. Sheth, S. S. Yim, F. L. Wu, H. H. Wang, *Science* **358**, 1457–1461 (2017).
4. S. D. Perli, C. H. Cui, T. K. Lu, *Science* **353**, aag0511 (2016).
5. W. Tang, D. R. Liu, *Science* **360**, eaap8992 (2018).
6. F. Farzadfard *et al.*, bioRxiv 263657 [Preprint]. 16 February 2018.
7. N. Roquet, A. P. Soleimany, A. C. Ferris, S. Aaronson, T. K. Lu, *Science* **353**, aad8559 (2016).
8. P. Siuti, J. Yazbek, T. K. Lu, *Nat. Biotechnol.* **31**, 448–452 (2013).
9. J. Bonnet, P. Yin, M. E. Ortiz, P. Subsoontorn, D. Endy, *Science* **340**, 599–603 (2013).
10. B. H. Weinberg *et al.*, *Nat. Biotechnol.* **35**, 453–462 (2017).
11. A. McKenna *et al.*, *Science* **353**, aaf7907 (2016).
12. S. T. Schmidt, S. M. Zimmerman, J. Wang, S. K. Kim, S. R. Quake, *ACS Synth. Biol.* **6**, 936–942 (2017).
13. K. L. Frieda *et al.*, *Nature* **541**, 107–111 (2017).
14. R. Kalhor, P. Mali, G. M. Church, *Nat. Methods* **14**, 195–200 (2017).
15. F. Farzadfard, S. D. Perli, T. K. Lu, *ACS Synth. Biol.* **2**, 604–613 (2013).
16. A. J. Meyer, T. H. Segall-Shapiro, C. A. Voigt, bioRxiv 285866 [Preprint]. 10 April 2018.
17. Y. Chen *et al.*, *Nat. Commun.* **9**, 64 (2018).
18. K. P. Kording, *PLOS Comput. Biol.* **7**, e1002291 (2011).
19. T. Kawashima, H. Okuno, H. Bito, *Front. Neural Circuits* **8**, 37 (2014).
20. A. M. Zador *et al.*, *PLOS Biol.* **10**, e1001411 (2012).
21. J. H. Hu *et al.*, *Nature* **556**, 57–63 (2018).
22. N. M. Gaudelli *et al.*, *Nature* **551**, 464–471 (2017).
23. K. Kretzschmar, F. M. Watt, *Cell* **148**, 33–45 (2012).
24. J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, *Dev. Biol.* **100**, 64–119 (1983).
25. J. Livet *et al.*, *Nature* **450**, 56–62 (2007).
26. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, *Nature* **533**, 420–424 (2016).
27. S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, *Science* **353**, aaf1175 (2016).
28. S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, *Nature* **547**, 345–349 (2017).

ACKNOWLEDGMENTS

Funding: This work was supported by the National Institutes of Health (P50 GM098792), the Office of Naval Research (N00014-13-1-0424), the NSF (MCB-1350625), the Defense Advanced Research Projects Agency, and NSF Expeditions in Computing program award 1522074. F.F. thanks the Schmidt Science Fellows Program, in partnership with the Rhodes Trust, for support.
Competing interests: F.F. and T.K.L. have filed patent applications on some works related to this Review.

10.1126/science.aat9249

REVIEW

Single-particle cryo-EM—How did it get here and where will it go

Yifan Cheng

Cryo–electron microscopy, or simply cryo-EM, refers mainly to three very different yet closely related techniques: electron crystallography, single-particle cryo-EM, and electron cryotomography. In the past few years, single-particle cryo-EM in particular has triggered a revolution in structural biology and has become a newly dominant discipline. This Review examines the fascinating story of its start and evolution over the past 40-plus years, delves into how and why the recent technological advances have been so groundbreaking, and briefly considers where the technique may be headed in the future.

Physicist and Nobel laureate Richard Feynman once famously stated, "It is very easy to answer many fundamental biological questions; you just look at the thing!" (*1*). Indeed, the central idea behind structural biology is that once we are able to "look" at "things" in great enough detail to discern their atomic structures, we will naturally be able to answer how and why the components and players of complex biological processes work the way they do. True to this aim, structural biology has contributed substantially to major biological discoveries throughout history (*2*). It has also and will continue to facilitate developments of therapeutic agents to cure diseases or ameliorate pathological symptoms.

The major techniques available to structural biologists are x-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and electron microscopy (EM). Among them, x-ray crystallography contributes most of the atomic coordinates of biological macromolecules deposited in the Protein Data Bank (PDB) (*3*). In this method, structures are determined from diffraction patterns generated from well-ordered three-dimensional (3D) crystals of biological macromolecules. The resolutions of structures determined depend largely on the quality of the crystals; in short, obtaining well-ordered 3D crystals of sufficient size is usually a prerequisite for atomic structure determination of any biological macromolecules by using this technique. It works well for many proteins or stable complexes; however, for certain categories of biological macromolecules, growing large and well-ordered 3D crystals is a very difficult or impossible task. For example, crystallizing integral membrane proteins or large and dynamic complexes and machineries can be challenging. An extension of the x-ray crystallography is the x-ray free electron laser (XFEL), whose ultimate goal is to determine atomic structures without crystals but currently still requires a large amount of small crystals (*4*).

Can atomic structures of biological macromolecules be determined without crystallization? Or is it possible, as Feynman once suggested, to determine their structures by "looking" at them by using a powerful electron microscope? Early pioneers pursued this question in the 1970s and developed a new EM-based method known today as single-particle cryo–electron microscopy (cryo-EM) (*5*, *6*). At the beginning, the method yielded rather low-resolution results. Drawn to the promise of being able to study biological macromolecules without crystallizing them, however, the cryo-EM community dedicated itself to perfecting the technique over more than four decades, yielding steady improvements in both the technique's applicability and the resolution of its results. Gradually, it has become a major tool in structural biology, complementary to x-ray crystallography and widely used to study large macromolecular complexes that are difficult to be crystallized. A few years ago, some amazing technological breakthroughs further enabled routine atomic resolution structure determinations with this method. Today, single-particle cryo-EM is no longer a complementary technique but a dominant one, changing the field of structural biology in a profound and unprecedented way and facilitating major new discoveries.

A brief history of single-particle cryo-EM

It is actually quite difficult to "look" at biological macromolecules in an electron microscope. Determining their atomic structures from electron micrographs is even more complicated. First, EM images are 2D projections of biological macromolecules, but not their 3D structures. This was resolved by De Rosier and Klug, who demonstrated that a 3D structure can be reconstructed by combining 2D projection images of the same object along different directions (*7*).

Second, because of strong scattering, the electron beam has to be confined in a high vacuum, and all EM samples need to be placed inside said vacuum. This is not a problem for inorganic materials. But if one simply places a biological sample inside an electron microscope, vacuum-caused dehydration would destroy the sample's structural integrity. The seemingly impossible task of keeping protein samples hydrated in a high

Howard Hughes Medical Institute, Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, CA, USA.

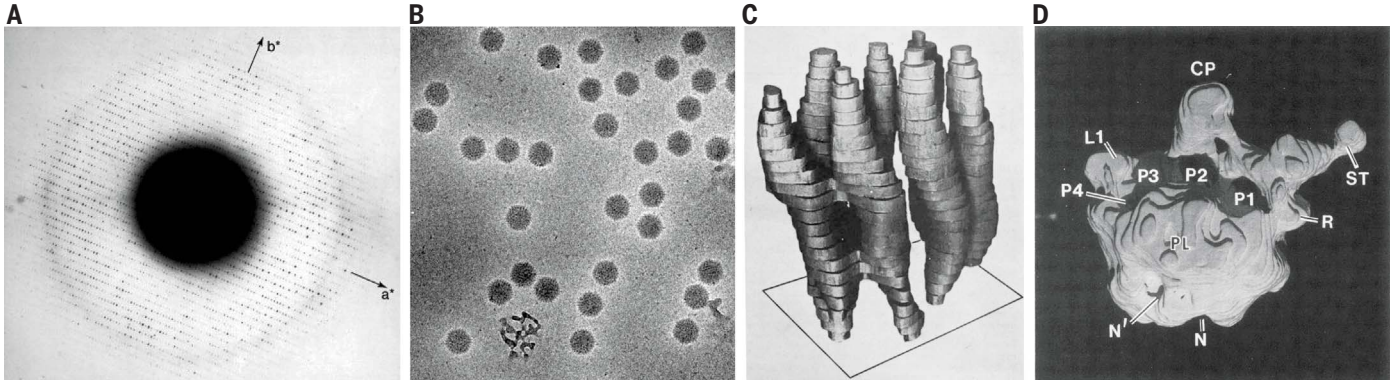


Fig. 1. Establishing single-particle cryo-EM. (**A**) An electron diffraction pattern of frozen hydrated catalase crystal. [Reprinted from (*9*) with permissions from Elsevier] The diffraction spots are visible at beyond 3-Å resolution. This experiment established the concept of cryo-EM. (**B**) An electron micrograph of frozen hydrated adenovirus particle recorded from a frozen hydrated grid prepared with plunge freezing. The micrograph is from the same dataset described in (*11*). [Reprinted from (*9*) with permissions from Elsevier] (**C**) The first 3D reconstruction of bacteriorhodopsin determined with electron crystallography. [Reprinted from (*15*) with permission from Springer *Nature*] (**D**) A 3D model of the 50*S* ribosome subunit determined with single-particle reconstruction of a negatively stained large ribosomal subunit from *Escherichia coli*. [Reprinted from (*22*) with permission from Wiley]

vacuum was accomplished by Taylor and Glaeser. They recorded better than 3-Å resolution electron diffraction patterns from frozen hydrated catalase crystals, demonstrating that the structural integrity of biological macromolecules in a high vacuum can be maintained through frozen hydration (Fig. 1A) (*8*, *9*). The practical implication of this approach, however, was not easy until a plunge freezing technique was developed by Dubochet and colleagues in the 1980s (*10*, *11*). They applied purified protein samples in solution to an EM grid covered with a thin layer of carbon holey film and blotted the grid with a filter paper, which removed most of the solution, and surface tension drove the remaining solution into a thin liquid film across holes in the carbon film. Plunging the grid rapidly into liquid ethane cooled by liquid nitrogen froze it into a thin layer of amorphous ice with the protein sample embedded within it in random orientations. After that, the frozen grid was transferred into an electron microscope and kept at near–liquid nitrogen temperature for imaging (Fig. 1B). This method is still used routinely without major changes, except that we now use a machine to blot and plunge grids with tunable parameters.

Third, radiation damage by the high-energy electron beam limits the total electron dose that can be used to image biological samples. The consequence is that images recorded with such a low electron dose have very poor signal-to-noise ratios (SNRs). Cooling the sample to liquid nitrogen or even liquid helium temperature can reduce the radiation damage and allow the sample to tolerate a higher electron dose (*12*, *13*) but still far from being able to directly visualize high-resolution details from raw micrographs. Henderson and Unwin overcame this problem by use of a crystallographic approach, averaging images of many identical proteins packed as 2D crystals (*14*). Images of glucose-embedded 2D crystals recorded with very low electron doses show no visible features, but their Fourier transforms show clear reflections. Similarly, electron diffraction from such 2D crystals at very low electron doses produced good-quality diffraction patterns. Combining the phases calculated from the Fourier transformations of images and the amplitudes obtained from diffractions produced a high-resolution projection map of the specimen (*14*). Combining data collected from specimens tilted at different angles produced a 3D reconstruction similar to the density map of an x-ray crystal structure. This approach produced the first structure of an integral membrane protein, initially at ~7-Å resolution (Fig. 1C) (*15*) and finally at atomic resolution (*16*). The method became known as electron crystallography, which relies on well-ordered 2D crystals. This method has produced atomic structures of several integral membrane proteins (*17*, *18*) and one soluble protein (*19*). The highest resolution achieved was 1.9 Å, resolving a lipid bilayer surrounding an aquaporin-0 (AQP0) water channel (*20*). But the difficulty of growing well-ordered 2D crystals hinders the broad application of the method.

"The resolution of a single particle cryo-EM structure depends on many factors..."

In parallel, Frank proposed an idea to determine protein structures without crystallization: computationally combining images of many individual protein particles of the same type (*5*). This conceptually novel idea was first tested by using protein samples that were negatively stained for EM observation (Fig. 1D) (*21*, *22*). The later combination of this approach with the plunge freezing sample preparation became what we now call "single-particle cryo-EM." It does not require growing proteins into crystals of any form. Instead, it determines structure by computationally aligning and combining cryo-EM images of many biological molecules randomly oriented within a thin layer of vitreous ice. A large number of images is needed to both enhance SNR and to provide all different views needed for 3D reconstruction. More detailed technical descriptions of this method can be found in many recent reviews, such as (*23*, *24*).

Furthermore, what electron micrographs record are projections of the specimen convoluted by a contrast transfer function (CTF). CTF is a sine function oscillating in a frequency-dependent manner that modulates both phase and amplitude of an image in frequency space (*25*). In addition to a number of microscope-dependent parameters, CTF is determined by how much off the focus an image is recorded—the so-called "defocus." To retain the highest resolution, images must be recorded very close to focus. Such images, however, have very limited contrast. This is not a problem for radiation-insensitive inorganic materials, which can be imaged with very high electron doses so as to generate sufficient contrast while retaining a high-resolution signal. However, images of radiation-sensitive frozen hydrated biological samples have to be recorded with a large defocus in order to generate sufficient contrast, which in turn dampens the high-resolution signal.

The resolution of a single-particle cryo-EM structure depends on many factors, including the resolution and contrast of individual particle images, accuracy of aligning these images with each other, obtaining a sufficient number of images from all necessary views of the macromolecule within a reasonable time frame, the conformational and compositional homogeneity of these particles, and access to powerful enough computers with which to process images efficiently. Less than optimal conditions for any of these factors—for example, instability of the electron microscope, relatively poor performance of the image-recording device and beam-induced sample motions, accuracy of classifying and aligning particle images limited by image quality and computational algorithms, and limited computer power to process very large numbers of particle images—can limit

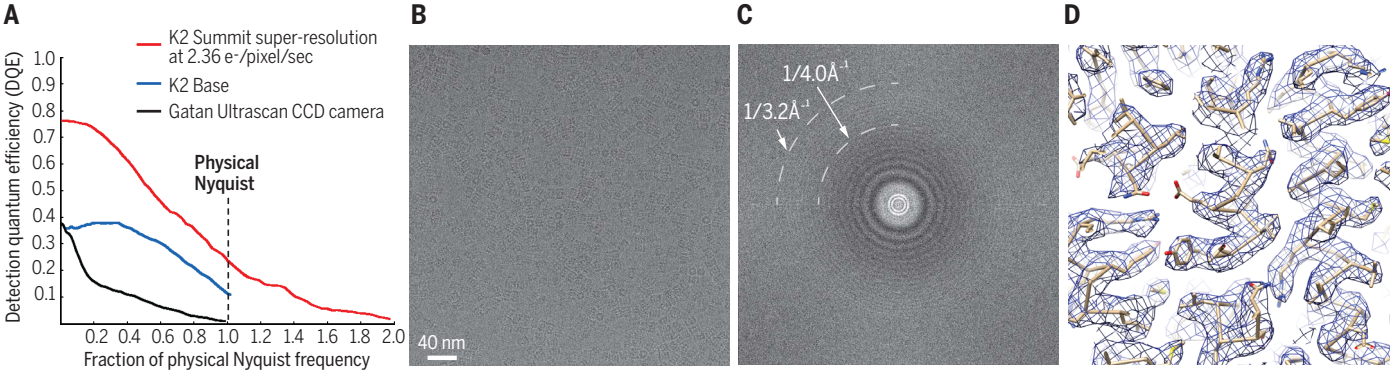


Fig. 2. Direct-electron-detection camera–enabled atomic structure determination. (**A**) Comparison of DQE curves of a scintillator-based CCD camera (black), and direct electron detection camera K2 operating in base mode (blue) and super-resolution counting mode (red) (*31*). (**B**) A typical electron micrograph of archaeal 20*S* proteasome (~700 kDa in molecular weight) recorded by use of a direct-electron-detection camera (*60*). (**C**) Fourier power spectrum calculated from the image in (B). [(B) and (C) reprinted from (*60*) with permission from Elsevier] High-resolution signal at ~3 Å resolution is clearly visible (*60*). (**D**) A portion of density map from the 3D reconstruction of archaeal 20*S* proteasome (*31*).

the resolution. Because of these many technical challenges, the resolution achieved was, for a long time, limited to levels far from sufficient for deriving de novo atomic models.

At a time when the resolution of some best reconstructions was at the 30- to ~50-Å range [for example, the mammalian 40*S* ribosome (*26*) and ryanodine receptor (*27*)], microscopists were nevertheless encouraged by the prediction that single-particle cryo-EM could, theoretically, achieve atomic resolution (*28*). The prediction was made by considering how much electron scattering a biological sample could tolerate, how much structural information or SNR such scattering can produce, and—in a perfect situation—how many images would be needed to produce a reconstruction at a given resolution. Although the claim was bold, the theories behind are solid, and this motivated the cryo-EM community to push the methodology toward its perfection. At last, 20 years after Henderson made the prediction, the necessary breakthrough came when direct electron detection cameras were developed and became commercially available (*29*). The expanded capabilities of the new cameras coupled with unprecedented and ever-increasing computational power fueled the development of new computational algorithms so that cryo-EM images could be reliably produced with sufficient quality for atomic structure determination (*30*, *31*). The resolution potential of single-particle cryo-EM at long last became a reality (*32*).

Transformative technological breakthroughs

The direct electron detection camera

Since the beginning of cryo-EM, electron micrographs were recorded on photographic films, which were subsequently processed in dark rooms and digitized for computational processing. For single-particle cryo-EM, the performance of film was sufficient to produce 3D reconstructions at subnanometer resolutions (*33*), at which α-helices are resolved, but atomic models cannot be built. It was difficult to push the resolution further because photographic film is particularly poor at retaining low-frequency signals. Images were therefore recorded with high defocuses in order to generate sufficient contrast for particle picking or alignment, at the price of losing high-resolution signal. For large icosahedral viruses, it was possible to record two images of the same specimen area, the first with a low defocus to retain the high-resolution signal and the second with a high defocus to generate sufficient contrast (*34*). With substantial effort, the resolution of some virus particles reached near atomic level (*35*), but the whole process is slow and tedious, limiting the throughput of structure determination.

In the late 1990s, charge-coupled device (CCD) cameras were introduced to record EM micrographs digitally. CCD cameras cannot detect electrons directly and require a phosphor scintillator to convert electrons into photon signals. Such conversion blurs a point event of a single electron striking the sensor into a blob of photons of much larger size and reduces high-resolution signals (*36*). Characterized by detective quantum efficiency (DQE), which measures the level of signals retained by a camera in spatial frequency (*36*, *37*), CCD cameras are not suitable for routine high-resolution structures determinations (Fig. 2A). The impact of introducing CCD cameras into cryo-EM was nonetheless important because it facilitated automated image acquisition (*38*).

A major breakthrough that elevated single-particle cryo-EM from "blobology" to a practical technique for atomic structure determination was the introduction of direct electron-detection cameras (*29*). Such cameras detect charges generated directly from electrons striking the camera sensor, thus localizing the electron with much greater precision and resulting in substantially higher DQE than that of scintillator-based cameras. These sensors also run at high frame rates, enabling cryo-EM images to be recorded as a stack of movie frames that each is recorded in a short period of time (*39*). Certain cameras can even count individual electron events on every single frame. The DQE of such single-electron-counting cameras is even higher (Fig. 2A) (*31*). Images recorded with direct-electron-detection cameras retain signal both at high frequency, for high-resolution structure determination, and at low frequency, for contrast required for image alignment (Fig. 2, B and C).

Being able to record images as movie stacks facilitated many new imaging approaches that are critical for maximizing the achievable resolution. Most importantly, it allows the correction of beam-induced image motion (*30*, *31*, *39*) and partially mitigates radiation damage (*31*, *40*, *41*), solving the two most difficult problems in cryo-EM. The use of the direct detection camera, improved motion correction, and the ability to record images at a high electron dose make the reconstruction of 3D density maps at atomic resolution possible for many proteins.

New image processing algorithms

The resolution of a reconstruction also depends on the conformational homogeneity of the sample and accuracy of image alignment of all particles used to reconstruct the density map. Image classification and alignment are thus the two most critical steps in the computational image processing. Because each particle image has a poor SNR, individual particles cannot be classified and assigned to a specific class and orientation with certainty, making a probabilistic approach better suited for image classification and alignment. Early attempts to use a maximum likelihood–based probabilistic approach in cryo-EM image processing were made in the late 1990s (*42*). Later, a Bayesian approach to cryo-EM reconstruction was described (*43*) and implemented in RELION (*44*), a user-friendly program that soon became very popular in single-particle cryo-EM image processing. This approach is more powerful and robust than traditional deterministic approaches, particularly in classifying a subset of particle images with homogeneous conformations out of a larger and more heterogeneous dataset. Coincidentally, this development happened at around the same time that direct-electron-detection cameras were becoming widely used in cryo-EM. Together, they allowed the full potential of single-particle cryo-EM—as theoretically predicted more than 20 years earlier—to be realized.

Automation in electron microscopy

Technically, single-particle cryo-EM had been a rather complicated and tedious technique. A very large number of high-quality cryo-EM images are required for each reconstruction. Such images were mainly collected by individuals who had many years of training and experience in operating complicated electron microscopes. The technical requirement on users was high because they needed a good understanding of not only the electron optic system but also many fundamental technical issues related to cryo-EM data acquisition. They also needed to be very patient, sitting in front of a microscope for many hours repeating the same procedure in order to collect large numbers of micrographs. The level of expertise required made single-particle cryo-EM inaccessible to the wider structural biology community. Fortunately, pioneered by Carragher and Potter, who recognized this problem early on, automation of high-quality data acquisition was developed (*38*). The electron microscope itself also evolved to become better suited for automated data acquisition. Now, many cryo-EM facilities are operated in a way similar to x-ray synchrotron beamlines—supported by a few highly trained staff scientists. In such facilities, regular users with minimal training can also acquire high-quality cryo-EM data, even remotely, by using automated procedures.

A new era of structural biology

Thanks to these technological breakthroughs in single-particle cryo-EM, structural biology has entered a new era. Structures of many difficult crystallization targets are now within reach. One such area is integral membrane proteins. A specific example is the transient receptor potential (TRP) ion channel. The TRP channel superfamily contains seven subfamilies with a total of 27 members in humans (*45*, *46*). Each of these channels plays different physiological roles; some of them are drug targets for treatment of various human diseases (*47*). With the exception of some small domains, attempts to crystallize any member of the TRP channel superfamily had failed (*48*). This lack of structural information ended when atomic structures of the TRPV1 ion channel—a capsaicin receptor that plays a physiological role in sensing heat and activating pain pathways—were determined by means of single-particle cryo-EM in three different functional states (*49*, *50*). This discovery demonstrated the power of single-particle cryo-EM and showed that it rivaled x-ray crystallography in determining atomic structures of challenging protein complexes that resist crystallization. The TRPV1 structures prompted many crystallographers to think seriously about cryo-EM, and many quickly seized the opportunity to apply it to their favorite difficult targets. With the roadblock of crystallization removed, atomic structures of integral membrane proteins are now being determined at a rapid pace. In less than 5 years,

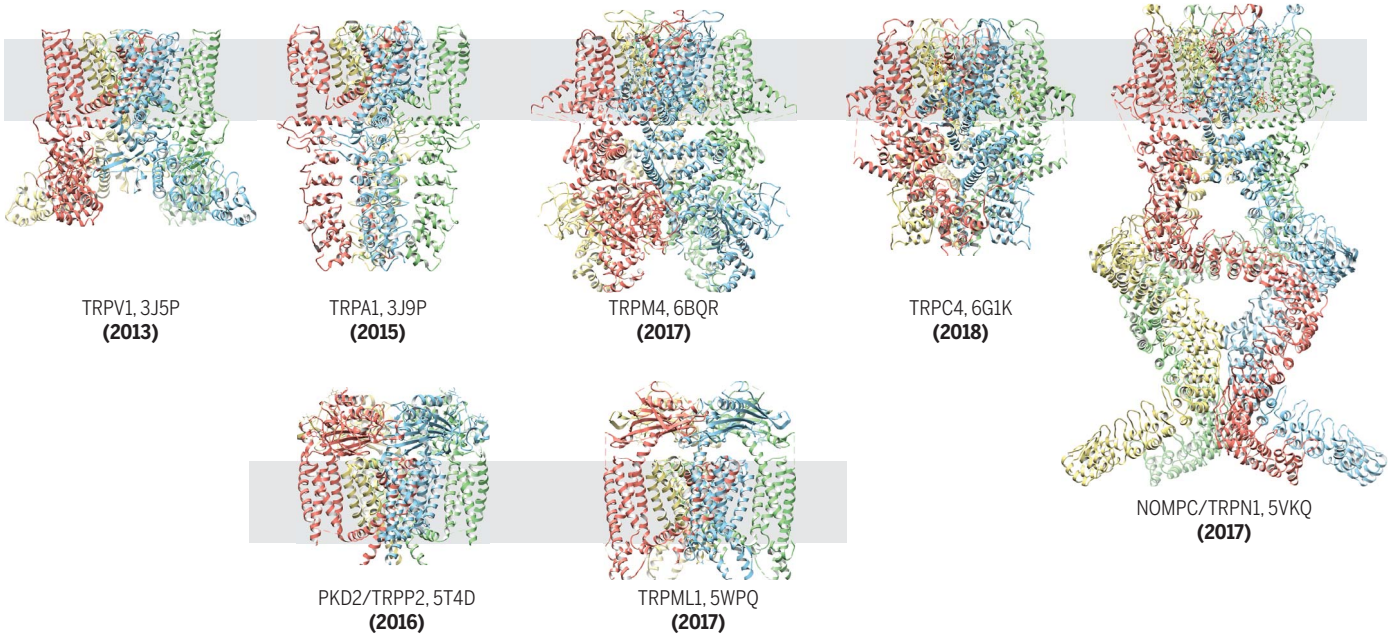


Fig. 3. Single-particle cryo-EM enables atomic structure determination of TRP channels. Ribbon diagrams of atomic structures from each subfamily of TRP channel superfamily. They are TRPV1 (*49*), TRPA1 (*61*), TRPM4 (*62*), TRPC4 (*63*), NOMPC (also named TRPN1) (*64*), PKD2 (or TRPP2) (*65*), and TRPML (*66*). The rapid pace of integral membrane protein structure determination is enabled with single-particle cryo-EM and is unprecedented.

atomic structures of at least one member of each of the seven TRP channel subfamilies have now been determined (Fig. 3).

Single-particle cryo-EM is also game-changing for structural studies of many large and dynamic complexes and machineries that were impossible for crystallization. A traditional approach was to fit crystal structures of individual components or domains into lower-resolution cryo-EM density maps of the whole complex (*51*). Now, atomic structures of many large complexes are determined directly by using single-particle cryo-EM. An excellent example is the spliceosome complex. Past efforts produced atomic structures of some small components (*52*), as well as low-resolution cryo-EM structures of the whole complex (*53*). Now, in only a few short years, atomic structures of the spliceosome in different functional states have been determined (*54*, *55*).

One can easily call forth many more examples to illustrate how the recent technological breakthroughs in single-particle cryo-EM have changed how we tackle complex biological problems. The rapid pace of advancement in structural biology is unprecedented. It has also attracted major pharmaceutical companies, with many hoping to implement cryo-EM into structure-based drug discovery and optimization.

The future

Single-particle cryo-EM continues to move forward fast, with many new technologies being developed. One such example is the Volta phase plate, which is a thin carbon film placed in the back focal plane of the microscope's objective lens. It adds a phase shift to the CTF so that an image recorded at near focus has good contrast (*56*), facilitating the study of very small proteins (*57*). Another example is the development of a new sample preparation technology that is fundamentally different from the current blotting method (*58*). The new method promises many benefits, including reducing the total amount of sample needed from microliters to nanoliters.

As impressive as these innovations have been, there are still many technical details that can be improved. Regularly achieving resolutions close to 2 Å and beyond is one goal; improving robustness and throughput is another. Once very high resolutions can be achieved reliably, pharmaceutical companies will be able to routinely use EM to speed up structure-based drug discovery. The range of what can be studied by means of single-particle cryo-EM can also be expanded to include smaller or bigger targets with higher resolution, as well as more dynamic complexes or assemblies with irregular shapes.

There is still more that single-particle cryo-EM can offer to further biological discoveries beyond structure determination if we continue to push the boundary of this technology. For example, in theory, image classification can sort a cryo-EM dataset that contains particles of heterogeneous conformation or composition into multiple classes, each corresponding to a different functional state. By freezing cryo-EM grids at specific time points, it could be possible to derive structural information in a time-dependent manner (*59*). It may therefore be possible to understand and break down the complex cycles, movements, and processes of biological macromolecular complexes and machineries step by step, in complete atomic detail. In the future, it may also be possible to study both physiological and pathological protein states through affinity-purifying specific proteins directly from cells onto EM grids for single-particle cryo-EM studies.

The past 40 years have seen the evolution of single-particle cryo-EM from blobology to a routine and powerful method for atomic structure determination. The technique's new popularity is attracting not only more users but also talents from different fields, ranging from physics, material science, and mathematics, to machine learning. This infusion of new ideas and a larger community heralds an even brighter future for single-particle cryo-EM full of method development, collaboration, and most importantly, wonderful new biological discoveries.

REFERENCES AND NOTES

1. R. P. Feynman, in *Feynman and Computation*, A. J. G. Hey, Ed. (Perseus Books Cambridge, 1999), pp. 63–76.
2. Y. Shi, *Cell* **159**, 995–1014 (2014).
3. N. Jones, *Nature* **505**, 602–603 (2014).
4. Y. Kang *et al.*, *Nature* **523**, 561–567 (2015).
5. J. Frank, *Ultramicroscopy* **1**, 159–162 (1975).
6. J. Dubochet *et al.*, *Q. Rev. Biophys.* **21**, 129–228 (1988).
7. D. J. De Rosier, A. Klug, *Nature* **217**, 130–134 (1968).
8. K. A. Taylor, R. M. Glaeser, *Science* **186**, 1036–1037 (1974).
9. K. A. Taylor, R. M. Glaeser, *J. Struct. Biol.* **163**, 214–223 (2008).
10. J. Dubochet, J. J. Chang, R. Freeman, J. Lepault, A. W. McDowall, *Ultramicroscopy* **10**, 55–61 (1982).
11. M. Adrian, J. Dubochet, J. Lepault, A. W. McDowall, *Nature* **308**, 32–36 (1984).
12. H. Stark, F. Zemlin, C. Boettcher, *Ultramicroscopy* **63**, 75–79 (1996).
13. Y. Fujiyoshi, *Adv. Biophys.* **35**, 25–80 (1998).
14. P. N. Unwin, R. Henderson, *J. Mol. Biol.* **94**, 425–440 (1975).
15. R. Henderson, P. N. Unwin, *Nature* **257**, 28–32 (1975).
16. R. Henderson *et al.*, *J. Mol. Biol.* **213**, 899–929 (1990).
17. W. Kühlbrandt, D. N. Wang, Y. Fujiyoshi, *Nature* **367**, 614–621 (1994).
18. K. Murata *et al.*, *Nature* **407**, 599–605 (2000).

19. E. Nogales, S. G. Wolf, K. H. Downing, *Nature* **391**, 199–203 (1998).
20. T. Gonen *et al*., *Nature* **438**, 633–638 (2005).
21. J. Frank, W. Goldfarb, D. Eisenberg, T. S. Baker, *Ultramicroscopy* **3**, 283–290 (1978).
22. M. Radermacher, T. Wagenknecht, A. Verschoor, J. Frank, *EMBO J.* **6**, 1107–1114 (1987).
23. Y. Cheng, N. Grigorieff, P. A. Penczek, T. Walz, *Cell* **161**, 438–449 (2015).
24. R. Fernandez-Leiro, S. H. Scheres, *Nature* **537**, 339–346 (2016).
25. R. H. Wade, *Ultramicroscopy* **46**, 145–156 (1992).
26. S. Srivastava, A. Verschoor, M. Radermacher, R. Grassucci, J. Frank, *J. Mol. Biol.* **245**, 461–466 (1995).
27. M. Radermacher *et al*., *J. Cell Biol.* **127**, 411–423 (1994).
28. R. Henderson, *Q. Rev. Biophys.* **28**, 171–193 (1995).
29. G. McMullan, A. R. Faruqi, R. Henderson, *Methods Enzymol.* **579**, 1–17 (2016).
30. X. C. Bai, I. S. Fernandez, G. McMullan, S. H. Scheres, *eLife* **2**, e00461 (2013).
31. X. Li *et al*., *Nat. Methods* **10**, 584–590 (2013).
32. Y. Cheng, *Cell* **161**, 450–457 (2015).
33. B. Böttcher, S. A. Wynne, R. A. Crowther, *Nature* **386**, 88–91 (1997).
34. S. J. Ludtke, W. Chiu, *J. Struct. Biol.* **144**, 73–78 (2003).
35. X. Zhang, L. Jin, Q. Fang, W. H. Hui, Z. H. Zhou, *Cell* **141**, 472–482 (2010).
36. P. Mooney, *Methods Cell Biol.* **79**, 661–719 (2007).
37. E. Samei, M. J. Flynn, D. A. Reimann, *Med. Phys.* **25**, 102–113 (1998).
38. C. Suloway *et al*., *J. Struct. Biol.* **151**, 41–60 (2005).
39. A. F. Brilot *et al*., *J. Struct. Biol.* **177**, 630–637 (2012).
40. S. H. Scheres, *eLife* **3**, e03665 (2014).
41. T. Grant, N. Grigorieff, *eLife* **4**, e06980 (2015).
42. F. J. Sigworth, *J. Struct. Biol.* **122**, 328–339 (1998).
43. S. H. Scheres, *J. Mol. Biol.* **415**, 406–418 (2012).
44. S. H. Scheres, *J. Struct. Biol.* **180**, 519–530 (2012).
45. D. E. Clapham, *Nature* **426**, 517–524 (2003).
46. D. Julius, *Annu. Rev. Cell Dev. Biol.* **29**, 355–384 (2013).
47. M. M. Moran, *Annu. Rev. Pharmacol. Toxicol.* **58**, 309–330 (2018).
48. M. Li, Y. Yu, J. Yang, *Adv. Exp. Med. Biol.* **704**, 1–23 (2011).
49. M. Liao, E. Cao, D. Julius, Y. Cheng, *Nature* **504**, 107–112 (2013).
50. E. Cao, M. Liao, Y. Cheng, D. Julius, *Nature* **504**, 113–118 (2013).
51. G. C. Lander *et al*., *Nature* **482**, 186–191 (2012).
52. W. P. Galej, T. H. Nguyen, A. J. Newman, K. Nagai, *Curr. Opin. Struct. Biol.* **25**, 57–66 (2014).
53. M. D. Ohi, L. Ren, J. S. Wall, K. L. Gould, T. Walz, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 3195–3200 (2007).
54. S. M. Fica, K. Nagai, *Nat. Struct. Mol. Biol.* **24**, 791–799 (2017).
55. Y. Shi, *Nat. Rev. Mol. Cell Biol.* **18**, 655–670 (2017).
56. R. Danev, B. Buijsse, M. Khoshouei, J. M. Plitzko, W. Baumeister, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 15635–15640 (2014).
57. M. Khoshouei, R. Danev, J. M. Plitzko, W. Baumeister, *J. Mol. Biol.* **429**, 2611–2618 (2017).
58. V. P. Dandey *et al*., *J. Struct. Biol.* **202**, 161–169 (2018).
59. S. Kaledhonkar, Z. Fu, H. White, J. Frank, *Methods Mol. Biol.* **1764**, 59–71 (2018).
60. X. Li, S. Q. Zheng, K. Egami, D. A. Agard, Y. Cheng, *J. Struct. Biol.* **184**, 251–260 (2013).
61. C. E. Paulsen, J. P. Armache, Y. Gao, Y. Cheng, D. Julius, *Nature* **520**, 511–517 (2015).
62. H. E. Autzen *et al*., *Science* **359**, 228–232 (2018).
63. D. Vinayagam *et al*., *eLife* **7**, e36615 (2018).
64. P. Jin *et al*., *Nature* **547**, 118–122 (2017).
65. P. S. Shen *et al*., *Cell* **167**, 763–773.e11 (2016).
66. Q. Chen *et al*., *Nature* **550**, 415–418 (2017).

ACKNOWLEDGMENTS

This is a brief story of single-particle cryo-EM but not a comprehensive historical review. I apologize to many colleagues whose important works are not cited here. I thank many of my colleagues and all past and present members of my laboratory for sharing their thoughts about single-particle cryo-EM with me, and L. Wang for editing. The cryo-EM work in my laboratory is funded by various NIH grants: R01GM082893, R01GM098672, R01HL134183, P50GM082250 (Nevan Krogan), P01GM111126 (Robert Stroud), and S10OD020054 and S10OD021741. I am a Howard Hughes Medical Institute investigator.

10.1126/science.aat4346

REVIEW

Visualizing and discovering cellular structures with super-resolution microscopy

Yaron M. Sigal, Ruobo Zhou, Xiaowei Zhuang*

Super-resolution microscopy has overcome a long-held resolution barrier—the diffraction limit—in light microscopy and enabled visualization of previously invisible molecular details in biological systems. Since their conception, super-resolution imaging methods have continually evolved and can now be used to image cellular structures in three dimensions, multiple colors, and living systems with nanometer-scale resolution. These methods have been applied to answer questions involving the organization, interaction, stoichiometry, and dynamics of individual molecular building blocks and their integration into functional machineries in cells and tissues. In this Review, we provide an overview of super-resolution methods, their state-of-the-art capabilities, and their constantly expanding applications to biology, with a focus on the latter. We will also describe the current technical challenges and future advances anticipated in super-resolution imaging.

Fluorescence microscopy has been central in shaping our understanding of the molecular organization and interactions of biological systems. Its high molecular specificity and multicolor imaging capability allow direct visualization of interactions between specific molecular species, and its low invasiveness allows the study of living systems under physiological conditions. However, a main challenge in fluorescence microscopy was the limited spatial resolution set by the diffraction of light. This resolution limit, first described by Ernst Abbe in 1873, restricts the smallest objects that can be resolved by conventional light microscopes. As a result, objects separated by a distance smaller than approximately half of the wavelength of visible light, i.e., ~200 to 300 nm, are indistinguishable, making many molecular structures in cells unresolvable. The advent of super-resolution imaging methods has shattered this limit. In this Review, we will provide an overview of the methods that surpass the diffraction limit in the far field, with emphasis on the new biological insights afforded by these methods.

Overview of super-resolution imaging methods

The key to overcoming the diffraction limit lies in the ability to distinguish molecules that reside within the same diffraction-limited volume. This has been achieved by two main categories of approaches. The first category accomplishes this in a spatially coordinated manner by using patterned illumination to differentially modulate the fluorescence emission of molecules within the diffraction-limited volume and thereby achieve separate detection of these molecules. The pioneering method in this category is stimulated emission depletion (STED) microscopy (*1*, *2*), subsequently generalized to reversible saturable optical linear fluorescence transitions (RESOLFT) (*3*). STED and RESOLFT overcome the diffraction limit by accompanying a focused excitation beam with a spatially patterned "depletion" beam, typically in a donut shape, which serves to counteract excitation through either stimulated emission (STED) (*1*, *2*) or other types of fluorescence transitions, such as photoswitching (RESOLFT) (*3*). As a result, only molecules at the very center of the donut-shaped beam (where the laser intensity is near zero) can emit light, thus creating a region of fluorescence emission that is much smaller than a typical focal spot of the light microscope. The reverse strategy is also possible, with the donut beam serving as patterned activation rather than depletion, limiting the emission-free region instead of emission region to the center of the beam (*4*). Scanning these beams across the sample then generates an image with a resolution much higher than the diffraction limit. Various other illumination patterns can also be used to increase the spatial frequency of the emission region and hence the image resolution (*4*). For example, in structured illumination microscopy (SIM), the sample is excited by a series of standing waves with different orientations or phases to increase the spatial frequency detectable by the microscope (*5*). Because the standing-wave pattern is itself limited by diffraction, the linear form of SIM only extends the diffraction limit by a factor of 2, whereas the nonlinear form of SIM (NL-SIM) overcomes the diffraction limit by using the nonlinear or saturated response of fluorophores to further increase the spatial frequency of the emission pattern (*5*), similar to STED and RESOLFT (*4*). Unlike STED and RESOLFT, which generate super-resolution images directly from the recorded raw data, SIM and NL-SIM require additional computational treatment to reconstruct final images (*4*, *5*).

Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology, Department of Physics, Harvard University, Cambridge, MA 02138, USA.
*Corresponding author. Email: zhuang@chemistry.harvard.edu

The second category of methods achieves the separation of molecules by stochastically turning on individual molecules within the diffraction-limited volume at different time points, including stochastic optical reconstruction microscopy (STORM) (*6*) and (fluorescence) photoactivated localization microscopy [(F)PALM] (*7*, *8*), and subsequent variations of these approaches (*9*, *10*). When isolated in space, the positions of individual molecules can be determined to nanometer or even subnanometer precision by localizing the center positions of their images (*11*–*13*). However, molecules within the same diffraction-limited volume generate overlapping images, which is the fundamental cause of the diffraction limit in resolution. STORM and PALM overcome this limit by switching on only a stochastic subset of fluorescent molecules within a field of view at any given time such that their images do not substantially overlap, allowing their positions to be localized with high precision; these molecules are then switched off (or bleached) and a stochastically different subset of molecules are switched on and localized—iterating this process allows a super-resolution image to be constructed from numerous molecular localizations accumulated over time (*6*–*8*). Such stochastic activation of molecules is typically achieved by using photoswitchable dyes or fluorescent proteins (*6*–*10*). A variety of photoswitchable probes have been used for this approach, in some cases leading to the creation of different acronyms subsequently, but the imaging principle is the same as that for STORM and PALM. In addition to using photoswitchable probes, transient binding of fluorescent probes can also be used to stochastically "turn on" fluorescent signals in space and time, as in point accumulation for imaging in nanoscale topography (PAINT) (*14*).

Recently, a new super-resolution imaging method named MINFLUX has been developed that combines strengths from both categories of approaches, by using stochastic switching of individual molecules to enable the separate detection of nearby molecules, along with patterned illumination, such as a donut-shaped beam, to achieve ultrahigh-precision localization of individual molecules by detecting local emission minima (*15*).

In addition to the above methods, which directly overcome the diffraction limit optically, a different form of super-resolution microscopy, expansion microscopy (ExM), has been recently developed, which increases the image resolution effectively through physical expansion of samples (*16*). In ExM, the specimen is embedded in a gel with the labeling probes attached to the gel. The sample is then digested to leave only the labeling probes attached to the gel followed by gel expansion to increase the probe separation, allowing super-resolution images to be taken with diffraction-limited microscopes.

Super-resolution technologies are constantly expanding, including both variations of the above approaches and other distinct methods, such as fluctuation-based methods and computer-vision–based methods. Owing to the limited space of this short review and its focus on biological applications, we cannot describe all methods here but

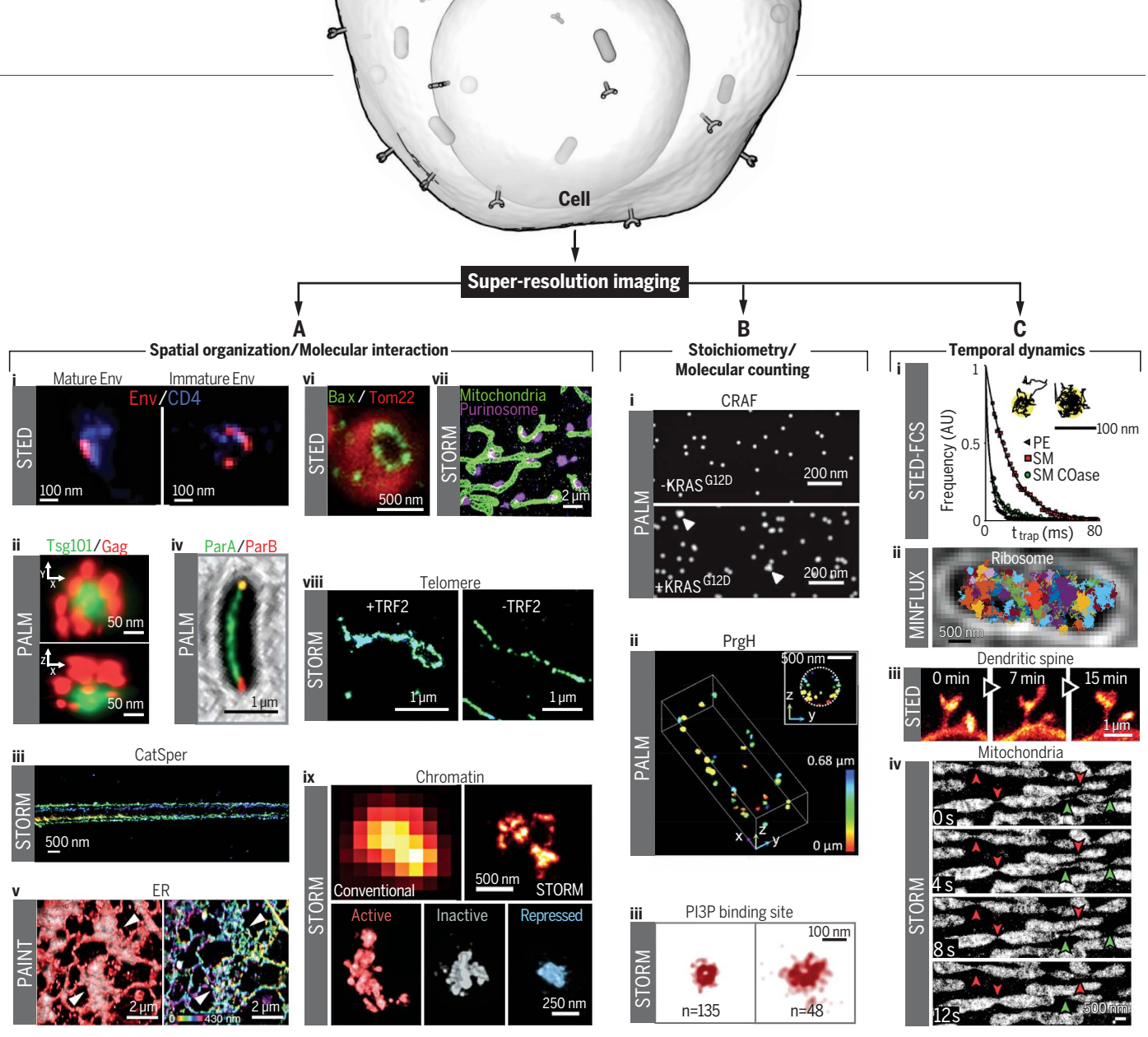


Fig. 1. Quantitative biological insights from three directions of applications of super-resolution imaging. (**A**) (**i**) STED images showing distinct distribution patterns of the envelope protein Env (red) in mature (left) and immature (right) HIV-1 particles attached to the cell, overlaid with the cell surface HIV-1 receptor CD4 (blue). (**ii**) PALM images showing the organization of ESCRT-I subunit Tsg101 (green) in an HIV assembly site marked by HIV Gag proteins (red) in lateral (top) and axial (bottom) views. (**iii**) 3D STORM images of a sperm-specific calcium channel (CatSper1) showing four linear domains along the sperm flagella. The *z*-position information is color-coded. (**iv**) PALM image on a bacterial cell showing the distribution of the ParA ATPase (green) with the ParB DNA binding protein (red) localized to the cell poles, for the coordination of chromosome segregation and cell division. (**v**) Left: Overlay of PAINT (red) and diffraction-limited (gray) images of the ER obtained using lattice light-sheet microscopy. Right: PAINT image from the left panel, but color-coded by the *z*-position information. White arrowheads indicate areas that appear as sheets in diffraction-limited images but are resolved as connected tubular structures in super-resolution images. (**vi**) STED image of the proapoptotic cell-death mediator Bax (green) showing ring structures in apoptotic mitochondria marked by Tom22 (red). (**vii**) STORM image showing interactions between mitochondria (green) and purinosomes marked by the core protein FGAMS (magenta). (**viii**) Comparison of STORM images of telomeric DNA in mouse embryo fibroblasts in the presence (left) and absence (right) of the shelterin protein TRF2 that is required for t-loop formation. (**ix**) Top: Comparison of diffraction-limited (left) and 3D STORM (right) images for DNA in a chromatin domain in the nucleus of *Drosophila* Kc167 cells. Bottom: Differential DNA compaction of transcriptionally active (red), inactive (gray), or polycomb-repressed (blue) epigenetic domains visualized using STORM. (**B**) (**i**) PALM images of proto-oncogene cRAF clusters on the cell plasma membrane, with (bottom) and without (top) coexpression of $KRAS^{G12D}$, which induces cRAF clustering. (**ii**) 3D PALM image of molecular clusters with various sizes formed by a secretion system protein PrgH near the membrane of a bacterial cell. (**iii**) STORM images of endocytic vesicles displaying distinct vesicle size and phosphatidylinositol 3-phosphate (PI3P) content. The number of PI3P binding sites on each vesicle (*n*) is indicated. (**C**) (**i**) Durations (t_{trap}) for three lipid types—phosphoethanolamine (PE, gray), sphingomyelin (SM, red), and sphingomyelin after cholesterol depletion (SM COase, green)—that are differentially trapped in ~20-nm nanodomains at the plasma membrane, which are detected and distinguished by STED-FCS and confocal single-molecule tracking. (**ii**) Single-particle tracking of a 30*S* ribosomal subunit protein in a bacterial cell by using MINFLUX. Trajectories of individual molecules are shown in different colors. (**iii**) Time-lapse STED images of a region of the somatosensory cortex of a living mouse with enhanced yellow fluorescent protein (EYFP)–labeled neurons, showing dynamics of dendritic spines. (**iv**) Time-lapse STORM images showing fission (green arrowheads) and fusion (red arrowheads) events of mitochondria, with thin tubular structures connecting neighboring mitochondria as fission and fusion intermediates. Figures are modified from the following sources: (A) i (*49*); ii (*50*); iii (*51*), with permission from Elsevier; iv (*52*), with permission from Springer Nature; v (*35*); vi (*54*); vii (*56*); viii (*59*), with permission from Elsevier; ix (*60*), with permission from Springer Nature; (B) i (*66*); ii (*67*); iii (*64*); (C) i (*69*); ii (*15*); iii (*71*); iv (*34*).

refer interested readers to other reviews (*4*, *9*, *10*) for additional coverage on super-resolution technologies.

Imaging capabilities of super-resolution microscopy

Three-dimensional (3D) imaging

The 3D nature of biological structures calls for super-resolution in all three dimensions. For methods based on stochastic activation of single molecules, such as STORM and PALM, achieving 3D super-resolution imaging requires high-precision localization not only in the *xy* plane, but also in the *z* direction along the optical axis. This was first achieved by astigmatism imaging [by using a cylindrical lens to create *z*-dependent point-spread-function (PSF)] (*17*), followed by various other approaches including bifocal plane imaging (*18*), PSF engineering (*19*), and interferometry (*20*), among others (*9*, *10*, *21*). In STED and RESOLFT, isotropic 3D super-resolution imaging was achieved by generating a depletion illumination pattern to counteract excitation in all directions surrounding the focal point—for example, by using a donut-shaped STED beam in conjunction with two opposing objectives (*4*, *22*).

Image resolution

Both the methods based on patterned illumination, like STED, RESOLFT, and NL-SIM, and the methods based on single-molecule switching and localization, like PALM and STORM, are diffraction-unlimited, and thus do not have a theoretical resolution limit. In practice, however, many factors can influence the achievable resolution, including the excitation and detection schemes, and the photophysical properties and size of fluorescent probes, as well as the labeling and sampling density of these probes. In biological applications, resolutions achieved by these methods are typically in the range of 10 to 70 nm, with sub–10 nm resolution achieved in some cases (*9*, *10*).

For the patterned-illumination–based methods, the spatial frequency (or sharpness) of the final emission pattern determines the image resolution. For example, in STED and RESOLFT, the donut-shaped depletion beam limits the fluorescence emission zone to the very center of the donut beam. The stronger the depletion light, the narrower this emission zone and the higher the achievable image resolution (*4*). However, strong illumination can lead to substantial photobleaching, phototoxicity, and enhanced background noise. Hence, the resolutions typically achieved are tens of nanometers, although resolution as high as a few nanometers has also been demonstrated by using probes with ultrahigh photostability, such as diamond nitrogen-vacancy centers (*4*). With isoSTED, isotropic 3D resolution of ~30 nm has been demonstrated (*4*, *22*). Combining patterned illumination with photoswitchable probes, RESOLFT (*3*, *4*) has also achieved ~30-nm isotropic 3D resolution (*23*). Similarly, by combining sinusoidal patterned illumination and photoswitchable probes, and using additional computational image reconstruction, NL-SIM has demonstrated ~45- to 60-nm resolution in 2D using saturated depletion (SD NL-SIM) or patterned activation (PA NL-SIM) (*24*, *25*). PA NL-SIM has been extended to 3D with the help of lattice light sheet microscopy (*26*), providing a resolution of ~120 to 230 nm in 3D (*25*).

For single-molecule-switching–based methods, such as STORM and PALM, the resolution depends on the photophysical properties of the fluorophores. Although many fluorophores exhibit blinking or switching behavior, only those with sufficient brightness and proper on-off switching kinetics yield high-quality images (*27*). The achievable image resolution depends on the number of photons detected from individual molecules, known as the photon budget. Typical experiments with bright photoswitchable dyes provide ~20- to 30-nm *xy* resolution, whereas the resolution is worse for fluorescent proteins because of their lower photon budget. The resolution is often worse in the *z* direction, but the use of interferometry (*20*, *28*, *29*) or specially engineered PSF (*30*) can improve the *z* resolution to become equal to or even better than the *xy* resolution. For example, interferometry can provide <10-nm *z* resolution, though a more complicated imaging setup is needed (*20*, *28*, *29*). In general, the resolution in both *xy* and *z* directions can be increased by improving the photon budget of the fluorophores. For example, the development of ultrabright photoactivatable dyes allowed a resolution as high as a few nanometers to be achieved on biological structures using STORM (*31*). More recently, using stochastic binding of dye-labeled DNA probes, DNA-PAINT also achieved similar image resolution on DNA-origami nanostructures (*32*). However, the time required to detect such a large number of photons for each molecule substantially increased the acquisition time per image. The recently developed MINFLUX thus represents an important advance in that it uses patterned excitation to drastically increase the localization precision (or reduce the number of photons required to reach a set precision), achieving an impressive localization precision of ~1 nm with an orders-of-magnitude lower photon budget (*15*). In ExM, the resolution depends on the number of rounds of sample expansion and the expansion factor per round, and a resolution of ~25 nm has been demonstrated with two rounds of expansion (*33*).

Other factors can also limit the final image resolution, such as the size of the fluorescent probes and the labeling density, which affect all super-resolution methods. For single-molecule–based approaches, methods that increase the number of times each target is sampled can also increase the final resolution if the resolution is limited by sampling density. This includes using fluorophores that undergo many on-off switching cycles (*27*), diffusible probes that can sample multiple locations of the target (*34*), and PAINT approaches that sample the target numerous times using reversible probe binding (*35*).

Finally, thick samples pose additional challenges, including reduced localization precision for out-of-focus molecules, optical aberration, and light scattering, as well as increase in background noise, all of which can lower image resolution. Various PSF engineering methods have been developed to allow high localization precision over substantially longer focal depths (*19*, *21*, *30*). Adaptive optics have been used to correct for aberrations in super-resolution imaging of thick samples (*29*), and light-sheet illumination provides an effective optical-sectioning approach to reduce background in thick-sample super-resolution imaging (*36*). Tissue clearing methods can reduce not only aberration but also scattering and are particularly powerful for thick-sample imaging (*37*). Alternatively, serial physical sectioning has also been used to reconstruct super-resolution images over large volumes of tissue (*38*).

Live-cell imaging, temporal resolution, and phototoxicity

Several super-resolution methods have demonstrated live-cell imaging. As scanning-based techniques, STED and RESOLFT can image a relatively small field of view (FOV) with very high temporal resolution, and thus have the impressive capability of probing millisecond dynamics of cellular structures at the spatial resolution of tens of nanometers (*39*). Although the time resolution decreases with increasing FOVs, highly parallelized RESOLFT with 100,000 intensity minima, effectively akin to 100,000 tightly spaced donut patterns, allows subsecond time resolution for large FOVs (*40*).

As widefield imaging methods, the time resolutions of STORM and PALM do not change as rapidly with the FOV size. Subsecond time resolution at ~20- to 30-nm spatial resolution has been achieved for large FOVs in live-cell imaging by STORM using fast-switching dyes (*41*) and fast scientific-CMOS cameras (*42*). Several recently developed algorithms to localize a high density of molecules with overlapping images (*43*) can further increase the time resolution of these methods. In addition, the single-particle–tracking mode of PALM, STORM, and PAINT (*44*, *45*) allows movement of individual molecules to be tracked with time resolutions of milliseconds to tens of milliseconds at high molecular concentrations. The ability of MINFLUX to achieve high localization precision with a minimal photon budget has led to a drastic increase in the tracking time resolution of molecules in live cells to the submillisecond scale (~100 μs), coupled with a corresponding increase in the number of snapshots possible for each molecule before photobleaching (*15*).

For live-cell imaging, in addition to achieving high spatiotemporal resolution, it is also important to reduce photobleaching and phototoxicity to prolong the overall duration of imaging and to minimize perturbations to the biological systems. Because of the trade-offs between spatial and temporal resolutions and between spatiotemporal resolution and phototoxicity/imaging duration, it is possible to reduce the spatial resolution of both the patterned-illumination-based methods, such as STED and RESOLFT, and the single-molecule-switching–based methods, such as STORM and PALM, to trade for higher time resolution, or lower phototoxicity and longer imaging duration.

In addition, adaptively changing the intensity of the STED beam based on the presence or absence of fluorophores, as in DyMIN, results in a substantial reduction in photobleaching and phototoxicity (*46*). By using photoswitching instead of stimulated emission, RESOLFT requires a much lower light intensity than STED, and thus drastically reduces phototoxicity in live-cell super-resolution imaging (*4*). When the spatial resolution requirement is not particularly high (~100 nm), SIM is a popularly used live-cell imaging method because of its capability for high-speed widefield imaging with low phototoxicity. The recently reported PA NL-SIM demonstrated live-cell imaging with ~60-nm spatial resolution and subsecond time resolution over large FOVs and tens of time points (*25*). In general, using light-sheet illumination for optical sectioning can also reduce phototoxicity in imaging (*36*). The recently developed lattice light sheet microscopy (*26*) further decreases phototoxicity and improves optical sectioning (to ~300 nm) compared to previous light-sheet schemes, and has been used in conjunction with super-resolution approaches to improve their volumetric live-cell imaging capability (*25*, *26*, *35*).

Quantitative biological insights offered by super-resolution imaging

Super-resolution imaging has transformed our understanding of biological systems and the applications are rapidly expanding, prohibiting comprehensive descriptions in a short review. Instead, we will highlight in this section the types of quantitative insights that can be obtained by super-resolution imaging, with representative examples illustrating each case (Fig. 1). In the next section, we will provide more detailed descriptions of a few examples to further illustrate the power of super-resolution imaging (Figs. 2 to 4).

Spatial organization and molecular interaction of cellular structures

The nanometer-scale resolution afforded by super-resolution imaging has substantially advanced our

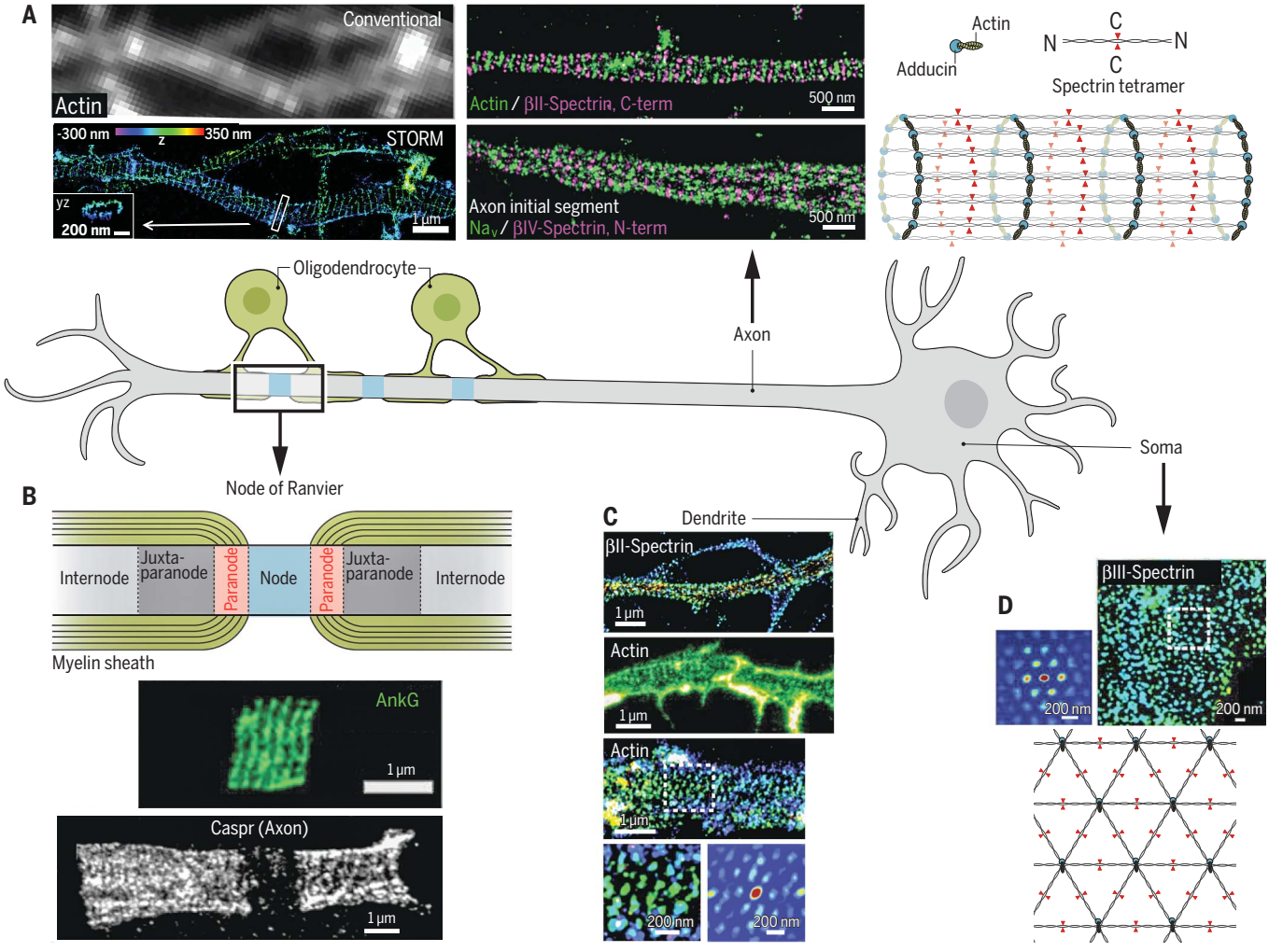


Fig. 2. The membrane-associated periodic skeleton (MPS) in neurons discovered by super-resolution imaging. (**A**) Quasi-1D periodic MPS observed in axons by using STORM. Left: Comparison of diffraction-limited (top) and 3D-STORM (bottom) images of actin in axons. STORM image shows the periodic distribution of actin rings along the axon that is obscured by diffraction-limited imaging. Middle: Two-color STORM images showing the periodic distributions of and spatial relationship between actin, spectrins (βII- and βIV-spectrin), and voltage-gated sodium channels (Na_v). Right: Schematic of the 1D MPS structure showing the organization of actin, spectrin tetramers, and adducin. Modified from (*47*). (**B**) Top: Schematic of a node of Ranvier. Center: STED image showing the periodic distribution ankyrin-G (AnkG) on the 1D MPS structure at a node of Ranvier. Modified from (*74*). Bottom: STED image showing the periodic distribution of the adhesion molecule Caspr on the 1D MPS structure observed flanking a node of Ranvier. Modified from (*76*). (**C**) MPS structures observed in dendrites: Top: 1D MPS in a dendritic region observed by STORM imaging of βII-spectrin. Modified from (*72*). Upper middle: 1D MPS observed in a dendritic region by STED imaging of actin. Modified from (*73*). Lower middle: 2D polygonal lattice-like arrangement of MPS components observed in a dendritic region by STORM imaging of actin. Bottom: A magnified region of the STORM image (left) and the corresponding 2D autocorrelation analysis (right) are shown. Modified from (*77*). (**D**) 2D MPS observed on the soma of neuron by STORM imaging of βIII-spectrin (top right) along with a 2D autocorrelation analysis of the boxed region (top left). Bottom: Schematic of the 2D MPS structure. Modified from (*77*).

ability to interrogate the spatial organization of molecular structures in cells (*9*, *10*). In addition, multicolor super-resolution imaging has allowed molecular interactions to be examined at unprecedented resolution (*9*, *10*). With these abilities, super-resolution imaging has provided new insights into numerous cellular structures, and even led to discoveries of previously unknown cellular structures, such as the membrane-associated periodic skeleton (MPS) in neurons (*47*) detailed in the next section.

At the cell surface, membrane proteins such as receptors, channels, vesicle scission proteins, and viral fusion proteins have been investigated by various super-resolution approaches and are often found to assume functionally important spatial organizations. For example, it was shown that the SNARE complex component syntaxin-1 is densely packed within discrete clusters that are regulated by the lipid composition (*48*). The HIV envelope protein (Env) was observed to reorganize upon maturation, which is important for viral entry (*49*) (Fig. 1A, i), whereas the ESCRT complex is localized to the virus budding site and plays an important role in HIV budding (*50*) (Fig. 1A, ii). The calcium channel CatSper was shown to adopt a linear-domain organization along the sperm tail together with other signaling and scaffolding molecules, playing an important role in calcium signaling and sperm activity (*51*) (Fig. 1A, iii). In the cytoplasm, super-resolution imaging has provided new insights into the organization of cytoskeleton structures and membrane organelles, as well as other molecular assemblies. In addition to the discovery of the MPS in neurons (*47*), as will be detailed in the next section, novel organization has also been observed for other cytoskeletal structures, such as the ParA/ParB system in bacteria (*52*) (Fig. 1A, iv) and focal adhesions connecting the cytoskeleton to the plasma membrane (*53*). For membrane organelles, super-resolution imaging has revealed, for example, densely packed and dynamic tubular structures in endoplasmic reticulum (ER) sheets (*35*) (Fig. 1A, v), ring structures of Bax on apoptotic mitochondria (*54*, *55*) (Fig. 1A, vi), and synergistic interactions between mitochondria and purinosomes (*56*) (Fig. 1A, vii). In addition to protein structures, super-resolution imaging has also provided new insights into RNA distributions and interactions in cells (*57*, *58*). In the cell nucleus, super-resolution imaging has revealed interesting organizations of DNA and DNA-interacting proteins, such as the TRF2-dependent telomere loop (t-loop) formation important for DNA end protection (*59*) (Fig. 1A, viii), distinct chromatin organization and compaction in different epigenetic states (*60*) (Fig. 1A, ix), and cell-type–dependent organizations of nucleosomes (*61*).

Stoichiometry of molecular complexes

Although measuring stoichiometry by spatially resolving individual subunits within molecular complexes is still challenging, the ability to activate and localize individual molecules by PALM and STORM has triggered growing interest in stoichiometric characterizations within intact cells. However, it is important to note that the number of measured single-molecule localizations is not equivalent to the number of molecules because of two complications. The first arises from imperfect labeling. New labeling approaches, such as using gene editing to label endogenous proteins with rapidly maturing, monomeric fluorescent proteins or with protein or peptide tags that can be conjugated to dyes with near 100% efficiency, can help mitigate this challenge. The second complication arises from complex fluorophore switching: No dye or fluorescent protein has the ability to give precisely one localization per molecule because fluorophores blink (multiple times), and most fluorophores also have an inactivatable fraction. Multiple methods have been developed to combat this problem, including calibrations of fluorophores using standards of known stoichiometry or quantification and modeling of blinking properties (*62*–*64*). STED has also been used to quantify the number of molecules based on coincident photon detection (*65*). These methods have been applied to quantifying, for example, the number of proteins in flagellar motors (*62*), receptor complexes (*63*), kinase complexes (*66*) (Fig. 1B, i), and secretion machinery (*67*) (Fig. 1B, ii), as well as the number of lipid binding sites in endocytic vesicles (*64*) (Fig. 1B, iii).

Temporal dynamics of cellular structures

Super-resolution imaging has enhanced our ability to extract dynamic information of cellular structures, allowing the mobility of biomolecules and the shape or structural dynamics of molecular complexes and organelles to be tracked with higher accuracy. For example, STED imaging has been used in combination with fluorescence correlation spectroscopy (FCS) to study the diffusion properties of molecules on the membrane. The drastic reduction in the region of fluorescence emission by the STED beam has allowed the detection of membrane nanodomains <20 nm in size, within which different lipid molecules show distinct diffusion properties (*68*, *69*) (Fig. 1C, i). Super-resolution imaging has also enhanced our ability to perform single-particle tracking (SPT) in live cells. Conventional SPT experiments require a low labeling density for the molecule of interest to avoid signal overlap between molecules. Stochastically turning on only a subset of labeled molecules at a given time, as in PALM, STORM, and PAINT, allows SPT at much higher molecular concentrations at the endogenous expression level (*34*, *44*, *45*), facilitating the studies of gene expression, protein–nucleic acid interaction, and dynamic processes on cell membranes. With its unique capabilities, MINFLUX has allowed the tracking of ribosomes in bacterial cells with unprecedented spatiotemporal resolution, achieving a localization precision of <50 nm with a time resolution of ~100 μs (*15*) (Fig. 1C, ii).

In addition, various super-resolution microscopy methods have been used to measure structural and shape dynamics of molecular assemblies, organelles and small cellular compartments, such as the dynamics of neuronal processes and dendritic spines in tissue (*70*) and even in live animals (Fig. 1C, iii) (*71*), fission and fusion dynamics of mitochondria (*34*) (Fig. 1C, iv), and structural dynamics of ER (*35*).

Super-resolution studies of specific molecular assemblies

The membrane-associated periodic skeleton in neurons

Super-resolution imaging enabled the discovery of the membrane-associated periodic skeleton (MPS) in neurons, which was initially observed in the axons by STORM imaging (*47*). In the MPS, short actin filaments, capped by actin-capping proteins, such as adducin, are organized into repetitive, ring-like structures that wrap around the circumference of the axon; adjacent actin rings are connected by spectrin tetramers, forming a long-range quasi-1D periodic structure with a periodicity of ~180 to 190 nm underneath the axonal membrane (*47*) (Fig. 2A).

The MPS spans the entire axon shaft, in both myelinated and unmyelinated axonal segments, including the axon initial segments (AIS) and nodes of Ranvier where action potentials are generated and amplified, respectively (*47*, *72*–*76*) (Fig. 2, A and B). This structure was observed in all neuronal types examined, including excitatory and inhibitory neurons in both central and peripheral nervous systems (*74*, *75*), and is evolutionarily conserved across diverse animal species (*75*). Subsequent to its discovery in axons, this 1D periodic structure was also observed in dendrites by both STORM and STED (*72*, *73*) (Fig. 2C), but the formation propensity and development rate of MPS appear to be lower in dendrites than in axons (*77*). In addition, a 2D polygonal lattice structure formed by MPS components was observed in the soma and dendrites (Fig. 2D), resembling the membrane skeleton structure observed in erythrocytes (*77*).

This highly ordered submembrane skeletal structure can play diverse functional roles in neurons. It provides flexible mechanical support for axons that is likely critical for axon stability under mechanical stress (*47*); indeed, axons tend to break in spectrin-deficient animals under movement-induced stress (*78*). The MPS was also implicated in mechanosensation (*79*). Moreover, the MPS organizes membrane proteins, such as ion channels and adhesion molecules, into periodic distributions along axons (*47*, *72*, *73*, *76*), potentially influencing the generation and propagation of action potentials, and other signaling pathways in axons. The MPS also influences axon and dendrite morphology (*80*), is important for the formation of the AIS and nodes of Ranvier (*72*, *76*, *80*), and may also act as a diffusion barrier at the AIS (*81*). Disruption of the MPS causes widespread neurodegeneration and a range of neurological impairments (*80*), and mutations of MPS components are implicated in various neurodegenerative diseases. The discovery of the MPS, which escaped detection by previous imaging methods, demonstrates the power of

super-resolution imaging for uncovering new cellular structures.

Molecular organization in synapses

Neuronal synapses are typically only several hundred nanometers in size but contain elaborate protein machineries to orchestrate neurotransmitter-mediated signal transmission; hence, the structural interrogation of synapses requires high spatial resolution and has benefited from extensive super-resolution imaging efforts. For example, STED has revealed the spatial organization of several important components within the *Drosophila* neuromuscular junction, including the clustered organization of Ca^{2+} channels, as well as the organization of scaffolding proteins required for both Ca^{2+} channel clustering and synaptic vesicle tethering at the presynaptic active zone (*82*, *83*) (Fig. 3A). STORM imaging has mapped the spatial organization of many proteins in the pre- and postsynaptic terminals, which show oriented organization of presynaptic scaffolding proteins, laminar organization of postsynaptic density proteins, and synapse-to-synapse variability in the lateral distributions of neurotransmitter receptors (*84*) (Fig. 3B).

In addition, recent super-resolution studies revealed that the neurotransmitter receptors and postsynaptic scaffolding proteins adopt activity-dependent clustered organization (*85*, *86*) (Fig. 3C). Such clustered organization also extends across the synaptic cleft, giving rise to "nanocolumns" formed by spatially aligned presynaptic vesicle fusion sites and postsynaptic receptor clusters (*87*) (Fig. 3D). This nanocolumn organization provides a mechanism for the coordination of synaptic vesicle release and neurotransmitter receptor response.

Super-resolution studies of synapses have been recently extended to both proteomic-scale analysis of synaptic structures and circuit-scale analysis of synapse distributions. For example, STED has been used to image numerous protein components in the presynaptic terminals, creating a model of an "average" synaptic bouton (*88*). A

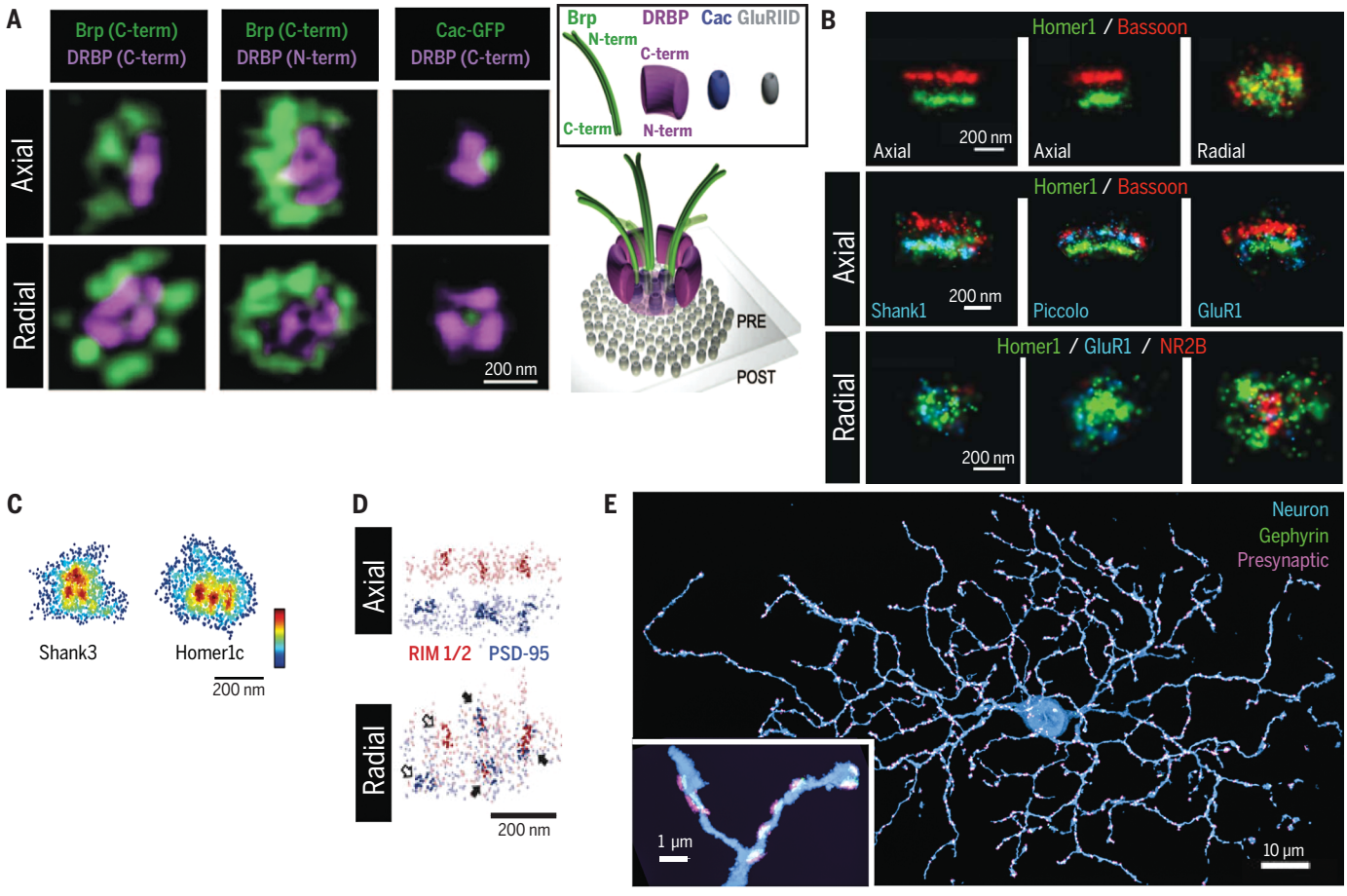


Fig. 3. Super-resolution imaging of synaptic structures. (**A**) Left: 3D STED images of the presynaptic active zone including Bruchpilot (Brp) and *Drosophila* RIM binding protein (DRBP), as well as the voltage-gated calcium channel Cacophony (Cac) at *Drosophila* neuromuscular junction synapses. Both axial (top) and radial (bottom) projections are shown. Right: Schematic of the active zone showing positions and orientations of components of the active zone cytomatrix including Brp, DRBP, and Cac in relation to the postsynaptic glutamate receptor (GluRIID) determined by using STED. Modified from (*83*). (**B**) Top: 3D STORM images of presynaptic protein Bassoon (red) and postsynaptic protein Homer (green). Two orthogonal axial views (left and middle) and the radial view (right) are shown. Center: Axial views of three synapses. In addition to Bassoon (red) and Homer1 (green), a third color (blue) was used to map the positions of additional postsynaptic (Shank1, left; GluR1, right) and presynaptic (Piccolo, middle) components at synapses. Bottom: Radial views of three example synapses showing differential abundance and spatial distribution of neurotransmitter receptors, NR2B and GluR1. Modified from (*84*) with permission from Elsevier. (**C**) Radial projections of PALM images showing the clustered organization of postsynaptic proteins Shank3 (left) and Homer1c (right). Modified from (*85*) with permission from Elsevier. (**D**) STORM and PALM images show that areas of higher protein density (darker colors) of both presynaptic (RIM1/2, red) and postsynaptic (PSD-95, blue) components are often trans-synaptically aligned to form "nanocolumns" (indicated by filled arrows). Both axial (top) and radial (bottom) projections are shown. Modified from (*87*) with permission from Springer Nature. (**E**) STORM maximum intensity projection of a retinal ganglion cell (blue) with associated synapses marked by postsynaptic scaffolding protein gephyrin (green) and presynaptic proteins (Bassoon, Piccolo, Munc13-1, and ELKS) (magenta), reconstructed from ultrathin serial sections. Inset shows a magnified view of a region of dendrite. Modified from (*38*) with permission from Elsevier.

volumetric STORM platform has been developed to determine the entire synaptic fields of neurons (*38*) (Fig. 3E), providing synaptic connectivity at the neural circuit scale.

Protein complexes with structural symmetry

On the shorter length scale of individual protein complexes, it is possible to obtain higher-resolution reconstructions from many super-resolution images through particle averaging in a way that is similar to electron microscopy (EM) reconstruction, especially for structures with well-defined symmetry. Two notable examples are centrioles and nuclear pore complexes (NPCs).

STED and STORM, the latter in combination with particle averaging, have been used to visualize the ninefold symmetry of centrioles (*89*, *90*) (Fig. 4A). In addition to resolving this symmetric arrangement, super-resolution imaging has also been used to map the 3D organization of several centriolar proteins and determine the order of protein recruitment during centriole formation (*91*, *92*). At centrosomes, centrioles are surrounded by the less structured pericentriolar material (PCM), and the radial distribution of proteins within the centrosome and PCM have also been mapped by STORM and SIM (*93*, *94*) (Fig. 4, B and C).

Similarly, STORM imaging showed the eightfold radial symmetry of NPCs (*95*) (Fig. 4D). In combination with particle averaging, STORM allowed the positions of seven nucleoporin components to be determined with ~1-nm precision, which in turn allowed the orientation of the Nup107-160 subcomplex within the pore to be determined (*96*). These super-resolution pictures allowed discrimination between contradictory models of the structural organization of the NPC scaffold (*96*) (Fig. 4E).

Outlook

Super-resolution fluorescence microscopy has transformed understanding of the structure and function of many biological systems. However, challenges are still present, and to maximize the impact of super-resolution microscopy, further technological advancements are still needed.

The spatial resolution achieved by super-resolution microscopy in biological systems typically ranges from 10 to 70 nm, larger than most biomolecules. Achieving true molecular-scale resolution (~1 nm) would allow molecular interactions and conformations to be directly probed inside cells, but remains a challenging task. In principle, the two main categories of optical approaches to overcome the diffraction limit, including the patterned-illumination–based methods represented by STED, RESOLFT, and NL-SIM and the single-molecule-switching–based methods represented by STORM, PALM, and PAINT, can both achieve unlimited high resolution. However, practical factors, such as the requirement of increasing illumination intensity (in the former category) and increasing fluorophore photon budget (in the latter category) for higher resolution, limit the resolution that can be achieved. Reinspection of the fundamental principles of super-resolution methods can lead to powerful new innovations and concepts, as demonstrated by MINFLUX, which combines strengths from both approaches and achieves ultrahigh, previously inaccessible, resolutions. In addition, these methods can be combined with ExM, an orthogonal approach that achieves resolution increase through physical sample expansion, potentially leading to a direct multiplication in the fold increases in resolution that are separately achievable by individual methods. However, it is worth noting that the final image resolution is also limited by probe size and labeling density. Thus, to ultimately benefit from the ultrahigh resolution, parallel development in probes and labeling methods is needed to allow molecules in cells to be labeled with small-molecule probes with high efficiency.

Furthermore, although super-resolution imaging has demonstrated subsecond and even millisecond time resolution in some cases, owing to the trade-off between spatial and temporal resolutions, the limited photon budget of the fluorophores, and phototoxicity to samples, live-cell imaging with high spatiotemporal resolution for a long period of time remains difficult and an active area of development. In addition, in vivo super-resolution imaging deep inside tissues remains challenging, notwithstanding considerable efforts combating tissue-induced background, aberration, and light scattering.

Another challenge, but also an exciting new direction, is to increase the number of molecular species that can be simultaneously imaged. Cells contain thousands of distinct genes and other molecules that act collectively to give rise to behavior and function, yet multicolor imaging usually only allows simultaneous visualization of a few different molecular species. Recent advances have broken new ground in this direction, and genomic-scale imaging is now within reach. For example, single-cell transcriptome-imaging methods have allowed RNAs of 1000 or more genes to be simultaneously imaged in individual cells by using multiplexed fluorescence in situ hybridization (FISH) (*97*, *98*) or in situ sequencing (*99*, *100*). A similar level of multiplexity

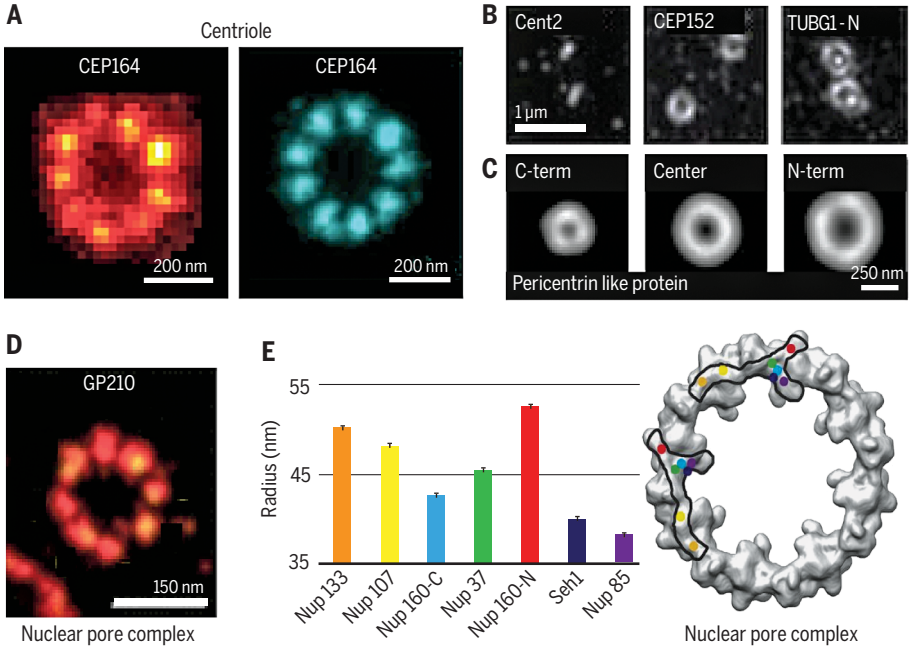


Fig. 4. Super-resolution visualizing molecular complexes: Centriole-containing complexes and the nuclear pore complex. (**A**) Single STED image (left) and particle-averaged STORM image (right) of the centriolar protein (CEP164), showing a radial ninefold symmetry. STED image modified from (*89*) with permission from Elsevier; STORM image modified from (*90*) with permission from Springer Nature. (**B**) Three centriolar and pericentriolar proteins—Cent2, CEP152 and γ-tubulin (TUBG1), imaged using SIM showing a concentric organization of the pericentriolar matrix. Modified from (*93*) with permission from Springer Nature. (**C**) Average distribution of C-terminus, central domain, and N-terminus of the pericentrin-like protein (PLP) demonstrating a radial, spoke-like orientation for PLP through the pericentriolar matrix as determined by SIM. Modified from (*94*) with permission from Springer Nature. (**D**) Single STORM image of the nucleoporin protein GP210 showing an eightfold symmetry within the nuclear pore complex of *Xenopus* oocytes. Modified from (*95*). (**E**) Left: Radial distribution of several nucleoporins including Nup133, Nup107, Nup160 (C-terminus), Nup37, Nup160 (N-terminus), Seh1, and Nup85, that comprise the Y-shaped Nup107-160 complex determined using STORM and particle averaging. Right: A projection of the electron density of the cytoplasmic ring of the NPC determined by EM is overlaid with two possible arrangements of the Nup107-160 complex, determined by super-resolution imaging. Each protein is represented by a colored dot corresponding to the color and radius in the graph (left). Modified from (*96*).

may be achievable for DNA and proteins in the future. Combination of these approaches with super-resolution microscopy could potentially allow genomic-scale super-resolution imaging. Technologically, a major challenge in genomic-scale imaging is molecular crowding, which can prevent resolution of neighboring molecules by conventional imaging, and super-resolution microscopy provides a promising solution. Biologically, the ability to image all molecules in a complex molecular machinery or in a whole signaling pathway, and ultimately at the whole-genome scale, will provide a comprehensive picture of the molecular basis of cellular behavior and function. It is exhilarating to imagine how such a picture of a cell, with all molecules imaged at a resolution that allows direct inference of molecular interactions, would open new opportunities for understanding life at the molecular level.

REFERENCES AND NOTES

1. S. W. Hell, J. Wichmann, *Opt. Lett.* **19**, 780–782 (1994).
2. T. A. Klar, S. W. Hell, *Opt. Lett.* **24**, 954–956 (1999).
3. M. Hofmann, C. Eggeling, S. Jakobs, S. W. Hell, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 17565–17569 (2005).
4. C. Eggeling, K. I. Willig, S. J. Sahl, S. W. Hell, *Q. Rev. Biophys.* **48**, 178–243 (2015).
5. R. Heintzmann, M. G. L. Gustafsson, *Nat. Photonics* **3**, 362–364 (2009).
6. M. J. Rust, M. Bates, X. Zhuang, *Nat. Methods* **3**, 793–796 (2006).
7. E. Betzig *et al*., *Science* **313**, 1642–1645 (2006).
8. S. T. Hess, T. P. Girirajan, M. D. Mason, *Biophys. J.* **91**, 4258–4272 (2006).
9. B. Huang, H. Babcock, X. Zhuang, *Cell* **143**, 1047–1058 (2010).
10. S. J. Sahl, S. W. Hell, S. Jakobs, *Nat. Rev. Mol. Cell Biol.* **18**, 685–701 (2017).
11. W. E. Moerner, M. Orrit, *Science* **283**, 1670–1676 (1999).
12. A. Yildiz *et al*., *Science* **300**, 2061–2065 (2003).
13. A. Pertsinidis, Y. Zhang, S. Chu, *Nature* **466**, 647–651 (2010).
14. A. Sharonov, R. M. Hochstrasser, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 18911–18916 (2006).
15. F. Balzarotti *et al*., *Science* **355**, 606–612 (2017).
16. F. Chen, P. W. Tillberg, E. S. Boyden, *Science* **347**, 543–548 (2015).
17. B. Huang, W. Wang, M. Bates, X. Zhuang, *Science* **319**, 810–813 (2008).
18. M. F. Juette *et al*., *Nat. Methods* **5**, 527–529 (2008).
19. S. R. Pavani *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2995–2999 (2009).
20. G. Shtengel *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 3125–3130 (2009).
21. A. von Diezmann, Y. Shechtman, W. E. Moerner, *Chem. Rev.* **117**, 7244–7275 (2017).
22. R. Schmidt *et al*., *Nat. Methods* **5**, 539–544 (2008).
23. U. Böhm, S. W. Hell, R. Schmidt, *Nat. Commun.* **7**, 10504 (2016).
24. E. H. Rego *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, E135–E143 (2012).
25. D. Li *et al*., *Science* **349**, aab3500 (2015).
26. B. C. Chen *et al*., *Science* **346**, 1257998 (2014).
27. G. T. Dempsey, J. C. Vaughan, K. H. Chen, M. Bates, X. Zhuang, *Nat. Methods* **8**, 1027–1036 (2011).
28. D. Aquino *et al*., *Nat. Methods* **8**, 353–359 (2011).
29. F. Huang *et al*., *Cell* **166**, 1028–1040 (2016).
30. S. Jia, J. C. Vaughan, X. Zhuang, *Nat. Photonics* **8**, 302–306 (2014).
31. J. C. Vaughan, S. Jia, X. Zhuang, *Nat. Methods* **9**, 1181–1184 (2012).
32. M. Dai, R. Jungmann, P. Yin, *Nat. Nanotechnol.* **11**, 798–807 (2016).
33. J. B. Chang *et al*., *Nat. Methods* **14**, 593–599 (2017).
34. S. H. Shim *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 13978–13983 (2012).
35. J. Nixon-Abell *et al*., *Science* **354**, aaf3928 (2016).
36. A. K. Gustavsson, P. N. Petrov, W. E. Moerner, *Opt. Express* **26**, 13122–13147 (2018).
37. V. Gradinaru, J. Treweek, K. Overton, K. Deisseroth, *Annu. Rev. Biophys.* **47**, 355–376 (2018).
38. Y. M. Sigal, C. M. Speer, H. P. Babcock, X. Zhuang, *Cell* **163**, 493–505 (2015).
39. J. Schneider *et al*., *Nat. Methods* **12**, 827–830 (2015).
40. A. Chmyrov *et al*., *Nat. Methods* **10**, 737–740 (2013).
41. S. A. Jones, S. H. Shim, J. He, X. Zhuang, *Nat. Methods* **8**, 499–505 (2011).
42. F. Huang *et al*., *Nat. Methods* **10**, 653–658 (2013).
43. D. Sage *et al*., *Nat. Methods* **12**, 717–724 (2015).
44. S. Manley *et al*., *Nat. Methods* **5**, 155–157 (2008).
45. G. Giannone *et al*., *Biophys. J.* **99**, 1303–1310 (2010).
46. J. Heine *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9797–9802 (2017).
47. K. Xu, G. Zhong, X. Zhuang, *Science* **339**, 452–456 (2013).
48. G. van den Bogaart *et al*., *Nature* **479**, 552–555 (2011).
49. J. Chojnacki *et al*., *Science* **338**, 524–528 (2012).
50. S. B. Van Engelenburg *et al*., *Science* **343**, 653–656 (2014).
51. J.-J. Chung *et al*., *Cell* **157**, 808–822 (2014).
52. J. L. Ptacin *et al*., *Nat. Cell Biol.* **12**, 791–798 (2010).
53. P. Kanchanawong *et al*., *Nature* **468**, 580–584 (2010).
54. L. Große *et al*., *EMBO J.* **35**, 402–413 (2016).
55. R. Salvador-Gallego *et al*., *EMBO J.* **35**, 389–401 (2016).
56. J. B. French *et al*., *Science* **351**, 733–737 (2016).
57. J. Fei *et al*., *Science* **347**, 1371–1374 (2015).
58. J. R. Moffitt, S. Pandey, A. N. Boettiger, S. Wang, X. Zhuang, *eLife* **5**, e13065 (2016).
59. Y. Doksani, J. Y. Wu, T. de Lange, X. Zhuang, *Cell* **155**, 345–356 (2013).
60. A. N. Boettiger *et al*., *Nature* **529**, 418–422 (2016).
61. M. A. Ricci, C. Manzo, M. F. García-Parajo, M. Lakadamyali, M. P. Cosma, *Cell* **160**, 1145–1158 (2015).
62. S.-H. Lee, J. Y. Shin, A. Lee, C. Bustamante, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17436–17441 (2012).
63. M. Renz, B. R. Daniels, G. Vámosi, I. M. Arias, J. Lippincott-Schwartz, *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2989–E2997 (2012).
64. E. M. Puchner, J. M. Walter, R. Kasper, B. Huang, W. A. Lim, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 16015–16020 (2013).
65. H. Ta *et al*., *Nat. Commun.* **6**, 7977 (2015).
66. X. Nan *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18519–18524 (2013).
67. Y. Zhang, M. Lara-Tejero, J. Bewersdorf, J. E. Galán, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 6098–6103 (2017).
68. C. Eggeling *et al*., *Nature* **457**, 1159–1162 (2009).
69. S. J. Sahl, M. Leutenegger, M. Hilbert, S. W. Hell, C. Eggeling, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6829–6834 (2010).
70. J. Tønnesen, V. V. G. K. Inavalli, U. V. Nägerl, *Cell* **172**, 1108–1121.e15 (2018).
71. S. Berning, K. I. Willig, H. Steffens, P. Dibaj, S. W. Hell, *Science* **335**, 551 (2012).
72. G. Zhong *et al*., *eLife* **3**, e04581 (2014).
73. E. D'Este, D. Kamin, F. Göttfert, A. El-Hady, S. W. Hell, *Cell Reports* **10**, 1246–1251 (2015).
74. E. D'Este *et al*., *Sci. Rep.* **6**, 22741 (2016).
75. J. He *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6029–6034 (2016).
76. E. D'Este, D. Kamin, F. Balzarotti, S. W. Hell, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E191–E199 (2017).
77. B. Han, R. Zhou, C. Xia, X. Zhuang, *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6678–E6685 (2017).
78. M. Hammarlund, E. M. Jorgensen, M. J. Bastiani, *J. Cell Biol.* **176**, 269–275 (2007).
79. M. Krieg, A. R. Dunn, M. B. Goodman, *Nat. Cell Biol.* **16**, 224–233 (2014).
80. C. Y. Huang, C. Zhang, D. R. Zollinger, C. Leterrier, M. N. Rasband, *J. Neurosci.* **37**, 11323–11334 (2017).
81. D. Albrecht *et al*., *J. Cell Biol.* **215**, 37–46 (2016).
82. R. J. Kittel *et al*., *Science* **312**, 1051–1054 (2006).
83. K. S. Y. Liu *et al*., *Science* **334**, 1565–1569 (2011).
84. A. Dani, B. Huang, J. Bergan, C. Dulac, X. Zhuang, *Neuron* **68**, 843–856 (2010).
85. H. D. MacGillavry, Y. Song, S. Raghavachari, T. A. Blanpied, *Neuron* **78**, 615–622 (2013).
86. D. Nair *et al*., *J. Neurosci.* **33**, 13204–13224 (2013).
87. A. H. Tang *et al*., *Nature* **536**, 210–214 (2016).
88. B. G. Wilhelm *et al*., *Science* **344**, 1023–1028 (2014).
89. L. Lau, Y. L. Lee, S. J. Sahl, T. Stearns, W. E. Moerner, *Biophys. J.* **102**, 2926–2935 (2012).
90. X. Shi *et al*., *Nat. Cell Biol.* **19**, 1178–1188 (2017).
91. L. Gartenmann *et al*., *Curr. Biol.* **27**, R1054–R1055 (2017).
92. V. Hamel *et al*., *Curr. Biol.* **27**, 2486–2498.e6 (2017).
93. S. Lawo, M. Hasegan, G. D. Gupta, L. Pelletier, *Nat. Cell Biol.* **14**, 1148–1158 (2012).
94. V. Mennella *et al*., *Nat. Cell Biol.* **14**, 1159–1168 (2012).
95. A. Löschberger *et al*., *J. Cell Sci.* **125**, 570–575 (2012).
96. A. Szymborska *et al*., *Science* **341**, 655–658 (2013).
97. K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, *Science* **348**, aaa6090 (2015).
98. S. Shah *et al*., *Cell* **174**, 363–376.e16 (2018).
99. X. Wang *et al*., *Science* **361**, eaat5691 (2018).
100. J. H. Lee *et al*., *Science* **343**, 1360–1363 (2014).

ACKNOWLEDGMENTS

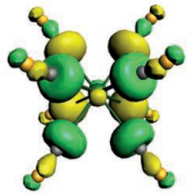
Funding: This work is in part supported by the National Institutes of Health. X.Z. is a Howard Hughes Medical Institute Investigator.
Competing interests: None declared.

10.1126/science.aau1044

RESEARCH

Alkaline earth metals can slide their electrons from s to d orbitals

Wu et al., *p. 912*



IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

Yellow fever is moving from the forests to the cities.

YELLOW FEVER

Arbovirus risk in Brazil

Despite the existence of an effective vaccine for yellow fever, there are still almost 80,000 fatalities from this infection each year. Since 2016, there has been a resurgence of cases in Africa and South America—and this at a time when the vaccine is in short supply. The worry is that yellow fever will spread from the forests to the cities, because its vector, *Aedes* spp. mosquitoes, are globally ubiquitous. Faria *et al.* integrate genomic, epidemiological, and case distribution data from Brazil to estimate patterns of geographic spread, the risks of virus exposure, and the contributions of rural versus urban transmission (see the Perspective by Barrett). Currently, the yellow fever epidemic in Brazil seems to be driven by infections acquired while visiting forested areas and indicates spillover from susceptible wild primates. —CA

Science, this issue p. 894; see also p. 847

INDUCED SEISMICITY

Seismic limits for hard and soft rock

Induced earthquakes from oil, gas, and geothermal energy exploration projects can damage infrastructure and concern the public. However, it remains unclear how far away from an injection site an earthquake can still be triggered. Goebel and Brodsky looked at 18 different earthquake-producing injection sites around the world to address this issue. Injecting fluid into softer layers increased the range for seismic hazard, whereas harder basement rock better confined the fluid. These findings should be considered when regulating and managing projects with the potential to induce seismicity. —BG

Science, this issue p. 899

SOLAR CELLS

Perovskite/CIGS tandem cells

Tandem solar cells can boost efficiency by using more of the available solar spectrum. Han *et al.* fabricated a two-terminal tandem cell with an inorganic-organic hybrid perovskite top layer and a Cu(In,Ga)Se_2 (CIGS) bottom layer. Control of the roughness of the CIGS surface and the use of a heavily doped organic hole transport layer were crucial to achieve a 22.4% power conversion efficiency. The unencapsulated tandem cells maintained almost 90% of their efficiency after 500 hours of operation under ambient conditions. —PDS

Science, this issue p. 904

GENOME STRUCTURE

The structure of the genome

Beyond the sequence of the genome, its three-dimensional structure is important in regulating gene expression. To understand cell-to-cell variation, the structure needs to be understood at a single-cell level. Chromatin conformation capture methods have allowed characterization of genome structure in haploid cells. Now, Tan *et al.* report a method called Dip-C that allows them to reconstruct the genome structures of single diploid human cells. Their examination of different cell types highlights the tissue dependence of three-dimensional genome structures. —VV

Science, this issue p. 924

PHOTO: AMI VITALE/GETTY IMAGES; GRAPHIC: *WU ET AL.*

Published by AAAS

ECOLOGY

Fisheries management and human adaptation

Finding effective ways to mitigate the future impacts of climate change on fisheries is critical. Previous efforts have not incorporated alternative human responses to climate change. These could limit or exacerbate ecosystem changes that affect fish stocks and population locations. Gaines *et al.* analyzed four fisheries management approaches that address fish stock–productivity adaptation and/or range-shift adaptation. They then applied these management scenarios to 915 species stocks worldwide. Implementing proactive and adaptive fishery management approaches would bring about higher global profits (154%), harvest (34%), and biomass (60%) as compared with strategies reflecting no management changes. Addressing both range shifts and productivity changes lead to greater benefits as compared with targeting one challenge alone. —PJB

Sci. Adv. 10.1126/sciadv.aao1378 (2018).

PAIN

A dual-targeting painkiller

Opioids are among the most effective treatments for severe pain. Their pain-relieving effects are mediated by activation of the mu opioid receptor (MOR). Unfortunately, selective MOR agonists induce diverse side effects, including respiratory depression, tolerance, hyperalgesia, and dependence. Recently, activation of the nociceptin/orphanin FQ peptide receptor (NOR) has been reported to enhance MOR agonist–induced analgesia without producing side effects. Ding *et al.* developed a bifunctional MOR/NOR agonist called AT-121, which showed potent analgesic effects in nonhuman primates without inducing hyperalgesia, respiratory depression, or dependence. Bifunctional MOR/NOR agonists might thus represent a safe and effective pharmacological tool for treating severe pain. —MM

Sci. Transl. Med. **10**, eaar3483 (2018).

MOLECULAR MOTORS

Tiny cargos ferried along a track

Control of molecules at the nanometer scale requires motors that convert potential energy into movement. Qing *et al.* designed a small molecule that could hop along a track of cysteine residues within a membrane-embedded protein pore. The direction of processive movement along the track was reversible, driven by an applied potential across the membrane. Cargos were attached to a carrier motor, and their position and chemical identity read out from changes in the current through the pore. These features enabled repeat observations of a single molecule as it moved back and forth on the track. —MAF

Science, this issue p. 908

CLIMATE CHANGE

Warming, crops, and insect pests

Crop responses to climate warming suggest that yields will decrease as growing-season temperatures increase. Deutsch *et al.* show that this effect may be exacerbated by insect pests (see the Perspective by Riegler). Insects already consume 5 to 20% of major grain crops. The authors' models show that for the three most important grain crops—wheat, rice, and maize—yield lost to insects will increase by 10 to 25% per degree Celsius of warming, hitting hardest in the temperate zone. These findings provide an estimate of further potential climate impacts on global food supply and a benchmark for future regional and field-specific studies of crop-pest-climate interactions. —AMS

Science, this issue p. 916; see also p. 846

IN OTHER JOURNALS

Edited by **Caroline Ash** *and* **Jesse Smith**

Conifer pollen from Yellowstone lake sediments reveals vegetation shifts over 18,000 years.

PALEOECOLOGY

Climate change in a mountain ecosystem

Fossil pollen in lake sediments provides valuable records of past vegetation patterns and offers a baseline for assessing how vegetation responds to climate change. To assess how the vegetation composition and distribution of a mountain system has varied with climate change over 18,000 years, Iglesias *et al.* studied pollen sequences from lakes in the Greater Yellowstone Ecosystem of the United States. They found complex patterns, with long-term stability in some plant communities and rapid change in others. The present-day mixed conifer forest cover, known to be vulnerable to climate warming, is now more compressed in its elevation range than in previous postglacial millennia. These data provide a context for assessing future responses to climate change. —AMS

J. Biogeogr. **45**, 1768 (2018).

NEUROSCIENCE

Degrees of stress in neurodegeneration

In the neurodegenerative disorder amyotrophic lateral sclerosis (ALS), the nuclear protein called transactive response DNA binding protein of 43 kDa (TDP-43) accumulates in stress granules within the cytoplasm of neurons and glia and is linked to disease pathology. McGurk *et al.* report that TDP-43 binds to poly(ADP-ribose) (PAR), which triggers phase separation of TDP-43 and its subsequent recruitment to stress granules. Under short-term stress, phosphorylated TDP-43, which is considered a hallmark of disease, is unexpectedly excluded

PHOTO: DESIGN PICS/DAVID PONTON/GETTY IMAGES

Exposure to disruptive behavior, like bullying, during primary school can have lifelong effects.

EDUCATION

Disruptive classmates, long-term harm

Children who are behaviorally disruptive during primary school can have harmful impacts on their classmates into adulthood. Carrell *et al.* use data from Florida, USA, to show that a child who experiences domestic violence at home (a well-recognized proxy for that child demonstrating disruptive behavior such as bullying) can lower their classmates' secondary-school math and reading test scores, lower their likelihood of enrolling in college, and reduce earnings in their mid-20s by 3%. Differential exposure to such classmates accounts for roughly 5% of the rich-poor earnings gap in adulthood. —BW

Amer. Econ. Rev. www.aeaweb.org/articles?id=10.1257/aer.20160763 (2018).

from stress granules. This finding indicates that the granules initially prevent phosphorylated TDP-43 aggregation unless stress is prolonged. This work also points to an approach to ALS treatment by inhibition of PAR polymerase (PARP) to reduce PAR production. For instance, a small-molecule inhibitor of PARP that prevents cancer-cell proliferation also blocks cytoplasmic TDP-43 aggregation. —LC

Mol. Cell 10.1016/j.molcel.2018.07.002 (2018).

REPROGRAMMING

FACTs behind control of cell fate

As animals develop, their cells become progressively less plastic and follow defined functional destinies. Kolundzic *et al.* used a genetic screen of the worm *Caenorhabditis elegans* to uncover proteins that prevent cells from straying from their intended fate. They found that the histone chaperone FACT plays a regulatory part in an unexpected way: It is non-repressive and also promotes gene expression. FACT acts as a barrier to cell reprogramming by stabilizing gene expression and thereby safeguarding cell identity. A germline-specific isoform of FACT ensures that cells with intestinal and germline programming confirm their fate and do not adopt a neuronal role. Furthermore, depletion of FACT in human fibroblasts enhances production of induced pluripotent stem cells, indicating that a conserved mechanism is at work to channel cell fate in animals. —BAP

Dev. Cell 10.1016/j.devcel.2018.07.006 (2018).

FRAMEWORK MATERIALS

Transversal zigzag linkers

The linkers for metal-organic frameworks are usually bidentate molecules (for example, dicarboxylic acids) connected by an organic group to create a linear or, in some cases, a bent geometry like isophthalic acid. Guillerm *et al.* explored the effect of a "zigzag" linker, *trans,trans*-muconic acid (tmuc), that forces an offset of inorganic building blocks. Reaction with $ZrCl_4$ formed the metal-organic framework $Zr_6O_4(OH)_4(tmuc)_8(H_2O)_8$, which had an eight-connected **bcu** topology, a subset of the 12-connected **fcu** topology seen with linear linkers. This **bcu** topology was maintained with linkers of even larger transversal width, such as azobenzene-3,3′-dicarboxylic acid. —PDS

J. Am. Chem. Soc. **140**, 10153 (2018).

MICROBIOLOGY

Impermanent permafrost

Permafrost constitutes a quarter of Earth's surface and about half the buried ancient carbon. Thaw releases water, and, together with higher temperatures, this promotes microbial respiration. Thus, permafrost melt during global warming represents a threat for escalating greenhouse gas release. Müller *et al.* extracted 2-meter core samples from Svalbard permafrost in Norway for 16*S* ribosomal RNA gene analysis. Sampling at 3-centimeter intervals, they noted distinctive strata of microbial communities. On thawing and subsequent incubation, each community showed different metabolic rates and different CO_2 fluxes. Within 24 hours, thawing the deepest permafrost layers released most CO_2, but over a longer term, most CO_2 was produced under shallow aerobic conditions. These Svalbard mineral soils also have high iron availability. Intimate knowledge of the microbial, as well as the physicochemical, conditions prevailing in any specific permafrost area is needed to accurately estimate CO_2 emission during anthropogenic climate warming. —CA

Environ. Microbiol. 10.1111/1462-2920.14348 (2018).

QUANTUM COMPUTATION

Trapped ions tackle chemistry

Some of the most likely first applications of future quantum computers may be in quantum chemistry. Even with currently available quantum computers consisting of just a few qubits, it is possible to address certain simple problems, but most of the development has occurred in systems using superconducting qubits. Hempel *et al.* used up to four qubits encoded by trapped ions to calculate the ground-state energies of two simple molecules, H_2 and LiH. They made use of a hybrid classical-quantum method called the variational quantum eigensolver, which relegates parts of the computation such as preprocessing and optimization to a classical computer. —JS

Phys. Rev. X **8**, 031022 (2018).

PHOTO: FATCAMERA/GETTY IMAGES

ALSO IN *SCIENCE* JOURNALS

Edited by **Stella Hurtley**

CANCER

Looping together genes in cancer

A subset of human cancers are characterized by aberrant fusion of two specific genes. In some cases, the activity of the resultant fusion protein drives tumor growth. Most fusion genes in cancer appear to arise from simple reciprocal chromosomal translocations. Anderson *et al.* found that the characteristic fusion gene in a bone and soft tissue tumor called Ewing sarcoma is produced by a far more complicated mechanism (see the Perspective by Imielinski and Ladanyi). In nearly half of the tumors examined, the fusion gene was created by the formation of dramatic genomic loops that disrupt multiple genes. These complex rearrangements occur in early replicating and transcriptionally active regions of the genome and are associated with poor prognosis. —PAK

Science, this issue p. 891; see also p. 848

SIGNAL TRANSDUCTION

Dynamics of cell signaling and decoding

Defects in cellular signaling pathways, like those in some cancer cells, are often thought to result in increased or decreased steady-state signals that promote or inhibit cell proliferation. But Bugaj *et al.* show that dynamic changes in the duration or frequency of a signal can also alter cellular responses (see the Perspective by Kolch and Kiel). They took precise control of signaling in cultured human or mouse cells with a light-controlled mechanism for activating and inactivating the guanosine triphosphate Ras. Known cancer mutations in components of the Ras-activated signaling pathway or inhibitors of particular pathway components altered signal timing and readouts. The modified dynamics changed transcriptional outcomes and could inappropriately support cell proliferation. The ability to probe responses of signaling networks in this way may enhance understanding of biological regulation and reveal new therapeutic targets. —LBR

Science, this issue p. 892; see also p. 844

BIOTECHNOLOGY

Lineage tracing in mouse using CRISPR

A homing guide RNA (hgRNA) that directs CRISPR-Cas9 to its own DNA locus can diversify its sequence and act as an expressed genetic barcode. Kalhor *et al.* engineered a mouse line carrying 60 independent loci of hgRNAs, thus generating a large number of unique barcodes in various embryonic and extraembryonic tissues in fully developed mice. This method demonstrates lineage tracing from the very first branches of the development tree up to organogenesis events and was used to elucidate embryonic brain patterning. —SYM

Science, this issue p. 893

ORGANOMETALLICS

Carbonyls in the s block

Conventional wisdom in chemistry distinguishes transition metals from other elements by their use of d orbitals in bonding. Wu *et al.* now report that alkaline earth metals can slide their electrons from s- to d-orbital bonding motifs as well (see the Perspective by Armentrout). Calcium, strontium, and barium all form coordination complexes with a cubic arrangement of eight carbonyl ligands and an 18-electron valence shell. The compounds were characterized in frozen neon matrices by vibrational spectroscopy and in gas phase by mass spectrometry. —JSY

Science, this issue p. 912; see also p. 849

CLIMATE CHANGE

Future predictions from paleoecology

Terrestrial ecosystems will be transformed by current anthropogenic change, but the extent of this change remains a challenge to predict. Nolan *et al.* looked at documented vegetational and climatic changes at almost 600 sites worldwide since the last glacial maximum 21,000 years ago. From this, they determined vegetation responses to temperature changes of 4° to 7°C. They went on to estimate the extent of ecosystem changes under current similar (albeit more rapid) scenarios of warming. Without substantial mitigation efforts, terrestrial ecosystems are at risk of major transformation in composition and structure. —AMS

Science, this issue p. 920

EVOLUTION

Venoms yield their secrets

Venoms can be deadly but, in the right hands, can perform important therapeutic functions. Research into potential drug leads from venoms has, however, been hampered by an inability to study venoms from animals that are small, rare, or difficult to maintain in the laboratory. In a Perspective, Holford *et al.* highlight recent studies that have used genomics and other omics technologies to study venoms from a wide range of organisms, shedding light on the evolutionary biology of venoms. The work is also providing important leads for the development of therapeutics and eco-friendly insecticides. —JFU

Science, this issue p. 842

CELL DEATH

Casting NETs

Gasdermin D (GSDMD), a pore-forming protein, has emerged as a key downstream effector in pyroptosis, a form of cell death induced by intracellular lipopolysaccharide. Sollberger *et al.* found that GSDMD was activated in neutrophils during the generation of neutrophil extracellular traps (NETs). NETs are composed of chromatin and antimicrobial proteins and are cast by dying neutrophils in a process termed NETosis. While carrying out a chemical screen to identify molecules that block NETosis, the authors identified a pyrazolo-oxazepine scaffold–based molecule that binds GSDMD. Chen *et al.* also report a role for GSDMD in NETosis, and Rathkey *et al.* report necrosulfonamide to be an inhibitor of GSDMD. —AB

Sci. Immunol. **3**, eaar6689, eaar6676, eaat2738 (2018).

GPCR SIGNALING

Melatonin meets diabetes

Some of the single-nucleotide polymorphisms associated with type 2 diabetes occur in the gene that encodes the melatonin receptor MT_2, a G protein–coupled receptor (GPCR). Karamitri *et al.* measured the spontaneous and melatonin-stimulated signaling of 40 different MT_2 variants. MT_2 variants with defective melatonin-stimulated G protein signaling and reduced spontaneous β-arrestin recruitment were associated with the greatest risk for type 2 diabetes. These data may aid in developing specific type 2 diabetes treatments based on a patient's MT_2 variant. —JFF

Sci. Signal. **11**, eaan6622 (2018).

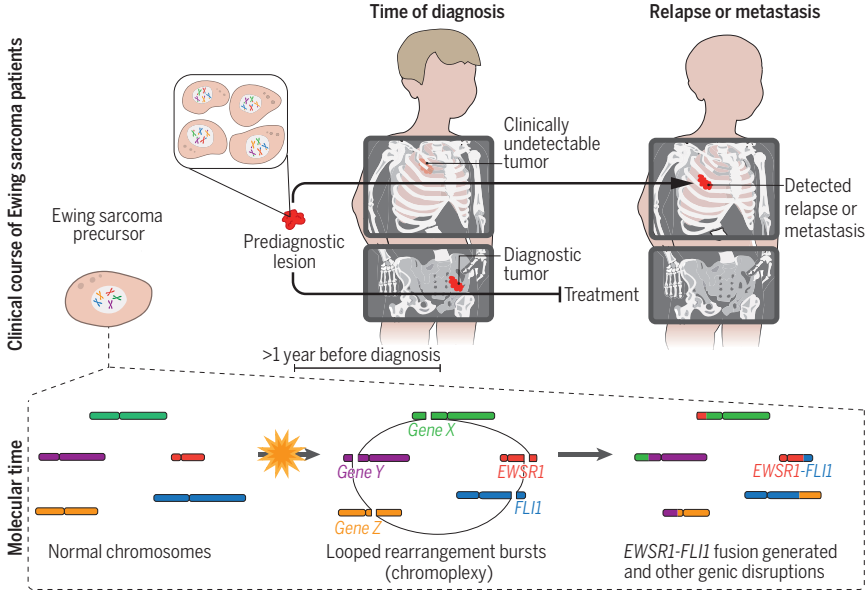
RESEARCH ARTICLE SUMMARY

CANCER

Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors

Nathaniel D. Anderson, Richard de Borja*, Matthew D. Young*, Fabio Fuligni*, Andrej Rosic, Nicola D. Roberts, Simon Hajjar, Mehdi Layeghifard, Ana Novokmet, Paul E. Kowalski, Matthew Anaka, Scott Davidson, Mehdi Zarrei, Badr Id Said, L. Christine Schreiner, Remi Marchand, Joseph Sitter, Nalan Gokgoz, Ledia Brunga, Garrett T. Graham, Anthony Fullam, Nischalan Pillay, Jeffrey A. Toretsky, Akihiko Yoshida, Tatsuhiro Shibata, Markus Metzler, Gino R. Somers, Stephen W. Scherer, Adrienne M. Flanagan, Peter J. Campbell, Joshua D. Schiffman, Mary Shago, Ludmil B. Alexandrov, Jay S. Wunder, Irene L. Andrulis, David Malkin†, Sam Behjati†, Adam Shlien†

INTRODUCTION: Gene fusions are often disease-defining events in cancer. The mutational processes that give rise to fusions, their timing relative to initial diagnosis, and whether they change at relapse are largely unknown. Mutational processes leave distinct marks in the tumor genome, meaning that DNA sequencing can be used to reconstruct how fusions are generated. A prototypical fusion-driven tumor is Ewing sarcoma (ES), a bone cancer predominantly affecting children and young adults. ES is defined by fusions involving *EWSR1*, a gene encoding an RNA binding protein, and genes encoding E26 transformation-specific (ETS) transcription factors such as *FLI1*. We sought to reconstruct the genomic events that give rise to *EWSR1-ETS* fusions in ES and chart their evolution from diagnosis to relapse.



Timing of mutations in a patient with ES. The schematic shows genetic alterations in tumors at prediagnosis, diagnosis, and relapse. In many cases, the fusion gene that drives tumorigenesis (*EWSR1-FLI1* or *EWSR1-ERG*) emerges via a sudden burst of genomic rearrangements involving multiple chromosomes and genes. This event, called chromoplexy (indicated by the starburst), happens early in the evolution of the disease in a prediagnostic lesion. After this event, the diagnostic and relapsed tumors evolve in parallel. In this model, the clone that would ultimately become the relapsed tumor was already present at the time of initial diagnosis, although it was undetectable.

RATIONALE: We studied the processes underpinning gene fusions in ES using the whole-genome sequences of 124 primary tumors. We determined the timing of the emergence of *EWSR1* fusions relative to other mutations. To measure ongoing mutation rates and evolutionary trajectories of ES, we studied the genomes of primary tumors, tumors at relapse, and metastatic tumors.

RESULTS: We found that *EWSR1-ETS*, the key ES fusion, arises in 42% of cases via complex, loop-like rearrangements called chromoplexy, rather than by simple reciprocal translocations. Similar loops forming canonical fusions were found in three other sarcoma types. Timing the emergence of loops revealed that they occur as bursts in early replicating DNA, as a primary event in ES development. Additional gene disruptions are generated concurrently with the fusions within the loops. Chromoplexy-generated *EWSR1* fusions appear to be associated with an aggressive form of the disease and a higher chance of relapse. Numerous mutations present in every cell of the primary were absent at relapse, demonstrating that the primary and relapsed diseases evolved independently. This divergence occurs after formation of an ancestral clone harboring *EWSR1* fusions. Importantly, we determined that divergence of the primary tumor and the future relapsed tumor occurs 1 to 2 years before initial diagnosis, as estimated from the number of cell division–associated mutations.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aam8419

CONCLUSION: Our findings provide insights into the pathogenesis and natural history of human sarcomas. They reveal complex DNA rearrangements to be a mutational process underpinning gene fusions in a large proportion of ES. Similar observations in other fusion-defined sarcoma types indicate that this process operates more generally. Such complex rearrangements occur preferentially in early replicating and transcriptionally active genomic regions, as evidenced by the additional genes disrupted. *EWSR1* fusions arising from chromoplexy correlated with worse clinical outcomes. Formation of the *EWSR1* fusion genes is a primary event in the life history of ES. We found evidence of a latency period between this seeding event and diagnosis. This is in keeping with the often-indolent nature of symptoms before clinical disease presentation. ■

The list of author affiliations is available in the full article online.
*These authors contributed equally to this work.
†Corresponding author. Email: adam.shlien@sickkids.ca (A.S.); sb31@sanger.ac.uk (S.B.); david.malkin@sickkids.ca (D.M.)
Cite this article as N. D. Anderson *et al.*, *Science* **361**, eaam8419 (2018). DOI: 10.1126/science.aam8419

RESEARCH ARTICLE

CANCER

Rearrangement bursts generate canonical gene fusions in bone and soft tissue tumors

Nathaniel D. Anderson[1,2], Richard de Borja[1]*, Matthew D. Young[3]*, Fabio Fuligni[1]*, Andrej Rosic[1], Nicola D. Roberts[3], Simon Hajjar[1]†, Mehdi Layeghifard[1], Ana Novokmet[1]†, Paul E. Kowalski[1], Matthew Anaka[1], Scott Davidson[4], Mehdi Zarrei[5], Badr Id Said[1], L. Christine Schreiner[1], Remi Marchand[1], Joseph Sitter[1], Nalan Gokgoz[6], Ledia Brunga[1], Garrett T. Graham[7], Anthony Fullam[3], Nischalan Pillay[8,9], Jeffrey A. Toretsky[7], Akihiko Yoshida[10], Tatsuhiro Shibata[11,12], Markus Metzler[13], Gino R. Somers[2,14], Stephen W. Scherer[1,5,15,16], Adrienne M. Flanagan[9,10], Peter J. Campbell[3,17], Joshua D. Schiffman[18], Mary Shago[2,4], Ludmil B. Alexandrov[19], Jay S. Wunder[20,21], Irene L. Andrulis[6,15], David Malkin[1,22,23]‡, Sam Behjati[3,24]‡, Adam Shlien[1,2,4]‡

Sarcomas are cancers of the bone and soft tissue often defined by gene fusions. Ewing sarcoma involves fusions between *EWSR1*, a gene encoding an RNA binding protein, and E26 transformation-specific (ETS) transcription factors. We explored how and when *EWSR1-ETS* fusions arise by studying the whole genomes of Ewing sarcomas. In 52 of 124 (42%) of tumors, the fusion gene arises by a sudden burst of complex, loop-like rearrangements, a process called chromoplexy, rather than by simple reciprocal translocations. These loops always contained the disease-defining fusion at the center, but they disrupted multiple additional genes. The loops occurred preferentially in early replicating and transcriptionally active genomic regions. Similar loops forming canonical fusions were found in three other sarcoma types. Chromoplexy-generated fusions appear to be associated with an aggressive form of Ewing sarcoma. These loops arise early, giving rise to both primary and relapse Ewing sarcoma tumors, which can continue to evolve in parallel.

Genomic rearrangements (structural variants) are a ubiquitous source of somatic mutation in human cancer. They arise from breaks in chromosomes, which are then aberrantly rejoined. Rearrangements may occur in isolation or in the context of complex genomic catastrophes that shatter chromosomes [chromothripsis (*1*)] or join chromosomes in chains or loop structures [chromoplexy (*2*)]. Rearrangements can generate cancer-driving mutations through several mechanisms, including the formation of gene fusions. Typically, fusions are fashioned by translocations that are often reciprocal. An exception is the prostate cancer fusion gene, *TMPRSS2-ERG*, which can occur through chromoplexy (*2*).

Oncogenic gene fusions are particularly common in leukemia and bone and soft tissue tumors (*3*), often acting as the sole driver mutation and delineating clinically relevant tumor entities and subgroups. In leukemia, recombination-activating gene (RAG)–mediated structural variation has been identified as the leading mutational process that creates canonical gene fusions and drives oncogenesis through translocations and deletions (*4*). Here we sought to investigate processes and timing of oncogenic fusions in human bone and soft tissue tumors.

The starting point of our investigation was Ewing sarcoma (ES), a bone and soft tissue cancer predominantly diagnosed in adolescents and young adults. It represents the prototypical fusion-driven sarcoma, defined by fusions between *EWSR1*, a gene encoding an RNA binding protein, and genes encoding E26 transformation-specific (ETS) transcription factors, including *FLI1* and *ERG* (*5*). Although the downstream consequences of *EWSR1-ETS* fusion genes are well established (*6*), the timing and mechanism by which they arise are unknown.

Burden and signatures of substitution mutations in ES

We sequenced either the gene-containing portions or the whole genomes of 50 ES tumors and their matched normal DNA (complete sequencing details in table S1). We used a conventional analysis pipeline to call somatic substitutions and rearrangements, with additional custom software to remove recurrent artifacts and sources of false positives (see methods and fig. S1). Overall, and consistent with previous reports (*7*–*10*), the ES genome is genetically quiet, with few somatic substitutions identified (median of <1 mutation/Mb; Fig. 1A).

We next asked if the collection of all mutations, when considered together, could help highlight consistent mutagenic processes in ES. We extracted mutational signatures using an established method that allows for the discovery of new signatures. Despite the young age of patients with ES, and overall low number of mutations, we found that the tumors contained at least seven distinct signatures, all of which matched patterns seen in adult cancers [Catalogue of Somatic Mutations in Cancer (COSMIC) signatures 1, 2, 5, 8, 13, 18, and 31; Fig. 1B and fig. S2A] (*11*, *12*). Two of these (signatures 1 and 5) were nearly universal and associated with patient age. Signature 1 generated a steady rate of seven mutations per gigabase per year, which is similar to that of adult ovarian and breast cancers (fig. S2B) (*13*). An overview of the somatic architecture and mutational signatures of each tumor in our discovery cohort is shown in Fig. 1, A to C (left panels, Toronto cohort).

Chromoplexy rearrangement loops are common in aggressive ES

Having observed few small mutations, we then focused our attention on structural rearrangements. We applied a bespoke analysis tool to detect clustered rearrangements from whole-genome data, defined as having an inter-rearrangement

[1]Program in Genetics and Genome Biology, The Hospital for Sick Children, Toronto, Ontario, Canada. [2]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. [3]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, UK. [4]Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada. [5]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, Ontario, Canada. [6]Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Toronto, Ontario, Canada. [7]Department of Oncology and Pediatrics, Georgetown University, Washington, DC, USA. [8]University College London Cancer Institute, Huntley Street, London, UK. [9]Histopathology, Royal National Orthopaedic Hospital NHS Trust, Stanmore, Middlesex, UK. [10]Department of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tokyo, Japan. [11]Division of Cancer Genomics, National Cancer Center Research Institute, Tokyo, Japan. [12]Laboratory of Molecular Medicine, Human Genome Center, The Institute of Medical Sciences, The University of Tokyo, Minato-ku, Tokyo, Japan. [13]Department of Pediatrics and Adolescent Medicine, University Hospital Erlangen, Erlangen, Germany. [14]Department of Pathology, Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada. [15]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [16]The McLaughlin Centre, University of Toronto, Toronto, Ontario, Canada. [17]Department of Haematology, University of Cambridge, Cambridge, UK. [18]Departments of Pediatrics and Oncological Sciences, Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, USA. [19]Department of Cellular and Molecular Medicine and Department of Bioengineering and Moores Cancer Center, University of California, La Jolla, San Diego, CA, USA. [20]University Musculoskeletal Oncology Unit, Mount Sinai Hospital, Toronto, Ontario, Canada. [21]Division of Orthopaedic Surgery, Department of Surgery, University of Toronto, Toronto, Ontario, Canada. [22]Division of Hematology-Oncology, The Hospital for Sick Children, Toronto, Ontario, Canada. [23]Department of Pediatrics, University of Toronto, Ontario, Canada. [24]Department of Paediatrics, University of Cambridge, Cambridge, UK.

*These authors contributed equally to this work. †Deceased

‡Corresponding author. Email: adam.shlien@sickkids.ca (A.S.); sb31@sanger.ac.uk (S.B.); david.malkin@sickkids.ca (D.M.)

distance of <10 kilobase pairs (kbp) (see methods). Using a computational data structure that modeled adjacent breakpoints as vertices and interconnected rearrangements as edges in a graph, we uncovered several distinct configurations of rearrangement clusters (Fig. 1D). As expected, one configuration of rearrangement clusters was a result of reciprocal rearrangements, where there is an equal exchange of genetic material and overlapping breakpoints. These were isolated rearrangements that occurred without additional breakpoints nearby. A second configuration, seen in 14/24 tumor genomes, was a distinctive pattern of focal clustered events with nearly overlapping junctions, organized as closed loops (distance of <30 bp; Fig. 1D, red distribution). That is, if one follows these complex rearrangements across their multiple constituent chromosomes, one is ultimately brought back to the point of departure. Importantly, the loops were nearly always centered on *EWSR1-ETS* (Fig. 1E). These abutting rearrangements that occur in a loop resemble a pattern of chromoplexy, akin to the loops of the prostate cancer fusion gene, *TMPRSS2-ERG*. Of note, the *EWSR1-ERG* fusion was always generated by a complex mechanism, whereas *EWSR1-FLI1* arose with or without this mechanism (fig. S3A). This is likely due to the opposite gene orientation of *EWSR1* relative to *ERG* on their respective chromosome arms. A simple two-chromosome break rearrangement cannot place the genes in the correct transcriptional orientation, necessitating more complex chromosomal rearrangements for fusion formation. Other than this difference, chromoplexy in *ERG*- and *FLI1*-driven tumors was very similar (fig. S3B).

In all cases, we resolved the breakpoints and found positions largely consistent with type I or type II ES (*14*). In the most complex case of chromoplexy, up to 18 genes were brought together with the canonical fusion on the same derivative chromosome (fig. S3C; the full list of genes affecting all samples is shown in fig. S3D). We validated chromoplectic looped rearrangements by deep sequencing or by cytogenetic analysis using standard G-banding and spectral karyotyping (methods and fig. S4). Using RNA sequencing, we found that chromoplectic loops universally disrupted the reciprocal fusion (*FLI1-EWSR1*); 52% of the cancers with simple rearrangements expressed the reciprocal fusion, but none of the chromoplectic tumors expressed it (n = 27, fig. S5). For further validation of chromoplexy in ES, we reanalyzed a published, independent cohort of 100 ES patients, whose genomes had been sequenced, using our informatics pipeline (*10*). The somatic architecture and mutational signature of the validation cohort is shown in Fig. 1, A to C (right panels, validation cohort). Both cohorts harbored copy number profiles consistent with previous reports (fig. S6) (*10*). With this series, the aggregated prevalence of chromoplectic *EWSR1-ETS* gene fusions was 42% (52/124).

Patients with relapsed ES have a poor survival rate, and new prognostic markers are needed. We evaluated the association between chromoplexy, patient outcomes, and known markers of worse prognosis. We found that higher overall genomic complexity, a marker of aggressive ES (*10*, *15*), was almost completely explained by chromoplectic rearrangements (Fig. 1F). By contrast, there was no difference in the burden of nonchromoplectic rearrangements. Similarly, *TP53* mutations, another established marker of poor prognosis (*10*, *16*), were enriched in chromoplexy ES (16 versus 3%, $P < 0.05$, Fisher's exact test). There was no enrichment for *CDKN2A* or *STAG2* mutations (fig. S7). Finally, and consistent with the above, patients with chromoplexy ES were more likely to relapse (54 versus 30%, $P < 0.05$, Fisher's exact test), strongly suggesting that it marks a more aggressive variant of ES.

Chromoplexy generates the key fusion in other bone and soft tissue tumors

We next widened our search across four different benign and malignant bone and soft tissue tumor types for which canonical gene fusions have been identified (table S2). We subjected 13 tumors to high- or low-coverage whole-genome sequencing, and RNA sequencing where feasible. In three tumor types—chondromyxoid fibroma, synovial sarcoma, and phosphaturic mesenchymal tumors—we found that chromoplectic rearrangements (occurring in a similar looped formation) did indeed generate canonical gene fusions (Fig. 2). Furthermore, in one of the chondromyxoid fibroma cases, the fusion emerged from chromothripsis across seven different chromosomes (fig. S8, CMF 2). Chromothripsis was seen in five ES cases, of which four involved the canonical fusion. Taken together, these findings in human bone and soft tissue tumors show that canonical gene fusions are frequently caused by complex rearrangement processes, predominantly chromoplexy but also chromothripsis.

We examined the microanatomy of chromoplexy fusion loops at base-pair resolution, comparing ES to a published series of prostate cancers (*2*). *EWSR1-ETS* ES loops were less complex than *TMPRSS2-ERG* prostate cancer loops, with fewer rearrangements and individual loops involved in their generation (2 to 10 rearrangements in 1 to 2 loops compared with up to 130 rearrangements in up to 25 loops in prostate cancer). This may be a consequence of the ES genome having a shorter time frame to mutate compared to the prostate cancer genome. Consistent with this proposition, multiple independent chromoplexy loops can exist in older prostate cancers, compared to the one simple loop seen in ES (*17*). In contrast to ES, where chromoplexy is virtually synonymous with the disease-defining fusion, several chromoplexy fusion loops occur in prostate cancer without necessarily forming the *TMPRSS2-ERG* fusion. When a loop was present in ES, it almost always generated the *EWSR1-ETS* fusion (47/52 cases, 90%) (Fig. 3, A and B, and figs. S9 and S10).

Transcriptional disruptions are associated with chromoplexy

These loops also led to targeted disruptions or fusions between genes brought together directly through chromoplexy (n = 168 gene disruptions, and n = 47 fusions; Fig. 3C). Given that chromoplexy appeared to mark an aggressive form of ES, we wondered if its gene expression program was globally different—above and beyond the immediate, focal structural consequences listed above. We identified 504 differentially expressed genes in chromopletic ES compared to simple ES ($P < 0.001$, Student's t test; Fig. 3D). Gene set enrichment analysis of well-curated pathways (*18*) uncovered a significant enrichment of dysregulated genes in established cancer hallmark pathways (table S3).

Both prostate cancer and ES loops were characterized by focal intrachromosomal rearrangements, so-called deletion bridges (*2*), that acted as local mediators of large-scale loops (illustrated in fig. S11). We found deletion bridges in ~60% (30/52) of chromoplectic ES. In contrast to the bridges observed in prostate cancer, more than a third of bridges are utilized in ES in a highly consistent manner. That is, if a deletion bridge was found in one component of the loop, it would occur on all chromosomes. For example, in sample 2226, we observed 13 rearrangements spanning three chromosomes, all of which involved deletion bridges. These bridged chromoplectic rearrangements fused *EWSR1-FLI1* and disrupted the neighboring gene, *AP1B1*, as well as the known cancer gene *ARID1B*. Thus, deletion bridges can create further oncogenic disruptions.

We also observed a remarkable pattern of splicing, whereby the transcriptional machinery further refined the looped rearrangements found in the genome. In chondromyxoid fibromas with chromoplectic *GRM1* fusions (3/4 cases), the rearrangement breakpoint did not actually reside within the *GRM1* gene body. Rather, the breakpoint was found in the upstream gene *SHPRH* within a narrow window (fig. S8). Thus, chromoplexy together with conventional splicing leads to the promoter swap that is characteristic of this tumor [see (*19*)]. Interestingly, we also observed the transcriptional generation of gene fusions in ES. Examination of the transcriptomic consequences of loops showed that genes that were unconnected at the DNA level were brought together, in cis, at the mRNA level. This included examples of the *EWSR1-ETS* fusion itself (Fig. 3D and fig. S12). In the cases reported here, no direct rearrangement links *EWSR1* and *FLI1*; however, they are linked via two rearrangements to a third locus. In this way, chromoplexy generates the canonical driver gene through a chromoplexy scaffolding event.

Chromoplexy occurs early in tumor evolution through a replication-associated mechanism

We next studied the timing of chromoplexy rearrangements in tumor evolution. Chromoplexy may arise from a one-off sudden event, generating many breakpoints simultaneously, or through stepwise progressive bursts of mutations in succession (*2*). To differentiate between these two modes of evolution, we used DNA copy number profiling associated with the breakpoints of chromoplexy rearrangements to assess the copy

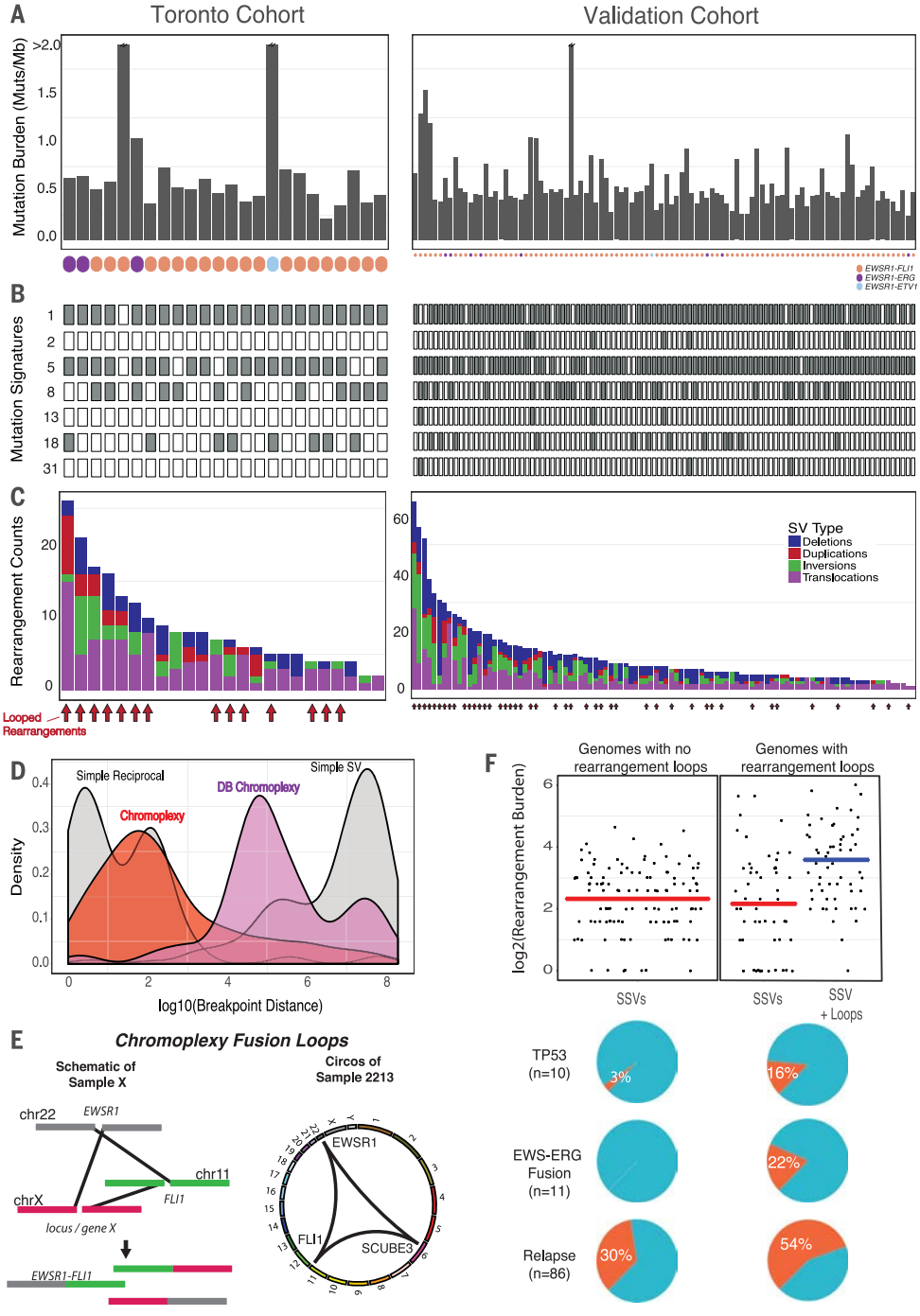


Fig. 1. Mutation landscape of ES. (**A** to **C**) The initial cohort consisted of 50 individuals with primary ES tumors, from which 23 tumors were whole-genome sequenced (Toronto cohort, left). One rearrangement screen sample (sample 4462) is included in this figure. The validation cohort consisted of individuals from which 100 ES whole genomes were sampled, from Tirode *et al*. (*10*) (right). (A) Somatic mutation burden for ES. The mutation burdens of all genome samples are shown. Three outlier samples with >2 mutations/Mb are clipped, as indicated by the hatch marks. (B) ES mutation signatures. Mutation signature analysis, defined by the proportion of 96 possible trinucleotides, identified common mutation patterns in most samples (age-associated, clock-like signature 1). Other signatures included 2, 5, 8, 13, 18, and 31. Signatures 2 and 13 are associated with the activity of the AID/APOBEC family of cytidine deaminase, whereas signature 5 is also clock-like in some cancers but less so in ES (*11*, *13*). Signatures 8 and 18 have an unknown molecular etiology; however, it has been suggested that signature 18 is caused by reactive oxygen species (*43*). Signature 31 is believed to be the result of exposure to platinum-based antineoplastic therapy (*24*). (C) Rearrangement profiles for ES. Shown are the burden of structural variants (SVs)—including deletions (blue), duplications (red), inversions (green), and translocations (purple)—in individual ES genomes. Samples with chained complex rearrangements (looped rearrangements) are highlighted by red arrows (14/24 for the Toronto cohort and 38/100 for the validation cohort; aggregated prevalence, 52/124). (**D**) Rearrangement breakpoint clusters. The aggregated density distributions of the genomic distance between consecutive rearrangement breakpoints are shown. Reciprocal breakpoints are close together (~10^2 bp) because there is an equal exchange of genetic material arising from a single break on each chromosome. Chromoplectic rearrangements (red) overlap this range owing to the proximity of breakpoints involved in looped rearrangements. Deletion-bridge (DB) chromoplexy (purple) is a looped rearrangement cluster in which a deletion spans two breakpoints, resulting in breakpoint distances that are farther apart (illustrated in fig. S11). Noncomplex breakpoints (simple SVs) are far apart (~10^8 bp). (**E**) Schematic diagram of chromoplexy fusion loops. Illustrative example of chromoplexy in ES shows three chromosomes undergoing double-strand breakage, shuffling, and religation in an aberrant configuration. This phenomenon generates the canonical fusion, *EWSR1-FLI1* (*ERG* or *ETV1*), and disrupts a third locus, X, in a one-off burst of rearrangements. In reality, up to eight chromosomes may be disrupted in this looping pattern. A representative genome-wide Circos plot depicting genomic rearrangements in an ES tumor (from the discovery cohort), which are organized in a loop. (**F**) Genomic correlates and clinical impact of looped rearrangements. In genomes without rearrangement loops, only simple structural variants (SSVs) exist with an average rearrangement burden of seven rearrangements per sample. This rate is similar to the background SSV rate (determined by removing rearrangements involved in a loop) in genomes with rearrangement bursts (compare the two red lines). The additional complexity of looped rearrangements results in higher genomic instability in these tumors. The most common genomic alterations include somatic *TP53* mutations, which are rare but enriched in patients with complex genomes (top pie chart, orange fraction; $P < 0.05$, Fisher's exact test). *EWS-ERG* fusions are also rare, as they represent 10% of all ES diagnoses; however, all *EWS-ERG* fusion ES tumors are either chomothriptic or chromoplectic (middle pie chart, orange fraction). Lastly, patients with complex genomes tend to relapse (bottom pie chart, orange fraction; $P < 0.05$, Fisher's exact test). All the markers of aggressive disease (high genomic instability, somatic *TP53* mutations, and relapse) are present in tumors with complex genomes.

number of neochromosomes. A low number of copy number states (three or fewer) is associated with a one-off mutational event, because breakage and ligation can only involve a small number of chromosomes inside a cell at any given time (*20*, *21*). By contrast, stepwise progression would result in multiple copy number states owing to the possibility of copy number alterations arising within older copy number alterations. Chromoplectic breakpoints involve many chromosomes and are not associated with any copy number alterations (fig. S13). That is, these looped rearrangements across the genome are balanced. In addition, using a bespoke algorithm, we found that the allele frequency of chromoplectic breakpoints was higher than that of simple structural rearrangements, providing further evidence that these breakpoints occurred together and early in tumor development (methods and fig. S14). Given their extremely tight clustering, low number of copy number–state transitions, and consistent clonal variant-allele frequency, *EWSR1-ETS* loops are likely to have arisen from singular bursts of rearrangements.

We then examined whether genomic regions of loop breakpoints share properties that might predispose these regions to simultaneous rearrangement. We performed a comprehensive analysis of 38 genomic properties, including adjacency to histone marks, association with replication timing, and proximity to genes and repetitive or transposable elements (table S4). Of these properties, early replicating DNA and features consistent with this were the most strongly associated with chromoplexy loops ($P < 1.0 \times 10^{-36}$, Mann-Whitney *U* rank sum test and Benjamini-Hochberg correction; Fig. 4, A and B). In notable contrast, neither nonlooped simple breakpoints of ES nor simulated simple breakpoints were significantly associated with replication timing or, indeed, any other feature (see methods). Replication timing is known to be strongly correlated with gene activity, chromatin accessibility, and nuclear position (*22*). Accordingly, chromoplectic breakpoint positions were also strongly associated with high gene density and high GC content (fig. S15A). Conversely, lamina-associated domains, which are enriched in late-replication regions and repressive chromatin environments, were found to be negatively associated with chromoplectic rearrangements. These significant associations were upheld when breakpoints directly residing in *EWSR1*, *FLI1*, or *ERG* were removed from the analyses. Of note, the same associations were found for looped rearrangements of ETS-positive (ETS+) prostate cancers but not for simple prostate cancer rearrangements (fig. S15B). Of further interest, we noted that the genes affected by chromoplexy were among the most highly expressed in ES across all patients (top 20%; fig. S16). Most expressed genes are found in early replicating DNA (*23*). These data are consistent with the proposed model of chromoplexy in which DNA is colocalized in transcription hubs, allowing for multiple genes from many chromosomes to be broken, shuffled, and aberrantly ligated (*2*).

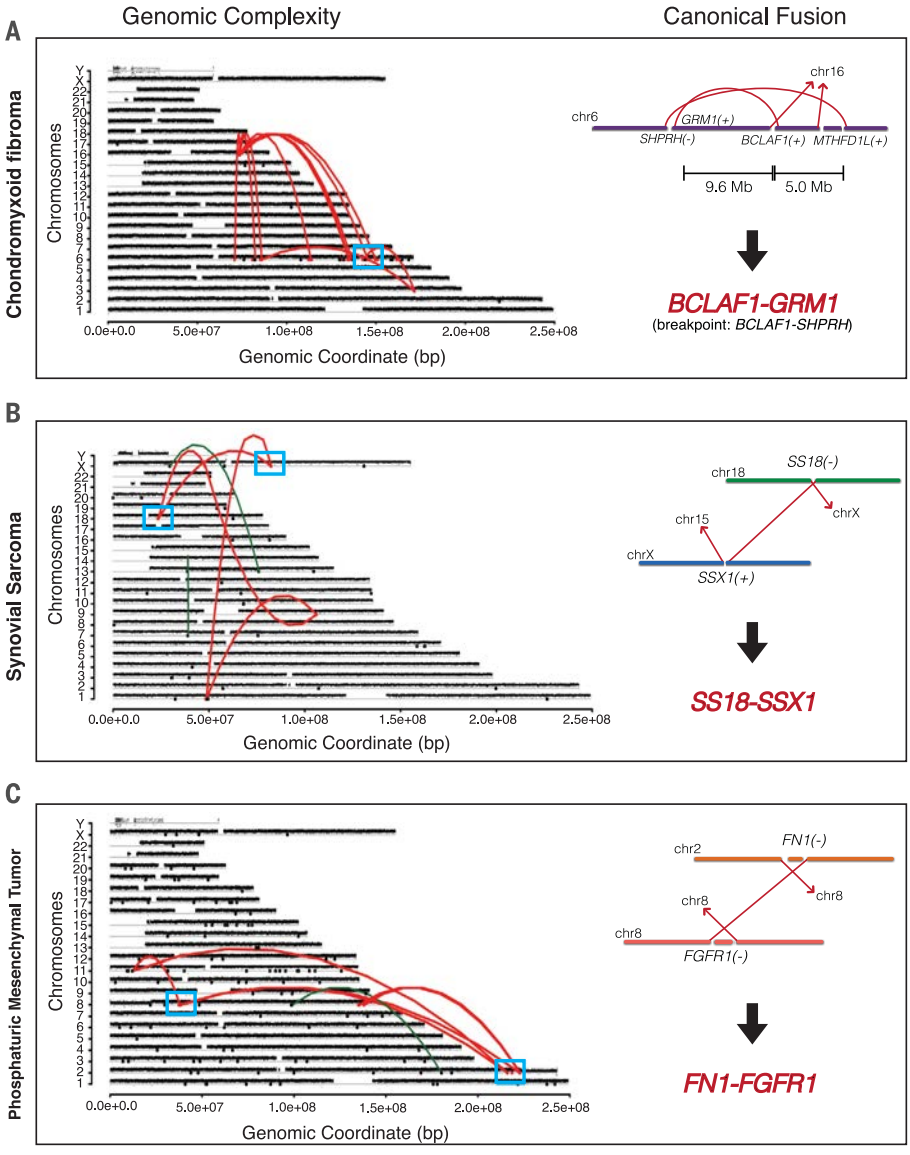


Fig. 2. Genomic catastrophes are common in sarcomas. Copy number profile for fusion-driven sarcomas with chromoplexy are shown. Rearrangements are colored red, and the loci with the canonical fusion are highlighted (blue box) and enlarged on the right. (**A**) Chondromyxoid fibroma (CMF) with chromoplexy. The genomic breakpoint lies in the upstream *SHPRH* gene, and the *BCLAF1-GRM1* fusion was detected by RNA sequencing. Additional complex CMFs, which also show a *SHPRH* genomic breakpoint but with the *GRM1* fusion found in the RNA, can be found in fig. S8. (**B**) Synovial sarcoma with chromoplexy. Chromoplexy generating the *SS18-SSX1* pathognomonic canonical fusion is shown. (**C**) Phosphaturic mesenchymal tumor (PMT) with chromoplexy. Genome sequencing of PMTs revealed deletion bridges occurring across the genome at chromoplectic loci, generating the canonical *FN1-FGFR1* fusion.

Mutation patterns of relapsed and metastatic ES

As noted above, chromoplexy arises early in the evolutionary history of ES and portends a worse prognosis and possible relapse. The genetic makeup of relapsed ES is largely unknown because standard of care for ES does not typically involve rebiopsy of the cancer when the disease returns or has metastasized. Therefore, whether further mutations—chromoplectic or otherwise—emerge at relapse is unclear, because few samples have been available. We obtained samples from six relapse, metastatic, or secondary tumors and performed whole-genome sequence analysis as well as full mutation and signature analyses (Fig. 5A). This included two primary-metastasis pairs, two primary-relapse pairs, one unpaired metastasis, and one tumor in which ES was secondary to a different cancer. Notably, every relapse or metastatic tumor

contained the chromoplexy-associated fusion, whether it was from a metastasis at the time of diagnosis or a relapse arising later (Fig. 5B). The pattern of point mutations was also distinct. There was an enrichment of signatures 8 and/or 18, in addition to the clocklike signature seen at diagnosis, suggesting that new processes drive relapse and metastatic ES (Fig. 5B). For example, in one patient's tumor, we found a pronounced increase of COSMIC signature 31, which has been recently associated with exposure to platinum therapy in chronic myelomonocytic leukemia (*24*). Notably, our patient had been treated with carboplatin for retinoblastoma 3 years before diagnosis of ES. At least three other patients in the validation cohort had a similar signature in their ES, which may also have been treatment induced (Fig. 1B).

Early divergence and parallel evolution of ES tumors

A predominant model for tumor progression posits that metastases originate directly from the primary tumor. The metastatic lesions may have acquired new mutations, but, because they are thought to be derived by linear clonal evolution, most of the genomic properties of the primary tumor will be found in the metastasis (*25*).

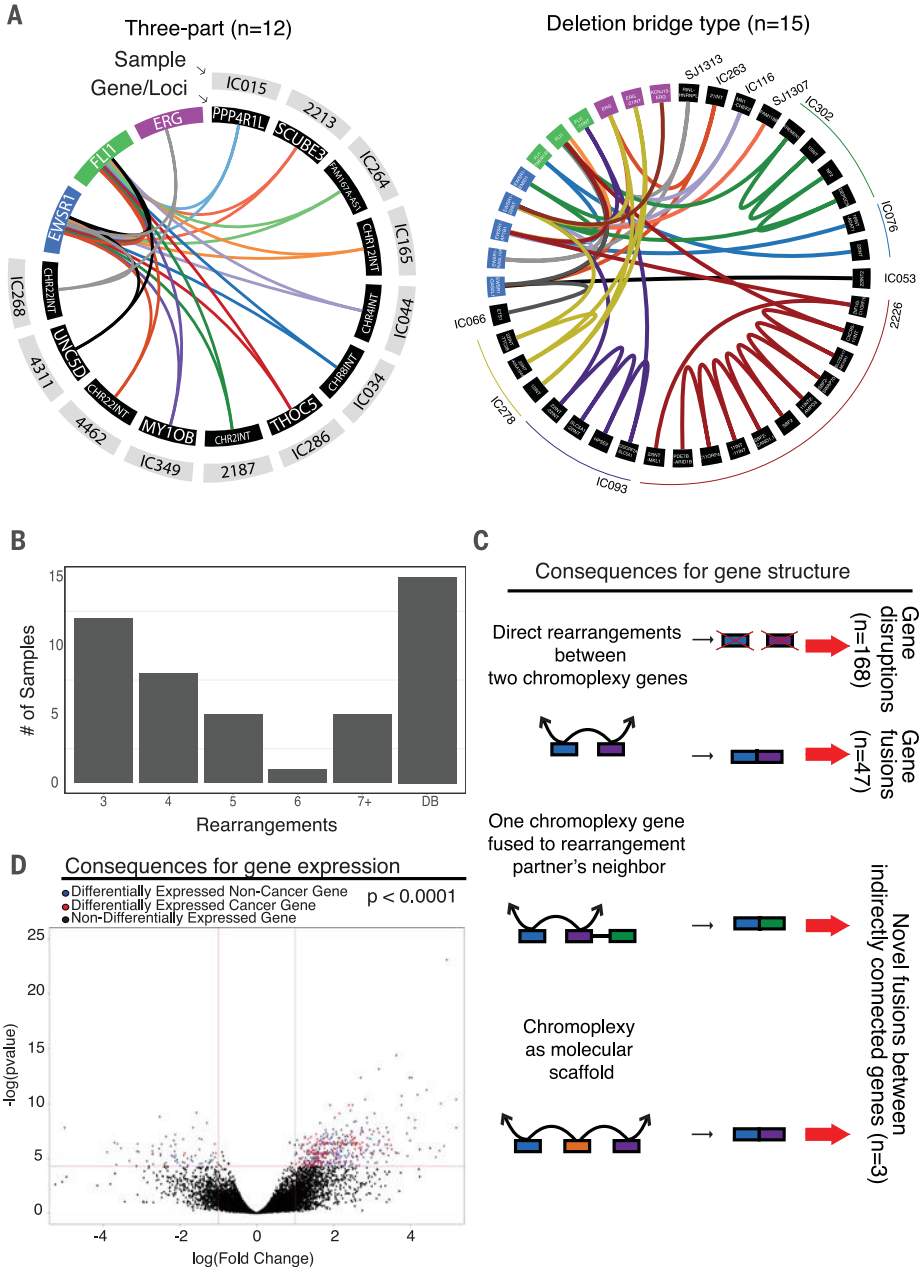


Fig. 3. Characterizing chromoplexy loops that generate *EWSR1-ETS* in ES. (**A**) Patterns of looped rearrangements. Chromoplexy circos webs demonstrate that patterns of looped rearrangements are conserved across samples, whereas different genes or loci are affected in each cancer (black panels). In each web, individual samples are indicated by different colors (and named in the gray panel). In all cases, central to chromoplexy fusion loops were the key driver genes: *EWSR1* (blue), *FLI1* (green), and *ERG* (purple). The most frequent patterns of chromoplexy in ES are those with a three-way looping structure as well as the presence of deletion bridges. An enlarged Circos web can be found in fig. S9 for readability. Three samples have structures only involving *EWSR1*, *FLI1*, and adjacent loci. Sample 4004 has deletion-bridge chromoplexy and is described in fig. S3C. (**B**) Summary of chromoplexy types. A bar chart showing the number of rearrangements in a loop (*x* axis) and the number of samples with that rearrangement pattern is shown. Other chromoplexy web structures can be found in fig. S10. (**C**) Transcriptional consequences of chromoplexy: Gene disruptions and fusions. There are three mechanisms of gene dysregulation via RNA fusion when chromoplexy occurs. The first involves two genes (blue and purple boxes) brought together by chromoplectic rearrangements (black lines with arrowheads), leading to gene disruptions (first scenario, shown at the top) and in-frame fusions (second scenario). This was detected in the 3/10 cases where there was genome (with chromoplexy) and transcriptome sequencing available. When RNA sequencing was not available, these were predicted to cause fusions (n = 47, excluding the *EWSR1-ETS* driver) and gene disruptions by fusing genes in opposite transcriptional orientation or fusing a gene to an intergenic sequence (n = 168). The second mechanism involves two chromoplexy genes brought together by a rearrangement at the genomic level, but one of the partner's neighboring genes (green box) is transcriptionally fused to the other chromoplexy partner in its place (third scenario). This is also the predominant mechanism of *GRM1* fusion generation in chondromyxoid fibromas (fig. S8). The final mechanism of gene dysregulation occurs when chromoplexy facilitates the production of a fusion by acting as a molecular scaffold (fourth scenario; also illustrated in fig. S12). Two genes are both rearranged to a third locus (orange) and are then transcriptionally fused together. No direct genomic link exists between these two genes. These phenomena can only be detected if both whole-genome and RNA sequencing are available. (**D**) Transcriptional consequences of chromoplexy: Gene expression. Volcano plot illustrating the differential gene expression in chromoplexy versus nonchromoplexy ES, revealing 504 differentially expressed genes. Points greater than 1 or less than −1 and above 1.3 (as indicated by the red lines) are genes that are significantly differentially expressed (blue dots). Red dots highlight genes that are differentially expressed and involved in a cancer hallmark pathway.

A different model was suggested for ES, on the basis of mutation data from two primary-metastasis pairs whose exomes were sequenced (*8*). Specifically, it was proposed that metastases diverged from the primary tumor early, although the timing of this divergence was not established. We compared coding, noncoding, and structural rearrangements across the genome within four ES pairs. As is the case in most tumor types, relapse and metastatic ES tumors acquired many new mutations (on average, 50% were unique, or "private," mutations). A notably high number of clonal mutations from the primary ES were lost in the relapse (average 20%), confirming that the latter diverged early, evolving in parallel. For example, a disruptive clonal *PTEN* inversion was found in all tumor cells of one primary ES but was absent from the relapse (Fig. 5C). We also confirmed the same model of parallel evolution in one additional primary-metastatic pair, profiled using microarrays (fig. S17). The clinical implication of this model is that one should be searching for therapeutically targetable mutations arising in parallel with those in the primary ES, using biomarkers like circulating tumor DNA, because these mutations would not necessarily be present in the primary tumor.

To determine when the divergence of the lethal clone occurred, we used the number of COSMIC signature 1 mutations, which emerge at a steady rate in ES (see methods and fig. S18). We first confirmed our approach by comparing the number of signature 1 mutations between established time intervals, such as the dates of diagnosis and recurrence. In all cases, the observed number of mutations was extremely close (75 to 90%) to what would be expected (fig. S18). Using the established rate, we calculated the amount of time between the divergence of the primary and relapse or metastatic tumors. Notably, the common ancestor in ES clonally diverges 1 to 2 years before diagnosis. Therefore, the cells that give rise to the primary and relapse tumor can exist in the patient years before diagnosis, providing a window for early cancer detection and surveillance. ES is often difficult to diagnose, and the time to diagnosis is notoriously long (*26*). These findings provide a plausible biological mechanism for this latency.

Discussion

Our analyses reveal rearrangement bursts (chromoplectic loops) as a source of gene fusion in human bone and soft tissue tumors. ESs with complex karyotypes are associated with a poorer prognosis than those with simpler karyotypes (*27*), and here we show chromoplexy as the mechanism in 42% of tumors. It is possible that the chromoplectic tumor's additional gene disruptions and fusions contribute to the difference in patient survival. Our whole-genome sequence data support a model in which there is an early clone of ES, containing *EWSR1-ETS* and chromoplexy, arising at least 1 year before diagnosis, which gives rise to both the primary and metastatic or relapse tumors (Fig. 5D). Whether the bursts described here are chance events or driven by specific mutational processes, akin to the RAG machinery operative in leukemia, remains to be established. As an increasing and diverse number of tumor genome sequences become available, we may be able to define further rearrangement processes that underlie fusion genes and thus unravel the causes of fusion-driven human cancers.

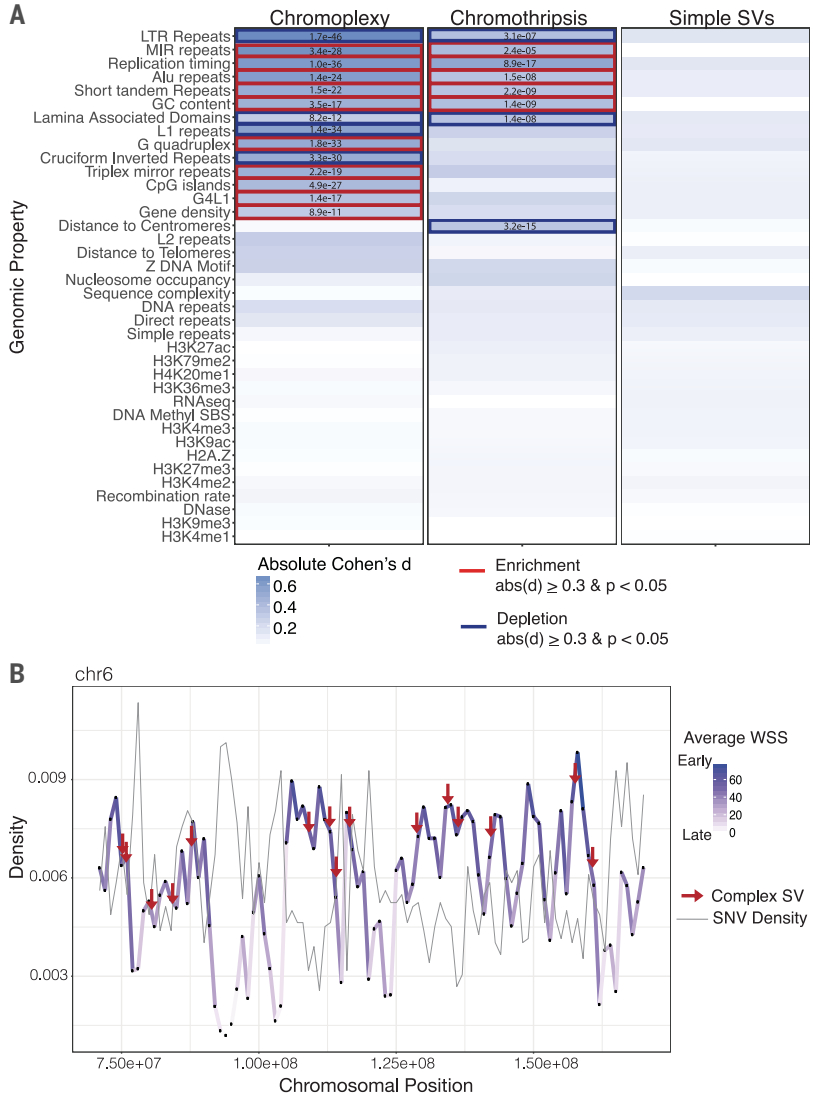


Fig. 4. Early replicating DNA and chromoplexy. (**A**) Heatmap of genomic property associations. The genomic properties listed in table S4 were calculated for all rearrangements in both cohorts. Complex rearrangements (chromoplexy and chromothripsis), exclusively, are strongly associated with early replication timing and other genomic features consistent with this feature (gene density, CpG density, Alu density, etc.). Table values are indicative of false discovery rate–corrected *P* values compared to a million random points in the genome. Blue highlights in the figure are indicative of a Cohen's *d* greater than or equal to 0.3. Bold boxes indicate a positive (red, enrichment) or negative (blue, depletion) association with the feature. All features were evaluated in 1-kb bins across the genome. For feature density metrics, associations were calculated in 1-Mb sliding windows centered in 1-kb bins. (**B**) Density distribution of the average wavelet-smoothed signal (WSS) and single-nucleotide variations (SNVs) on a representative chromosome. The average WSS, of replication timing, is plotted for a subset of chromosome 6 to illustrate changes between early and late replication timing and the co-association with mutations in ES. The positional variation of replication timing across the chromosome is depicted as changes in density and color. Point mutations peak in late-replicating regions (dip in WSS, light purple), whereas complex rearrangements peak in regions of early replication timing (peak in WSS, dark purple).

Materials and methods

Patient and sample collection

ES tumor and matched blood samples were collected from the Hospital for Sick Children (SickKids) and Mount Sinai Hospital in Toronto, Canada, in accordance with each institution's Research Ethical Board (REB) guidelines. Detailed clinical information (age at presentation,

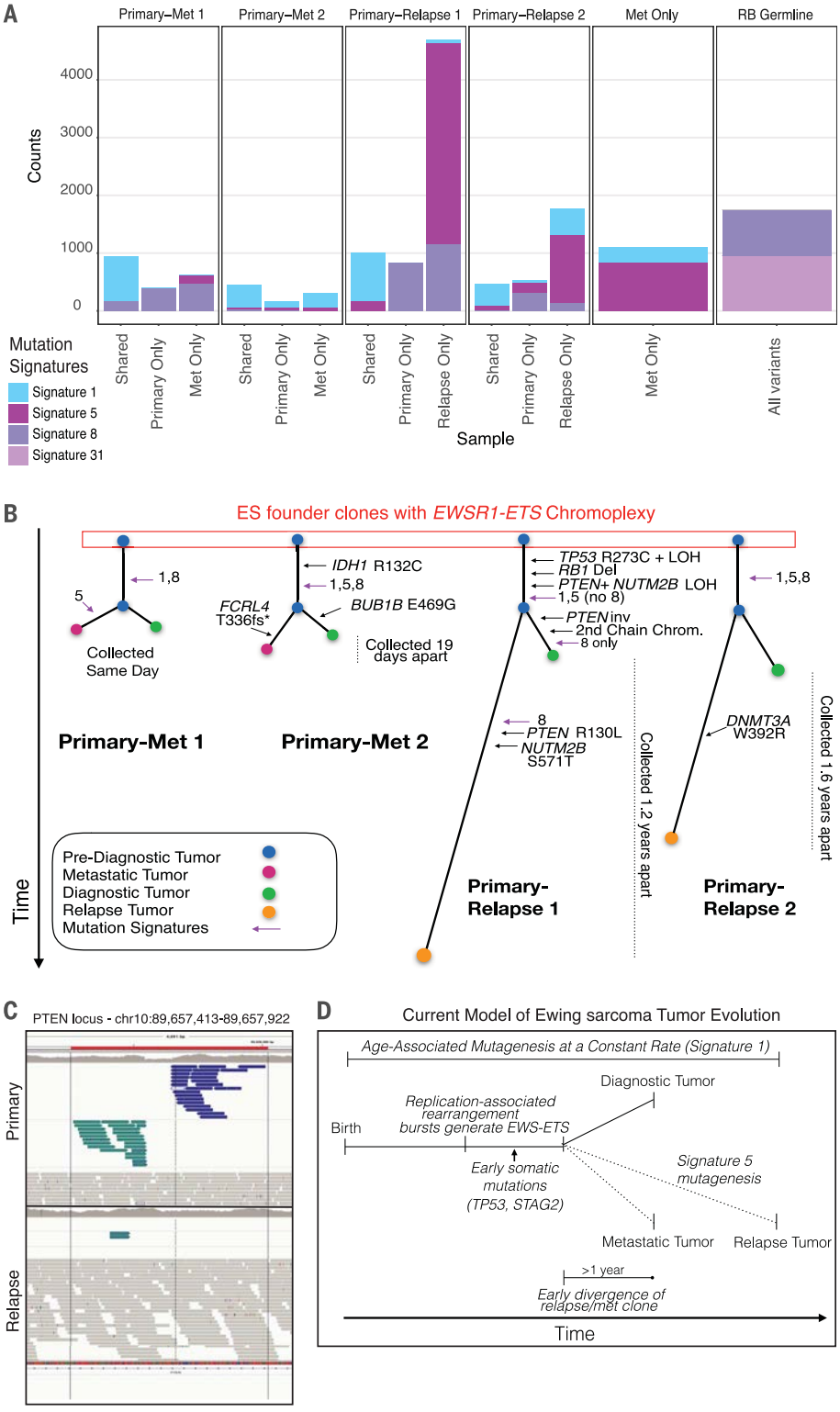


Fig. 5. Mutation signatures of relapse and metastatic ES tumors. (**A**) Prevalence of mutation signatures in relapse and metastatic (Met) tumors. Shared and private mutations for four primary-metastatic or relapse pairs are shown (first four columns). Signatures 1 and 5 are common throughout, with signature 5 contributing considerably to the mutations that arise at relapse. Signature 8 was also common throughout the cohort. One metastatic tumor (no paired primary) is also shown to have similar mutation signature patterns as other metastatic or relapse tumors. Lastly, a secondary ES tumor to a primary retinoblastoma (germline *RB1* mutation identified) was also sequenced in this cohort. This patient harbored the rare signature 31, which likely resulted from the patient's prior exposure to carboplatin for their primary retinoblastoma (the only patient to receive this treatment in the Toronto cohort). (**B**) Phylogenetic trees of primary-relapse and primary-metastatic ES. Using the shared and private mutations, we identified the mutational order in ES. Known cancer-driver mutations (*IDH1*, *TP53*, etc.) arise early (shared branches). LOH, loss of heterozygosity; del, deletion; inv, inversion. (**C**) A clonal *PTEN* inversion. A *PTEN* inversion was found in the primary but not in the relapse tissue, suggesting that the inversion arose after early divergence of a common clonal ancestor. However, a pathogenic *PTEN* SNV can be found in the relapse tissue. Together, these point toward parallel, convergent evolution on this gene. (**D**) Proposed model of ES tumor evolution. After birth, signature 1 is operative in all somatic tissues throughout life. ES patients' cells experience a replication-associated burst of rearrangements that generates the canonical fusion driver. Early somatic cancer gene mutations occur before clonal bifurcation. This occurs 1 to 2 years before an ES diagnosis; thus, the cells that would give rise to the relapse existed years before diagnosis. Signature 5 contributes substantially to the number of mutations seen at relapse.

gender, tumor site, stage, etc.) were obtained from the corresponding institutional tumor banks (table S5). Overall, the patients' clinical features and demographics were typical of ES: the average age of diagnosis was 14.8 years (2.8 to 36.6 years); the male to female ratio was 1.38:1; and 14 patients had relapsed, with 13 having died from their disease. Additional samples (n = 3) were also obtained from Universitätsklinikum Erlangen, Erlangen, Germany. All metastatic or relapse ES tumors were collected from the SickKids tumor bank or the SickKids cancer sequencing program (KiCS). Detailed information on KiCS is available at https://www.kicsprogram.com.

Of the 25 high-coverage genomes sequenced, *EWSR1-ETS* fusions were detected in all patients except for a 37-year-old individual who was instead found to have a *FUS-ERG* translocation. This patient's gene expression profile (by RNA-seq) was also discrepant, so they were removed from subsequent analyses (fig. S19). One additional genome was removed due to poor sequencing quality. We also performed low pass (~10X) rearrangement screens on 19 ES samples.

However, as we required breakpoint resolution, all but one of the rearrangement screens were excluded from this study due to insufficient coverage (see table S1, orange row). Taken together, our discovery cohort consisted of individuals from which 23 standard genomes (30X to 60X) and one rearrangement screen genome (20X) were sequenced. The validation cohort consisted of individuals from which 119 tumor-normal samples were sequenced by Tirode *et al.* (*10*), which we downloaded from the European Genome-phenome Archive (accessions: EGAS00001000855 and EGAS00001000839). Of these, 19 patient samples were omitted either because the *EWSR1-ETS* fusion was not detected by our pipeline and manual inspection of the aligned reads, or because they harbored an excess of artefactual small inversions or deletions.

Code availability

Custom code described here is available at github.com/shlienlab.

High-throughput sequencing and alignment

Exome, genome, and transcriptome (RNA-seq) sequencing were performed using established protocols on Illumina instruments. For exome and genomes, paired-end FASTQ files were aligned to the human genome (hg19/GRCh37) using BWA-MEM (v.0.7.8); Picard MarkDuplicates (v.1.108) was used to mark PCR duplicates. Indel realignment and base quality scores were recalibrated using the Genome Analysis Toolkit (v.2.8.1).

Detection of high-quality somatic substitutions and rearrangements

We detected somatic mutations using established tools [MuTect2 (part of GATK v.3.5) (*28*) and Delly v.0.7.1 (*29*)]. To evaluate and validate our WGS substitution pipeline, we used a "gold standard" cancer genome tumor-normal dataset, COLO829 (*30*). Using this somatic reference standard, we determined our precision to be 0.885 and our sensitivity to be 0.971. Copy number was detected for genomes and rearrangement screens using BIC-seq v.1.2.1 (*31*). When no matched normal was available (in the case of rearrangement screens), an ES normal was used. We then developed custom code to increase specificity of putative substitution and rearrangement detection, as follows:

1. Somatic and depth filter. No mutation should exist in the matched-normal sequence. For substitutions, we removed common single-nucleotide polymorphisms (SNPs) as previously described (*32*) and a required >10X coverage at the mutated locus (10-kb window), in both tumor and normal. For rearrangements, this filter required ≥4 discordant read-pairs in the tumor. We then directly interrogated the normal BAM file, at each putative somatic rearrangement to ensure no germline variants existed near the breakpoint, on either side of the rearrangement.

2. Panel of normals filtering. To remove common germline variants, we created a panel of normal, nonneoplastic, samples that had been sequenced using the same technology and to a similar depth of coverage (n = 133). We removed any putative rearrangements if present in ≥2 normals. For rearrangements, breakpoints must exist on both sides of the junction within a 1-kb window. We found that as we increased the number of normals in our panel, our specificity increased (fig. S1C).

3. Quality-control filtering. Putative rearrangements were removed if supported by reads with MAPQ < 30. Putative rearrangements were removed if they met any two of the following criteria:

(i) Non-unique mapping. <70% of the reads at the locus map uniquely.

(ii) Multimapping clusters. At the same locus (200 bp up and downstream), a pattern of multiple overlapping groups of discordant reads whose paired-ends align to different chromosomes (>3 reads in each group, mapping to >4 chromosomes). Seen in both the tumor and paired normal.

(iii) High depth. Excessively high depth alignments in difficult-to-align regions of the genome, as described (*33*). We apply a maximum depth threshold of d + 4*sqrt (d), where d is the average normal mean read depth of the chromosome in the corresponding normal.

(iv) Low-complexity regions. Overlap with a highly repetitive sequence (using DUST (*34*) with score >60).

Mutation signature extractions and analysis

First, a de novo extraction was performed on the catalog of ES point mutations to produce novel consensus mutational signatures. These signatures were deciphered using a previously described computational framework that optimally explains the proportion of each mutation type found in the catalog and then estimates the contribution of each signature to the mutation catalog (*11*). Overall, we identified 11 consensus mutational signatures. Four of these signatures were previously found to be attributed to sequencing artifacts. We then compared our true consensus mutational signatures to the previously curated COSMIC list and quantified their similarity using a cosine similarity as previously done (*13*). We report >0.9 cosine similarity between the Ewing signatures and the COSMIC list.

Validation by targeted custom-capture sequencing

A custom targeted-capture enrichment system was designed to capture 1 Mb of DNA (Nextera, Illumina) with custom probes for the whole of *EWSR1*, *FLI1*, and *ERG* genes as well as the exons of *TP53*, *STAG2*, and *ATRX*. We also targeted known complex breakpoints from the discovery cohort, achieving between 900- to 1000-fold coverage. We reasoned that paired-end sequencing would capture any locus joined to the three core genes, even if the panel did not specifically target it. In this way, we validated rearrangements in samples where chromoplexy was already known from the whole genome and uncovered new instances in samples that had not been whole-genome sequenced (n = 7 and 4, respectively). Each tumor had three or four rearrangements validated using the panel. All had the same breakpoint (as found by the whole-genome sequence) and were found to harbor looped rearrangements are on the same derivative chromosomes.

Validation by FISH, G-banding, or spectral karyotyping

We further validated these looped rearrangements by karyotyping ESs using standard G-banding as well as spectral karyotyping (n = 17 and 3; fig. S4). By cytogenetics we found additional complexity—beyond the canonical chr22-chr11 translocation—in eight cases. Of these, six tumors had been sequenced and found to be complex. Additionally, there were five cases for which chromoplexy was detected by genome sequencing yet not found by cytogenetics techniques, indicating that routine cytogenetics may miss chromosomal complexity present in these genomes due to the nature of these submicroscopic complexities (fig. S20).

Timing of rearrangements using breakpoint allele fraction

To determine the timing of the chromoplectic loops, we developed a tool to accurately measure the breakpoint allele fraction (BAF) of each rearrangement. The BAF is the proportion of reads containing a rearrangement breakpoint divided by the total number of reads, analogous to the variant allele fraction (VAF) for point mutations (illustrated in fig. S14A). This is analogous to the variant-allele frequency of substitution mutations and, similarly, can be used to infer the relative order of rearrangement mutations. The tool accurately counts all reads supporting each rearrangement, even if these had not been used to nominate the rearrangement in the first place. From the raw aligned reads, we first collected all split reads near the breakpoint (within 20 bp) from one side of the rearrangement. Next, we extracted the clipped sequence (i.e., the non-aligned portion) from these reads and attempted to map it to the other side of the rearrangement (within 70 bp of the breakpoint) using a Smith-Waterman algorithm (*35*). Clipped sequences shorter than 5 bp were discarded, as were those that failed to map to the other side of the rearrangement (≤80% similarity). Since the retained sequences can map at slightly different position, due to microhomology near the breakpoint, we considered all those close to one another as supportive of the same rearrangement. Overall, we found that most rearrangements are supported by remapped reads that are less than 10 bp apart. Finally, the total number of split and realigned reads were divided by the average coverage between the two breakpoints per side of each rearrangement. This allowed us to arrive at an accurate measure of the breakpoint allele fraction. To validate our tool, we applied it to a curated list of known

polymorphic copy number variants (CNVs) (*36*). As expected, the BAF of germline CNV deletions followed a bimodal distribution with peaks at 0.5 and 1.0, for heterozygous and homozygous rearrangements, respectively (fig. S14B, green line). We then compared the BAF of somatic rearrangements in chains to those without. Chained rearrangements had higher BAFs than simple structural variants (fig. S14B, red versus blue line), confirming that chromoplectic rearrangements are in fact earlier.

Detection of breakpoint clusters of chained rearrangements

Using their interbreakpoint distance, we identified rearrangements within 10 kbp of one another. Using these, we created an undirected graph in which two rearrangement breakpoints within 10 kbp of one another (a breakpoint cluster) were represented as a vertex and connected to other breakpoint clusters (rearrangements are edges in the graph). We selected connected components of the graph and identified components with greater than one vertex as interconnected rearrangements. In most of our cases, these interconnected rearrangements formed chains or loops, where one could follow the edges around the graph and return to the initial vertex of departure. These were further filtered for reciprocal rearrangements or overlapping intrachromosomal rearrangements. Chromoplexy rearrangements were validated by manual inspection and using the ChainFinder algorithm (*2*).

Association of rearrangements with genomic features

We formally evaluated the association of rearrangement position with 38 properties of the human genomes (table S4). We separately evaluated each of these associations in 1-kb bins across the genome. Feature density properties were calculated as densities in various sliding windows (1 kb, 10 kb, 100 kb, and 1 MB) centered on each 1-kb bin or as the $\log_2$ distance, as indicated in table S4. The positions of ES rearrangements were compared to a million random positions that had been uniformly sampled from regions of the genome where confident genotypes could be determined (i.e., the "callable" genome). We limited our analysis to chromosomes 1 to 22 and X. To test for significant associations between our rearrangements and these genomic properties, we performed a Mann Whitney *U* rank sum test and Benjamin and Hochberg FDR correction to raw *P* values. We used the Cohen's *d* metric to determine the effect size between the two groups to account for differences in sample size. We applied an absolute Cohen's *d* cut-off of 0.3, a medium effect size (*37*, *38*). Genomic properties were considered significantly different between rearrangements and random positions if absolute (d) ≥ 0.3 and the corrected $P < 0.05$.

Detection of gene fusions

We detected gene fusions in regions of genomic complexity using an approach that integrates multiple independent fusion algorithms, and then removed those found in normal tissue. Putative fusions were validated by de novo assembly. A total of 1277 normal (nonneoplastic) samples from 43 different tissues were obtained from the NHGRI GTEx consortium (database version 4) and used to remove artifacts. All fusions were visually inspected if one or both genes involved chromoplexy or were adjacent (up to 1 Mbp). Fusions were further filtered by quality of the realigned transcript, breakpoint coverage, and gene expression.

Detection of gene expression

Gene expression for fusions, differential gene expression analysis, and principal component analysis was performed using HT-seq (*39*) to count the reads aligning to every gene. PCR duplicates and reads mapping to ribosomal RNA, miRNA, and small nucleolar RNA were removed. We used the trimmed mean of M-value (TMM) method in the EdgeR package to perform normalization on genes with at least 1 read per million bases in at least three samples (*40*, *41*). Differential expression analysis in chromoplexy versus nonchromoplexy samples was performed using a generalized linear model (GLM) likelihood ratio test, taking in consideration different sources of variation like batch, gender and age. *P* values for the GLM test were adjusted for multiple testing using the Benjamini and Hochberg method for controlling the false discovery rate (FDR). Differentially expressed genes in chromoplexy versus nonchromoplexy were considered statistically significant if FDR ≤ 0.05 and absolute value of log(fold change) ≥ 1. Pathway analysis was performed on genes differentially expressed in samples with and without chromoplexy using Gene Set Enrichment Analysis (GSEA) software (javaGSEA v2.2.4). Cancer gene signatures were selected from the hallmark collection from the Molecular Signature DataBase (MsigDB) (*18*). Enrichment scores for the hallmark pathways were considered statistically significant if FDR < 0.01.

Evaluation of replication timing in prostate cancer rearrangements

We obtained prostate cancer rearrangements, including chained and others, from the Baca *et al.* publication [supplemental tables S3C and S5 from (*2*)]. Samples were annotated as "ETS+" or "ETS−" using supplemental table S1 from (*2*). ETS+ fusions include any ETS fusion detected by sequencing (including *ERG* and *ETV1*). Using this list, we performed the same test for genomic property enrichment as we did in ESs.

Molecular inversion probe (MIP) microarray

Raw MIP data from three additional primary-metastatic ES pairs were obtained from the Huntsman Cancer Institute, Salt Lake City, Utah (*42*). The original source material was clinically archived, formalin-fixed paraffin-embedded (FFPE) scrolls that were retrieved from three individual patients diagnosed with ES. Primary tumor samples were from diagnostic biopsies taken before chemotherapy. The raw MIP data from the completed assay was loaded into Nexus Copy Number (BioDiscovery, Inc., El Segundo, California) for copy number detection using default settings.

REFERENCES AND NOTES

1. P. J. Stephens *et al.*, Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* **144**, 27–40 (2011). doi: 10.1016/j.cell.2010.11.055; pmid: 21215367
2. S. C. Baca *et al.*, Punctuated evolution of prostate cancer genomes. *Cell* **153**, 666–677 (2013). doi: 10.1016/j.cell.2013.03.021; pmid: 23622249
3. F. Mitelman, B. Johansson, F. Mertens, The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer* **7**, 233–245 (2007). doi: 10.1038/nrc2091; pmid: 17361217
4. E. Papaemmanuil *et al.*, RAG-mediated recombination is the predominant driver of oncogenic rearrangement in *ETV6-RUNX1* acute lymphoblastic leukemia. *Nat. Genet.* **46**, 116–125 (2014). doi: 10.1038/ng.2874; pmid: 24413735
5. S. Sankar, S. L. Lessnick, Promiscuous partnerships in Ewing's sarcoma. *Cancer Genet.* **204**, 351–365 (2011). doi: 10.1016/j.cancergen.2011.07.008; pmid: 21872822
6. H. Kovar, Blocking the road, stopping the engine or killing the driver? Advances in targeting EWS/FLI-1 fusion in Ewing sarcoma as novel therapy. *Expert Opin. Ther. Targets* **18**, 1315–1328 (2014). doi: 10.1517/14728222.2014.947963; pmid: 25162919
7. A. S. Brohl *et al.*, The genomic landscape of the Ewing sarcoma family of tumors reveals recurrent *STAG2* mutation. *PLOS Genet.* **10**, e1004475 (2014). doi: 10.1371/journal.pgen.1004475; pmid: 25010205
8. B. D. Crompton *et al.*, The genomic landscape of pediatric Ewing sarcoma. *Cancer Discov.* **4**, 1326–1341 (2014). doi: 10.1158/2159-8290.CD-13-1037; pmid: 25186949
9. M. S. Lawrence *et al.*, Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013). doi: 10.1038/nature12213; pmid: 23770567
10. F. Tirode *et al.*, Genomic landscape of Ewing sarcoma defines an aggressive subtype with co-association of *STAG2* and *TP53* mutations. *Cancer Discov.* **4**, 1342–1353 (2014). doi: 10.1158/2159-8290.CD-14-0622; pmid: 25223734
11. L. B. Alexandrov *et al.*, Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013). doi: 10.1038/nature12477; pmid: 23945592
12. L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, P. J. Campbell, M. R. Stratton, Deciphering signatures of mutational processes operative in human cancer. *Cell Reports* **3**, 246–259 (2013). doi: 10.1016/j.celrep.2012.12.008; pmid: 23318258
13. L. B. Alexandrov *et al.*, Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015). doi: 10.1038/ng.3441; pmid: 26551669
14. A. Zoubek *et al.*, Variability of EWS chimaeric transcripts in Ewing tumours: A comparison of clinical and molecular data. *Br. J. Cancer* **70**, 908–913 (1994). doi: 10.1038/bjc.1994.419; pmid: 7524604
15. C. M. Hattinger *et al.*, Prognostic impact of chromosomal aberrations in Ewing tumours. *Br. J. Cancer* **86**, 1763–1769 (2002). doi: 10.1038/sj.bjc.6600332; pmid: 12087464
16. P. M. Neilsen, K. I. Pishas, D. F. Callen, D. M. Thomas, Targeting the p53 pathway in Ewing sarcoma. *Sarcoma* **2011**, 746939 (2011). doi: 10.1155/2011/746939; pmid: 21197471
17. C. S. Cooper *et al.*, Analysis of the genetic phylogeny of multifocal prostate cancer identifies multiple independent clonal expansions in neoplastic and morphologically normal prostate tissue. *Nat. Genet.* **47**, 367–372 (2015). doi: 10.1038/ng.3221; pmid: 25730763
18. A. Liberzon *et al.*, The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* **1**, 417–425 (2015). doi: 10.1016/j.cels.2015.12.004; pmid: 26771021
19. K. H. Nord *et al.*, *GRM1* is upregulated through gene fusion and promoter swapping in chondromyxoid fibroma. *Nat. Genet.* **46**, 474–477 (2014). doi: 10.1038/ng.2927; pmid: 24658000
20. A. Malhotra *et al.*, Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. *Genome Res.* **23**, 762–776 (2013). doi: 10.1101/gr.143677.112; pmid: 23410887

21. J. O. Korbel, P. J. Campbell, Criteria for inference of chromothripsis in cancer genomes. *Cell* **152**, 1226–1236 (2013). doi: 10.1016/j.cell.2013.02.023; pmid: 23498933
22. R. S. Hansen *et al*., Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 139–144 (2010). doi: 10.1073/pnas.0912402107; pmid: 19966280
23. J. Sima, D. M. Gilbert, Complex correlations: Replication timing and mutational landscapes during cancer and genome evolution. *Curr. Opin. Genet. Dev.* **25**, 93–100 (2014). doi: 10.1016/j.gde.2013.11.022; pmid: 24598232
24. J. Merlevede *et al*., Mutation allele burden remains unchanged in chronic myelomonocytic leukaemia responding to hypomethylating agents. *Nat. Commun.* **7**, 10767 (2016). doi: 10.1038/ncomms10767; pmid: 26908133
25. J. W. Gray, Evidence emerges for early metastasis and parallel evolution of primary and metastatic tumors. *Cancer Cell* **4**, 4–6 (2003). doi: 10.1016/S1535-6108(03)00167-3; pmid: 12892707
26. J. F. Brasme, M. Chalumeau, O. Oberlin, D. Valteau-Couanet, N. Gaspar, Time to diagnosis of Ewing tumors in children and adolescents is not associated with metastasis or survival: A prospective multicenter study of 436 patients. *J. Clin. Oncol.* **32**, 1935–1940 (2014). doi: 10.1200/JCO.2013.53.8058; pmid: 24841977
27. P. Roberts *et al*., Ploidy and karyotype complexity are powerful prognostic indicators in the Ewing's sarcoma family of tumors: A study by the United Kingdom Cancer Cytogenetics and the Children's Cancer and Leukaemia Group. *Genes Chromosomes Cancer* **47**, 207–220 (2008). doi: 10.1002/gcc.20523; pmid: 18064647
28. K. Cibulskis *et al*., Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013). doi: 10.1038/nbt.2514; pmid: 23396013
29. T. Rausch *et al*., DELLY: Structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics* **28**, i333–i339 (2012). doi: 10.1093/bioinformatics/bts378; pmid: 22962449
30. D. W. Craig *et al*., A somatic reference standard for cancer genome sequencing. *Sci. Rep.* **6**, 24607 (2016). doi: 10.1038/srep24607; pmid: 27094764
31. R. Xi *et al*., Copy number variation detection in whole-genome sequencing data using the Bayesian information criterion. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1128–E1136 (2011). doi: 10.1073/pnas.1110574108; pmid: 22065754
32. A. Shlien *et al*., Combined hereditary and somatic mutations of replication error repair genes result in rapid onset of ultra-hypermutated cancers. *Nat. Genet.* **47**, 257–262 (2015). doi: 10.1038/ng.3202; pmid: 25642631
33. H. Li, Toward better understanding of artifacts in variant calling from high-coverage samples. *Bioinformatics* **30**, 2843–2851 (2014). doi: 10.1093/bioinformatics/btu356; pmid: 24974202
34. A. Morgulis, E. M. Gertz, A. A. Schäffer, R. Agarwala, A fast and symmetric DUST implementation to mask low-complexity DNA sequences. *J. Comput. Biol.* **13**, 1028–1040 (2006). doi: 10.1089/cmb.2006.13.1028; pmid: 16796549
35. P. Rice, I. Longden, A. Bleasby, EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet.* **16**, 276–277 (2000). doi: 10.1016/S0168-9525(00)02024-2; pmid: 10827456
36. M. Zarrei, J. R. MacDonald, D. Merico, S. W. Scherer, A copy number variation map of the human genome. *Nat. Rev. Genet.* **16**, 172–183 (2015). doi: 10.1038/nrg3871; pmid: 25645873
37. J. Cohen, The statistical power of abnormal-social psychological research: A review. *J. Abnorm. Soc. Psychol.* **65**, 145–153 (1962). doi: 10.1037/h0045186; pmid: 13880271
38. J. Cohen, A power primer. *Psychol. Bull.* **112**, 155–159 (1992). doi: 10.1037/0033-2909.112.1.155; pmid: 19565683
39. S. Anders, P. T. Pyl, W. Huber, HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015). doi: 10.1093/bioinformatics/btu638; pmid: 25260700
40. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi: 10.1093/bioinformatics/btp616; pmid: 19910308
41. D. J. McCarthy, Y. Chen, G. K. Smyth, Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012). doi: 10.1093/nar/gks042; pmid: 22287627
42. M. S. Jahromi *et al*., Molecular inversion probe analysis detects novel copy number alterations in Ewing sarcoma. *Cancer Genet.* **205**, 391–404 (2012). doi: 10.1016/j.cancergen.2012.05.012; pmid: 22868000
43. C. Pilati *et al*., Mutational signature analysis identifies *MUTYH* deficiency in colorectal cancers and adrenocortical carcinomas. *J. Pathol.* **242**, 10–15 (2017). doi: 10.1002/path.4880; pmid: 28127763

ACKNOWLEDGMENTS

We dedicate this manuscript to Simon Hajjar, a brilliant and fearless friend, colleague, and student in our lab, and to Ana Novokmet, a friend and colleague who held all of us to high standards and was committed to helping children with cancer. Simon and Ana both died from cancer as this work neared completion. We thank The Centre for Applied Genomics (TCAG) NGS and Biobanking facility for sequencing services. We thank J. Lees, N. Alon, G. Collord, and R. Arnold for their contributions toward this work. We thank N. Tan for her work on the illustration used in the manuscript's summary page. **Funding**: This research project was conducted with support from C^{17} and partially funded by Ewings Cancer Foundation of Canada and Childhood Cancer Canada Foundation. A.S. and D.M. received financial support from the SickKids Foundation through the Garron Family Cancer Centre. N.D.A is personally supported by an NSERC Canada Graduate Scholarship (Master's) and a SickKids Restracomp Award. S.B. receives personal Fellowships from Wellcome and the St. Baldrick's Foundation. **Author contributions:** A.S., S.B., and D.M. designed the study. N.D.A., R.D., M.D.Y., M.L., A.R., N.D.R., F.F., S.H., P.E.K., B.I., L.C.S., R.M., and J.S. performed experiments. N.D.A., R.D., M.D.Y., A.R., F.F., M.A., M.S., and L.B.A. collected and analyzed data. L.B., S.D., N.G., A.N., A.F., N.P., A.Y., T.S., M.M., G.R.S., S.W.S., A.M.F., M.S., J.S.W., I.L.A., P.J.C., J.D.S., D.M., S.B., and A.S. contributed reagents, tissue, and clinical data. N.D.A., A.S., and S.B. wrote the manuscript. M.Z., P.E.K., G.T.G., J.A.T., M.S., I.L.A., D.M., S.B., and A.S. provided technical support and conceptual advice. A.S. oversaw the study. All authors have approved the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw sequencing data has been deposited at the European Genome-phenome Archive (EGA) under accession number EGAS00001003062.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/eaam8419/suppl/DC1
Figs. S1 to S20
Tables S1 to S5
References

9 June 2017; resubmitted 19 March 2018
Accepted 13 July 2018
10.1126/science.aam8419

RESEARCH ARTICLE SUMMARY

SIGNAL TRANSDUCTION

Cancer mutations and targeted drugs can disrupt dynamic signal encoding by the Ras-Erk pathway

L. J. Bugaj, A. J. Sabnis, A. Mitchell, J. E. Garbarino, J. E. Toettcher*, T. G. Bivona*, W. A. Lim*

INTRODUCTION: Signaling pathways, such as the Ras-Erk (extracellular signal-regulated kinase) pathway, encode information through both their amplitude and dynamics. Differences in signal duration and frequency can lead to distinct cellular output decisions. Thus, temporal signals must be faithfully transmitted from the plasma membrane (Ras) to the nucleus (Erk) to properly control the cell's response. Because the Ras-Erk pathway regulates important cell decisions such as proliferation, changes to dynamic signal transduction properties could result in improper cell decisions and dysfunction. However, it has been difficult to examine whether corruption of signal transmission dynamics is associated with diseases such as cancer.

RATIONALE: We used optogenetic stimulation of the Ras-Erk pathway to quantitatively screen whether cancer mutations and drug treatments alter the fidelity of dynamic signal transmission. Most cancer-associated mutations in the Ras-Erk pathway are thought to drive cancer by inducing constitutive pathway activation—a high basal amplitude of activity. We explored whether cancer cells might also have altered dynamic properties that could contribute to disease. We used live-cell microscopy and new high-throughput optogenetic devices to systematically measure cell responses to a broad range of dynamic input stimulus patterns. We could detect subtle but important perturbations in pathway signal transmission properties by monitoring how these upstream stimulus patterns (generated by use of Ras-activating optoSOS) altered pathway output at the downstream levels of signaling, gene expression, and cell proliferation.

RESULTS: We found that cells that harbor particular B-Raf mutations (in the kinase P-loop) exhibit substantially corrupted dynamic signal transmission properties. In particular, the kinetics of Ras-Erk pathway inactivation are substantially slowed (half-time for signal decay is 10-fold longer). In these cancer cells, the active Erk output signal remains abnormally high for ~20 min after Ras input activity (optoSOS) is withdrawn (compared with 1 to 2 min for normal cells). Mutants or drugs that enhance B-Raf dimerization led to similar slow pathway deactivation. We could pinpoint B-Raf as the node responsible for altered transmission by using a combination of small molecular inhibitors and optogenetic stimulation at alternative input points.

Elongated pathway decay kinetics resulted in physiologically important cellular misinterpretation of dynamic inputs. In response to pulsatile inputs with intermediate frequencies, the perturbed cells responded with transcriptional profiles typically observed with sustained inputs. This signal misinterpretation propagated to proliferative decisions, resulting in aberrant cell-cycle entry in response to otherwise nonproliferative pulsatile inputs. These changes in pathway transmission shift the threshold of temporal input patterns that can drive cell proliferation, so that a space of inert input patterns that are normally filtered by the pathway can now drive proliferation.

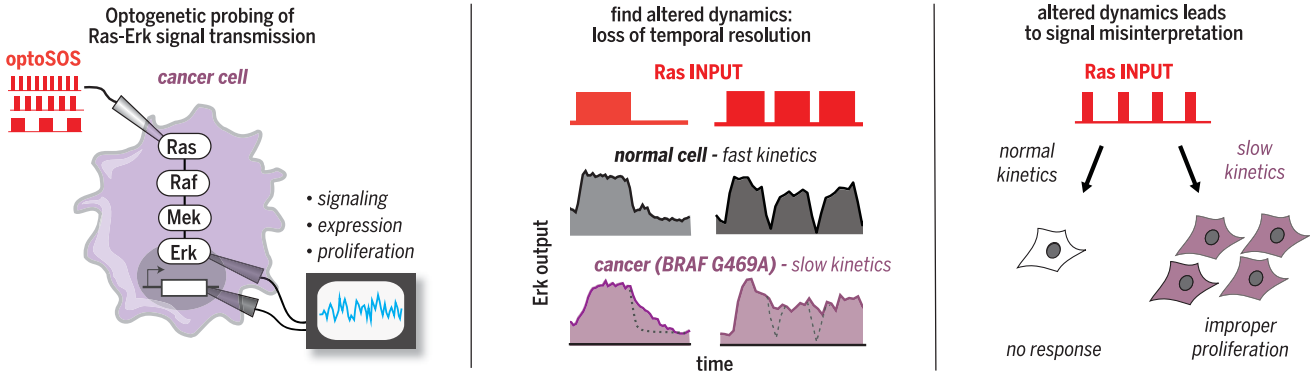
ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aao3048

CONCLUSION: Cancer mutations and targeted drugs can corrupt dynamic transmission properties in signaling pathways, shifting cellular response thresholds and changing cell decisions in a potentially pathological manner. Optogenetic approaches, especially in a high-throughput format, can be a powerful tool with which to systematically profile how a cell transmits and interprets information. We anticipate that further understanding the landscape of such functional alterations may help us mechanistically understand, stratify, and treat diseases that involve corrupted cellular decision-making. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: wendell.lim@ucsf.edu (W.A.L.); toettcher@princeton.edu (J.E.T.); trever.bivona@ucsf.edu (T.G.B.)
Cite this article as L. J. Bugaj *et al.*, *Science* **361**, eaao3048 (2018). DOI: 10.1126/science.aao3048

Optogenetic profiling of cancer cells reveals perturbed signal transmission dynamics that can drive improper proliferation. Optogenetic stimulation of Ras allows precise profiling of the fidelity of Ras-Erk pathway signaling in normal and cancer cells. We found that cancer cells with certain *BRAF* mutations have dramatically altered signal transmission dynamics compared with normal cells. These altered dynamics lead to a loss of temporal input resolution, so that the cancer cell may now misinterpret nonproliferative pulsatile input patterns as a trigger to proliferate.



RESEARCH ARTICLE

SIGNAL TRANSDUCTION

Cancer mutations and targeted drugs can disrupt dynamic signal encoding by the Ras-Erk pathway

L. J. Bugaj[1]*, A. J. Sabnis[2,3], A. Mitchell[1]†, J. E. Garbarino[4], J. E. Toettcher[1]‡§, T. G. Bivona[1,3,5]§, W. A. Lim[1,3,4,6]§

The Ras-Erk (extracellular signal-regulated kinase) pathway encodes information in its dynamics; the duration and frequency of Erk activity can specify distinct cell fates. To enable dynamic encoding, temporal information must be accurately transmitted from the plasma membrane to the nucleus. We used optogenetic profiling to show that both oncogenic B-Raf mutations and B-Raf inhibitors can cause corruption of this transmission, so that short pulses of input Ras activity are distorted into abnormally long Erk outputs. These changes can reshape downstream transcription and cell fates, resulting in improper decisions to proliferate. These findings illustrate how altered dynamic signal transmission properties, and not just constitutively increased signaling, can contribute to cell proliferation and perhaps cancer, and how optogenetic profiling can dissect mechanisms of signaling dysfunction in disease.

Signaling through the Ras-Erk (extracellular signal-regulated kinase) pathway controls diverse cell decisions, including survival, differentiation, and proliferation (*1*). A cell's fate is determined in part by the dynamics of Ras-Erk signals, which can be encoded by different receptors or cellular contexts (Fig. 1A) (*2*–*6*). Thus, the cell must be able to accurately transmit dynamic signal patterns and then decode them to make proper decisions (*7*–*15*). To understand how the cell transmits and decodes dynamic information, we recently developed optogenetic methods with which to interrogate cells with precisely controlled Ras inputs (Fig. 1B) (*16*). Our study revealed that the Ras-Raf-Mek-Erk protein kinase cascade acts as a high-fidelity transmission system, accurately transmitting dynamic signals with time scales ranging from minutes to hours to the nucleus [through nuclear localization of phosphorylated Erk (ppErk)] (Fig. 1A). In turn, downstream transcriptional networks then decode and integrate Erk dynamics to yield distinct cellular responses (*2*, *16*–*19*).

Given the functional importance of Ras-Erk dynamics, we realized that changes in how the pathway transmits and decodes signals could potentially lead to cellular malfunction and disease. Mutations within the Ras-Erk pathway underlie a large proportion of human tumors (*20*), and these mutations are commonly thought to drive cancer phenotypes through constitutive proliferative signaling. However, cancer phenotypes might also result from the corruption of proper dynamic signal transmission and decoding. Such changes could result in misinterpretation of dynamic environmental signals that might, for example, instruct cells to proliferate in response to normally nonproliferative inputs. Detecting potential defects in signal transmission and filtering requires appropriate tools that have only recently become available. We applied optogenetic profiling to identify alterations in Ras-Erk signaling dynamics within cancer cells, and we showed how these changes can result in inappropriate cellular decision-making.

Optogenetic profiling of Ras-Erk signal transmission in lung cancer cells

We examined Ras-Erk signaling in five patient-derived non–small cell lung cancer (NSCLC) cell lines with endogenous, validated oncogenes in the Ras-Erk pathway [in the epidermal growth factor receptor (EGFR), Ras, and B-Raf] (Cell lines and putative driver mutations are listed in table S1). As controls, we examined two normal human lung epithelial cell lines (Beas2B and 16HBE) and mouse NIH 3T3 fibroblasts. To probe how these cells processed dynamic Ras signals, we transduced each cell line with optoSOS, a genetically encoded light-activatible probe for toggling Ras activity in living cells (Fig. 1B). The optoSOS system relies on the light-dependent dimerization of PhytochromeB (PhyB) and phytochrome-interacting factor 6 (PIF), which associate when exposed to red (650 nm) light and dissociate when exposed to far-red (750 nm) light. PhyB was tethered to the membrane, and PIF, fused to the Ras-activating Son of Sevenless Homolog 2 (SOS2) catalytic domain (SOS2cat), was expressed in the cytoplasm. Therefore, light could be used to reversibly recruit SOS2cat to the membrane and thus dynamically modulate Ras activity. We tracked signal transmission from Ras to Erk through live-cell microscopy by co-expressing a blue fluorescent protein (BFP)–Erk2 reporter, which accumulates in the nucleus upon activation (Fig. 1C) (*21*). For more high-throughput and long-term analysis, we developed the optoPlate, a device for optogenetic illumination in microwell plates (Fig. 1C and fig. S1). This device allowed us to stimulate cells dynamically across a large parameter space and analyze multiple cellular outputs over time through fixed-cell fluorescence microscopy.

Identification of cancer cell line with substantially altered Ras-Erk signaling dynamics

One of the five lung cancer cell lines, H1395, had altered dynamic signal transmission properties (analysis of all the cancer cell lines is provided in fig. S2, A and B). When H1395 cells were subjected to various pulsatile optoSOS activation patterns, Erk activity (BFP-Erk2 nuclear localization) responded sluggishly whenever we switched optoSOS on or off (Fig. 2A). In particular, Erk activity took longer to diminish after optoSOS was switched off: The deactivation half-life ($t_{1/2}$) of Erk in H1395 cells was ~20-fold longer than that observed in normal NIH 3T3 control cells (H1395 $t_{1/2}$ = 21 min; NIH 3T3 $t_{1/2}$ = 1 min). We confirmed these slow dynamics by means of Western blot for ppErk (Fig. 2B and fig. S3A). Thus, in H1395 cells, instead of switching off immediately after Ras input stops, Erk continues to signal.

The slow responsiveness and decay of Erk activity may degrade the H1395 cell's ability to resolve distinct high-frequency dynamic patterns of Ras stimulation. We applied optoSOS pulse trains that resemble naturally observed pathway dynamics (*2*–*4*) to normal lung epithelial cells (16HBE) and H1395 cancer cells. Both cell types resolved input pulses spaced far apart (at 40-min intervals) (Fig. 2C and fig. S3B). However, as the input pulses were spaced progressively closer, the H1395 cells failed to distinguish the higher-frequency pulses. Ultimately, the H1395 cells failed to perceive gaps in the signal and produced a constant Erk response (Fig. 2C, 5′ OFF condition). Thus, compared with normal cells, H1395 cells have an impaired ability to perceive dynamic pathway input (Fig. 2D).

[1]Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA 94158, USA. [2]Department of Pediatrics, University of California, San Francisco, San Francisco, CA 94158, USA. [3]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA 94158, USA. [4]Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, CA 94158, USA. [5]Division of Hematology and Oncology, University of California, San Francisco, San Francisco, CA 94158, USA. [6]Center for Systems and Synthetic Biology, University of California, San Francisco, San Francisco, CA 94158, USA. *Present address: Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA. †Present address: Program in Systems Biology, University of Massachusetts Medical School, Worcester, MA 01655, USA. ‡Present address: Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA. §Corresponding author. Email: wendell.lim@ucsf.edu (W.A.L.); toettcher@princeton.edu (J.E.T.); trever.bivona@ucsf.edu (T.G.B.)

B-Raf P-loop mutation (G469A) slows kinetics of signal decay

The H1395 cell line harbors a B-Raf G469A mutation. We thought that the slow OFF-kinetics of Erk could be caused by either defects in switching off Erk or Mek (such as defects in downstream phosphatase function) or by defects in switching off the mutant B-Raf. To find nodes in the pathway that may cause the slow dynamics, we repeated our optoSOS stimulation studies, but when switching off optoSOS, we also added inhibitors to block particular steps in the pathway—either the Mek inhibitor U0126 (which rapidly shuts off signal flow from Mek to Erk) or mutant B-Raf inhibitor PLX-8394 (which rapidly shuts off signal flow from mutant B-Raf to Mek) (Fig. 3A and fig. S4A) (*22*). When either of these inhibitors was added concurrently with optoSOS inactivation in H1395 cells, active phospho-Erk (ppErk) decayed rapidly (Fig. 3A and fig. S4B), suggesting that normal Mek and Erk dephosphorylation activity was intact and that the source of extended signal decay lay upstream. Together, these experiments indicated that slow ppErk OFF-kinetics might emanate from mutant B-Raf.

To further test whether the mutant B-Raf accounted for the slow OFF-kinetics of the pathway, we performed optogenetic profiling in which we linked the light-induced input to a different node. We used an optogenetic tool called optoBRaf. OptoBRaf is activated through inducible membrane recruitment of PIF fused to wild-type B-Raf, which stimulates its signaling to endogenous Mek (Fig. 3A, bottom) (*23*). OptoBRaf enabled us to stimulate the H1395 cells in a manner that bypassed the B-Raf G469A mutant. We observed rapid ppErk deactivation kinetics with optoBRaf stimulation, which was again consistent with a model in which the B-Raf G469A mutant is directly responsible for the altered pathway dynamics (Fig. 3A, bottom, and fig. S4C). This experiment produced a rebound in ppErk signal after the initial rapid decay. This may be caused by relief of negative feedback of ppErk onto mutant B-Raf (*24*, *25*) because repeating this experiment in the presence of PLX-8394 eliminated this rebound. We also observed extension of Ras-Erk kinetics in Beas2B normal lung epithelial cells (lacking endogenous mutations in Ras, Raf, Mek, or Erk) engineered to express exogenous B-Raf G469A, and this extension was reversed in the presence of the B-Raf inhibitor PLX-8394 (fig. S4, D, E and F). Together, these results implicate the mutant B-Raf G469A as a kinetics-altering node in H1395 cancer cells.

Lagging pathway kinetics of B-Raf G469A mutants were specific to mutation position because another member of our cancer cell panel—HCC364—carried a B-Raf V600E mutation and showed wild-type, fast pathway kinetics (fig. S2B). G469 lies in the P-loop of B-Raf, which normally associates with the activation loop to maintain the B-Raf kinase domain in an inactive, autoinhibited conformation (*26*). Normal activation of B-Raf requires release of this autoinhibition, which both frees the kinase domain and promotes the activating homo- or heterodimerization of

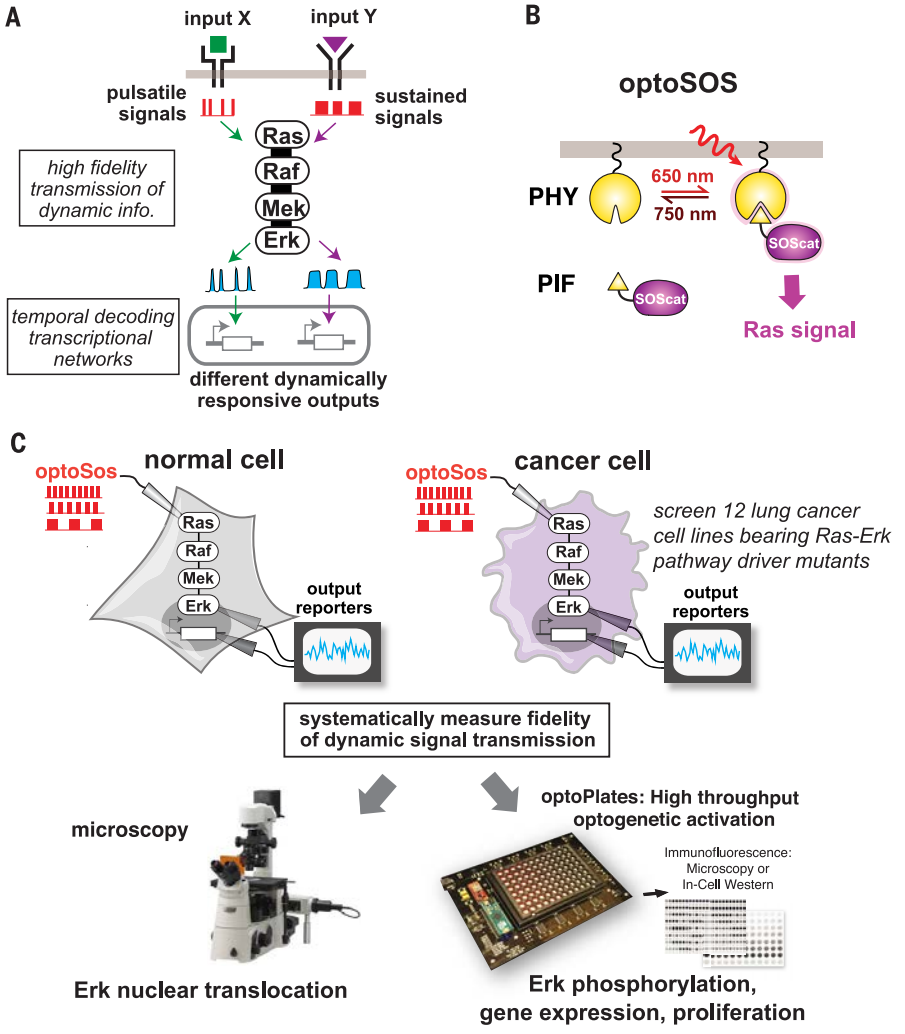


Fig. 1. Probing dynamic signal transduction and filtering in cancer cells. (**A**) Environmental stimuli can induce different dynamic patterns of Erk activity, which are then interpreted by downstream transcriptional circuits to specify cell behavior. (**B**) OptoSOS is an optogenetic method for Ras activation that enables probing of how cells filter and respond to dynamic Ras inputs. The light-inducible PhyB-PIF heterodimer drives membrane recruitment of the Ras-activating SOS2 catalytic domain, which activates Ras at the membrane. Red light (650 nm) induces PhyB-PIF dimerization, whereas far-red light (750 nm) dissociates the dimer. (**C**) We tested the hypothesis that some cancer cells may inappropriately filter dynamic Ras-Erk signals. We examined how dynamic optogenetic inputs were interpreted by normal or cancer cells through a combination of live-cell microscopy, high-throughput optogenetic stimulation (fig. S1), and immunofluorescence.

B-Raf. Oncogenic P-loop mutations both disrupt the inactive conformation and enhance dimerization of B-Raf (*26*–*28*). Several P-loop mutations impair B-Raf activity yet can also be oncogenic, likely because they enhance C-Raf transactivation through B-Raf–C-Raf dimerization (*26*).

Enhanced B-Raf dimerization induced by the P-loop G469A mutation might also cause the delayed OFF-kinetics. If so, delayed OFF-kinetics should be reversed in the presence of mutations that disrupt B-Raf dimerization, such as the R509H mutation (*29*, *30*). Indeed, Beas2B cells transiently transfected with *BRAF* G469A showed elongated ppErk decay kinetics, whereas cells transfected with the *BRAF* G469A/R509H double mutant showed wild-type kinetics (fig. S5).

Other B-Raf P-loop mutant lines would also be expected to show extended ppErk inactivation kinetics. Indeed, we searched for additional lung cancer cell lines driven by B-Raf P-loop mutations and found that they also showed slow ppErk response to input Ras pulses (fig. S6A). Specifically, the Cal12T and H1666 lung cancer cell lines, each expressing endogenous B-Raf G466V (a distinct oncogenic P-loop mutation from G469A), showed a similar sluggish response. As in H1395 cells, this slowed response was reversed in the presence of the Raf inhibitor PLX-8394. Unlike the activating G469A mutation, G466V decreases catalytic activity (*26*). Nevertheless, both mutants extended Ras-induced ppErk kinetics, which may reflect their shared propensity for enhanced

dimerization with C-Raf. Although we readily found examples of altered kinetics in *BRAF*-mutant cells, a screen of cell lines with diverse Ras mutants revealed no similar examples (fig. S7, A and B), suggesting that B-Raf is a more sensitive point for altering ppErk signaling dynamics.

Paradox-activating drugs that perturb B-Raf dimerization also alter Ras-Erk kinetics

Drugs that enhance B-Raf homo- or heterodimerization would also be expected to yield extended Erk inactivation kinetics. We therefore examined ppErk kinetics in the presence of so-called paradox-activating B-Raf inhibitors. Although designed to inhibit mutant B-Raf activity, this family of drugs paradoxically activates Raf-Mek-Erk signaling by enhancing B-Raf dimerization with C-Raf (*29*, *31*, *32*). This paradoxical pathway activation can actually stimulate cancer formation in certain patients receiving these drugs (*33*–*36*). We found that both vemurafenib and SB590885—two B-Raf inhibitors in this class—extended the otherwise fast Erk kinetics in both wild-type fibroblasts (Fig. 3B) and Beas2B lung epithelial cells (fig S6B). Although mechanistically similar, vemurafenib and SB590885 are chemically distinct and had differing dose-dependent effects on ppErk signal kinetics (fig. S8). By contrast, the B-Raf inhibitor PLX-8394, which does not enhance Raf dimerization, had no effect on Erk kinetics. As with G469A mutant-extended kinetics, drug-extended kinetics were reversed with Mek inhibition, indicating that the drugs extended kinetics through a mechanism upstream of Mek activation (fig. S9).

Increasing Raf dimerization with B-Raf inhibitors can enhance active Ras nanocluster formation, resulting in increased gain between Ras and Erk but no change in the dynamics of nanocluster formation (*37*, *38*). Our results confirm increased gain from Ras to Erk (fig. S8) and are thus consistent with this mechanism. Further, because slow ppErk kinetics emerge despite fast Ras nanocluster decay dynamics (*38*), we conclude that the sustained ppErk signal originates downstream of Ras activation and cluster formation at the level of Raf activation. In total, these data support a model in which B-Raf P-loop mutations that enhance homo- or heterodimerization cause a lag in the dynamics of shutting off overall Raf activity (Fig. 3C).

Modeling how slow Ras-Erk dynamics could alter cell decisions

We examined how such altered Ras-Erk transmission properties could affect downstream cellular decision-making. Changes in the dynamic response of the Ras-Erk pathway fundamentally change how the cell filters dynamic inputs. The wild-type Ras-Erk pathway filters signals shorter than ~4 min (the pathway loses ability to transmit more transient changes) while faithfully transmitting longer ones ranging from minutes to hours (*16*). To examine the consequences of changing these filtering parameters, we constructed a simple model that integrates a low-pass filter with

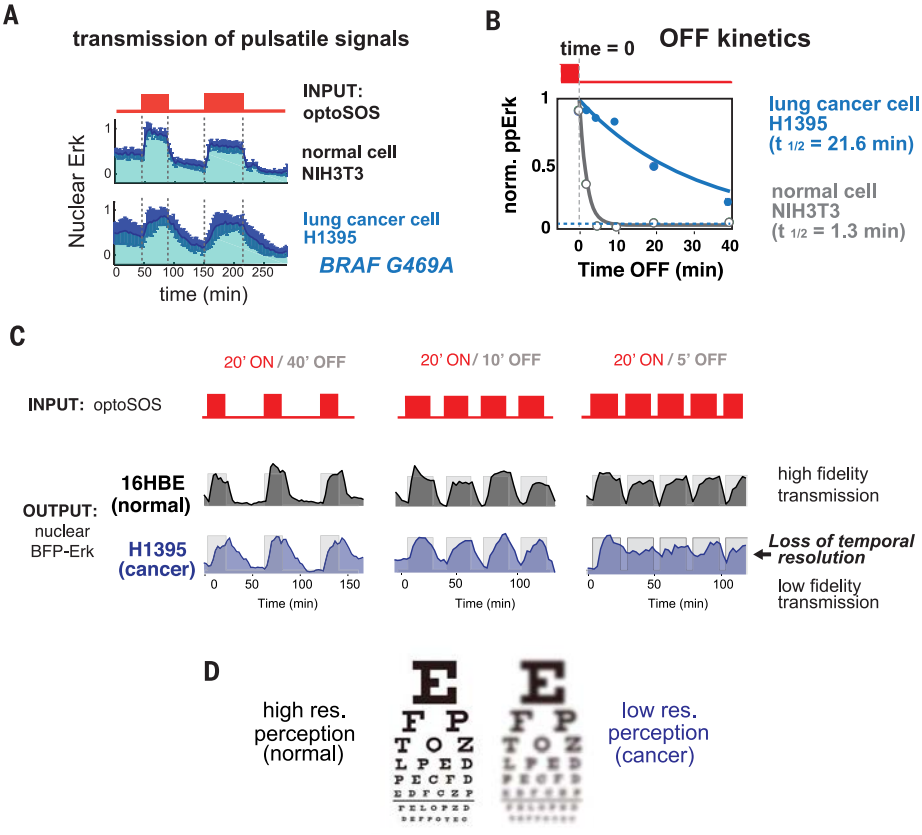


Fig. 2. B-Raf mutant H1395 cells have an impaired transmission of pulsatile Ras signals. (**A**) H1395 cells (bottom) showed extended kinetics of activation and inactivation in response to defined Ras input pulses. By contrast, NIH 3T3 cells (top) and the other cells in our cell line panel exhibited rapid kinetics (fig. S2B). Traces represent quantitation from live-cell imaging of nuclear BFP-Erk2 reporter accumulation. Traces were normalized between 0 and 1 and represent the mean ± 1 SD of 15 and 14 cells for 3T3 and H1395 cells, respectively. (**B**) Inactivation kinetics for H1395 and NIH 3T3 cells were confirmed through Western blot (blots are available in fig. S3A). Western blot quantification of ppErk is shown and fitted to single exponential decay. The dashed blue line depicts basal amount of ppErk from unstimulated cells. (**C**) Loss of fidelity in dynamic signal transduction in H1395 cells was observed through live-cell microscopy. 16-HBE (normal) and H1395 (cancer) cells were subjected to various dynamic patterns of input signal. (Three input conditions are shown. All six input conditions are shown in fig. S3B). As optoSOS input frequency increased, the H1395 cancer cells progressively lost their response to the gaps in the signal, whereas the normal cells did not. Traces represent the mean of five cells. Individual traces can be seen in fig. S3B. (**D**) Changes in the cell's signal perception are analogous to cellular "blurred vision" for external stimuli. (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; G, Gly; and V, Val. In the mutants, other amino acids were substituted at certain locations; for example, G469A indicates that glycine at position 469 is replaced by alanine.)

downstream transcription and resultant cell fate commitment (fig. S10A). In this model, the cell senses whether the intensity of a ppErk signal rises above an activation threshold, above which ppErk-dependent transcription begins (fig. S10B).

We compared this system with fast (normal) and slow system responses. In the model, slowing the system response can sustain otherwise transient signal activity in response to dynamic inputs (fig. S10B). In particular, pulse trains of short pulses were poorly resolved and interpreted as a stronger, more continuous input. Over a sufficiently long integration time, cellular decisions for downstream outputs such as gene expression and proliferation could differ substantially between the fast and slow pathway models (fig. S10, C and D).

Because oncogene- and drug-extended kinetics are often accompanied by increased basal signaling (figs. S4, E to F; S6B; S8; and S9), we examined the effects of increased basal signaling by changing the activation threshold in our model. Increased basal signaling is equivalent to a lower activation threshold. In the model, although increased basal signaling (lower threshold) minimally sensitized cells to proliferate under fast ppErk kinetics, slow ppErk kinetics dramatically increased the proliferative response (fig. S10D). Thus, although increased signaling and extended kinetics may synergize to control cellular response, our model

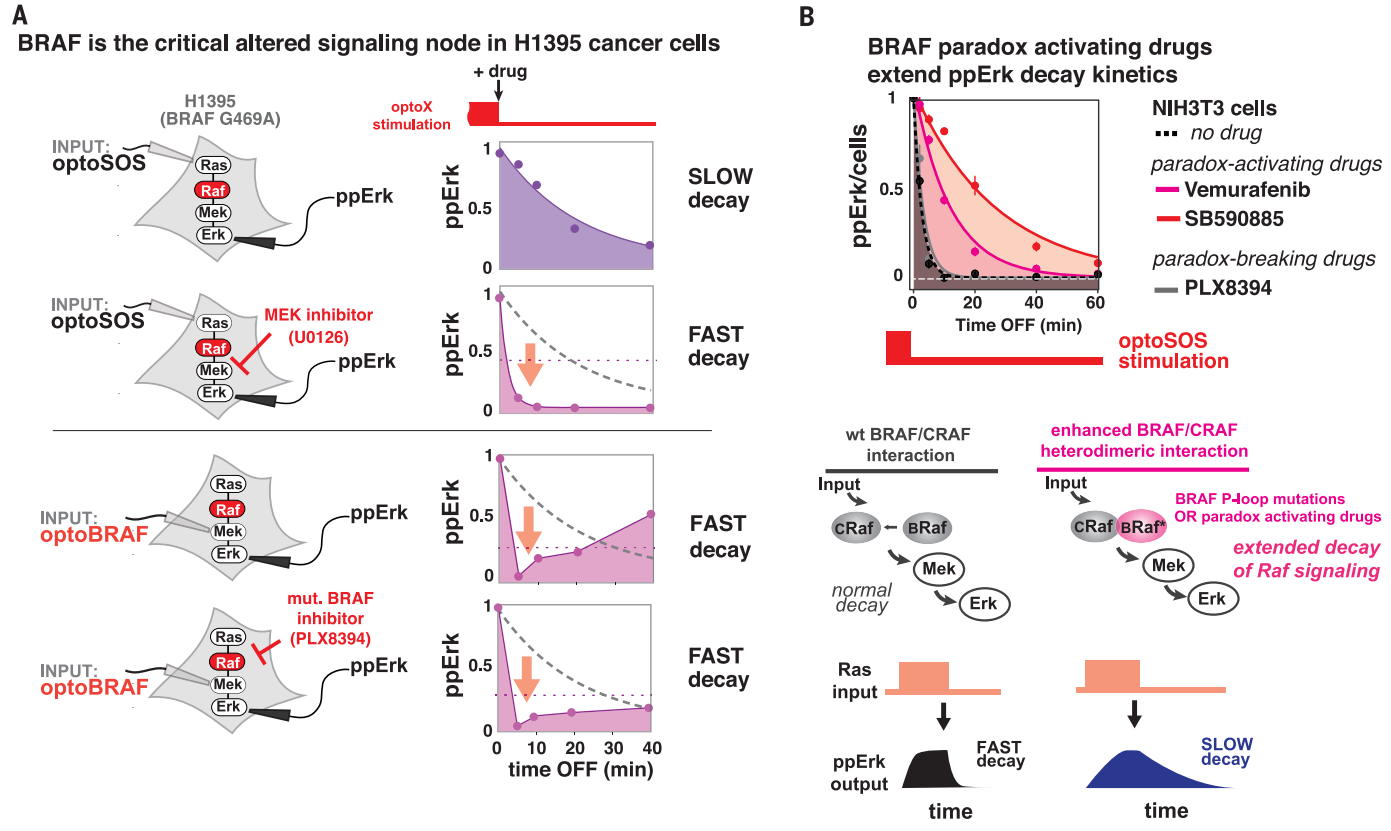


Fig. 3. B-Raf P-loop mutations and drugs that perturb Raf dimerization both extend Ras-Erk pathway kinetics. (**A**) OptoSOS and optoBRaf coupled with MEK inhibition (U0126) and mutant–B-Raf inhibition (PLX-8394) were used to isolate B-Raf as a network node that can extend Erk kinetics. Plots show quantification of Western blot data (blots are available in fig. S4, A, B, and C). Normal wild-type decay is indicated with the gray dashed line; basal signaling level (no opto-stimulation) is indicated by the purple dotted line. (**B**) Treatment of NIH 3T3s with paradoxically activating B-Raf inhibitors vemurafenib and SB590885 also extended ppErk decay kinetics. Datapoints show means ± 95% confidence interval (CI) of mean single-cell ppErk immunofluoresecnce from three replicates. (**C**) Our data support a model in which P-loop B-Raf mutations or paradoxically activating drugs can both enhance the Ras-induced dimerization potential of B-Raf and C-Raf, thus altering the kinetic properties of pathway activation and inactivation.

predicts that kinetics can have a dominant role in downstream cellular behavior.

Slow pathway dynamics alter transcriptional responses

To experimentally test whether changes in signal transmission dynamics could alter gene expression decisions in cells, we measured the amounts of several downstream output proteins in response to optoSOS inputs. We sought an experimental model in which we could isolate the effects of altered Ras-Erk kinetics in a well-controlled cell line that lacks potentially confounding mutations. Thus, we compared the responses of wild-type NIH 3T3 cells in the presence and absence of 100 nM SB590885 (paradox inhibitor of B-Raf). This concentration of drug extended Ras-Erk pathway decay kinetics and minimally increased basal ppErk levels (figs. S8, A and B, and S11A). Cells were seeded and serum-starved in 384-well plates and, in the presence or absence of drug, exposed to various dynamic input patterns with the optoPlate (Fig. 4A and fig. S11B). After stimulating the cells over several hours, cells were fixed and immunostained for Erk-dependent transcriptional targets.

We measured the expression of two immediate early gene targets, cJun and early growth response protein 1 (EGR1), and the cell-cycle regulator Cyclin D1 (Fig. 4A). These are targets that show strong dependence on the dynamics of input signals. In normal NIH 3T3s, cJun and Cyclin D1 expression are strongly induced with continuous optoSOS stimulation and are not induced by pulsatile, transient stimulation. By contrast, EGR1 shows a peak of expression with continuous stimulation but then shows an adaptive decrease in expression, potentially mediated by negative feedback (*39*). EGR1 expression stayed higher with pulsed stimulation (30 min on, 30 min off), probably because such pulsatile stimulation prevents the accumulation of maximal negative feedback (Fig. 4B and fig. S11, C to H).

These dynamically responsive Ras-Erk gene targets showed changes in regulation with altered Ras-Erk signal transmission. In all cases, when we performed dynamic stimulation studies in the presence of the drug SB590885, the pulsed-input response shifted to more closely resemble that of the normal constant-input response. Upon SB590885 addition, both cJun and Cyclin D1 accumulated in response to normally subthreshold pulsed stimulation. Conversely, pulsed stimulation in the presence of SB590885 yielded the adaptive response of EGR1 normally observed with a constant Ras input (Fig. 4B and fig. S11, C to H). Together, these results show that altering Ras-Erk transmission kinetics can change how cells filter dynamic signals, altering expression patterns of the genes that control important cell decisions.

Slow pathway dynamics shift thresholds for inducing proliferation

Because the Ras-Erk pathway is a key driver of proliferation, and because we observed differential expression of the cell-cycle regulator Cyclin D1, we tested whether slowed Ras-Erk transmission kinetics might corrupt proper control of proliferation by driving cell-cycle entry in response to what are normally nonproliferative dynamic input patterns. Dynamic Erk signals are linked to cell-cycle control in vivo, but the physiological parameters of Erk dynamics are not well defined (*2*–*4*, *40*). We therefore tested how cells responded to a range of dynamic Ras inputs, both in the presence and absence of SB590885-induced delay in ppErk kinetics. We examined a set of signal patterns that, after 19 hours of stimulation, could drive cell-cycle entry, as assayed by means of DNA incorporation of 5-ethynyl-2-deoxyuridine (Edu) during S-phase (Fig. 4C and fig. S12, A, B, and C).

Cells in each microwell received either no signal, a constant signal, or a periodic signal. The time in the ON phase (ON interval) and OFF phase (OFF interval) of the periodic signal was systematically varied between microwells.

The resulting heatmap of proliferation as a function of the dynamic stimulation pattern is shown in Fig. 4C (no-signal and constant-signal conditions are depicted in fig. S12D; heatmaps are shown in Fig. 4C and fig S12E). In normal 3T3s, we delineated a "growth regime" of input signals that drive the strongest proliferation. This showed that Ras-Erk–induced cell-cycle entry required strong and sustained signals: those largely consisting of long ON intervals separated by short OFF intervals. Slowing Ras-Erk transmission kinetics with drug dramatically expanded this growth regime, increasing proliferation across a range of otherwise nonproliferative Ras input patterns (Fig. 4C and figs. S12F and S13).

Conclusion: Mutants that alter how a cell perceives signals can contribute to disease

Some cancer mutations or targeted drugs can alter a cell's dynamic signal transmission and filtering properties, and such changes can reshape how a cell perceives or misperceives its environment, potentially contributing to disease phenotypes (Fig. 4D). Such signal misinterpretation might contribute to hyperproliferation, but Ras-Erk signaling functions in many cell behaviors, including cell survival and migration, and thus defective signal transduction may plausibly affect these behaviors in disease as well. Functional profiling of intact signaling networks with optogenetics provides a powerful method with which to uncover and understand such alterations in cellular decision-making (*41*). We anticipate that profiling more cancers and more pathways in this manner may uncover other types of dynamic signaling phenotypes that could contribute to disease. The improved understanding we gain of how mutated signaling networks differentially process information may help us mechanistically understand cancer, autoimmunity, and other diseases that involve corrupted cellular decision-making and may provide new dynamically optimized strategies for targeting these diseases (*42*–*44*).

Materials and Methods

Plasmid constructs, viral packaging, and transduction

OptoSOS components Phy-mCh-CAAX (Addgene #50839), YFP-PIF-SOS2cat (Addgene # 50851), and BFP-Erk2 (Addgene #50848) were described previously (*16*, *45*). Phy-mCh-CAAX used the KRas4B-derived CAAX sequence KMSKDGKKKKKKAKTKCVIM, which is expected to be farnesylated. This sequence differs from the wild-type KRas4B CAAX only at the underlined alanine, which represents an S>A mutation. This mutation was made to prevent endogenous regulation at this residue, as previously reported (*46*). PAmCh-BRAF(G469A) was created by site-directed mutagenesis of PAmCh-BRAF, a generous gift from Eric Collison (UCSF), using the QuikChange Lightning kit (Agilent). For optoBRAF, a YFP-PIF-BRAF was created by PCR amplification of the YFP-PIF-SOS2cat vector backbone to exclude S0S2cat, PCR amplification of wtBRAF, and ligation using the In-Fusion enzyme cocktail (Clontech). Lentivirus was packaged by cotransfecting the transfer vector, pCMVdR8.91, and pMD2.G (Addgene # 12259) into Lenti-X 293T cells (Clontech) using the Fugene 6 HD transfection reagent. 48 hours after transfection, viral supernatant was harvested, sterile filtered through a 0.45 μm filter, and added to cells for infection. Unused supernatant was stored at -80°C. 72 hours after infection, transduced

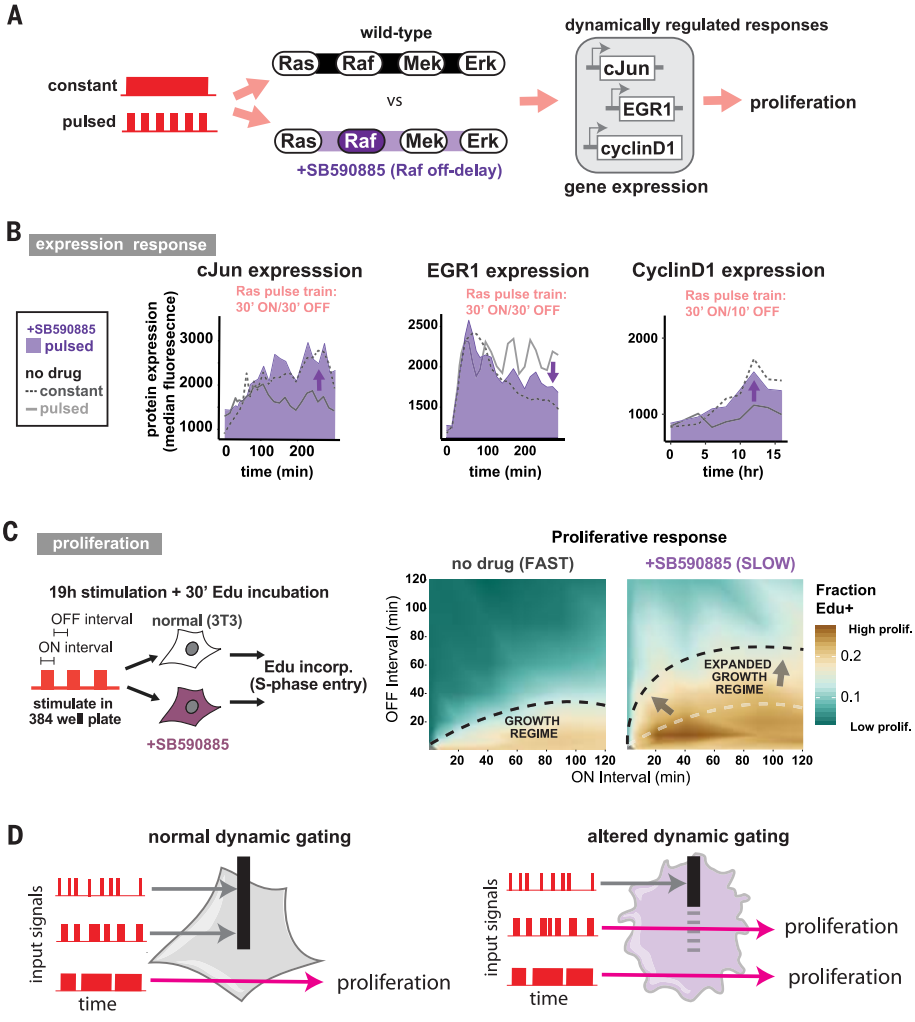


Fig. 4. Perturbation of Ras-Erk signaling dynamics can alter how cells make proliferative decisions. (**A**) Transcriptional decoding of dynamic signal inputs was examined in normal NIH 3T3 cells (fast pathway kinetics) or in cells treated with the kinetics-altering drug SB590885. Cells were stimulated with fixed-width signal pulses separated by various intervals. Expression of Erk targets and downstream cell-cycle entry were examined. (**B**) Altered Ras-Erk kinetics changed transcriptional output to dynamic Ras inputs. Immunofluorescence of cJun, EGR1, and Cyclin D1 expression time courses is depicted. Only expression in response to constant stimulus or a representative pulsed stimulus is shown.All input conditions tested are provided in fig. S11. Illumination was achieved with the optoPlate, and protein expression was assessed through single-cell immunofluorescence coupled with high-content imaging. Data points represent the median target fluorescence from 3000 to 4000 cells for each condition. (**C**) Extended Ras-Erk kinetics sensitized cells to proliferate under nonproliferative conditions. We used 384-well optoPlate illumination to examine proliferation of cells in response to a systematic scan of dynamic inputs. Normal and drug-treated cells were exposed to all combinations of six optoSOS pulse lengths (ON interval) and separated by seven pulse interval lengths (OFF interval) over 19 hours. Cells were then incubated with Edu for 30 min, fixed, stained, imaged, and analyzed. The percentage of cells incorporating Edu was plotted as an interpolated heatmap. Further analysis is available in figs. S12 and S13. The values used to generate the map represent means of biological quadruplicates. (**D**) Our data support the model that altering dynamic signal filtering properties can reshape the input-response map and may drive improper cellular behavior, such as hyperproliferation.

cells were sorted for expression using a FACS Aria Fusion (BD Biosciences).

Cell lines, cell culture, and inhibitors

All cell lines were maintained in standard tissue culture incubators at 37° C and 5% CO_2. NIH 3T3s (ATCC) were cultured in DMEM supplemented with 10% calf serum (HyClone) and 1% penicillin/streptomycin/glutamine (ThermoFisher # 10378016). Lenti-X 293T cells were cultured in DMEM High Glucose H-21 (UCSF Cell Culture Facility) supplemented with 10% FBS (UCSF Cell Culture Facility) and 1% penicillin/streptomycin (UCSF Cell Culture Facility). All other lines were cultured in RPMI (UCSF Cell Culture Facility) supplemented with 10% FBS and 1% penicillin/streptomycin/glutamine. The Mek inhibitors U0126 (Selleckchem #S1102) and trametinib (Selleckchem #S2673) and B-Raf inhibitors vemurafenib (Selleckchem #S1267) and SB590885 (Selleckchem #S2220) were obtained from Selleckchem. PLX-8394 was obtained as a gift from Plexxicon.

Optogenetic stimulation

For all optogenetic experiments, cells were supplemented with HPLC-purified phycocyanobilin (PCB, Frontier Scientific #P14137) at a concentration of 5 μM (3T3s) or 10 μM (all other cells). Cells were incubated in PCB for ~0.5-1 hour before optogenetic stimulation. For bulk Western blot experiments, cells were illuminated in a cell culture incubator with a custom built panel of either 650 nm or 750 nm LEDs for activation or inactivation of optoSOS, respectively. For 96- and 384-well In-Cell Western and immunofluorescence assays, optogenetic experiments were performed with a custom-built 96-well "optoPlate" illuminator with adapters accommodating either 96- or 384-well plates (see fig. S1). Briefly, a printed circuit board was designed using the Kicad software package and manufactured through PCBUnlimited (PCBUnlimited.com). The circuit board design allowed placement of 192 independently addressable LEDs, with two LEDs—one red (Vishay, VLMK31R1S2-GS18), one far-red (Marubeni, SMT780)—fitting under each well position. The LEDs shared a common anode, and each cathode was connected to one of 12 24-channel constant-current LED drivers (TLC5947, Texas Instruments). These drivers allow independent 12-bit grayscale control (0-4095) of each LED using pulse-width modulation. LED drivers were controlled by an on-board Arduino Micro microcontroller, which was programmed with custom script through the Arduino IDE. Custom adapters interfacing with 96- and 384-well plates were designed in the Autodesk Inventor program and printed on a Stratasys uPrint 3D printer.

Western blot

Cells were seeded in 6-well plates at a density of $1x10^5$ cells per well. After 24h, cells were starved in starvation medium (DMEM or RPMI media supplemented with 1% penicillin/streptomycin/glutamine, and 20 mM HEPES). Cells were lysed in ice-cold RIPA buffer supplemented with protease (cOmplete, Sigma #4693159001) and phosphatase inhibitors (PhosSTOP, Sigma #4906845001). After a 10 min centrifugation at 4°C, supernatants were supplemented with 5X Laemmli's sample buffer and were boiled for 10 min. SDS PAGE was performed in with in NuPAGE Bis-Tris gels (Invitrogen) using MES buffer (ThermoFisher #NP0002), and blots were transferred onto nitrocellulose membranes using the BioRad Trans-Blot Semi-Dry Transfer Cell. Transferred blots were then blocked with Odyssey blocking buffer (LI-COR #927-4000) and antibody stained using the Freedom Rocker liquid handling system. Western blots were imaged on a LI-COR Odyssey imager, and images were quantified using ImageJ. Phospho-Erk antibody was obtained from Cell Signaling Technologies (#4370), alpha-tubulin antibody was obtained from Santa Cruz Biotechnology (Santa Cruz, #23948, 1:1000), and IRDye conjugated secondary antibodies (#926-3221, #926-68020) were obtained from LI-COR.

96- and 384-well optogenetic experiments

Cell seeding, starvation, and illumination

96- or 384-well plates (Greiner #655087 and #781092) were coated in fibronectin (Millipore, #FC010,1:50 dilution in PBS) for 30 min in the incubator. Cells were seeded at 5000 or 1000 cells per well for 96- or 384-well experiments, respectively, and were spun down in the plate for 1 min at 100 × g immediately after seeding to obtain an even spatial distribution of cells. After 24h, cells were starved with starvation medium (basal medium with 1% penicillin/streptomycin/glutamine and 20 mM HEPES). To balance effective starvation while minimizing loss of cells, cells in 96-well experiments underwent one full medium replacement followed by 4X 70% replacements, performed manually. Cells in 384-well experiments underwent 7X 70% starvation washes with the Biomek FX liquid handling robot. After starvation for 24 hours (signaling experiments) or 36 hours (growth experiments), cells were supplemented with phycocyanobilin (PCB) by mixing a 2X PCB solution in starvation media, removing starvation media from the plate manually, and adding the appropriate amount of 2X PCB. Cells in PCB were then incubated in the dark for 30 min. Any additional drugs were added at the same time and in the same solution as the PCB, with the exception of Mek-i addition in the experiment described in fig. S9. All manipulations with cells in PCB were done under dim light settings, and cell-containing plates were covered with aluminum foil whenever possible to prevent unintended photoactivation. The plates were then placed onto a pre-programmed optoPlate device and exposed to the desired illumination profiles.

Cell fixation, immunostaining, and antibodies

Upon completion of the experiment, cells were immediately supplemented with 16% PFA to a final concentration of 4%PFA. After fixing for 10 min, the PFA-containing medium was manually aspirated with a multichannel pipette and cells were permeabilized with 0.5% Triton X-100 (Sigma) for 10 min followed by ice-cold 100% methanol at -20°C for 10 min. Cells were then blocked for 30 min at room temperature with Odyssey Blocking Buffer (LI-COR). Primary antibodies were then diluted in fresh blocking buffer: anti-ppErk (CST #4370, 1:200), anti-ppErk(CST # 4344, 1:50), anti-cJun (CST #9165, 1:100), anti-EGR1 (CST #4154, 1:800), anti-CyclinD1 (Abcam #ab134175, 1:100). Blocking buffer was removed and cells were incubated in primary antibody solutions overnight. Cells were washed 5X with PBS with 0.1% Tween-20 (Sigma). All washes were performed with a BioTek EL406 liquid handler. Cells were then incubated in secondary antibody solutions. For In-Cell Western experiments, IRDye 800CW-conjugated goat anti-Rabbit (Licor, #926-32211, 1:800) secondary was used, and CellTag700 (Licor, #926-41090, 1:2000) was used for normalization. For single cell immunofluorescence, Alexa-488 and Alexa-647 conjugated goat anti-Rabbit secondary antibodies (Jackson Immunoresearch, #111-545-003 and #111-605-003, 1:100) were used in conjunction with DAPI (Molecular Probes, #D1306, 300 nM) for nuclear labeling. After 1 hour of secondary antibody incubation, cells were washed 5X in PBS + 0.1% Tween-20.

Imaging

In-Cell Western: For In-Cell Western experiments, plates were imaged on the LI-COR Odyssey scanner. Intensity measurements for each well were exported using the integrated In-Cell Western analysis software and were further analyzed in R.

High content imaging: Single cell immunofluorescence and Edu labeling (below) was measured on the ThermoFisher Scientific ArrayScan XTI High Content Platform imager, and image quantitation was conducted through the integrated HCS Studio software. Briefly, cells were identified through segmentation of DAPI-stained nuclei, and parameters were specified to ensure proper segmentation of single cells. Mean nuclear intensities were then calculated for each cell for the fluorescence channels reporting on all targets except for ppErk. For ppErk, a 2-5 pixel-wide ring was drawn around the nucleus, and mean fluorescence intensity in this ring was recorded. The fluorescence measurements were then exported, and further analysis was conducted in R.

Edu proliferation assay

S-phase entry was assessed through Edu incorporation using the Click-iT Edu AlexaFluor 555 Imaging kit (ThermoFisher, #C10338). Cells were seeded and starved as described above in the "96- and 384-well optogenetics experiments" section. After completion of the desired illumination time, cells were supplemented with a 30 min Edu pulse (5 μM) for 30 additional minutes of illumination. Cells were then fixed and permeabilized as described, and Edu was conjugated to AlexaFluor 555 as per manufacturer instructions.

Live cell microscopy

Cells were seeded in 384 well glass bottom plates (Matrical, Inc. #MGB101-1-2-LG) that were pre-coated

with 50 uL of 20 μg/mL fibronectin (Millipore # FC010) for 1 hour. Upon seeding, cells were spun down at 100 × g for 1 min to promote an even distribution of cells on the well bottom. The following day, cells were starved with starvation medium (defined above). 1 hour before imaging, starvation medium was replaced with fresh starvation medium containing 5 μM PCB.

Confocal imaging was conducted on a Nikon Eclipse TI inverted microscope with a Yokagawa CSU-X1 spinning disk confocal unit, a 20× PlanApo TIRF 1.49 NA objective, and an EM-CCD camera (Andor). Environmental control was maintained with a humidified environmental chamber at 37°C and 5% CO_2 (In Vivo Scientific). BFP, YFP, and mCherry were imaged with 405 nm, 488 nm, and 561 nm lasers (LMM5, Spectral Applier Research), respectively.

Cells were exposed to 650 nm and 750 nm light for optogenetic control as described previously (*16*). Briefly, a 650 nm LED was mounted into the epifluorescence illumination port, and its light intensity (voltage control) was regulated with custom Matlab scripts (*45*) controlling the analog output of a DT9812 board (Data Translation), which was connected to the LED. 750 nm light was applied by filtering bright-field light through a 750 nm longpass filter (FSQ-RG9, Newport) and controlling its timing through software control of the diascopic shutter.

Image analysis

BFP-Erk2 responses to variable Ras input pulses

For visualizing BFP-Erk2 responses to dynamic Ras activation, live cell imaging was analyzed with a combination of ImageJ and custom R scripts. Nuclear accumulation of BFP-Erk2 was measured by mean fluorescence intensity of an ROI within the cell nucleus in the BFP channel, and YFP-PIF-SOS2cat membrane translocation was measured by cytoplasmic depletion of YFP, as described previously (*16*). BFP and YFP traces for individual cells were then corrected for fluorescence drift. Traces underwent a linear transform by calculating a linear regression of all points, and then subtracting this fit from the original trace. Photobleaching was corrected by fitting an exponential decay envelope to each trace and dividing each trace by its envelope function. Traces were then normalized between 0 and 1, and the YFP-SOS2cat trace was inverted for visual clarity.

Edu and immunofluorescence analysis

For quantifying the fraction of Edu+ cells in a well, density plots of Edu intensities of cells in each well were constructed. The Edu distribution was bimodal, with a tight peak for Edu- cells and a broader peak for Edu+ cells. A custom segmentation algorithm classified the Edu- and Edu+ populations for each trace. Example peaks and segmentation are shown in Fig. S12C.

For single cell immunofluorescence, fluorescence distributions were generated for each well and the distribution median was extracted.

Model and fitting

Low pass filter of signal processing

To model how cells filter and respond to signals, we constructed a 3-step model. The first step describes signal filtering, the second describes signal perception, and the third describes cell fate decisions resulting from that perception.

To model low-pass signal filtering, we implemented a conceptually simple filter, a first-order RC circuit. This model consists of a voltage source V, resistor R, and capacitor C wired in series. In our example, these can be thought of as the input signal, signal transduction through the pathway, and the ability for the cell to hold that signal (e.g., the total abundance of a protein that can be activated), respectively. We were interested in measuring the voltage (signal) across the capacitor ($V_c = q/C$) as a function of dynamic inputs. Kirchoff's voltage law states that the sum of the voltages across each component in this loop equals 0

$$V - V_R - V_C = 0$$

$$V(t) - \frac{dq(t)}{dt}R - \frac{q(t)}{C} = 0$$

where $C = 1$, R was variably defined, and V was the dynamic model input. Dynamic inputs were achieved by changing the value of V between 0 (OFF) and 1 (ON) at defined intervals.

The time constants $\tau = RC$ define the signal kinetics in this model, which describe the speed of exponential rise and decay of the signal. To model signal processing changes in cancer, we changed the value of R, thus changing the value of τ. We started the simulation from a state of rest, where $q(t = 0) = 0$. We implemented the model in the R programming language (https://cran.r-project.org) using the deSolve package.

To model cellular perception of a filtered signal, we defined a signal intensity at which downstream transcriptional circuits are turned on. In fig. S10, B and C, we defined this threshold to be 30% based on estimates from previous studies (*2*, *47*), but fig. S10D shows that a broad range of threshold values gives qualitatively similar results.

Finally, we assumed that a cell's decision to proliferate was directly correlated to the cumulative signal the cell perceived by the end of the simulation. Many possible relationships exist between perceived signal and proliferation, but we chose a direct correlation due to its simplicity. We expect that a different relationship would also show that altered signal perception could lead to differential cell fate choices, though the set of inputs subject to misperception may change.

Curve fitting

Single exponential decay was fitted to normalized kinetics data using the *glm* function in R assuming a Gaussian error distribution and a "log" linkage. The curves describing proliferation as a function of duty cycle in Fig. S12F were fit to a Hill function.

REFERENCES AND NOTES

1. Y. D. Shaul, R. Seger, The MEK/ERK cascade: From signaling specificity to diverse functions. *Biochim. Biophys. Acta* **1773**, 1213–1226 (2007).
2. J. G. Albeck, G. B. Mills, J. S. Brugge, Frequency-modulated pulses of ERK activity transmit quantitative proliferation signals. *Mol. Cell* **49**, 249–261 (2013). doi: 10.1016/j.molcel.2012.11.002; pmid: 23219535
3. K. Aoki *et al*., Stochastic ERK activation induced by noise and cell-to-cell propagation regulates cell density-dependent proliferation. *Mol. Cell* **52**, 529–540 (2013). doi: 10.1016/j.molcel.2013.09.015; pmid: 24140422
4. T. Hiratsuka *et al*., Intercellular propagation of extracellular signal-regulated kinase activation revealed by in vivo imaging of mouse skin. *eLife* **4**, e05178 (2015). doi: 10.7554/eLife.05178; pmid: 25668746
5. C. J. Marshall, Specificity of receptor tyrosine kinase signaling: Transient versus sustained extracellular signal-regulated kinase activation. *Cell* **80**, 179–185 (1995). doi: 10.1016/0092-8674(95)90401-8; pmid: 7834738
6. N. Rauch, O. S. Rukhlenko, W. Kolch, B. N. Kholodenko, MAPK kinase signalling dynamics regulate cell fate decisions and drug resistance. *Curr. Opin. Struct. Biol.* **41**, 151–158 (2016). doi: 10.1016/j.sbi.2016.07.019; pmid: 27521656
7. M. Behar, H. G. Dohlman, T. C. Elston, Kinetic insulation as an effective mechanism for achieving pathway specificity in intracellular signaling networks. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 16146–16151 (2007). doi: 10.1073/pnas.0703894104; pmid: 17913886
8. C. Cohen-Saidon, A. A. Cohen, A. Sigal, Y. Liron, U. Alon, Dynamics and variability of ERK2 response to EGF in individual living cells. *Mol. Cell* **36**, 885–893 (2009). doi: 10.1016/j.molcel.2009.11.025; pmid: 20005850
9. R. A. Kellogg, S. Tay, Noise facilitates transcriptional control under dynamic inputs. *Cell* **160**, 381–392 (2015). doi: 10.1016/j.cell.2015.01.013; pmid: 25635454
10. S. D. M. Santos, P. J. Verveer, P. I. H. Bastiaens, Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate. *Nat. Cell Biol.* **9**, 324–330 (2007). doi: 10.1038/ncb1543; pmid: 17310240
11. N. Hao, E. K. O'Shea, Signal-dependent dynamics of transcription factor translocation controls gene expression. *Nat. Struct. Mol. Biol.* **19**, 31–39 (2011). doi: 10.1038/nsmb.2192; pmid: 22179789
12. J. E. Purvis *et al*., p53 dynamics control cell fate. *Science* **336**, 1440–1444 (2012). doi: 10.1126/science.1218351; pmid: 22700930
13. J. E. Purvis, G. Lahav, Encoding and decoding cellular information through signaling dynamics. *Cell* **152**, 945–956 (2013). doi: 10.1016/j.cell.2013.02.005; pmid: 23452846
14. J. Selimkhanov *et al*., Accurate information transmission through dynamic biochemical signaling networks. *Science* **346**, 1370–1373 (2014). doi: 10.1126/science.1254933; pmid: 25504722
15. L. Ashall *et al*., Pulsatile stimulation determines timing and specificity of NF-κB-dependent transcription. *Science* **324**, 242–246 (2009). doi: 10.1126/science.1164860; pmid: 19359585
16. J. E. Toettcher, O. D. Weiner, W. A. Lim, Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell* **155**, 1422–1434 (2013). doi: 10.1016/j.cell.2013.11.004; pmid: 24315106
17. L. O. Murphy, J. P. MacKeigan, J. Blenis, A network of immediate early gene products propagates subtle differences in mitogen-activated protein kinase signal amplitude and duration. *Mol. Cell. Biol.* **24**, 144–153 (2004). doi: 10.1128/MCB.24.1.144-153.2004; pmid: 14673150
18. L. O. Murphy, S. Smith, R.-H. Chen, D. C. Fingar, J. Blenis, Molecular interpretation of ERK signal duration by immediate early gene products. *Nat. Cell Biol.* **4**, 556–564 (2002). doi: 10.1038/ncb822; pmid: 12134156
19. F. Uhlitz *et al*., An immediate-late gene expression module decodes ERK signal duration. *Mol. Syst. Biol.* **13**, 928 (2017). doi: 10.15252/msb.20177554; pmid: 28468958
20. P. J. Roberts, C. J. Der, Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. *Oncogene* **26**, 3291–3310 (2007). doi: 10.1038/sj.onc.1210422; pmid: 17496923
21. W. R. Burack, A. S. Shaw, Live cell imaging of ERK and MEK: Simple binding equilibrium explains the regulated nucleocytoplasmic distribution of ERK. *J. Biol. Chem.* **280**, 3832–3837 (2005). doi: 10.1074/jbc.M410031200; pmid: 15546878

22. C. Zhang *et al*., RAF inhibitors that evade paradoxical MAPK pathway activation. *Nature* **526**, 583–586 (2015). doi: 10.1038/nature14982; pmid: 26466569
23. K. Zhang *et al*., Light-mediated kinetic control reveals the temporal effect of the Raf/MEK/ERK pathway in PC12 cell neurite outgrowth. *PLOS ONE* **9**, e92917 (2014). doi: 10.1371/journal.pone.0092917; pmid: 24667437
24. T. Brummer, H. Naegele, M. Reth, Y. Misawa, Identification of novel ERK-mediated feedback phosphorylation sites at the C-terminus of B-Raf. *Oncogene* **22**, 8823–8834 (2003). doi: 10.1038/sj.onc.1207185; pmid: 14654779
25. D. A. Ritt, D. M. Monson, S. I. Specht, D. K. Morrison, Impact of feedback phosphorylation and Raf heterodimerization on normal and mutant B-Raf signaling. *Mol. Cell. Biol.* **30**, 806–819 (2010). doi: 10.1128/MCB.00569-09; pmid: 19933846
26. P. T. Wan *et al*., Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF. *Cell* **116**, 855–867 (2004). doi: 10.1016/S0092-8674(04)00215-6; pmid: 15035987
27. M. J. Garnett, S. Rana, H. Paterson, D. Barford, R. Marais, Wild-type and mutant B-RAF activate C-RAF through distinct mechanisms involving heterodimerization. *Mol. Cell* **20**, 963–969 (2005). doi: 10.1016/j.molcel.2005.10.022; pmid: 16364920
28. Z. Yao *et al*., BRAF mutants evade ERK-dependent feedback by different mechanisms that determine their sensitivity to pharmacologic inhibition. *Cancer Cell* **28**, 370–383 (2015). doi: 10.1016/j.ccell.2015.08.001; pmid: 26343582
29. P. I. Poulikakos, C. Zhang, G. Bollag, K. M. Shokat, N. Rosen, RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF. *Nature* **464**, 427–430 (2010). doi: 10.1038/nature08902; pmid: 20179705
30. T. Rajakulendran, M. Sahmi, M. Lefrançois, F. Sicheri, M. Therrien, A dimerization-dependent mechanism drives RAF catalytic activation. *Nature* **461**, 542–545 (2009). doi: 10.1038/nature08314; pmid: 19727074
31. G. Hatzivassiliou *et al*., RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth. *Nature* **464**, 431–435 (2010). doi: 10.1038/nature08833; pmid: 20130576
32. S. J. Heidorn *et al*., Kinase-dead BRAF and oncogenic RAS cooperate to drive tumor progression through CRAF. *Cell* **140**, 209–221 (2010). doi: 10.1016/j.cell.2009.12.040; pmid: 20141835
33. P. B. Chapman *et al*., Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N. Engl. J. Med.* **364**, 2507–2516 (2011). doi: 10.1056/NEJMoa1103782; pmid: 21639808
34. K. T. Flaherty *et al*., Inhibition of mutated, activated BRAF in metastatic melanoma. *N. Engl. J. Med.* **363**, 809–819 (2010). doi: 10.1056/NEJMoa1002011; pmid: 20818844
35. J. Larkin *et al*., Vemurafenib in patients with BRAF(V600) mutated metastatic melanoma: An open-label, multicentre, safety study. *Lancet Oncol.* **15**, 436–444 (2014). doi: 10.1016/S1470-2045(14)70051-8; pmid: 24582505
36. F. Su *et al*., RAS mutations in cutaneous squamous-cell carcinomas in patients treated with BRAF inhibitors. *N. Engl. J. Med.* **366**, 207–215 (2012). doi: 10.1056/NEJMoa1105358; pmid: 22256804
37. T. Tian *et al*., Plasma membrane nanoswitches generate high-fidelity Ras signal transduction. *Nat. Cell Biol.* **9**, 905–914 (2007). doi: 10.1038/ncb1615; pmid: 17618274
38. K. J. Cho *et al*., Raf inhibitors target ras spatiotemporal dynamics. *Curr. Biol.* **22**, 945–955 (2012). doi: 10.1016/j.cub.2012.03.067; pmid: 22560614
39. J. Kumbrink, M. Gerlinger, J. P. Johnson, Egr-1 induces the expression of its corepressor nab2 by activation of the nab2 promoter thereby establishing a negative feedback loop. *J. Biol. Chem.* **280**, 42785–42793 (2005). doi: 10.1074/jbc.M511079200; pmid: 16260776
40. B. Sparta *et al*., Receptor level mechanisms are required for epidermal growth factor (EGF)-stimulated extracellular signal-regulated kinase (ERK) activity pulses. *J. Biol. Chem.* **290**, 24784–24792 (2015). doi: 10.1074/jbc.M115.662247; pmid: 26304118
41. L. J. Bugaj, G. P. O'Donoghue, W. A. Lim, Interrogating cellular perception and decision making with optogenetic tools. *J. Cell Biol.* **216**, 25–28 (2017). doi: 10.1083/jcb.201612094; pmid: 28003330
42. M. J. Lee *et al*., Sequential application of anticancer drugs enhances cell death by rewiring apoptotic signaling networks. *Cell* **149**, 780–794 (2012). doi: 10.1016/j.cell.2012.03.031; pmid: 22579283
43. S. H. Chen, W. Forrester, G. Lahav, Schedule-dependent interaction between anticancer treatments. *Science* **351**, 1204–1208 (2016). doi: 10.1126/science.aac5610; pmid: 26965628
44. M. Behar, D. Barken, S. L. Werner, A. Hoffmann, The dynamics of signaling as a pharmacological target. *Cell* **155**, 448–461 (2013). doi: 10.1016/j.cell.2013.09.018; pmid: 24120141
45. J. E. Toettcher, D. Gong, W. A. Lim, O. D. Weiner, Light-based feedback for controlling intracellular signaling dynamics. *Nat. Methods* **8**, 837–839 (2011). doi: 10.1038/nmeth.1700; pmid: 21909100
46. T. G. Bivona *et al*., PKC regulates a farnesyl-electrostatic switch on K-Ras that promotes its association with Bcl-XL on mitochondria and induces apoptosis. *Mol. Cell* **21**, 481–493 (2006). doi: 10.1016/j.molcel.2006.01.012; pmid: 16483930
47. G. Bollag *et al*., Clinical efficacy of a RAF inhibitor needs broad target blockade in BRAF-mutant melanoma. *Nature* **467**, 596–599 (2010). doi: 10.1038/nature09454; pmid: 20823850

ACKNOWLEDGMENTS

We thank G. Bollag and Plexxikon for providing PLX8394 and N. Frankel and G. Bollag for critical reading and comments on the manuscript. We also thank N. Repina (University of California, Berkeley) for helpful discussions on optogenetic device design, E. Chow for discussions on device construction, A. Meyer-Franke for technical help with high-content microscopy, and A. Basu (Wayne State University) for providing an alternative device for multiwell optogenetic experiments. **Funding:** This work was funded by the Arnold O. Beckman Postdoctoral Fellowship (L.J.B.); the European Molecular Biology Organization ALTF 419-2010 (A.M.); the UCSF Program for Breakthrough Biomedical Research Postdoctoral Research Award (A.M.); NIH DP2EB024247 (J.E.T.); St. Baldrick's Foundation and a Damon Runyon-Sohn Foundation fellowship 6P-13 (A.S.); NIH DP2 CA174497, R01CA169338, R01CA204302, and R01CA211052 (T.G.B); the Pew and Stewart Foundations (T.G.B.); the Howard Hughes Medical Institute (W.A.L); and NIH P50GM081879 and R01GM55040 (W.A.L.). **Author contributions:** L.J.B., J.E.T., T.G.B., and W.A.L. designed the study and experiments. L.J.B and J.E.T. generated cell lines. L.J.B, A.J.S., J.E.G., and J.E.T. performed Western blots. L.J.B. and J.E.T. performed and analyzed live-cell microscopy. L.J.B. designed and built the optoPlate, developed optoPlate assays, and performed and analyzed all optoPlate-based experiments. L.J.B. and A.M. developed and analyzed protocols for proliferation experiments. W.A.L. and T.G.B. oversaw and guided the project. L.J.B. and W.A.L. wrote the manuscript with an early draft by J.E.T. Manuscript review and editing was provided by all authors. **Competing interests:** A patent application (U.S. provisional patent application no. 62/362,768) has been filed on the design of the optoPlate device. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/eaao3048/suppl/DC1
Figs. S1 to S13
Table S1

7 July 2017; accepted 16 July 2018
10.1126/science.aao3048

RESEARCH ARTICLE SUMMARY

BIOTECHNOLOGY

Developmental barcoding of whole mouse via homing CRISPR

Reza Kalhor*, Kian Kalhor, Leo Mejia, Kathleen Leeper, Amanda Graveline, Prashant Mali, George M. Church*

INTRODUCTION: The remarkable development of a single cell, the zygote, into the full organism occurs through a complex series of division and differentiation events that resemble a tree, with the zygote at the base branching through lineages that end in the terminal cell types at the top. Characterizing this tree of development has long been a subject of interest, and the combination of modern genome engineering and sequencing technologies promises a powerful strategy in its service: in vivo barcoding. For in vivo barcoding, heritable random mutations are induced to accumulate during development and sequenced post hoc to reconstruct the lineage tree. Demonstrations thus far have largely focused on lower vertebrates and have used a barcoding element with a constrained window of activity for clonal tracing of individual cells or cell types. Implementation in mammalian model systems, such as the mouse, incurs unique challenges that require major enhancements.

RATIONALE: To address the complexity of mammalian development, we reasoned that multiple independent in vivo barcoding elements could be deployed in parallel to exponentially expand their recording power. Independence requires both an absence of cross-talk between the elements and an absence of interference between their mutation outcomes. A system with the potential to deliver on these requirements is homing CRISPR, a modified version of canonical CRISPR wherein the homing guide RNA (hgRNA) combines with CRISPR-Cas9 nuclease for repeated targeting of its own locus, leading to diverse mutational outcomes. Therefore, in mouse embryonic stem cells, we scattered multiple hgRNA loci with distinct spacers in the genome to serve as barcoding elements. With this arrangement, each hgRNA acts independently as a result of its unique spacer sequence, and undesirable deletion events between multiple adjacent cut sites are less likely. Using these cells, we generated a chimeric mouse with 60 hgRNAs as the founder of the MARC1 (Mouse for Actively Recording Cells 1) line that enables barcoding and recording of cell lineages.

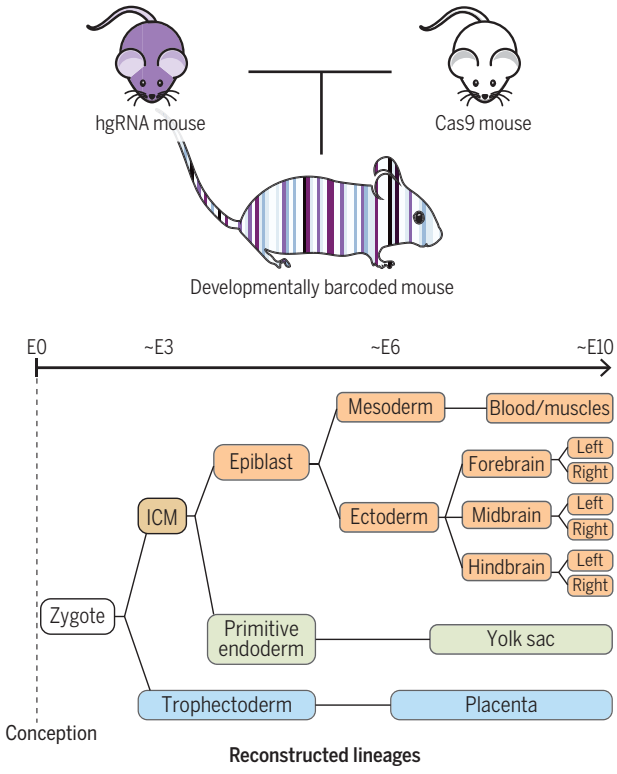
RESULTS: In the absence of Cas9, hgRNAs are stable and dormant; to initiate barcoding, we crossed MARC1 mice with Cas9 knock-in mice. In the resulting offspring, hgRNAs were activated, creating diverse mutations such that an estimated 10^{23} distinct barcode combinations can be generated with only 10 hgRNAs. Furthermore, hgRNAs showed a range of activity profiles, with some mutating soon after conception while others exhibited lower activity through most of the gestation period. This range resulted in sustained barcoding throughout gestation and recording of developmental lineages: Each cell inherits a set of unique mutations that are passed on to its daughter cells, where further unique mutations can be added. Consequently, at any stage in such developmentally barcoded mice, closely related cells have a more similar mutation profile, or barcode, than the more distant ones. These recordings remain embedded in the genomes of the cells and can be extracted by sequencing.

ON OUR WEBSITE

Read the full article at http://dx.doi.org/10.1126/science.aat9804

We used these recordings to carry out bottom-up reconstruction of the mouse lineage tree, starting with the first branches that emerged after the zygote, and observed robust reconstruction of the correct tree. We also investigated axis development in the brain by sequencing barcodes from the left and right side of the forebrain, midbrain, and hindbrain regions. We found that barcodes from the left and right sides of the same region were more closely related than those from different regions; this result suggests that in the precursor of the brain, commitment to the anterior-posterior axis is established prior to the lateral axis.

CONCLUSION: This system provides an enabling and versatile platform for in vivo barcoding and lineage tracing in a mammalian model system. It can straightforwardly create developmentally barcoded mice in which lineage information is prerecorded in cell genomes. Combining multiple independently acting molecular recording devices greatly enhances their capacity and allows for reliable information recovery and reconstruction of deep lineage trees. ■



Developmental barcoding and lineage reconstruction in mice. Crossing the MARC1 mouse line, which carries multiple hgRNAs, with a CRISPR-Cas9 mouse line results in developmentally barcoded offspring that record lineages in their cells. These recordings were extracted and used to reconstruct lineage trees. A combination of the trees extracted from different developmentally barcoded mice is shown. ICM, inner cell mass; E0, embryonic day 0.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: gchurch@genetics.med.harvard.edu (G.M.C.); kalhor@genetics.med.harvard.edu (R.K.)
Cite this article as R. Kalhor *et al.*, *Science* **361**, eaat9804 (2018). DOI: 10.1126/science.aat9804

RESEARCH ARTICLE

BIOTECHNOLOGY

Developmental barcoding of whole mouse via homing CRISPR

Reza Kalhor[1,2]*, Kian Kalhor[3], Leo Mejia[1], Kathleen Leeper[2], Amanda Graveline[2], Prashant Mali[4], George M. Church[1,2]*

In vivo barcoding using nuclease-induced mutations is a powerful approach for recording biological information, including developmental lineages; however, its application in mammalian systems has been limited. We present in vivo barcoding in the mouse with multiple homing guide RNAs that each generate hundreds of mutant alleles and combine to produce an exponential diversity of barcodes. Activation upon conception and continued mutagenesis through gestation resulted in developmentally barcoded mice wherein information is recorded in lineage-specific mutations. We used these recordings for reliable post hoc reconstruction of the earliest lineages and investigation of axis development in the brain. Our results provide an enabling and versatile platform for in vivo barcoding and lineage tracing in a mammalian model system.

In sexually reproducing multicellular eukaryotes, a single totipotent zygote remarkably develops into all cells of the full organism. This development occurs through a highly orchestrated series of differentiation events that take the zygote through many lineages as it divides to create all the different cell types (*1*). This path resembles a tree, with the zygote at the base of the trunk branching into stems of cell lineages that eventually end in the terminal cell types at the top of the tree (*2*, *3*). The ability to map this tree of development will have a far-ranging impact on our understanding of disease-causing developmental aberrations, our capacity to restore normal function in damaged or diseased tissues, and our capability to generate substitute tissues and organs from stem cells.

Tracing the lineage tree in non-eutelic higher eukaryotes with complex developmental pathways remains challenging. Clonal analysis, which entails cellular labeling and tracking with a distinguishable heritable marker, has been effective when evaluating a limited number of cells or lineages (*4*–*7*). Using more diverse presynthesized DNA sequences as markers, known as cellular barcoding, has allowed for analysis of larger cell numbers (*7*–*9*). What limits these approaches is the static nature of labeling that only allows analysis of a snapshot in time. Recent advances in genome engineering technologies, however, have enabled in vivo barcode generation (*10*, *11*). In this approach, a locus is targeted for rearrangement or mutagenesis such that a diverse set of outcomes is generated in different cells (*12*). As these barcodes can be generated over a sustained period of time, they drastically expand the scope of cellular barcoding strategies, promising deep and precise lineage tracing, from the single-cell to the whole-organism level (Fig. 1A) (*13*–*15*), and recording of cellular signals over time (*16*, *17*). Multiple studies establish proof of this principle in recording and lineage tracing, with demonstrations in cultured cells (*14*–*16*, *18*) and in lower vertebrates (*13*, *19*–*22*). However, no demonstrations have yet been carried out in mice, a model organism more relevant to human health in many aspects such as development. The challenges associated with work in mice can account for this discrepancy. Gestation in mice takes place inside the mother's womb, rendering genetic manipulation of individual zygotes or conceptuses difficult. Additionally, the longer gestation time in mice, together with the multitude of lineages that segregate throughout its development, demands sustained generation of highly diverse barcodes with minimal unwanted overwriting events to maximize the chance for successfully recording the events of interest.

Here, we deployed multiple independent barcoding loci in parallel for robust in vivo barcoding and lineage recording in mice. We created a mouse line that carries a scattered array of 60 genomically integrated homing CRISPR guide RNA (hgRNA) loci. hgRNAs are modified versions of canonical single guide RNAs (sgRNAs) (*23*) that target their own loci (Fig. 1B) to create a substantially larger diversity of mutants than canonical sgRNAs (Fig. 1C) and thus act as expressed genetic barcodes (*14*). Crossing this hgRNA line with a Cas9 line resulted in developmentally barcoded offspring because hgRNAs stochastically accumulate mutations throughout gestation, generating unique mutations in each lineage without deleting earlier mutations, in such a way that closely related cells have a more similar mutation profile, or barcode, than more distant ones. In developmentally barcoded mice, we extensively characterized the activity profile and mutant alleles of each hgRNA and carried out post hoc bottom-up reconstruction of the lineage tree in the early stage of development, starting with the first branches at its root and continuing through some of the germ layers. We also investigated lineage commitment with respect to the anterior-posterior and lateral axes in the brain.

Founder mouse with multiple hgRNA loci

We created a library of hgRNAs with four different transcript lengths, variable spacer sequences, and 10-base identifiers downstream of the hgRNA scaffold in a transposon backbone (Fig. 1D) (*24*). This library was transposed into mouse embryonic stem (mES) cells under conditions that would result in a high number of integrations per cell (Fig. 1E) (*24*). Transfected mES cells were injected into blastocysts, which were then implanted in surrogate females to generate chimeric mice. Of the 23 chimeric mice that resulted, eight males were more than 60% transgenic as assessed by their coat color (Fig. 1E). Five of the eight showed more than 20 total hgRNA integrations in their somatic genomes and were crossed with wild-type mice to determine the number of hgRNAs in their germlines. The chimera with the highest average number of germline hgRNAs, which were transmitted to its progeny, was selected for further studies and starting a line. We refer to this mouse as the MARC1 (Mouse for Actively Recording Cells 1) founder and its progeny as the MARC1 line. All results described below focus on the MARC1 founder and its progeny.

Sequence, genomic position, and inheritance of hgRNA loci

By sequencing the hgRNA loci in the MARC1 founder, we identified 60 different hgRNAs (Table 1 and table S1). Each hgRNA has a unique 10-base identifier and a different spacer sequence (table S1). We also sequenced the regions immediately flanking the transposed hgRNA elements (*24*), which allowed us to determine the genomic positions of 54 of the 60 hgRNAs (Fig. 1F, Table 1, and table S1), of which 26 are intergenic and 28 are located in an intron of a known gene (table S3) (*24*); none are located in an exon or are expected to disrupt the gene. We then crossed the MARC1 founder with multiple females and analyzed germline transmission and the inheritance pattern of these hgRNAs in the more than 100 resulting offspring. All 60 hgRNAs were transmitted through the germline, and the offspring carrying them were fertile, had normal litter sizes, and presented no morphological abnormalities. Of these 60 hgRNAs, 55 showed a Mendelian inheritance pattern, appearing in about 50% of the offspring (Table 1 and table S1). An additional 3 of the 60, all L30 hgRNAs, were detected in fewer than 20% of the offspring,

[1]Department of Genetics, Harvard Medical School, Boston, MA, USA. [2]Wyss Institute for Biologically Inspired Engineering at Harvard University, Boston, MA, USA. [3]Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran. [4]Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA.
*Corresponding author. Email: gchurch@genetics.med.harvard.edu (G.M.C.); kalhor@genetics.med.harvard.edu (R.K.)

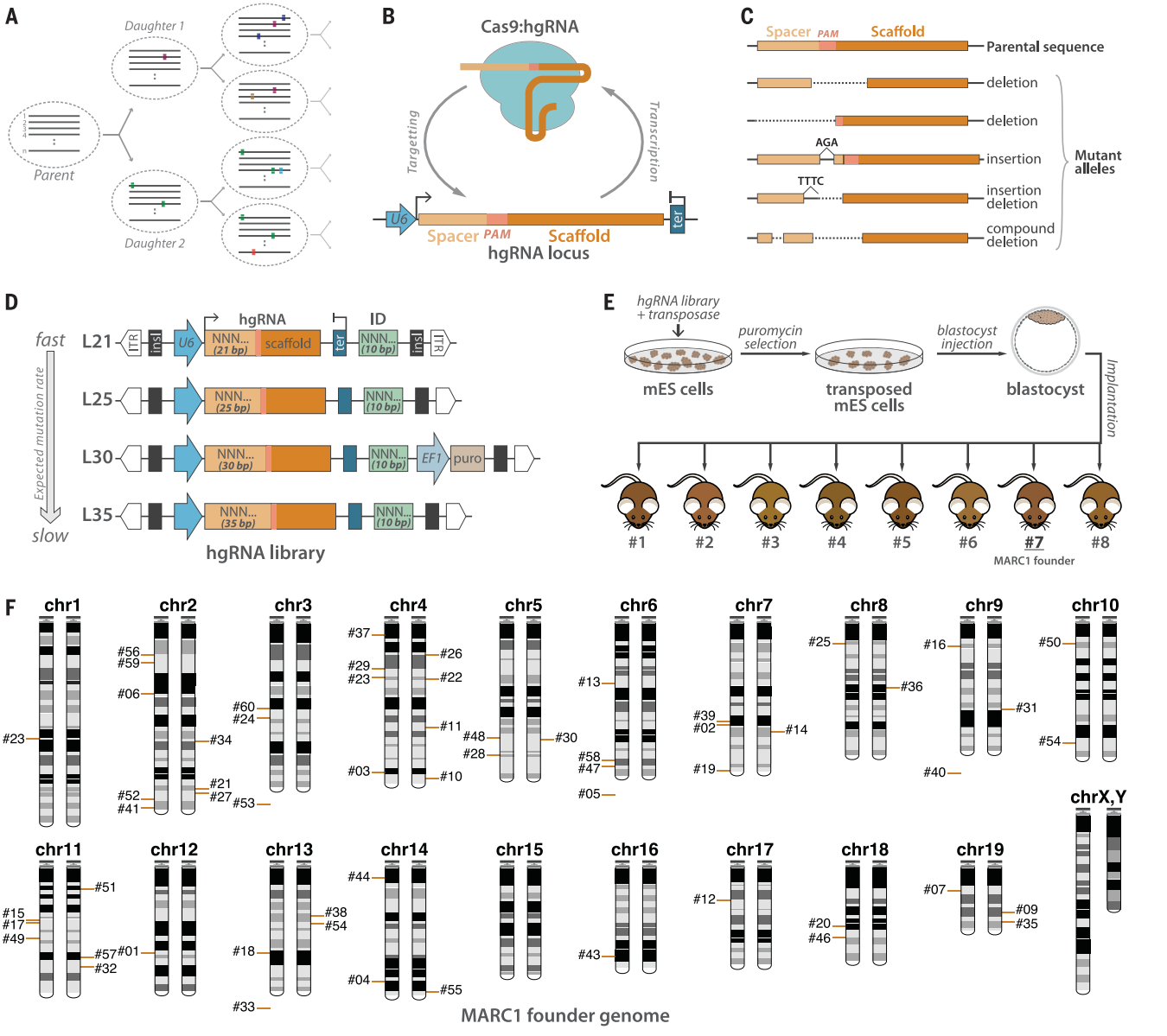


Fig. 1. In vivo barcoding with hgRNAs and strategy to generate a mouse with multiple hgRNA integrations. (**A**) Recording lineages using synthetically induced mutations in the genome. A number of loci (*n*) gradually accumulate heritable mutations as cells divide, thereby recording the lineage relationship of the cells in an array of mutational barcodes. Dashed ovals, cells; gray lines, an array of *n* mutating loci; colored rectangles, mutations. (**B**) Homing CRISPR system in which the Cas9-hgRNA complex cuts the locus encoding the hgRNA itself. As the NHEJ repair system repairs the cut (*63*), it introduces mutations in the hgRNA locus. (**C**) Example of mutations that are created in the hgRNA locus that can effectively act as barcodes. (**D**) Design of PiggyBac hgRNA library for creating a transgenic mouse. Four hgRNA sublibraries with 21, 25, 30, and 35 bases of distance between transcription start site (TSS) and scaffold PAM were constructed and combined. The spacer sequence (light orange box) and the identifier sequence (green box) were composed of degenerate bases. (**E**) Blastocyst injection strategy for producing hgRNA mice. The hgRNA library was transposed into mES cells. Cells with a high number of transpositions were enriched using puromycin selection and injected into E3.5 mouse blastocysts to obtain chimeras. Chimera 7 was chosen as the MARC1 founder. (**F**) Chromosomal position of all 54 hgRNAs whose genomic position was deciphered in the MARC1 founder (red bars). Bars on the left or right copy of the chromosome indicate the hgRNAs that are linked on the same homologous copy. hgRNAs whose exact genomic position is not known but whose chromosome can be determined on the basis of linkage are shown below the chromosome. ITR, PiggyBac inverted terminal repeats; insl, insulator; U6, U6 promoter; ter, U6 terminator; ID, identifier sequence; EF1, human elongation factor–1 promoter; puro, puromycin resistance.

which we attribute to the low detection rate of L30 hgRNAs due to the performance of the polymerase chain reaction primer used for these and only these three hgRNAs (*24*). The remaining two hgRNAs were transmitted to almost 75% of the offspring—a result best explained by the duplication of these hgRNAs to loci more than 50 cM away on the same chromosome or to loci on different chromosomes and confirmed by the genomic location data (table S1) (*24*).

We also compared the co-inheritance frequencies of MARC1 hgRNAs to those expected from Mendelian inheritance of independently segregating loci (fig. S1A). We found no mutually exclusive cosegregating groups of hgRNAs (fig. S1A), indicating that the entire germline in the MARC1 founder was derived from only one of the injected stem cells and is thus genetically homogeneous.

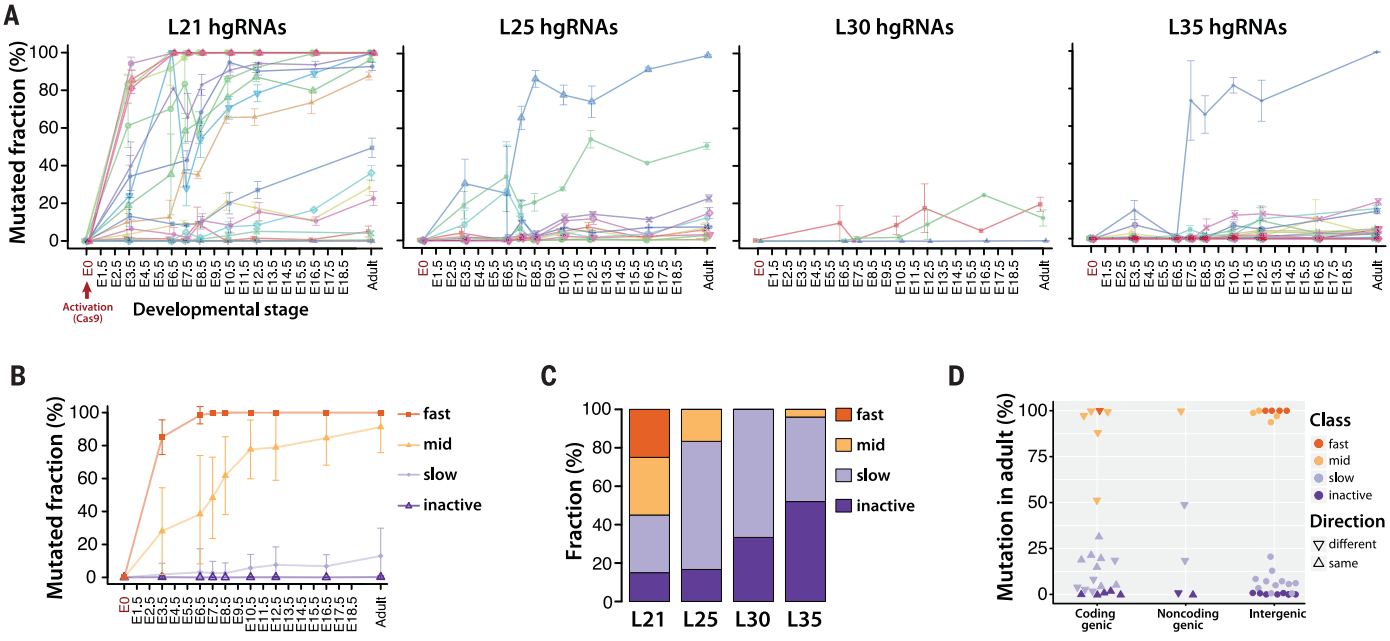


Fig. 2. Activity of MARC1 hgRNAs. (**A**) Activity profiles of all 60 hgRNAs in embryonic and adult progenies of the MARC1 founder crossed with Cas9 knock-in females, broken down by hgRNA length. The fraction of mutant (nonparental) spacer sequences in each hgRNA is measured. Lines connect the observed average mutation rates of one hgRNA. Means ± SEM are shown (*N* is different for each value; see Table 2). See table S2 for numerical values of the plot. (**B**) Average activity profiles of each hgRNA class in embryonic and adult progenies of the MARC1 founder crossed with Cas9 knock-in females. Means ± SEM are shown as representations of range of activity (*N* is different for each value; see Table 2). (**C**) Functional categorization of hgRNAs based on their activity profile in (A), broken down by length. (**D**) Position and transcription direction of hgRNAs with respect to all known coding and noncoding genes, annotated for their functional category. See table S3 for the genes in which hgRNAs are located; see fig. S3 for breakdown of this plot by hgRNA length.

Considering that every hgRNA detected in the somatic tissue of the MARC1 founder was also transmitted to its offspring, these results further suggest that almost all transgenic cells within this chimera were derived from one of the stem cells that were injected into its blastocyst, an observation consistent with previous studies (*25*, *26*). The co-inheritance analysis also revealed the groups of hgRNAs that deviate from an independent segregation pattern, suggesting that they are linked on a chromosome (fig. S1B). Close examination of this linkage disequilibrium allowed us to determine which linked hgRNAs were on different homologous copies of the same chromosome or were linked on the same copy of a chromosome (Fig. 1F and fig. S1C). Combined with the genomic location information that was obtained by sequencing, this co-inheritance analysis allowed us to decipher the cytogenetic location of most hgRNAs in the MARC1 founder with a high degree of confidence (Fig. 1F).

Activity of hgRNAs

We next studied the activity of MARC1's hgRNAs upon activation with Cas9. For that, we crossed the MARC1 founder with Rosa26-Cas9 knock-in females, which constitutively express the *Streptococcus pyogenes* Cas9 protein (*27*). Considering that major zygotic genome activation in the mouse occurs at the two-cell stage (*28*), hgRNA activation is expected soon after conception. We sampled these Cas9-activated offspring at various stages after conception to measure the fraction of mutated spacers for each hgRNA. In all, we gathered 190 samples from 102 animals in seven embryonic stages and the adult stage (Table 2). The results confirm that hgRNAs start mutating their loci soon after the introduction of Cas9 (Fig. 2A). However, the rate at which these mutations accumulate varied widely among the 60 MARC1 hgRNAs (Fig. 2A). On the basis of these activity levels, we classified hgRNAs into four categories with distinct activation profiles (Fig. 2B): (i) five "fast" hgRNAs that mutate in at least 80% of the cells in each sample by embryonic day 3.5 (E3.5) and in almost all cells by E8.5; (ii) 27 "slow" hgRNAs that mutate in only a minority of cells even in the adult stage; (iii) nine "mid" hgRNAs, intermediate between fast and slow, that accumulate mutations throughout embryonic development and are mutated in almost all cells only in later embryonic or adult stages; and (iv) 19 hgRNAs that appear to be inactive, at least with this level of Cas9 expression, mutating in fewer than 2% of sampled cells even in the adult stage (table S2). Most mutations that were detected (about 80% for fast hgRNAs) are expected to render the hgRNA nonfunctional and thus prevent further changes (fig. S2).

Transcript length clearly affects hgRNA activity: A far higher fraction of L21 hgRNAs, which have the shortest possible transcript length, were active by comparison to L25, L30, and L35 hgRNAs, which are longer by 4, 9, and 14 bases, respectively (Fig. 2, A and C). Furthermore, all fast hgRNAs were L21 hgRNAs, whereas in longer hgRNAs the inactive proportion appeared to grow with increasing length (Fig. 2C). Beyond transcript length, we found that the variation in activity among hgRNAs with an identical length (Fig. 2A) is far more than would be expected solely on the basis of differences in their spacers (*14*), which suggests that genomic location may play a substantial role. Although we detected no significant difference between the activity of hgRNAs that are in intergenic regions relative to those within known genes (Wilcoxon $P > 0.1$), among hgRNAs that have landed within known coding and noncoding genes, those that transcribe in the same direction as the gene had a lower activity than those that transcribe in the opposite direction (Wilcoxon $P < 0.05$; Fig. 2D, fig. S3, and table S3). These observations suggest that hgRNA activity is affected by both genomic location and interplay with endogenous elements.

Diversity and composition of hgRNA mutants

We next analyzed the diversity produced by MARC1 hgRNAs by considering all observed mutant spacer alleles in MARC1 × Cas9 offspring (table S4). Only a handful of mutant spacer alleles were detected for each hgRNA in each sample (Fig. 3A and fig. S4A). However, when combining mutant spacers from all offspring, on average, more than 200 distinct mutant spacers for each fast hgRNA and more than 300 for each mid hgRNA were observed (Fig. 3B and

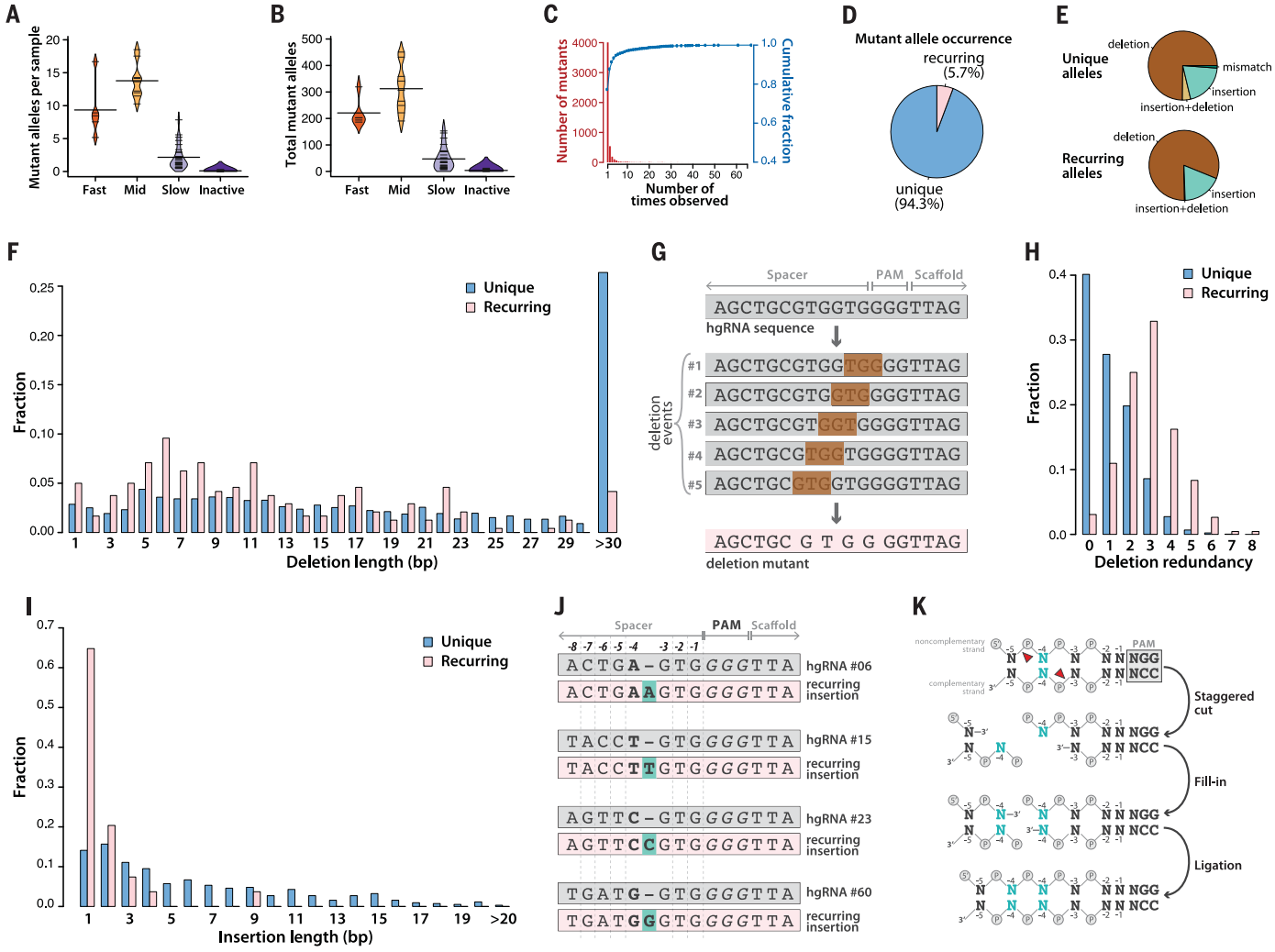


Fig. 3. Diversity of mutant hgRNA alleles in offspring of MARC1 × Cas9 cross. (**A**) Beanplots of the number of mutant spacer alleles observed in each mouse for each hgRNA category. Short horizontal lines mark the average for each hgRNA in the category; long horizontal lines mark the average of all the hgRNAs in the category. See fig. S4A for a separate plot for each hgRNA. (**B**) Beanplots of the total number of mutant spacer alleles observed for each hgRNA in all mice. See fig. S4B for a separate plot for each hgRNA. (**C**) Histogram (red bars) and cumulative fraction (blue connected dots) of the number of mice in which each mutant allele was observed, combined for all hgRNAs. See fig. S5 for a separate plot for each hgRNA. (**D**) Relative ratio of recurring mutant spacer alleles (fig. S5) (*24*) to the unique alleles. (**E**) Mutation types in unique (top) and recurring (bottom) spacer alleles. See tables S4 and S5 for the sequences and alignment of all mutants and recurring mutants, respectively, and fig. S6 for a separate plot for each hgRNA. (**F**) Distribution of deletion length for unique and recurring mutant spacer alleles. Deletions larger than 30 bp have been aggregated. (**G**) Schematic representation of how five distinct deletion events can lead to the same mutant spacer allele. (**H**) Distribution of deletion redundancy—that is, the number of independent simple deletion events in the parental spacer allele that would lead to the same observed deletion mutant—for unique and recurring spacer alleles. Simple deletion is defined as deletion of a contiguous stretch of bases without creating insertions or mismatches. Redundancy of 0 represents non–simple mutant alleles, which involve insertions, mismatches, or noncontiguous deletions. (**I**) Distribution of insertion length for unique and recurring mutant spacer alleles. Insertions of 20 bp or longer have been aggregated. (**J**) Four observed examples of recurring single-base insertions, involving duplication of the −4 position, for four different hgRNAs. (**K**) Schematic representation of how a single-base staggered overhang generated by Cas9 can lead to duplication of the −4 position.

fig. S4B). Furthermore, about 80% of all mutant spacer alleles were unique observations in a single offspring (Fig. 3C and fig. S5), which suggests that the mutant alleles observed with our sampling level constitute only a minority of all mutant spacers possible. These results indicate that each hgRNA can produce hundreds of mutant alleles.

Notably, although most mutant spacer alleles appeared in only a single sample, about 6% recurred in multiple MARC1 × Cas9 offspring (Fig. 3, C and D, and fig. S5). To understand this phenomenon, we compared the nature of unique and recurring mutant alleles (tables S4 and S5). We observed that insertions or deletions (indels) underlie the vast majority of alleles in both unique and recurring mutations (Fig. 3E and fig. S6, A and B). The exact nature of these indel mutations, however, differs. First, short deletions of 23 base pairs (bp) or fewer are enriched in the recurring alleles (Fig. 3F). Interestingly, these mutant alleles can be identical results of multiple different simple deletions in the parental spacer sequence (Fig. 3, G and H), which implies that this group of recurring mutations can result from distinct mutagenesis events that lead to the same sequence. Second, single-base insertions are drastically enriched among recurring insertion mutants (Fig. 3I). A closer examination of these single-base insertions revealed that many follow the same

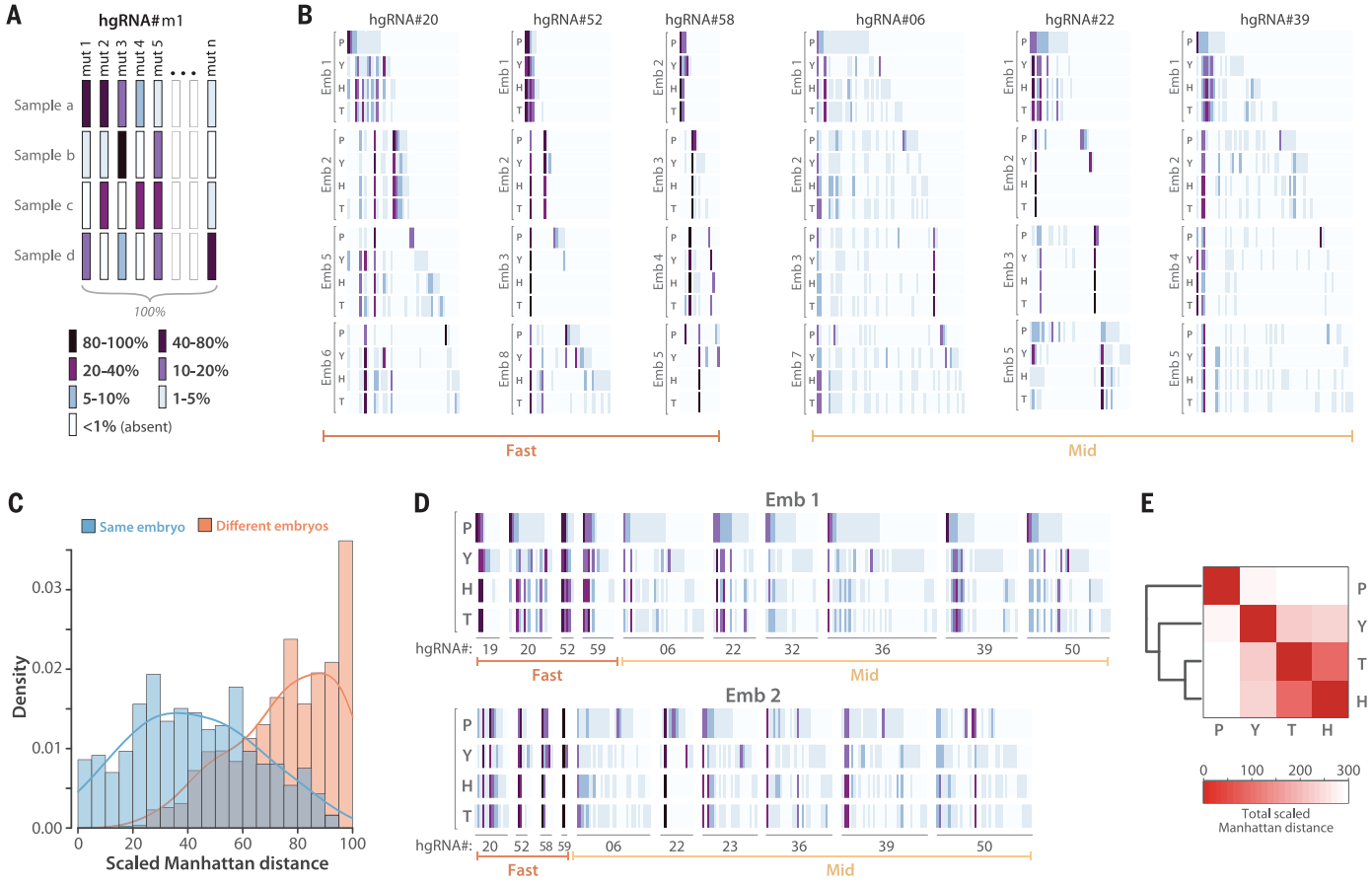


Fig. 4. In vivo barcoding in mouse embryos. (**A**) Barcode depiction for each hgRNA in each sample. Each column corresponds to an observed mutant spacer; each row corresponds to a sample. The color of each block represents the observed frequency of the corresponding mutant spacer in the corresponding sample. (**B**) In vivo–generated barcodes of three fast and three mid hgRNAs in eight embryos from a MARC1 × Cas9 cross. Four tissues were sampled from each embryo: the placenta (P), the yolk sac (Y), the head (H), and the tail (T). Embryos 1 and 2 were obtained at E16.5, whereas embryos 3 to 8 were obtained at E12.5 (Table 2). For each hgRNA, the results for a maximum of four embryos are shown. Full barcodes for all hgRNAs are in fig. S7. The color code is as shown in (A). Only mutant alleles with a maximum abundance of more than 1% are shown. (**C**) Histogram of the scaled Manhattan distances (L1) between the barcodes of all possible sample pairs for each hgRNA, broken down by sample pairs belonging to the same embryo (blue) and pairs belonging to different embryos (orange). (**D**) The complete barcode, composed of the concatenation of all hgRNA barcodes, for embryos 1 and 2. (**E**) Heat map of the average Manhattan distance between the full barcodes of placenta, yolk sac, head, and tail samples in all eight embryos. For a separate map for each embryo, see fig. S8.

pattern: duplication of the base at the −4 position relative to the protospacer adjacent motif (PAM) (Fig. 3J). In fact, this type of insertion was recurring in 34 of the 41 active hgRNAs. This observation can be best explained by Cas9 cutting at the −4 position of the noncomplementary strand and at the −3 position of the complementary strand, thus creating a staggered end with a 5′ overhang, which is then filled in on both ends and ligated (Fig. 3K). Therefore, our results suggest that Cas9 can produce staggered cuts, and that the nature of these cuts, together with the sequence of the target site, affects the eventual outcome of nonhomologous end joining (NHEJ) repair.

Developmental hgRNA barcodes

The results thus far indicate that MARC1 hgRNAs accumulate mutations upon activation with Cas9 nuclease after conception. We next queried whether these mutations indeed reflect developmental events. For simplicity, we focused on fast and mid hgRNAs in eight post-E12 MARC1 × Cas9 offspring for which four different tissues had been sampled (Table 2). The sampled tissues were the placenta, the yolk sac, the head, and the tail. The barcode was defined for each hgRNA in each sample as the frequency vector of the relative abundances of all observed mutant alleles (Fig. 4A). For the 32 samples under consideration (eight conceptuses with four samples each), these barcodes showed diverse and complex patterns, with each sample having a unique barcode but with varying degrees of similarity to other samples (Fig. 4B and fig. S7). To compare the hgRNA barcodes between samples, we used a scaled Manhattan distance (L1) of their frequency vectors, such that a distance of 100 would indicate a completely nonoverlapping set of mutant alleles and a distance of 0 would indicate a complete overlap of mutant alleles with identical relative frequencies (*24*). Pairwise comparison of all hgRNA barcodes among all samples (Fig. 4C) showed that more than 99% of barcode pairs have a scaled Manhattan distance of more than 5, indicating unique barcoding of each sample by each hgRNA. Furthermore, barcodes from different tissues of the same embryo were more similar to each other (median distance = 41) and more distinct from different embryos (median distance = 78) (Fig. 4C), which suggests that barcodes may record information about the history of samples relative to one another.

To further evaluate this recording of sample histories, we created a "full" barcode for each sample by combining the barcodes generated by each of its hgRNAs (Fig. 4D) and compared the distance between these barcodes in the four tissues obtained from each embryo (Fig. 4E and fig. S8). The results show higher similarity between the head and tail samples, which together are the most different from placenta. The samples

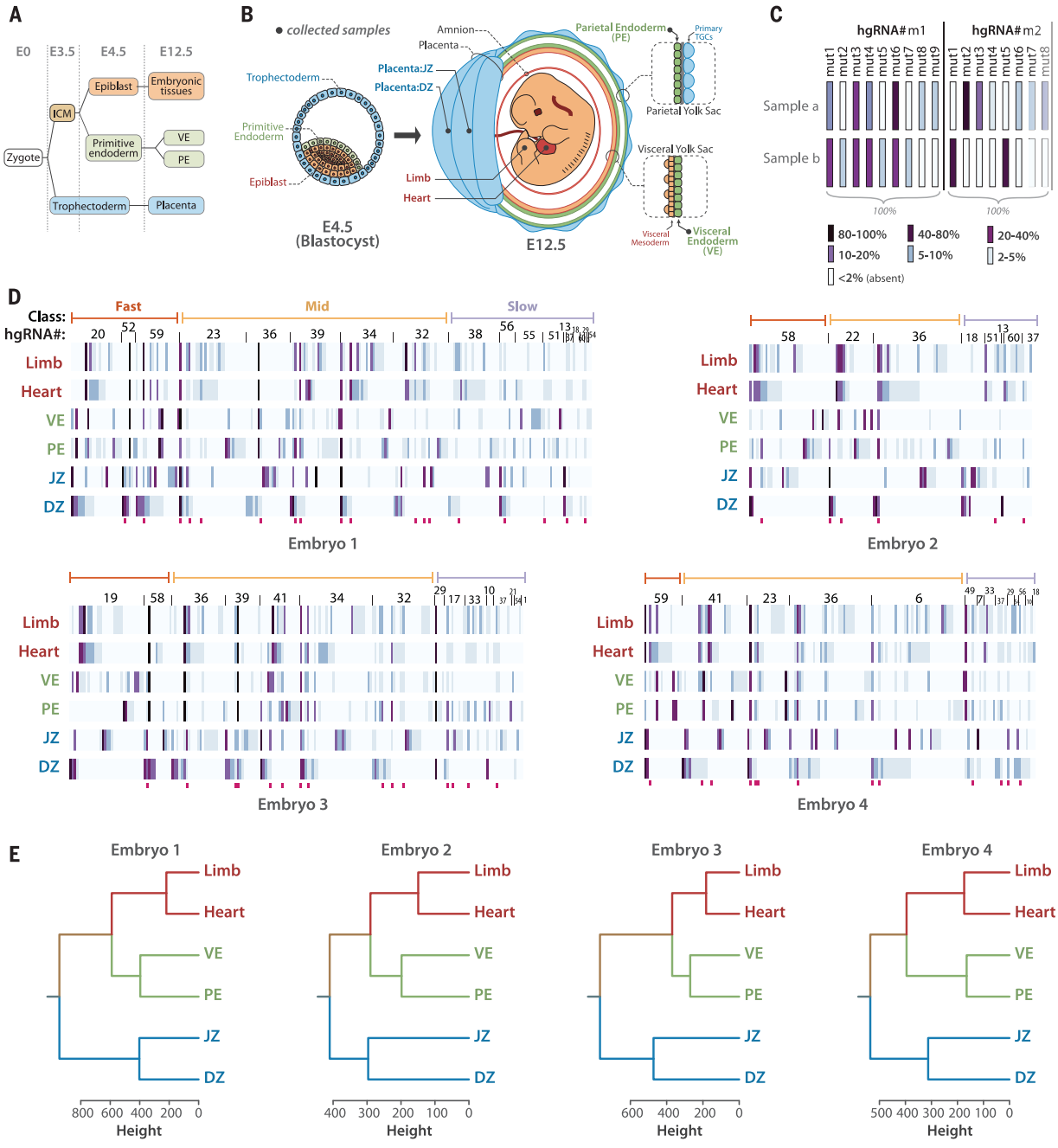


Fig. 5. Lineage derivation based on hgRNA-generated developmental barcodes. (**A**) Summary of the earliest lineages in mouse. (**B**) Schematic representation of a blastocyst and an E12.5 mouse conceptus, color-coded according to the origin of tissues in the blastocyst. Black dots show the positions and tissues of the samples obtained from E12.5 conceptuses. (**C**) Summary of how hgRNA barcodes were compiled for each sample. Each bar represents a mutant spacer of an hgRNA, and its color represents its abundance relative to other mutant spacers of the same hgRNA in the same sample. (**D**) Full hgRNA barcodes for all samples from the four mouse embryos analyzed. The barcode is annotated in (C). Only mutant alleles with a maximum abundance of more than 2% are shown. Deep pink bars below each map mark highly recurring alleles that have been observed in more than 60% of all mice analyzed in Table 2. See table S6 for a numerical version of each barcode map. (**E**) Lineage tree for each embryo calculated from the barcodes in (D).

obtained here represent mixed and overlapping lineages. However, considering that the head and tail are derived from the inner cell mass (ICM) whereas the placenta is mostly derived from the trophectoderm (*29*–*31*), these results suggest that hgRNA barcodes of different tissues embody their lineage histories.

First lineage tree from barcode recordings

We next assessed whether accurate lineage trees can be constructed de novo from developmentally barcoded mice. To assess this potential, we focused on the tree of the first lineages in development. The first lineage segregation events in mammals are the differentiation of blastomeres into the trophectoderm and ICM before E3.5, followed by differentiation of the ICM into the primitive endoderm and epiblast by E4.5 (Fig. 5A) (*29*). To reconstruct this lineage tree, we used developmentally barcoded E12.5 conceptuses and sampled two distinct tissues from each of

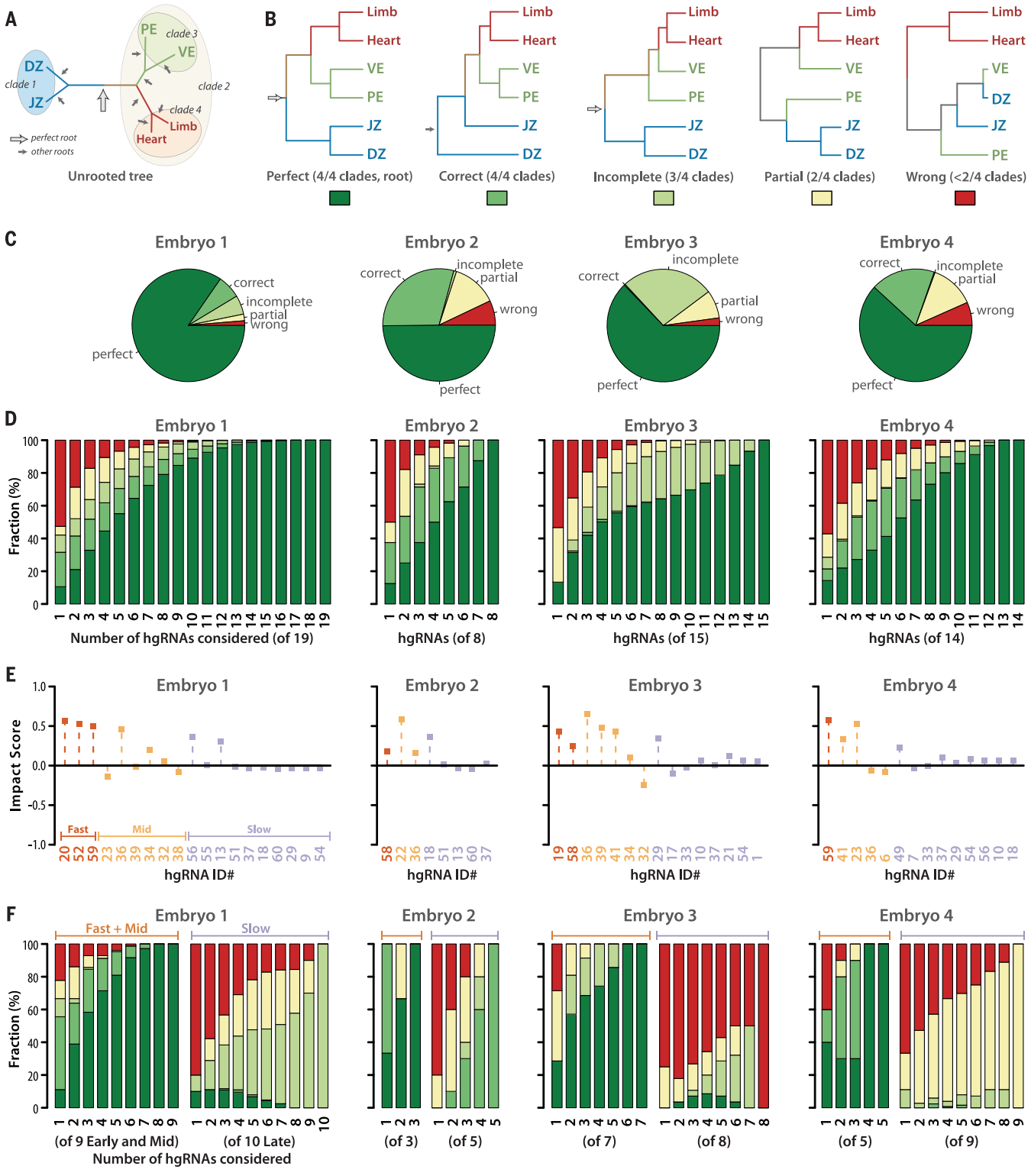


Fig. 6. Lineage tree derivation robustness and contribution of each hgRNA. (**A**) The correct unrooted tree topology for the earliest lineages in mouse. Arrows indicate all possible roots. The empty arrow indicates the perfect root. (**B**) The perfect rooted topology and an example from each of the other topology classifications. The colored boxes below each topology constitute the color key for the remaining panels of the figure. (**C**) For each of the four embryos analyzed, distribution of tree calculation outcomes from all possible subsets of hgRNAs ($2^n - 1$ non-null subsets for an embryo with n hgRNAs). (**D**) Distribution of tree calculation outcomes when including only m of the n hgRNAs in each embryo (${}^nC_m$ combinations, $1 \le m \le n$). Color code is as described in (B). See also figs. S9 and S10 for all combinations included and excluding each hgRNA. (**E**) Impact score of each hgRNA in the early lineage tree of each embryo. (**F**) Distribution of tree calculation outcomes when only including k of the $n_{fast} + n_{mid}$ fast and mid hgRNAs (left side of each panel, ${}^{n_{fast}+n_{mid}}C_k$ combinations) or k of the n_{slow} slow hgRNAs (right side of each panel, ${}^{n_{slow}}C_k$ combinations). Color code is as described in (B).

three lineages: the decidual zone (DZ) and the junctional zone (JZ) of the placenta, which are descendants of the trophectoderm (*30*, *31*); the parietal endoderm (PE) and visceral endoderm (VE) of the yolk sac, which are descendants of the primitive endoderm; and the heart and a limb bud of the embryo proper, which are descendants of the epiblast (*29*) (Fig. 5B). We then assembled the full barcode for each sample (Fig. 5, C and D, and table S6) and, using their Manhattan distances, clustered them to form a tree for each embryo (Fig. 5E) (*24*). Remarkably, despite the differences in the number and composition of hgRNAs inherited, the resulting tree perfectly matched the expected lineage in all four embryos, showing that the DZ and JZ form one clade of the tree while the other clade comprises two subclades, one with PE and VE and the other with the heart and limb bud (Fig. 5E). These results demonstrate that accurate lineage trees can be constructed from developmentally barcoded mice.

We next evaluated the robustness of lineage tree derivation from hgRNA barcodes by calculating the tree topology with only parts of the full barcodes. For a bifurcating tree with six tips (limb, heart, VE, PE, JZ, and DZ; Fig. 6A), 945 distinct rooted topologies are possible (*32*). Only a single one of these 945 tree topologies perfectly matches the expected lineage tree; we refer to this topology as "perfect" (Figs. 5E and 6B). Another eight topologies would be correct if unrooted—that is, if all four clades are correctly assigned but the root is misplaced because a branch other than the one connecting the (DZ, JZ) clade to the ((PE, VE), (heart, limb)) clade is the longest (Fig. 6A). We refer to these topologies as "correct" (Fig. 6B). If three, two, or fewer than two of the four clades have been assigned correctly, we consider the topologies as "incomplete," "partial," and "wrong," respectively (Fig. 6B). With these distinctions, we evaluated the trees generated with all possible non-null subsets of the hgRNAs in each embryo. The results show that, depending on the embryo, 60% to 85% of all possible hgRNA subsets result in a perfect or correct topology (Fig. 6C), which compares favorably to the ~1% chance of randomly finding such topologies. With only three hgRNAs, more than 50% of all derived trees have a correct or perfect topology for each embryo (Fig. 6D). Furthermore, calculated topologies improve with increasing the number of hgRNAs (Fig. 6D). Combined, these results show that lineage tree derivation from in vivo–generated hgRNA barcodes is robust and that the use of a higher number of hgRNAs results in more reliable outcomes.

We then examined the contribution of each hgRNA to deriving the correct tree topology for each embryo. We defined the "impact score" of an hgRNA in each embryo's early lineage tree as the difference between the fraction of all correct and perfect topologies in which the hgRNA was considered and the fraction of all wrong and partial topologies in which the hgRNA was considered (Fig. 6E and figs. S9 and S10) (*24*). As such, an impact score of +1 would indicate that whenever the hgRNA was included in tree derivation, a correct or a perfect topology was obtained, and no such topologies were obtained without that hgRNA. An impact score of −1 would indicate that when the hgRNA was included in tree derivation, only partial or wrong topologies were obtained. Values between +1 and −1 define the range between those entirely constructive or destructive outcomes, with an impact score of 0 indicating that the likelihood of obtaining a correct topology is the same with or without the hgRNA. Impact scores for hgRNAs in our four embryos show a positive average contribution by all three active hgRNA classes with slow hgRNAs, which are largely unmutated early in development (Fig. 2B), having an average impact close to 0, and mid and fast hgRNAs, which are active early in development (Fig. 2B), having increasingly positive impacts on the derivation of the correct tree (Fig. 6E). In fact, only three fast and mid hgRNAs suffice to obtain a correct or perfect topology in more than 90% of all derived trees (Fig. 6F). By contrast, exclusive use of slow hgRNAs does not recover the early lineage tree as reliably (Fig. 6F). Combined, these results suggest that active mutagenesis during a differentiation event allows it to be recorded. They also suggest that when the developmental stage in which a lineage differentiates is known, hgRNA activity profiles (Fig. 2A) can aid in choosing the appropriate hgRNAs such that correct trees can be reliably obtained with just a few hgRNAs.

Interestingly, when only slow hgRNAs were considered in tree construction for early lineages, increasing the number of hgRNAs still resulted in improved outcomes (Fig. 6F). This observation suggests that even when hgRNAs have low activity levels at the time an event is being recorded, partial recordings from multiple hgRNAs can be combined to obtain a more complete recording. As another example, four different hgRNAs from embryo 3 predict a partial tree when considered on their own, yet the perfect tree is derived when all four are considered together (fig. S11), further supporting the integrability of hgRNA recordings.

In two of our lineage-analyzed embryo samples (Fig. 5), we noted several hgRNAs in which all ICM-derived tissue samples (PE, VE, heart, limb) were dominated by a single mutant allele, whereas the corresponding trophectoderm-derived tissue samples (DZ, JZ) displayed a more uniform distribution of multiple mutant alleles (Fig. 7). These profiles suggest that in these embryos, these hgRNAs mutated as the trophectoderm and ICM lineages differentiated, and that fewer blastomeres led to the ICM than to the trophectoderm. These observations are consistent with previously reported observations (*33*, *34*) and suggest that hgRNA mutation profiles could be used to measure both the relationship between lineages and the relative number of cells that seed lineages.

Axis development in the brain

We next used developmentally barcoded mice to address lineages above the first lineages in the tree, with a focus on the establishment of the anterior-posterior (A-P) axis versus the lateral (L-R) axis in the brain. Patterning of the nervous system and its progenitors starts in gastrulation (E6.5) when the embryo has radial symmetry (*35*, *36*). By E8.5, both A-P and L-R axes are established in the neural tube (Fig. 8A); however, it remains unclear which axis is established first (*37*, *38*). At a morphological level they appear concurrently (*39*), and previous single-cell labeling and tracing experiments carried out ex vivo do not adequately address the issue (*40*). We analyzed two developmentally barcoded mice in the adult stage. In one, we dissected the left and right cortex and cerebellum, while in the other we additionally dissected the tectum. The cortex, tectum, and cerebellum respectively originate from embryonic forebrain (prosencephalon), midbrain (mesencephalon), and hindbrain (rhombencephalon) vesicles of the neural tube (Fig. 8A). From each region, two samples of neuronal nuclei were sorted (*24*). We also obtained

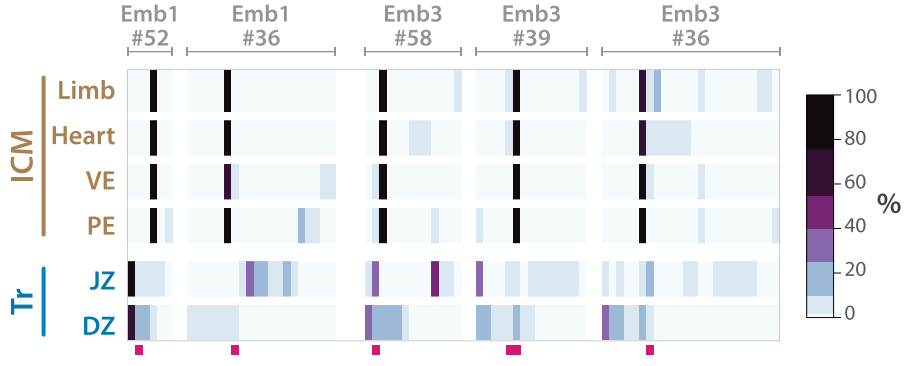


Fig. 7. Trophectoderm and ICM barcodes show differences in their number of mutant hgRNA alleles. Five barcodes from two embryos in Fig. 5D are shown that distinguish trophectoderm-derived and ICM-derived samples. Deep pink bars below each map mark highly recurring alleles that have been observed in more than 60% of all mice analyzed in Table 2. See table S6 for a numerical version of each barcode map. Only mutant alleles with a maximum abundance of more than 2% are shown.

samples of the blood and muscle from each mouse, both mesoderm-derived, to serve as outgroups. We then assembled the full barcode for each sample and applied clustering as before (Fig. 8, B and C, and fig. S12). In addition to segregating the mesoderm- and ectoderm-derived cells, the results clearly show that neurons from the left side of each brain region are more closely related to neurons from the right side of the same region than they are to neurons from either of the other two regions. Considering that no extensive migration of neuronal cell bodies between the regions sampled here has been reported (*41*), these results suggest that commitment to the A-P axis is established before commitment to the L-R axis in development of the central nervous system.

Similar to the first lineage tree analysis above (Fig. 6), we evaluated the robustness of the brain axis tree derivation as well as the contribution of each hgRNA in mouse 2 (Fig. 8, D and E) (*24*). We assigned topologies with all three left and right sample pairs placed closest to one another as correct, and those with two, one, or zero pairs placed as incomplete, partial, and wrong, respectively. We then calculated the distribution of tree derivation outcomes with all possible subsets of active hgRNAs in mouse 2 (Fig. 8D). The results show that half of the combinations with only three hgRNAs derive a correct or partially correct topology, a ratio that only improves when including more hgRNAs. We also calculated the impact score of each hgRNA (Fig. 8E) (*24*). Relative to impact scores for the first lineage tree (Fig. 6E), we found smaller contributions by fast hgRNAs, which would be expected for lineages that segregate much later in development. Taken together, these results demonstrate that lineages across diverse developmental times are recorded in our developmentally barcoded mice and can be extracted.

Discussion

In this study, we created an hgRNA mouse line for in vivo barcoding and used it to generate developmentally barcoded mice in which lineage information is recorded in cell genomes and can be extracted and reconstructed. Our strategy to create the MARC1 line was designed to address challenges associated with in vivo barcoding in a mouse model. First, genetic manipulation of individual mouse embryos is more challenging than that of lower vertebrates. Therefore, a line with genomically integrated, stable, and heritable barcoding elements that can be activated by simply crossing with other lines is powerful, versatile, and shareable (*24*). Second, tracking development in mice demands that the system be capable of generating a great many barcodes with little overwriting or deletion (*14*). As such, we scattered hgRNAs throughout the genome instead of using a contiguous array, circumventing large deletion events that can occur with multiple adjacent cut sites (*42*–*44*) and can remove prior recordings. In fact, we estimate that less than 1% of all mutations resulted in unidentifiable alleles by removing amplification primer binding sites or all unique sequences (*24*). The scarcity of these unwanted deletion events led to a great success rate in analyzing barcoded mice (4/4 in Fig. 5, 2/2 in Fig. 8). Furthermore, as hgRNA loci in this scattered array accumulate mutations independently, their mutant alleles combine exponentially to create a large diversity of barcodes. Consequently, the 41 active MARC1 hgRNAs can in theory combine to create more than 10^{74} different barcodes [$\prod_{i=1}^{41} n_i$, where n_i is the total number of observed mutant alleles for hgRNA i in this study, which is likely an underestimation (see above)]. Even only five fast and five mid hgRNAs can combine for roughly

Table 1. hgRNAs in the MARC1 founder male. TSS-to-PAM length, observed inheritance probability, and chromosome number for all 60 hgRNAs. See tables S1 to S3 for more details.

ID	Length (bp)	Inheritance (%)	Location	Class	ID	Length (bp)	Inheritance (%)	Location	Class
1	21	49.6	chr12	Slow	**31**	35	58.4	chr9	Inactive
2	35	55.2	chr7	Inactive	**32**	24	44	chr11	Mid
3	35	55.2	chr4	Inactive	**33**	21	55.2	chr13	Slow
4	21	29.6	chr14	Inactive	**34**	35	40	chr2	Mid
5	35	41.6	chr6	Inactive	**35**	25	52	chr19	Slow
6	21	54.4	chr2	Mid	**36**	21	45.6	chr8	Mid
7	34	43.2	chr19	Slow	**37**	35	51.2	chr4	Slow
8	30	16.8		Slow	**38**	21	38.4	chr13	Slow
9	35	51.2	chr19	Slow	**39**	25	54.4	chr7	Mid
10	25	42.4	chr4	Slow	**40**	35	45.6	chr9	Inactive
11	25	35.2	chr4	Inactive	**41**	21	38.4	chr2	Mid
12	35	48.8	chr17	Inactive	**42**	30	15.2		Slow
13	35	42.4	chr6	Slow	**43**	35	48.8	chr16	Slow
14	35	43.2	chr7	Inactive	**44**	21	51.2	chr14	Inactive
15	34	60.8	chr11	Slow	**45**	30	7.2		Inactive
16	35	47.2	chr9	Inactive	**46**	25	47.2	chr18	Slow
17	21	59.2	chr11	Slow	**47**	35	48.8	chr6	Inactive
18	25	50.4	chr13	Slow	**48**	35	34.4	chr5	Slow
19	21	39.2	chr7	Fast	**49**	35	58.4	chr11	Slow
20	21	49.6	chr18	Fast	**50**	21	44	chr10	Mid
21	35	40	chr2	Slow	**51**	25	36.8	chr11	Slow
22	21	42.4	chr4	Mid	**52**	21	53.6	chr2	Fast
23	21	80.8	chr1&4	Mid	**53**	35	47.2	chr3	Inactive
24	21	51.2	chr3	Inactive	**54**	35	73.6	chr10&13	Slow
25	25	48	chr8	Inactive	**55**	25	68.8	chr14	Slow
26	35	46.4	chr4	Inactive	**56**	21	55.2	chr2	Slow
27	21	39.2	chr2	Slow	**57**	34	44	chr11	Inactive
28	25	40	chr5	Slow	**58**	21	43.2	chr6	Fast
29	35	52.8	chr4	Slow	**59**	21	50.4	chr2	Fast
30	35	55.2	chr5	Inactive	**60**	25	51.2	chr3	Slow

10^{23} different barcodes ($200^5 \times 300^5$, where 200 and 300 are the observed average number of mutant alleles for fast and mid hgRNAs, respectively). This remarkable diversity is adequate for uniquely barcoding every one of the ~10^{10} cells in a mouse. Furthermore, assuming a perfect binary developmental tree as a first-order approximation, this diversity is adequate for uniquely marking all the ~2×10^{10} internal and terminal nodes of the mouse developmental tree.

Close analysis of the nature of mutant alleles in hgRNA barcodes showed the interplay between target site sequence and Cas9-induced double-strand breaks that determines the possible NHEJ outcomes (Fig. 3). Specifically, short indels underlie recurring NHEJ outcomes. Notable among these is a recurring duplication of the base at the −4 position relative to the PAM in a majority of active hgRNAs. The most likely explanation for this observation is Cas9 creating staggered cuts that produce a single-base 5′ overhang (Fig. 3K), because a terminal transferase activity would not duplicate the base adjacent to the cut site, and RuvC exonuclease activity on the noncomplementary strand would not result in an insertion at all. Whether Cas9 creates blunt or staggered ends in vivo has been a subject of debate. Our observation in mice, combined with a recent report in yeast (*45*) and previous in vitro and in vivo evidence (*44*, *46*–*50*), clarifies that Cas9 can create staggered ends as well as blunt ends, although the ratio of the two is unknown as of yet.

By crossing the MARC1 line with a line that constitutively expresses Cas9, we generated developmentally barcoded mice in which lineage information is recorded in the hgRNA barcodes. We were able to reconstruct parts of the lineage tree using these mice, with the first branches that emerge after the zygote, up to some of the germ layer, neuroectoderm, and the neural tube branches (Figs. 5 and 8). We find remarkable robustness and flexibility in these recordings (Figs. 6 and 8D). Specifically, there is overlap in recordings made by various hgRNAs, and therefore the derived lineage tree is robust to removing any part of the barcode. Furthermore, partial nonoverlapping recordings from different hgRNAs can be integrated to reconstruct a complete tree. Combined with evidence of sustained hgRNA

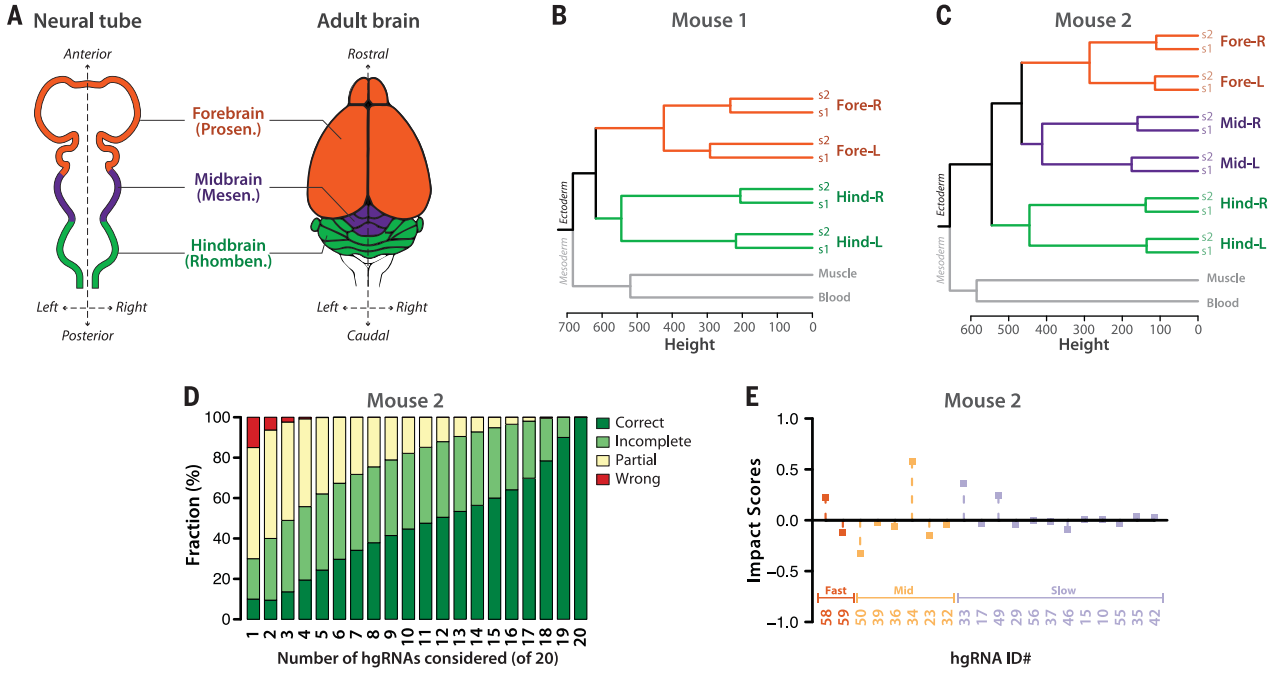


Fig. 8. The anterior-posterior axis is established before the left-right axis in the development of the brain. (**A**) Dorsal view of the neural tube and superior view of the adult brain in mouse. The primary brain vesicles in the neural tube and their corresponding structures in the adult brain are shown. (**B** and **C**) Calculated trees based on hgRNA barcodes in two adult mice. See fig. S12 for the full barcodes. (**D**) Distribution of tree calculation outcomes for mouse 2 when only including *m* of the *n* hgRNAs (^nC_m combinations). Only hgRNAs with at least a 7% mutation rate in one of the samples were considered. (**E**) Impact score of each hgRNA in the early lineage tree of mouse 2.

Table 2. Breakdown of all mice used for hgRNA activity analysis according to developmental stage and number of samples obtained per mouse.

Samples per mouse	Stage of development										Total mice	Total samples
	E3.5	E6.5	E7.5	E8.5	E10.5	E12.5	E14.5	E15.5	E16.5	Adult		
One	11	6	3	3	1	0	0	0	0	45	69	69
Two	0	0	4	6	1	0	0	0	0	0	11	22
Three	0	0	1	0	0	0	0	0	0	0	1	3
Four	0	0	0	0	7	6	0	0	2	0	15	60
Five	0	0	0	0	0	0	0	0	0	0	0	0
Six	0	0	0	0	0	6	0	0	0	0	6	36
Total	11	6	8	9	9	12	0	0	2	45	102	190

mutagenesis throughout gestation, these results suggest that developmentally barcoded mice embody information from various stages of development in embryonic and extraembryonic tissues. Extracting such information will be a matter of the type of question being investigated and an ability to isolate cells from relevant lineages. Another interesting possibility is to create different types of barcoded mice by crossing the MARC1 line with other *S. pyogenes* Cas9 lines. Among these are inducible Cas9s (*51*, *52*), ones with different activity levels (*53*), tissue- or lineage-specific versions based on Cre drivers (*27*, *53*), or base-editing Cas9s (*54*, *55*). Such barcoded mice may enhance the capabilities of the system, overcome its shortcomings (*24*), or better focus its potential on specific problems.

Our results provide a platform for in vivo barcoding and lineage tracing in the mouse. Although we have focused here on the recording aspect of in vivo developmental barcoding, more effective readout strategies—in particular, those with transcriptome-coupled single-cell readouts (*19*, *21*, *22*) or with in situ readouts (*56*)—will be necessary. Finally, in addition to lineage-tracing applications, this platform may also be applied to recording cellular signals over time (*16*, *17*, *57*–*59*) and uniquely barcoding each cell in a tissue or an organism for identification purposes, such as for connectome mapping in the brain (*60*–*62*).

Methods summary

All animal procedures were approved by the Harvard University Institutional Animal Care and Use Committee (IACUC). For embryonic samples, the MARC1 founder was crossed with a Cas9 knock-in female. Pregnant females were then dissected at the desired embryonic time points, designating noon of the day of vaginal plug detection as E0.5. For isolating neurons from adult barcoded female mice, brains were dissected into the regions of interest and homogenized. Nuclei were isolated from the homogenate by gradient ultracentrifugation, labeling with a NeuN antibody, and sorting the NeuN-positive fraction in flow cytometry. From all obtained samples, DNA was extracted and amplified with specific primers for hgRNA loci. The resulting amplicons were sequenced with paired ends and analyzed to identify the hgRNA itself according to the identifier sequence, and the mutant allele according to the spacer sequence. The sequencing results were processed and filtered to obtain a list of high-confidence unique spacer-identifier pairs observed in each sample and their respective abundances. For obtaining lineage trees, these lists were converted into frequency matrices and clustered hierarchically using Ward's criterion. All procedures for the experiments and data analyses are described in detail in the supplementary materials.

REFERENCES AND NOTES

1. M. H. Kaufman, *The Atlas of Mouse Development* (Elsevier, 1992).
2. J. E. Sulston, E. Schierenberg, J. G. White, J. N. Thomson, The embryonic cell lineage of the nematode Caenorhabditis elegans. *Dev. Biol.* **100**, 64–119 (1983). doi: 10.1016/0012-1606(83)90201-4; pmid: 6684600
3. D. L. Riddle, T. Blumenthal, B. J. Meyer, J. R. Priess, in *C. elegans II* (Cold Spring Harbor Laboratory Press, ed. 2, 1997).
4. D. A. Weisblat, R. T. Sawyer, G. S. Stent, Cell lineage analysis by intracellular injection of a tracer enzyme. *Science* **202**, 1295–1298 (1978). doi: 10.1126/science.725606; pmid: 725606
5. C. Walsh, C. L. Cepko, Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science* **255**, 434–440 (1992). doi: 10.1126/science.1734520; pmid: 1734520
6. S. M. Dymecki, H. Tomasiewicz, Using Flp-recombinase to characterize expansion of Wnt1-expressing neural progenitors in the mouse. *Dev. Biol.* **201**, 57–65 (1998). doi: 10.1006/dbio.1998.8971; pmid: 9733573
7. K. Kretzschmar, F. M. Watt, Lineage tracing. *Cell* **148**, 33–45 (2012). doi: 10.1016/j.cell.2012.01.002; pmid: 22265400
8. S. H. Naik, T. N. Schumacher, L. Perié, Cellular barcoding: A technical appraisal. *Exp. Hematol.* **42**, 598–608 (2014). doi: 10.1016/j.exphem.2014.05.003; pmid: 24996012
9. A. Gerrits *et al.*, Cellular barcoding tool for clonal analysis in the hematopoietic system. *Blood* **115**, 2610–2618 (2010). doi: 10.1182/blood-2009-06-229757; pmid: 20093403
10. M. B. Woodworth, K. M. Girskis, C. A. Walsh, Building a lineage from single cells: Genetic techniques for cell lineage tracking. *Nat. Rev. Genet.* **18**, 230–244 (2017). doi: 10.1038/nrg.2016.159; pmid: 28111472
11. J. Ma, Z. Shen, Y.-C. Yu, S.-H. Shi, Neural lineage tracing in the mammalian brain. *Curr. Opin. Neurobiol.* **50**, 7–16 (2018). doi: 10.1016/j.conb.2017.10.013; pmid: 29125960
12. I. D. Peikon, D. I. Gizatullina, A. M. Zador, In vivo generation of DNA sequence diversity for cellular barcoding. *Nucleic Acids Res.* **42**, e127 (2014). doi: 10.1093/nar/gku604; pmid: 25013177
13. A. McKenna *et al.*, Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science* **353**, aaf7907 (2016). doi: 10.1126/science.aaf7907; pmid: 27229144
14. R. Kalhor, P. Mali, G. M. Church, Rapidly evolving homing CRISPR barcodes. *Nat. Methods* **14**, 195–200 (2017). doi: 10.1038/nmeth.4108; pmid: 27918539
15. K. L. Frieda *et al.*, Synthetic recording and in situ readout of lineage information in single cells. *Nature* **541**, 107–111 (2017). doi: 10.1038/nature20777; pmid: 27869821
16. S. D. Perli, C. H. Cui, T. K. Lu, Continuous genetic recording with self-targeting CRISPR-Cas in human cells. *Science* **353**, aag0511 (2016). doi: 10.1126/science.aag0511; pmid: 27540006
17. R. U. Sheth, S. S. Yim, F. L. Wu, H. H. Wang, Multiplex recording of cellular events over time on CRISPR biological tape. *Science* **358**, 1457–1461 (2017). doi: 10.1126/science.aao0958; pmid: 29170279
18. S. T. Schmidt, S. M. Zimmerman, J. Wang, S. K. Kim, S. R. Quake, Quantitative Analysis of Synthetic Cell Lineage Tracing Using Nuclease Barcoding. *ACS Synth. Biol.* **6**, 936–942 (2017). doi: 10.1021/acssynbio.6b00309; pmid: 28264564
19. A. Alemany, M. Florescu, C. S. Baron, J. Peterson-Maduro, A. van Oudenaarden, Whole-organism clone tracing using single-cell sequencing. *Nature* **556**, 108–112 (2018). doi: 10.1038/nature25969; pmid: 29590089
20. G. P. Flowers, L. D. Sanor, C. M. Crews, Lineage tracing of genome-edited alleles reveals high fidelity axolotl limb regeneration. *eLife* **6**, e25726 (2017). doi: 10.7554/eLife.25726; pmid: 28917058
21. B. Spanjaard *et al.*, Simultaneous lineage tracing and cell-type identification using CRISPR-Cas9-induced genetic scars. *Nat. Biotechnol.* **36**, 469–473 (2018). doi: 10.1038/nbt.4124; pmid: 29644996
22. B. Raj *et al.*, Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat. Biotechnol.* **36**, 442–450 (2018). doi: 10.1038/nbt.4103; pmid: 29608178
23. P. Mali *et al.*, RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013). doi: 10.1126/science.1232033; pmid: 23287722
24. See supplementary materials.
25. C. L. Markert, R. M. Petters, Manufactured hexaparental mice show that adults are derived from three embryonic cells. *Science* **202**, 56–58 (1978). doi: 10.1126/science.694518; pmid: 694518
26. Z. Wang, R. Jaenisch, At most three ES cells contribute to the somatic lineages of chimeric mice and of mice produced by ES-tetraploid complementation. *Dev. Biol.* **275**, 192–201 (2004). doi: 10.1016/j.ydbio.2004.06.026; pmid: 15464582
27. R. J. Platt *et al.*, CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014). doi: 10.1016/j.cell.2014.09.014; pmid: 25263330
28. M. T. Lee, A. R. Bonneau, A. J. Giraldez, Zygotic genome activation during the maternal-to-zygotic transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581–613 (2014). doi: 10.1146/annurev-cellbio-100913-013027; pmid: 25150012
29. C. C. Lu, J. Brennan, E. J. Robertson, From fertilization to gastrulation: Axis formation in the mouse embryo. *Curr. Opin. Genet. Dev.* **11**, 384–392 (2001). doi: 10.1016/S0959-437X(00)00208-2; pmid: 11448624
30. J. Rossant, J. C. Cross, in *Mouse Development: Patterning, Morphogenesis, and Organogenesis*, J. Rossant, P. P. L. Tam, Eds. (Academic Press, 2002), pp. 155–180.
31. D. G. Simmons, J. C. Cross, Determinants of trophoblast lineage and cell subtype specification in the mouse placenta. *Dev. Biol.* **284**, 12–24 (2005). doi: 10.1016/j.ydbio.2005.05.010; pmid: 15963972
32. J. Felsenstein, *Inferring Phylogenies* (Sinauer Associates, 2004).
33. C. Y. Leung, M. Zhu, M. Zernicka-Goetz, Polarity in Cell-Fate Acquisition in the Early Mouse Embryo. *Curr. Top. Dev. Biol.* **120**, 203–234 (2016). doi: 10.1016/bs.ctdb.2016.04.008; pmid: 27475853
34. S. Anani, S. Bhat, N. Honma-Yamanaka, D. Krawchuk, Y. Yamanaka, Initiation of Hippo signaling is linked to polarity rather than to cell position in the pre-implantation mouse embryo. *Development* **141**, 2813–2824 (2014). doi: 10.1242/dev.107276; pmid: 24948601
35. R. S. Beddington, E. J. Robertson, Axis development and early asymmetry in mammals. *Cell* **96**, 195–209 (1999). doi: 10.1016/S0092-8674(00)80560-7; pmid: 9988215
36. Z. Molnár, D. J. Price, in *Kaufman's Atlas of Mouse Development Supplement*, R. Baldock, J. Bard, D. R. Davidson, G. Morriss-Kay, Eds. (Elsevier, 2016), pp. 239–252.
37. S. J. Arnold, E. J. Robertson, Making a commitment: Cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* **10**, 91–103 (2009). doi: 10.1038/nrm2618; pmid: 19129791
38. C. Kiecker, A. Lumsden, The role of organizers in patterning the nervous system. *Annu. Rev. Neurosci.* **35**, 347–367 (2012). doi: 10.1146/annurev-neuro-062111-150543; pmid: 22462542
39. C. R. Altmann, A. H. Brivanlou, Neural patterning in the vertebrate embryo. *Int. Rev. Cytol.* **203**, 447–482 (2001). doi: 10.1016/S0074-7696(01)03013-3; pmid: 11131523
40. K. A. Lawson, J. J. Meneses, R. A. Pedersen, Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* **113**, 891–911 (1991). pmid: 1821858
41. R. Ayala, T. Shu, L.-H. Tsai, Trekking across the brain: The journey of neuronal migration. *Cell* **128**, 29–43 (2007). doi: 10.1016/j.cell.2006.12.021; pmid: 17218253
42. H. J. Lee, E. Kim, J.-S. Kim, Targeted chromosomal deletions in human cells using zinc finger nucleases. *Genome Res.* **20**, 81–89 (2010). doi: 10.1101/gr.099747.109; pmid: 19952142
43. M. C. Canver *et al.*, Characterization of genomic deletion efficiency mediated by clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J. Biol. Chem.* **289**, 21312–21324 (2014). doi: 10.1074/jbc.M114.564625; pmid: 24907273
44. S. M. Byrne, L. Ortiz, P. Mali, J. Aach, G. M. Church, Multi-kilobase homozygous targeted gene replacement in human induced pluripotent stem cells. *Nucleic Acids Res.* **43**, e21 (2015). doi: 10.1093/nar/gku1246; pmid: 25414332
45. B. R. Lemos *et al.*, CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E2040–E2047 (2018). doi: 10.1073/pnas.1716855115; pmid: 29440496
46. G. Gasiunas, R. Barrangou, P. Horvath, V. Siksnys, Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* **109**, E2579–E2586 (2012). doi: 10.1073/pnas.1208507109; pmid: 22949671
47. M. Jinek *et al.*, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012). doi: 10.1126/science.1225829; pmid: 22745249
48. M. Jinek *et al.*, Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997 (2014). doi: 10.1126/science.1247997; pmid: 24505130
49. Z. Zuo, J. Liu, Cas9-catalyzed DNA Cleavage Generates Staggered Ends: Evidence from Molecular Dynamics

Simulations. *Sci. Rep.* **5**, 37584 (2016). doi: 10.1038/srep37584; pmid: 27874072
50. S. Q. Tsai *et al*., GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* **33**, 187–197 (2015). doi: 10.1038/nbt.3117; pmid: 25513782
51. L. E. Dow *et al*., Inducible in vivo genome editing with CRISPR-Cas9. *Nat. Biotechnol.* **33**, 390–394 (2015). doi: 10.1038/nbt.3155; pmid: 25690852
52. L. R. Polstein, C. A. Gersbach, A light-inducible CRISPR-Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11**, 198–200 (2015). doi: 10.1038/nchembio.1753; pmid: 25664691
53. S.-H. Chiou *et al*., Pancreatic cancer modeling using retrograde viral vector delivery and in vivo CRISPR/Cas9-mediated somatic genome editing. *Genes Dev.* **29**, 1576–1585 (2015). doi: 10.1101/gad.264861.115; pmid: 26178787
54. A. C. Komor, Y. B. Kim, M. S. Packer, J. A. Zuris, D. R. Liu, Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016). doi: 10.1038/nature17946; pmid: 27096365
55. N. M. Gaudelli *et al*., Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* **551**, 464–471 (2017). doi: 10.1038/nature24644; pmid: 29160308
56. J. H. Lee *et al*., Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014). doi: 10.1126/science.1250212; pmid: 24578530
57. A. H. Marblestone *et al*., Physical principles for scalable neural recording. *Front. Comput. Neurosci.* **7**, 137 (2013). doi: 10.3389/fncom.2013.00137; pmid: 24187539
58. S. L. Shipman, J. Nivala, J. D. Macklis, G. M. Church, CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature* **547**, 345–349 (2017). doi: 10.1038/nature23017; pmid: 28700573
59. S. Punthambaker, R. I. Hume, Potent and long-lasting inhibition of human P2X2 receptors by copper. *Neuropharmacology* **77**, 167–176 (2014). doi: 10.1016/j.neuropharm.2013.09.001; pmid: 24067922
60. D. Cai, K. B. Cohen, T. Luo, J. W. Lichtman, J. R. Sanes, Improved tools for the Brainbow toolbox. *Nat. Methods* **10**, 540–547 (2013). doi: 10.1038/nmeth.2450
61. J. M. Kebschull *et al*., High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. *Neuron* **91**, 975–987 (2016). doi: 10.1016/j.neuron.2016.07.036; pmid: 27545715
62. G. Church, A. Marblestone, R. Kalhor, in *The Future of the Brain*, G. Marcus, J. Freeman, Eds. (Princeton Univ. Press, 2014), pp. 50–64.
63. M. R. Lieber, T. E. Wilson, SnapShot: Nonhomologous DNA end joining (NHEJ). *Cell* **142**, 496–496.e1 (2010). doi: 10.1016/j.cell.2010.07.035; pmid: 20691907

ACKNOWLEDGMENTS

We thank R. Jaenisch for generously sharing reagents; J. Aach for critical reading of the manuscript and helpful comments; A. Vernet for assistance with animal husbandry, G. Cuneo for assistance in FACS; L. Wu and the Genome Modification Facility for blastocyst injections as well as helpful discussions; S. Kennedy, J. Young, B. Fields, and A. Reslow for generously sharing equipment and providing technical advice; and M. Warman, C. Tabin, C. Cepko, Y. Stelzer, S. Shipman, J. Macklis, B. Khalaj, N. Davidsohn, A. Ng, and R. Kohman for helpful discussions. **Funding:** Supported by NIH grants MH103910 and HG005550 (G.M.C.), the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior/Interior Business Center (DoI/IBC) contract number D16PC00008 (G.M.C.), Burroughs Wellcome Fund 1013926 (P.M.), and NIH grants R01HG009285, R01CA222826, and R01GM123313 (P.M.). **Author contributions:** R.K., P.M., and G.M.C. conceived the study; R.K. designed and carried out the experiments and analyzed the data; K.K. analyzed the genomic location data and formulated other analyses; L.M. and K.L. assisted in mouse genotyping, dissections, library preparations, and sequencing; A.G. supervised animal husbandry; R.K., K.K., and P.M. interpreted the data and wrote the manuscript with input from all other authors; and G.M.C. supervised the project. **Competing interests:** The authors declare no competing financial interest. G.M.C.'s competing financial interests can be found at v.ht/PHNc. **Data and materials availability:** All sequencing data are available in the Sequence Read Archive (SRP155997); all other data are available in supplementary materials, and other materials are available upon request.

SUPPLEMENTARY MATERIALS

26 April 2018; resubmitted 19 July 2018
Accepted 2 August 2018
Published online 9 August 2018
10.1126/science.aat9804

RESEARCH ARTICLE

YELLOW FEVER

Genomic and epidemiological monitoring of yellow fever virus transmission potential

N. R. Faria[1]*†, M. U. G. Kraemer[1,2,3]*, S. C. Hill[1]*, J. Goes de Jesus[4]*, R. S. Aguiar[5]*, F. C. M. Iani[6,7]*, J. Xavier[4], J. Quick[8], L. du Plessis[1], S. Dellicour[9], J. Thézé[1], R. D. O. Carvalho[7], G. Baele[9], C.-H. Wu[10], P. P. Silveira[5], M. B. Arruda[5], M. A. Pereira[6], G. C. Pereira[6], J. Lourenço[1], U. Obolski[1], L. Abade[1,11], T. I. Vasylyeva[1], M. Giovanetti[4,7], D. Yi[12], D. J. Weiss[13], G. R. W. Wint[1], F. M. Shearer[13], S. Funk[14], B. Nikolay[15,16], V. Fonseca[7,17], T. E. R. Adelino[6], M. A. A. Oliveira[6], M. V. F. Silva[6], L. Sacchetto[7], P. O. Figueiredo[7], I. M. Rezende[7], E. M. Mello[7], R. F. C. Said[18], D. A. Santos[18], M. L. Ferraz[18], M. G. Brito[18], L. F. Santana[18], M. T. Menezes[5], R. M. Brindeiro[5], A. Tanuri[5], F. C. P. dos Santos[19], M. S. Cunha[19], J. S. Nogueira[19], I. M. Rocco[19], A. C. da Costa[20], S. C. V. Komninakis[21,22], V. Azevedo[7], A. O. Chieppe[23], E. S. M. Araujo[4], M. C. L. Mendonça[4], C. C. dos Santos[4], C. D. dos Santos[4], A. M. Mares-Guia[4], R. M. R. Nogueira[4], P. C. Sequeira[4], R. G. Abreu[24], M. H. O. Garcia[24], A. L. Abreu[25], O. Okumoto[25], E. G. Kroon[7], C. F. C. de Albuquerque[26], K. Lewandowski[27], S. T. Pullan[27], M. Carroll[28], T. de Oliveira[4,17,29], E. C. Sabino[20], R. P. Souza[19], M. A. Suchard[30,31], P. Lemey[9], G. S. Trindade[7], B. P. Drumond[7], A. M. B. Filippis[4], N. J. Loman[8], S. Cauchemez[15,16]*, L. C. J. Alcantara[4,7]*†, O. G. Pybus[1]*†

The yellow fever virus (YFV) epidemic in Brazil is the largest in decades. The recent discovery of YFV in Brazilian *Aedes* species mosquitos highlights a need to monitor the risk of reestablishment of urban YFV transmission in the Americas. We use a suite of epidemiological, spatial, and genomic approaches to characterize YFV transmission. We show that the age and sex distribution of human cases is characteristic of sylvatic transmission. Analysis of YFV cases combined with genomes generated locally reveals an early phase of sylvatic YFV transmission and spatial expansion toward previously YFV-free areas, followed by a rise in viral spillover to humans in late 2016. Our results establish a framework for monitoring YFV transmission in real time that will contribute to a global strategy to eliminate future YFV epidemics.

Yellow fever (YF) is responsible for 29,000 to 60,000 deaths annually in South America and Africa (*1*) and is the most severe mosquito-borne infection in the tropics (*2*). Despite the existence of an effective YF vaccine since 1937 (*3*), an estimated >400 million unvaccinated people live in areas at risk of infection (*4*). Yellow fever virus (YFV) is a member of the *Flaviviridae* family and is classified into four genotypes: East African, West African, South American I, and South American II (*5*–*9*). In the Americas, YFV transmission occurs mainly via the sylvatic cycle, in which nonhuman primates (NHPs) are infected by tree-dwelling mosquito vectors such as *Haemagogus* spp. and *Sabethes* spp. (*10*, *11*). YFV transmission can also occur via an urban cycle, in which humans are infected by *Aedes* spp. mosquitoes that feed mostly on humans (*12*, *13*).

Brazil has recently experienced its largest-recorded YF outbreak in decades, with 2043 confirmed cases and 676 deaths since December 2016 (supplementary text and fig. S1) (*14*). The last YF cases in Brazil attributed to an urban cycle were in Sena Madureira, in the northern state of Acre, in 1942 (*15*). An intensive eradication campaign eliminated *Aedes aegypti* and YF from Brazil in the 1950s (*16*), but the vector became reestablished in the 1970s and *Aedes* spp. mosquitoes are now abundant across most of Brazil (*17*). The consequences of a reignition of urban cycle transmission in Brazil would be serious, as an estimated 35 million people in areas at risk for YFV transmission in Brazil remain unvaccinated (*4*). New surveillance and analytical approaches are therefore needed to monitor this risk in real time.

Yellow fever virus outbreak in Brazil, 2016–2017

Between December 2016 and the end of June 2017, there were 777 polymerase chain reaction (PCR)–confirmed human cases of YF across 10 Brazilian states—mostly in Minas Gerais (MG) (60% of cases), followed by Espírito Santo (32%), Rio de Janeiro (3%), and São Paulo (3%) (*18*). The fatality ratio of severe YF cases was estimated at 33.6%, comparable to previous outbreaks (*19*, *20*). Despite the exceptional magnitude and rapid expansion of the outbreak, little is known about its genomic epidemiology. Further, it is uncertain how the virus is spreading through space, as well as between humans and NHPs, and analytical insights into the contribution of the urban cycle to ongoing transmission are lacking.

To characterize the 2017 YFV outbreak in Brazil, we first compared time series of confirmed cases in humans (n = 683) and NHPs (n = 313) reported until October 2017 by public health institutes in MG, the epicenter of the outbreak (Fig. 1, A and B, and fig. S2). The time series are strongly associated (cross-correlation coefficient = 0.97; $P < 0.001$). Both peak in late January 2017, and we estimate that human cases lag behind those in NHPs by 4 days (table S1). NHP cases are geographically more dispersed

[1]Department of Zoology, University of Oxford, Oxford, UK. [2]Computational Epidemiology Lab, Boston Children's Hospital, Boston, MA, USA. [3]Department of Pediatrics, Harvard Medical School, Boston, MA, USA. [4]Laboratório de Flavivírus, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil. [5]Laboratório de Virologia Molecular, Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. [6]Laboratório Central de Saúde Pública, Instituto Octávio Magalhães, FUNED, Belo Horizonte, Minas Gerais, Brazil. [7]Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil. [8]Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK. [9]Department of Microbiology and Immunology, Rega Institute, KU Leuven, Leuven, Belgium. [10]Department of Statistics, University of Oxford, Oxford, UK. [11]The Global Health Network, University of Oxford, Oxford, UK. [12]Department of Statistics, Harvard University, Cambridge, MA, USA. [13]Malaria Atlas Project, Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK. [14]Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK. [15]Mathematical Modelling of Infectious Diseases and Center of Bioinformatics, Institut Pasteur, Paris, France. [16]CNRS UMR2000: Génomique Évolutive, Modélisation et Santé, Institut Pasteur, Paris, France. [17]KwaZulu-Natal Research, Innovation and Sequencing Platform (KRISP), School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, South Africa. [18]Secretaria de Estado de Saúde de Minas Gerais, Belo Horizonte, Minas Gerais, Brazil. [19]Núcleo de Doenças de Transmissão Vetorial, Instituto Adolfo Lutz, São Paulo, Brazil. [20]Instituto de Medicina Tropical e Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil. [21]Retrovirology Laboratory, Federal University of São Paulo, São Paulo, Brazil. [22]School of Medicine of ABC (FMABC), Clinical Immunology Laboratory, Santo André, São Paulo, Brazil. [23]Coordenação de Vigilância Epidemiológica do Estado do Rio de Janeiro, Rio de Janeiro, Brazil. [24]Departamento de Vigilância das Doenças Transmissíveis da Secretaria de Vigilância em Saúde, Ministério da Saúde, Brasília-DF, Brazil. [25]Secretaria de Vigilância em Saúde, Coordenação Geral de Laboratórios de Saúde Pública, Ministério da Saúde, Brasília-DF, Brazil. [26]Organização Pan - Americana da Saúde/Organização Mundial da Saúde - (OPAS/OMS), Brasília-DF, Brazil. [27]Public Health England, National Infections Service, Porton Down, Salisbury, UK. [28]NIHR HPRU in Emerging and Zoonotic Infections, Public Health England, London, UK. [29]Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa. [30]Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles, CA, USA. [31]Department of Biomathematics and Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA.
*These authors contributed equally to this work.
†Corresponding author. Email: nuno.faria@zoo.ox.ac.uk (N.R.F.); luiz.alcantara@ioc.fiocruz.br (L.C.J.A.); oliver.pybus@zoo.ox.ac.uk (O.G.P.)

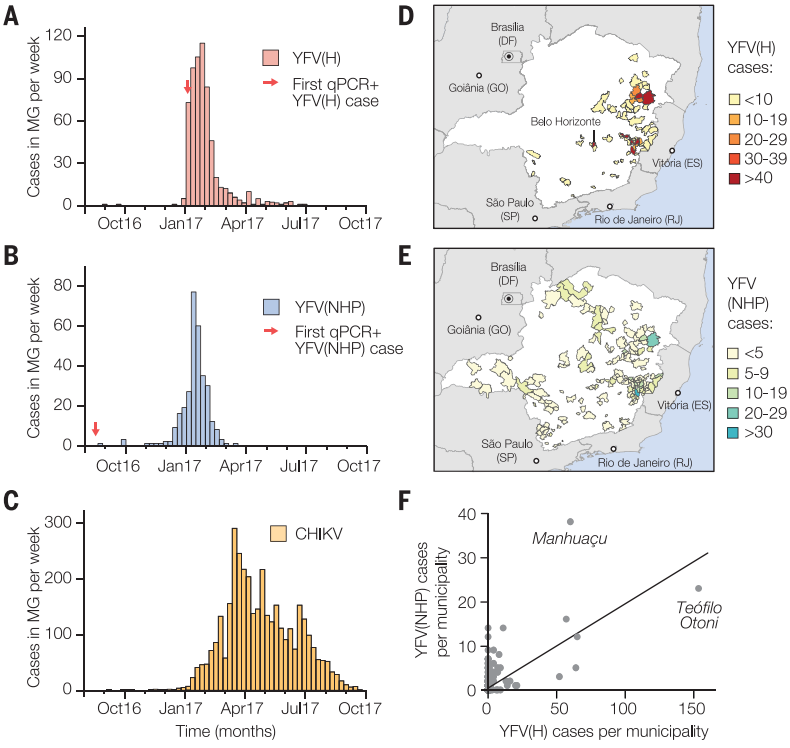


Fig. 1. Spatial and temporal epidemiology of YFV and CHIKV in Minas Gerais (MG). (**A**) Time series of human (H) YFV cases in MG (676 cases across 61 municipalities)—confirmed by serology, reverse transcription quantitative PCR (RT-qPCR), or virus isolation—during the first YFV epidemic wave (August 2016 to October 2017). (**B**) Same as in (A) but showing NHP YFV cases (313 cases across 90 municipalities) confirmed by RT-qPCR. (**C**) Same as in (A) but showing human CHIKV cases (3668 cases across 129 municipalities). (**D**) Geographic distribution of human YFV cases in MG. (**E**) Geographic distribution of NHP YFV cases in MG. Figure S3 shows the corresponding geographic distribution of CHIKV cases. (**F**) Association between the number of human and NHP cases in each municipality of MG (Pearson's $r = 0.62$; $P < 0.0001$; nonparametric Spearman's rank $\rho = 0.32$; $P < 0.05$).

in MG than human cases, which are more concentrated in the Teófilo Otóni and Manhuaçu municipalities (Fig. 1, D and E). Despite this, the numbers of human and NHP cases per municipality are positively correlated (Fig. 1F).

To establish whether human cases are acquired in proximity to potential sources of sylvatic infection, we estimated the distance between the municipality of residence of each human case and the nearest habitat of potential transmission, determined by using the enhanced vegetation index (EVI) (*21*) (supplementary materials). The average minimum distance between areas with EVI > 0.4 and the residence of confirmed human cases is only 5.3 km. In contrast, a randomly chosen resident of MG lives, on average, ≥51 km away from areas with EVI > 0.4. Similarly, human YFV cases reside, on average, 1.7 km from the nearest NHP case, whereas the mean minimum distance of a randomly chosen MG resident to the nearest NHP case is 39.1 km. This is consistent with YF infection risk being greatest for people who reside or work in forested areas where sylvatic transmission occurs. We find that most human cases (98.5%) were reported in municipalities with estimated YFV vaccination coverage above the 80% threshold recommended by the World Health Organization (WHO). On average, human cases would need to travel 65 km from their place of residence to reach an area where vaccination coverage is <80% (*4*).

Risk of YFV urban transmission

YFV was detected in *Aedes albopictus* mosquitoes caught in MG in January 2017 (*22*). Further, experiments suggest that *Aedes* spp. mosquitoes from southeast Brazil can transmit Brazilian YFV, although perhaps less effectively than vectors from elsewhere in the country (*23*, *24*). It is therefore important to investigate whether YFV cases in MG occur where and when *Aedes* spp. vectors are active. To do so, we analyzed confirmed chikungunya virus (CHIKV) cases from MG (Fig. 1C).

CHIKV is transmitted by the urban mosquitoes *Ae. aegypti* and *Ae. albopictus* (*25*). There were 3755 confirmed CHIKV cases in MG during January 2015 to October 2017. The CHIKV epidemic in MG in 2017 began later and lasted longer than the YFV outbreak (Fig. 1C), consistent with the hypothesis that YFV and CHIKV in the region are transmitted by different vector species. However, 29 municipalities with human YFV cases also reported CHIKV cases (Fig. 1D and fig. S3), indicating that YFV is indeed present in municipalities with *Aedes* mosquitoes. The mean

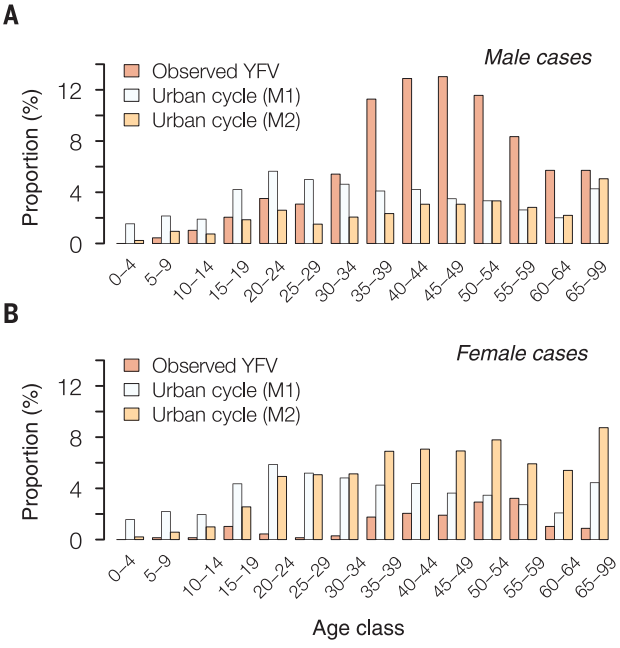


Fig. 2. Age and sex distribution of YFV cases in MG, 2016–2017. Red bars show the proportion of observed YFV cases in MG that occur in each age class, in (**A**) males and (**B**) females. These empirical distributions are different from those predicted under two models (M1, pale blue bars; M2, orange bars) of urban cycle transmission (see text for details).

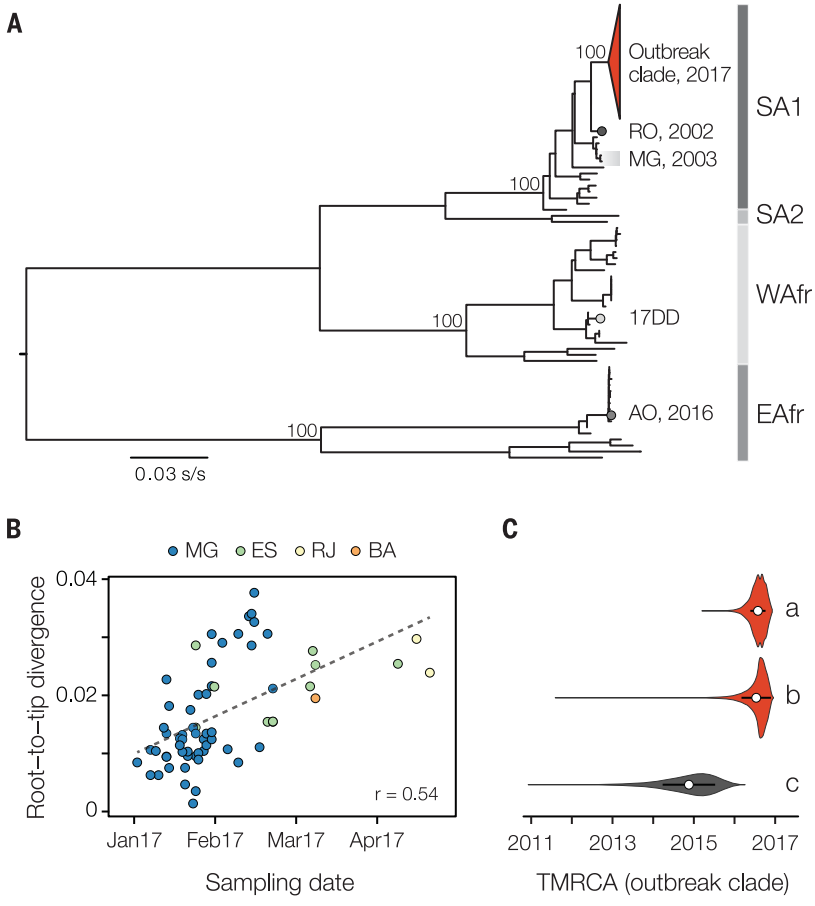


Fig. 3. Molecular phylogenetics of the Brazilian YFV epidemic. (**A**) Maximum likelihood phylogeny of complete YFV genomes showing the outbreak clade (red triangle) within the South American I (SA1) genotype (Fig. 4 and fig. S6). SA2, WAfr, and EAfr indicate the South America II, West Africa, and East Africa genotypes, respectively. For clarity, five YFV strains introduced to Venezuela from Brazil (*49*) are not shown. The scale bar is in units of substitutions per site (s/s). Node labels indicate bootstrap support values. RO 2002, strain BeH655417 from Roraima; MG 2003, two strains from the previous YF outbreak in MG in 2003; 17DD, the vaccine strain used in Brazil; AO 2016, YFV outbreak Angola in 2015–2016 (*13*). (**B**) Root-to-tip regression of sequence sampling date against genetic divergence from the root of the outbreak clade (fig. S6). Sequences are colored according to sampling location (MG, Minas Gerais; ES, Espírito Santo; RJ, Rio de Janeiro; BA, Bahia). (**C**) Violin plots showing estimated posterior distributions (white circles denote means) of the time of the most recent common ancestor (TMRCA) of the outbreak clade. Estimates were obtained using two different datasets (gray, SA1 genotype; red, outbreak clade) and under different evolutionary models: a, uncorrelated lognormal relaxed clock (UCLN) model with a skygrid tree prior with covariates specifically, the time series data shown in Fig. 1, A to C; also see fig. S7); b, UCLN model with a skygrid tree prior without covariates; c, fixed local clock model (see supplementary materials).

YFV vaccination rate in districts with both YFV and CHIKV cases is 72.6% (range = 61 to 78%) (*4*). Thus, relatively high vaccination rates in the locations in MG where YF spillover to humans occurs, and potentially lower vector competence (*23*, *24*), may ameliorate the risk of establishment of an urban YFV cycle in the state. However, adjacent urban regions (including São Paulo and Rio de Janeiro) have lower vaccination rates (*4*), receive tens of millions of visitors per year (*26*), and have recently experienced many human YFV cases (*20*). Thus, the possibility of sustained urban YFV transmission in southern Brazil and beyond necessitates continual virological and epidemiological monitoring.

We sought to establish a framework to evaluate routes of YFV transmission during an outbreak from the characteristics of infected individuals. Specifically, we assessed whether an outbreak is driven by sylvatic or urban transmission by comparing the age and sex distributions of observed YFV cases with those expected under an urban cycle in MG. For example, an individual's risk of acquiring YFV via the sylvatic cycle depends on their likelihood to travel to forested areas, an occurence that is typically highest among male adults (*27*). In contrast, under an urban cycle, we expect more uniform exposure across age and sex classes, similar to that observed for urban cases in Paraguay (*28*) and Nigeria (*29*).

The male-to-female sex ratio of reported YFV cases in MG is 5.7 (85% of cases are male), and incidence is highest among males aged 40 to 49 (Fig. 2). We compared this distribution to that expected under two models of urban cycle transmission (supplementary materials). In model M1, age and sex classes vary in vaccination status but are equally exposed to YFV, a scenario that is typical of arboviral transmission (*30*). Under model M1, predicted cases are characterized by a sex ratio ~1, and incidence peaks among individuals aged 20 to 25 (Fig. 2). In model M2, we assume that the pattern of YFV exposure among age and sex classes follows that observed for CHIKV. The sex ratio of reported CHIKV cases in MG is 0.49 (33% of cases are male) (fig. S4). Under model M2, predicted incidence is highest in females aged >30. The discrepancy between the observed distribution and that predicted under the two urban cycle models indicates that the YF epidemic in MG is dominated by sylvatic transmission. This method shows that age- and sex-structured epidemiological data can be used to qualitatively evaluate the mode of YFV transmission during an outbreak.

Genomic surveillance of the Brazilian YFV outbreak

During a YF outbreak, it is important to undertake virological surveillance to (i) track epidemic origins and transmission hotspots, (ii) characterize genetic diversity to aid molecular diagnostics, (iii) detect viral mutations associated with disease severity, and (iv) exclude the possibility that human cases are caused by vaccine reversion. We generated 62 complete YF genomes from infected humans (n = 33) and NHPs (n = 29) from the most affected Brazilian states, including MG (n = 51), Espírito Santo (n = 8), Rio de Janeiro (n = 2), and Bahia (n = 1) (Fig. 3 and table S3). We also report two genomes from samples collected in 2003 during a previous YFV outbreak in MG from 2002 to 2003 (*31*). Genomes were generated in Brazil using a combination of methods (tables S5 to S7); half were generated in MG using a MinION portable YFV sequencing protocol adapted from (*32*) (tables S4 and S5). This protocol was made publicly available in May 2017 after the completion of pilot sequencing experiments using a cultured vaccine strain (supplementary materials). Median genome coverages were similar for samples obtained from NHPs [99%; median cycle threshold value (Ct) = 11] and from human cases (99%; median Ct = 16) (tables S5 to S7).

To put the newly sequenced YFV genomes in a global context, we added our genomes to a pool of 61 publicly available genomes (*33*, *34*). We developed and applied an automated online phylogenetic tool to identify and classify YFV gene sequences (also publicly available, see supplementary materials). Phylogenies estimated by

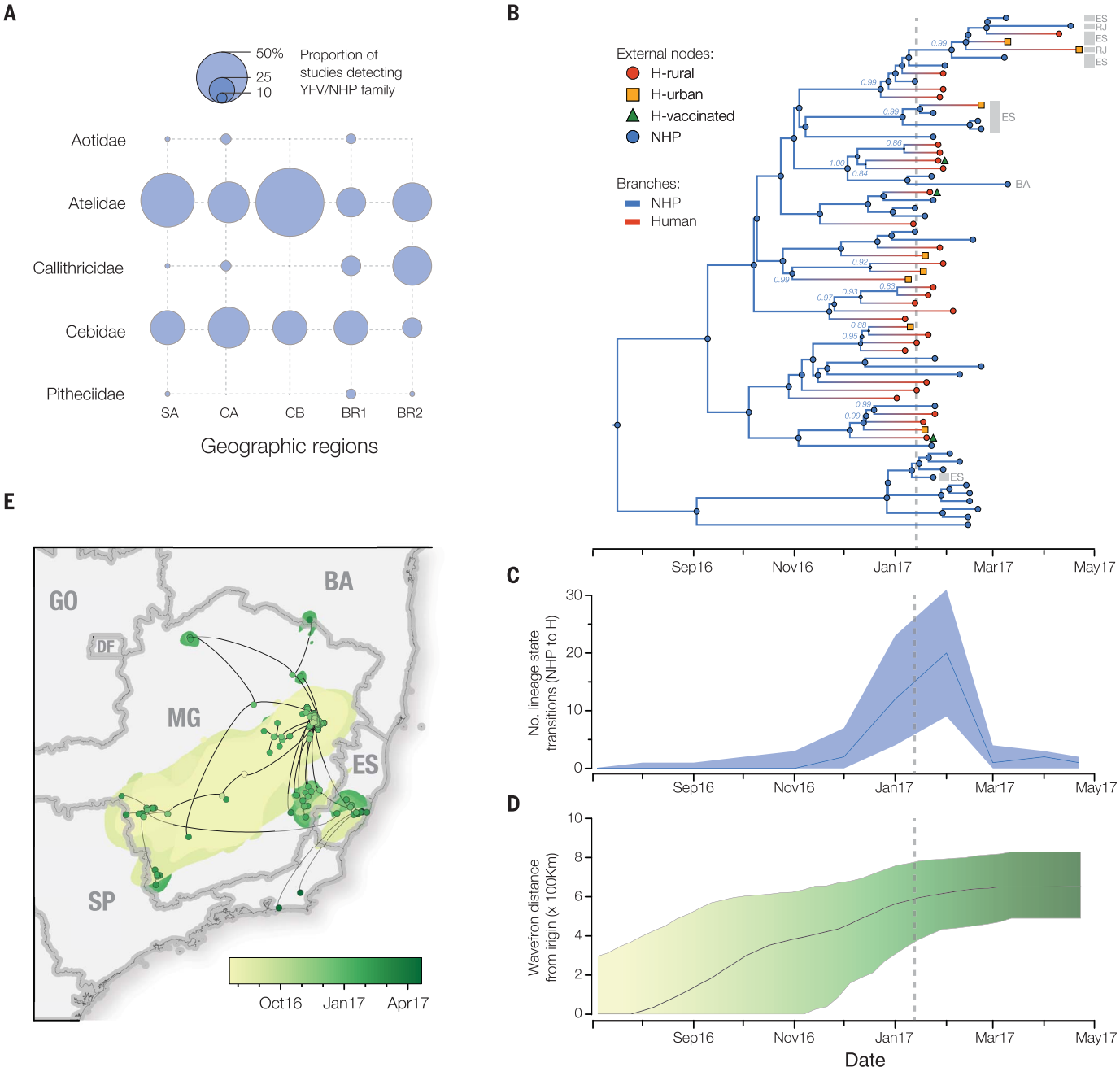


Fig. 4. Spatial and evolutionary dynamics of YFV outbreak. (**A**) Frequency of detection of YFV in NHPs in the Americas (*50*). Circle sizes represent the proportion of published studies (n = 15) that have detected YFV in each primate family and region. SA, South America (except Brazil); CA, Central America; CB, Caribbean; BR1, Brazil (before 2017); BR2, Brazil (this study). (**B**) Maximum clade credibility phylogeny inferred under a two-state (human and NHP) structured coalescent model. External node symbols denote sample type. Gray bars and labels indicate sample location (RJ, Rio de Janeiro; ES, Espírito Santo; BA, Bahia; others were sampled in MG). Internal nodes whose posterior state probabilities are >0.8 are annotated by circles. Node labels indicate posterior state probabilities for selected nodes. Internal branches are blue for NHPs and red for humans. Figure S8 shows a fully annotated tree. (**C**) Average number of YFV phylogenetic state transitions (from NHPs to humans) per month. Solid line, median estimate; shaded area, 95% BCI. (**D**) Expansion of the YFV epidemic wavefront estimated using a continuous phylogeographic approach (*35*). At each time point the plot shows the maximum spatial distance between phylogeny branches and the inferred location of outbreak origin. Solid line, median estimate; shaded area, 95% BCI. The dashed lines in (B) to (D) indicate when YF was declared a public health emergency in MG (13 January 2017). (**E**) Reconstructed spatiotemporal diffusion of the YFV outbreak. Phylogeny branches are arranged in space according the locations of phylogeny nodes (circles). Locations of external nodes are known, whereas those of internal nodes are inferred (*44*). DF, Distrito Federal; GO, Goiás; SP, São Paulo. Shaded regions represent 95% credible regions of internal nodes. Nodes and uncertainty regions are colored according to time.

this tool, along with maximum likelihood and Bayesian methods, consistently place the Brazilian outbreak strains in a single clade within the South America I (SA1) genotype with maximum statistical support (bootstrap = 100%; posterior probability > 0.99) (Fig. 3A and fig. S5).

The outgroup to the outbreak clade is strain BeH655417, a human case sampled in Alto Alegre, Roraima, north Brazil, in 2002. In contrast, isolates sampled during the previous outbreak in MG in 2003 are more distantly related to the outbreak clade within the SA1 genotype (Fig. 3A). Thus, the 2017 outbreak was more likely caused by a YFV strain introduced from an endemic area, possibly northern or center-west Brazil (*35*), than by the reemergence of a lineage that had persisted in MG. Although low sampling densities mean that this conclusion is provisional, similar scenarios have been suggested for previous Brazilian epizootics (*36*). The 14-year gap between the current outbreak and the date of the most closely related nonoutbreak strain agrees with the reported periodicity of YF outbreaks in northern Brazil (*37*), thought to be dictated by vector abundance and the accumulation of susceptible NHP hosts (*19*, *38*).

At least seven humans from MG with PCR-confirmed YFV received a YF vaccine before the onset of symptoms. To test that these occurrences were caused by natural infection, and not by vaccine reactivation, we sequenced the YFV genomes from three of these cases (Fig. 3A and table S3). Our phylogenetic analysis clearly shows that these represent natural infections caused by the ongoing outbreak and are conclusively not derived from the 17DD vaccine strain (which belongs to the West African YFV genotype) (Fig. 3A and fig. S6).

Unifying YFV epidemiology and molecular evolution

Virus genomes are a valuable source of information about epidemic dynamics (*39*) but are rarely used to investigate specific YFV outbreaks in detail. Here we show how a suite of three analytical approaches, which combine genetic, epidemiological, and spatial data, can provide insights into YFV transmission.

First, we used a Bayesian method (*40*) to explore potential covariates of fluctuations in the effective population size of the YFV outbreak in 2017. After finding that genetic divergence in the outbreak clade accumulates over the time scale of sampling (Fig. 3B and fig. S6), albeit weakly, we sought to determine which epidemiological time series best describe trends in inferred YFV effective population size. We found that effective population size fluctuations of the YFV outbreak are well explained by the dynamics of both human and NHP YFV cases (inclusion probability: 0.37 for human cases and 0.63 for NHP cases) (table S8). These two YFV time series explain the genetic diversity dynamics of the ongoing outbreak 10^3 times more effectively than CHIKV incidence (inclusion probability <0.001), which represents transmission by *Aedes* spp. vectors. One benefit of this approach is that epidemiological data contribute to estimation of the outbreak time scale. By incorporating YFV incidence data into evolutionary inference, we estimate the time of the most recent common ancestor (TMRCA) of the outbreak clade to be late July 2016 [95% Bayesian credible interval (BCI): March to November 2016] (Fig. 3C and fig. S7), consistent with the date of the first PCR-confirmed case of YFV in a NHP in MG (Jul 2016). The uncertainty around the TMRCA estimate is reduced by 30% when epidemiological and genomic data are combined, compared with genetic data alone (Fig. 3C).

Second, to better understand YFV transmission between humans and NHPs, we measured the movement of YFV lineages between the NHP reservoir and humans, using a phylogenetic structured coalescent model (*41*). Although previous studies have confirmed that YFV is circulating in five neotropical NHP families (Aotidae, Atelidae, Callitrichidae, Pitheciidae, and Cebidae) (Fig. 4A), thus far NHP YFV genomes during the 2017 outbreak have been recovered only from *Alouatta* spp. (family Cebidae) (*33*). In this analysis, we used the TMRCA estimate obtained above (Fig. 3C) to inform the phylogenetic time scale (Fig. 4B). All internal nodes in the outbreak phylogeny whose host state is well supported (posterior probability >0.8) are inferred to belong to the NHP population, consistent with an absence of urban transmission and in agreement with the large number of NHP cases reported in southeast Brazil (*20*). Despite this, we caution that hypotheses of human-to-human transmission linkage should not be tested directly using phylogenetic data alone, owing to the large undersampling of NHP infections. Notably, the structured coalescent approach reveals substantial changes in the frequency of NHP-to-human host transitions through time, rising from zero around November 2016 and peaking in February 2017 (Fig. 4C). This phylogenetic trend matches the time series of confirmed YFV cases in MG (Fig. 1, A and B), demonstrating that viral genomes, when analyzed using appropriate models, can be used to quantitatively track the dynamics of zoonosis during the course of an outbreak (*42*).

Third, we used a phylogenetic relaxed random walk approach to measure the outbreak's spatial spread (*43*) (supplementary materials and methods and table S9). When projected through space and time (Fig. 4, D and E, and movie S1), the phylogeny shows a southerly dissemination of virus lineages from their inferred origin in MG toward densely populated areas, including Rio de Janeiro and São Paulo (where YF vaccination was not recommended until July 2017 and January 2018, respectively). We estimate that virus lineages move, on average, 4.25 km/day (95% BCI: 2.64 to 10.76 km/day) (*44*). This velocity is similar when human YFV terminal branches are removed (5.3 km/day) and therefore most likely reflects YFV lineage movement within the sylvatic cycle and not the movement of asymptomatic infected humans. These rates are higher than expected given the distances typically travelled by NHPs in the region (*45*) and suggest the possibility that YFV lineage movement may have been aided by human activity—e.g., transport of infected mosquitoes in vehicles (*46*) or hunting or illegal trade of NHPs in the Atlantic forest (*47*, *48*). The epidemic wavefront (maximum distance of phylogeny branches from the inferred epidemic origin) expanded steadily between August 2016 and February 2017 at an estimated rate of ~3.3 km/day. Therefore, by the time YF was declared a public health emergency in MG (13 January 2017; dashed lines in Fig. 4, B to D), the epidemic had already expanded ~600 km (Fig. 4D) and caused >100 reported cases in both humans and NHPs (Fig. 1). Notably, the first detection in humans in December 2016 was concomitant with the outbreak's spatial expansion phase (Fig. 4D) and the rise in estimated NHP-to-human zoonoses (Fig. 4C); both were likely driven by an increase in the abundance of sylvatic vectors. Thus, the outbreak lineage appeared to circulate among NHPs in a widening geographic area for several months before human cases were detected.

Conclusion

Epidemiological and genomic surveillance of human and animal populations at risk is crucial for early detection and rapid containment of YFV transmission. The YFV epidemic in Brazil continues to unfold with an increase in cases since December 2017. Longitudinal studies of NHPs are needed to understand how YFV lineages disseminate across South America between outbreaks and how epizootics are determined by the dynamics of susceptible animals in the reservoir. To achieve the WHO's goal to eliminate YF epidemics by 2026, YF surveillance necessitates a global, coordinated strategy. Our results and analyses show that rapid genomic surveillance of YFV, when integrated with epidemiological and spatial data, could help anticipate the risk of human YFV exposure through space and time and monitor the likelihood of sylvatic versus urban transmission. We hope that the toolkit introduced here will prove useful in guiding YF control in a resource-efficient manner.

REFERENCES AND NOTES

1. T. Garske *et al.*, *PLOS Med.* **11**, e1001638 (2014).
2. C. I. Paules, A. S. Fauci, *N. Engl. J. Med.* **376**, 1397–1399 (2017).
3. M. Theiler, H. H. Smith, *J. Exp. Med.* **65**, 767–786 (1937).
4. F. M. Shearer *et al.*, *Lancet Infect. Dis.* **17**, 1209–1217 (2017).
5. M. R. Nunes *et al.*, *J. Virol.* **86**, 13263–13271 (2012).
6. J. P. Mutebi, H. Wang, L. Li, J. E. Bryant, A. D. T. Barrett, *J. Virol.* **75**, 6999–7008 (2001).
7. J. J. von Lindern *et al.*, *J. Gen. Virol.* **87**, 895–907 (2006).
8. E. Wang *et al.*, *Virology* **225**, 274–281 (1996).
9. G. J. Chang, B. C. Cropp, R. M. Kinney, D. W. Trent, D. J. Gubler, *J. Virol.* **69**, 5773–5780 (1995).
10. M. N. O. Segura, F. C. C. Castro, *Atlas de Culicídeos na Amazônia Brasileira* (Instituto Evandro Chagas Press, 2007), vol. 167.
11. J. da C. Cardoso *et al.*, *Emerg. Infect. Dis.* **16**, 1918–1924 (2010).
12. A. A. Grobbelaar *et al.*, *Emerg. Infect. Dis.* **22**, 1854–1855 (2016).
13. M. U. G. Kraemer *et al.*, *Lancet Infect. Dis.* **17**, 330–338 (2017).
14. O. Franco, *História da Febre-Amarela no Brasil* (Ministério da Saúde, DNERU, 1969).

15. R. A. G. B. Consoli, R. Lourenço de Oliveira, *Principais Mosquitos de Importância Sanitária no Brasil* (FioCruz, 1994).
16. J. Vainio, "Yellow fever" (World Health Organization, 1998).
17. M. U. Kraemer *et al*., *eLife* **4**, e08347 (2015).
18. Secretaria de Vigilância em Saúde, Ministério da Saúde, "Emergência epidemiológica de febre amarela no Brasil, no período de Dezembro de 2016 a Julho de 2017" (Boletim Epidemiológico vol. 48, 2017); http://portalarquivos2.saude.gov.br/images/pdf/2017/setembro/06/2017_027.pdf.
19. P. F. Vasconcelos *et al*., *J. Med. Virol.* **65**, 598–604 (2001).
20. Pan American Health Organization/World Health Organization, "Epidemiological update yellow fever 16 Feb 2018" (PAHO/WHO, 2018); www.paho.org/hq/index.php?option=com_topics&view=readall&cid=2194&Itemid=40784&lang=en.
21. F. M. Shearer *et al*., *Lancet Glob. Health* **6**, e270–e278 (2018).
22. Instituto Evandro Chagas, www.iec.gov.br/portal/descoberta/ (2018).
23. R. Lourenço-de-Oliveira, M. Vazeille, A. M. B. de Filippis, A. B. Failloux, *Trans. R. Soc. Trop. Med. Hyg.* **98**, 43–54 (2004).
24. D. Couto-Lima *et al*., *Sci. Rep.* **7**, 4848 (2017).
25. N. R. Faria *et al*., *PLOS Currents Outbreaks*, 10.1371/currents.outbreaks.c97507e3e48efb946401755d468c28b2 (2016).
26. N. R. Faria *et al*., *Sci. Rep.* **7**, 15216 (2017).
27. S. H. Tuboi, Z. G. A. Costa, P. F. da Costa Vasconcelos, D. Hatch, *Trans. R. Soc. Trop. Med. Hyg.* **101**, 169–175 (2007).
28. PAHO, "Outbreak of Yellow Fever in Paraguay" (PAHO, 2009).
29. A. Nasidi *et al*., *Trans. R. Soc. Trop. Med. Hyg.* **83**, 401–406 (1989).
30. S. D. Thiberville *et al*., *PLOS Negl. Trop. Dis.* **7**, e2004 (2013).
31. M. Ribeiro, C. M. F. Antunes, *Rev. Soc. Bras. Med. Trop.* **42**, 523–531 (2009).
32. J. Quick *et al*., *Nat. Protoc.* **12**, 1261–1276 (2017).
33. M. C. Bonaldo *et al*., *Mem. Inst. Oswaldo Cruz* **112**, 447–451 (2017).
34. A. Moreira-Soto *et al*., *Clin. Microbiol. Infect.* 10.1016/j.cmi.2018.01.026 (2018).
35. Secretaria de Vigilância em Saúde, Ministério da Saúde, "Reemergência da Febre Amarela Silvestre no Brasil, 2014/2015: Situação epidemiológica e a importância da vacinação preventiva e da vigilância intensificada no período sazonal" (Boletim Epidemiologico vol. 46, 2015); http://portalarquivos2.saude.gov.br/images/pdf/2015/outubro/19/2015-032—FA-ok.pdf.
36. P. F. C. Vasconcelos *et al*., *Emerg. Infect. Dis.* **10**, 1578–1584 (2004).
37. F. P. Câmara, A. L. B. B. Gomes, L. M. F. Carvalho, L. G. V. Castello, *Rev. Soc. Bras. Med. Trop.* **44**, 297–299 (2011).
38. P. F. Vasconcelos *et al*., *Cad. Saude Publica* **17** (suppl.), 155–164 (2001).
39. G. Dudas *et al*., *Nature* **544**, 309–315 (2017).
40. M. S. Gill, P. Lemey, S. N. Bennett, R. Biek, M. A. Suchard, *Syst. Biol.* **65**, 1041–1056 (2016).
41. T. G. Vaughan, D. Kühnert, A. Popinga, D. Welch, A. J. Drummond, *Bioinformatics* **30**, 2272–2279 (2014).
42. G. Dudas, L. M. Carvalho, A. Rambaut, T. Bedford, *eLife* **7**, e31257 (2018).
43. O. G. Pybus *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15066–15071 (2012).
44. P. Lemey, A. Rambaut, J. J. Welch, M. A. Suchard, *Mol. Biol. Evol.* **27**, 1877–1885 (2010).
45. L. Jung, I. Mourthe, C. E. V. Grelle, K. B. Strier, J. P. Boubli, *PLOS ONE* **10**, e0129789 (2015).
46. E. Flacio, L. Engeler, M. Tonolla, P. Müller, *Parasit. Vectors* **9**, 304 (2016).
47. A. Estrada *et al*., *Sci. Adv.* **3**, e1600946 (2017).
48. R. A. Nascimento, R. A. M. Montano, *Neotrop. Biol. Conserv.* **8**, 79–87 (2013).
49. A. J. Auguste *et al*., *Emerg. Infect. Dis.* **21**, 99–102 (2015).
50. P. R. Stephens *et al*., *Ecology* **98**, 1476 (2017).

ACKNOWLEDGMENTS

We thank FUNED-MG and the Brazilian YFV surveillance network for their contributions. N.R.F. thanks J. F. Drexler for sharing data, N. Trovão for discussions, F. Campos for proof imaging, and P. Fonseca for proofreading. We thank Oxford Nanopore Technologies for technical support. L.C.J.A. thanks QIAGEN for reagents and equipment. A.C.d.C. and E.C.S. thank Illumina, Zymo Research, Sage Science, and Promega for donation of reagents. **Funding:** This work was supported in part by CNPq 400354/2016-0 and FAPESP 2016/01735-2. N.R.F. is supported by a Sir Henry Dale Fellowship (204311/Z/16/Z), internal GCRF grant 005073, and John Fell Research Fund grant 005166. This research received funding from the ERC (grant agreement 614725-PATHPHYLODYN) and from the Oxford Martin School. M.U.G.K. acknowledges funding from a Branco Weiss Fellowship, administered by ETH Zurich, a Training Grant from the National Institute of Child Health and Human Development (T32HD040128), and the National Library of Medicine of the National Institutes of Health (R01LM010812 and R01LM011965). S.D. is funded by the Fonds Wetenschappelijk Onderzoek (FWO, Flanders, Belgium). G.B. acknowledges support from the Interne Fondsen KU Leuven/Internal Funds KU Leuven. A.C.d.C. is funded by FAPESP 2017/00021-9. B.B.N. and S.C. are supported by the EU's Horizon 2020 Programme through ZIKAlliance (grant 734548), the Investissement d'Avenir program, the Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases program (grant ANR-10-LABX-62-IBEID), the Models of Infectious Disease Agent Study of the National Institute of General Medical Sciences, the AXA Research Fund, and the Association Robert Debré. P.L. and M.A.S. acknowledge funding from the European Research Council (grant agreement 725422-ReservoirDOCS) and from the Wellcome Trust Collaborative Award 206298/Z/17/Z. P.L. acknowledges support from the Research Foundation, Flanders (Fonds voor Wetenschappelijk Onderzoek, Vlaanderen, G066215N, G0D5117N, and G0B9317N). **Author contributions:** N.R.F., L.C.J.A., S.C.H., A.M.B.F., M.U.G.K., S.C., and O.G.P. designed the study. S.C.H., J.J.G., R.S.d.A., F.C.M.I., J.X., R.D.O.C., J.T., M.G., L.C.J.A., and N.R.F. undertook fieldwork. S.C.H., J.Q., J.J.G., A.C.d.C., S.C.V.K., V.F., and T.d.O. undertook experiments. N.R.F., L.d.P., J.T., S.D., G.B., O.G.P., C.-H.W., T.I.V., and P.L. performed genetic analyses. M.U.G.K., S.C., S.F., J.L., U.O., L.A., D.Y., and N.R.F. performed epidemiological and cartographic analyses. B.N., F.M.S., and N.R.F. performed historical YFV review. N.R.F., M.U.G.K., L.C.J.A., S.C., and O.G.P. wrote the manuscript. E.C.S., J.T., L.d.P., R.P.S., P.L., C.F.C.d.A., R.S.d.A., and A.M.B.F. edited the manuscript. All other authors were involved in collection, processing, sequencing, and bioinformatics of samples and geographic data. All authors read and approved the contents of the manuscript. **Competing interests:** N.J.L. and L.C.J.A. received free-of-charge reagents in support of the project from Oxford Nanopore Technologies. A.C.d.C. and E.C.S. received reagents at no cost from Illumina, Zymo Research, Sage Science, and Promega. **Data and materials availability:** Raw data, code, and analysis files are available via the GitHub repository (https://github.com/arbospread/YFV-monitoring). See https://github.com/zibraproject/zika-pipeline/tree/master/schemes for MinION sequencing protocols. Genome sequences generated here are available under GenBank accession numbers MH018064 to MH018115 and MH484423 to MH484434.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/894/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S10
Tables S1 to S9
References (*51*–*107*)
Movie S1

31 March 2018; accepted 20 July 2018
Published online 23 August 2018
10.1126/science.aat7115

REPORT

INDUCED SEISMICITY

The spatial footprint of injection wells in a global compilation of induced earthquake sequences

Thomas H. W. Goebel*† and Emily E. Brodsky†

Fluid injection can cause extensive earthquake activity, sometimes at unexpectedly large distances. Appropriately mitigating associated seismic hazards requires a better understanding of the zone of influence of injection. We analyze spatial seismicity decay in a global dataset of 18 induced cases with clear association between isolated wells and earthquakes. We distinguish two populations. The first is characterized by near-well seismicity density plateaus and abrupt decay, dominated by square-root space-time migration and pressure diffusion. Injection at these sites occurs within the crystalline basement. The second population exhibits larger spatial footprints and magnitudes, as well as a power law–like, steady spatial decay over more than 10 kilometers, potentially caused by poroelastic effects. Far-reaching spatial effects during injection may increase event magnitudes and seismic hazard beyond expectations based on purely pressure-driven seismicity.

Human-induced seismicity close to geothermal, hydraulic fracturing, and wastewater disposal wells is a source of substantial seismic hazard. Such injection activity has led to many moderate-magnitude earthquakes and an exceptional increase in earthquake rates in parts of North America and Central Europe after ~2006 (*1*–*3*). The hazard from injection-induced earthquakes is particularly difficult to manage because the earthquakes frequently occur at large distances (>10 km) from the targeted injection zones (*4*–*6*).

A better understanding of the driving mechanisms of injection-induced seismicity is vital for improving seismic hazard assessment and mitigation. Traditionally, such hazard assessment has concentrated on pore-pressure increase in a volume hydraulically connected to the injection wells (*7*, *8*). Pore-pressure increase, which we here call the direct pressure effect, is thought to reduce the normal load on locked faults, thereby allowing sliding to occur.

As induced seismicity cases and studies have proliferated, the importance of additional mechanisms such as elastic and fully coupled poroelastic stresses have become increasingly clear (*9*–*11*). The pore-pressure increase in the injection zone is expected to load the surrounding rock matrix and result in a fully coupled fluid-solid stress field. Elasticity is an effective means of transmitting forces to great distances, and therefore the fully coupled poroelastic stress field can extend well beyond the fluid pressure increase in the hydraulically connected region. For example, large-scale, field-wide injection can perturb faults and induce earthquakes more than 30 km away (*11*). In controlled injection experiments that resolve coupled aseismic and seismic processes during fluid injection, induced earthquakes are absent within the pressurized zone but occur as a result of elastic stress changes in the surrounding rock volume (*12*). These observations require a reexamination of the controls on induced seismicity and suggest that the spatial reach of the earthquakes may be a useful discriminant of triggering mechanisms. The observations also suggest that induced earthquakes may extend farther than previously thought [i.e., at least to 30 km from wells (*4*, *11*)], and empirically measuring the spatial extent is necessary to develop appropriate mitigation and regulatory approaches.

Here we examine the distance of induced earthquakes from injection wells to understand the mechanical controls on the spatial extent of induced seismicity. We start by quantifying the shape of spatial decay and find two groups with distinct decay patterns. We then analyze migratory behavior within each group, as well as the relationship between operational parameters and the resultant spatial pattern of earthquakes. We examine differences in observed maximum magnitudes, which appear to be linked to the type of spatial pattern. Lastly, we discuss possible physical mechanisms that control the distinct behavior within the two groups, with a focus on potential differences in poroelastic properties between sedimentary and basement units.

We compiled measurements of the spatial decay of induced earthquakes from causative wells worldwide (*13*–*23*). We focused on isolated injection sites with well-recorded induced sequences, which are clearly connected to a single geothermal, scientific, or waste-fluid injection well (fig. S1; section S2 describes two exceptions with more than one well). This strategy eliminates some major regions, such as Oklahoma, where effects of field-wide injection are difficult to unravel. All of the data used are in the public domain (*24*). The initial 27 selected injection sites are predominantly located in intraplate regions with low seismic activity throughout the United States, central Europe, and Australia (Fig. 1A). We included one well from the northwestern edge of the Geysers geothermal field because of its isolated location and several-kilometer distance to other active injection wells during the analyzed time period from 2007 to 2009 (*16*).

Before determining the spatial seismicity decay from wells, we assessed the quality of the seismic record, focusing on spatial and magnitude information. We first selected events above the magnitude of completeness (M_c), determined by minimizing the misfit between the observed distribution and the Gutenberg-Richter relation (fig. S2). The datasets for Soultz-Sous-Foretz, France, and Fenton Hill, New Mexico, do not include magnitude information. For the Soultz site, magnitude estimates are available from a surface array for a subset of events, showing consistent results between borehole and magnitude-corrected surface catalogs. For the Fenton-Hill site, the entire seismic record was used.

In a second quality-control step, we tested whether the observed distributions exhibit significant deviations from random Gaussian location uncertainty and show significant spatial clustering close to wells at rates above the background activity, using a two-sided Kolmogorov-Smirnov test (fig. S4). The quality-control steps eliminated nine induced cases, leaving 18 sequences for further analysis.

For the high-quality sequences, we computed two-dimensional (2D) distances between wells and earthquakes at the average depth of the injection interval, taking into account the well trajectories. Relative horizontal location uncertainties of seismic events are on the order of tens of meters, whereas absolute location uncertainty ranges from 100 to 500 m (section S4), and average vertical location errors are more than 1 km. Because the vertical uncertainties are large, we initially focused on 2D distances and later compared the results with 3D distances. Seismicity distance fall-off from wells was calculated from areal densities determined by nearest-neighbor binning of the k closest events with a moving window of $k/2$, resulting in density estimates, $\rho = k/\Delta r$, where Δr is the distance between the first and kth event.

The seismicity density distributions can generally be described by two types of spatial decay: (i) sequences with an extended plateau close to the well, followed by an abrupt decay within less than 1 km, and (ii) sequences with steady, power law–like decay out to distances of more than 10 km (Fig. 1B). The shapes of the two types

Department of Earth and Planetary Sciences, University of California, Santa Cruz, CA, USA.
*Corresponding author. Email: tgoebel@ucsc.edu
†These authors contributed equally to this work.

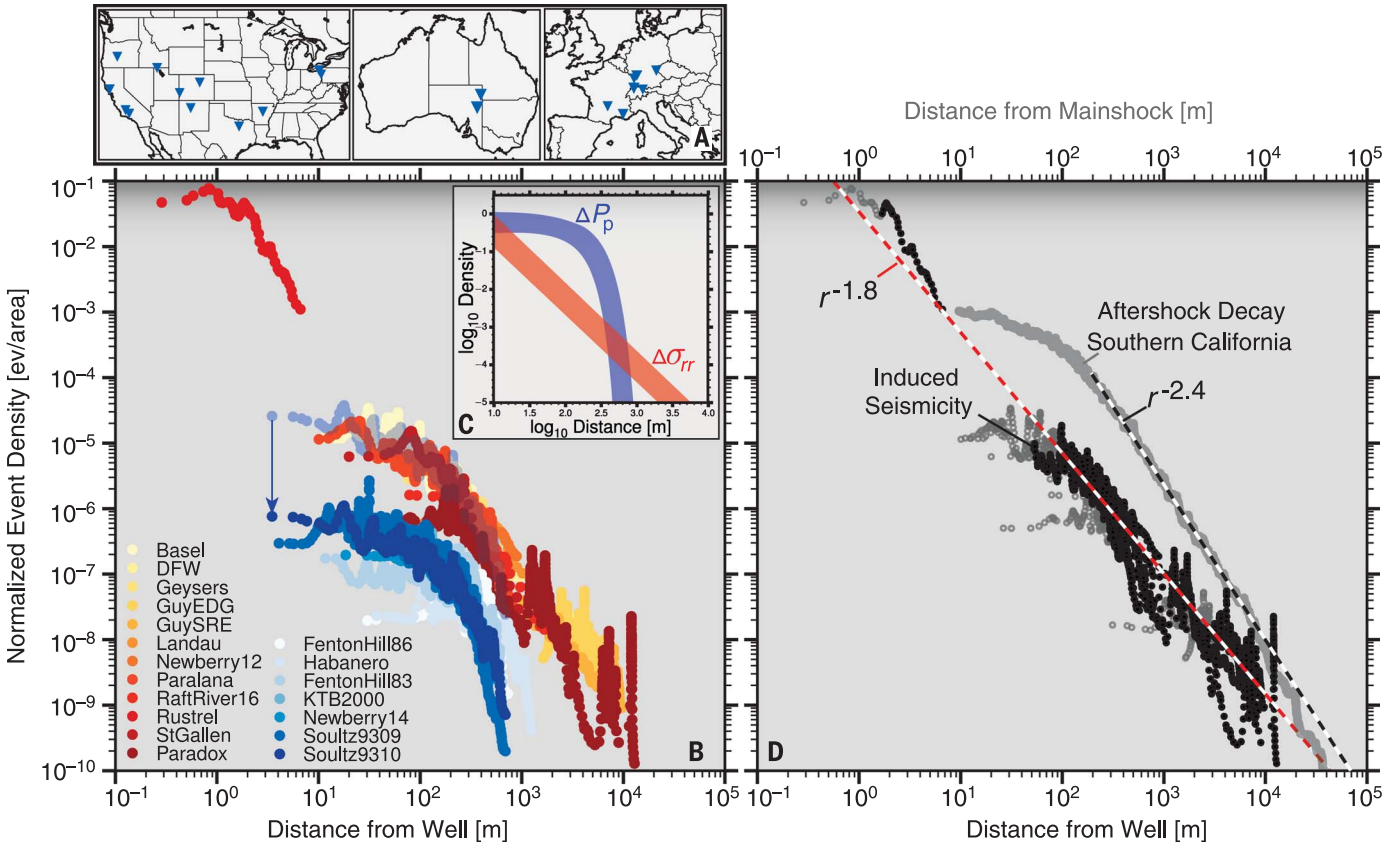


Fig. 1. The spatial decay of induced sequences can be classified as abrupt or steady. (**A**) Map of the injection sites (blue triangles) in the United States, Australia, and Europe (fig. S1). (**B**) Seismic density of all studied induced sequences, normalized by number of events above completeness. We show cases with abrupt decay in shades of blue and steady decay in shades of red. We shifted abrupt decays vertically (translucent blue dots show one example of the original vertical position; the blue arrow shows the shift) to allow for easier visualization. (**C**) Theoretical expectation of spatial density fall-off using Eq. 2; the blue curve shows the abrupt decay of pressure-dominated sites, and the red curve shows a power law decay for sites with strong poroelastic coupling. (**D**) Merged densities above r_c for sequences with steady decay (black markers) and power law fit (red dashed line) with ~$r^{-1.8}$, which is more gradual than spatial aftershock decay in California (gray markers and black dashed line).

of spatial decay remain consistent across a range of spatial scales and durations of injection operations, spanning hours (e.g., Rustrel) to years (e.g., Paradox), both for the complete datasets and for the subsets of events from individual sites (fig. S8). The consistency in the shapes of spatial decay suggests an underlying time-invariant process. We distinguished the two types of distance fall-offs quantitatively by fitting with the functional form

$$\rho = \rho_0 \frac{1}{\left[1 + (r/r_c)^{2\gamma}\right]^{1/2}} \quad (1)$$

where ρ is seismic density, ρ_0 describes the short-distance density plateau, r_c is the corner distance, and γ is an exponent describing the abruptness of the decay at larger distances. We chose this functional form to capture both differences in near-well density plateaus and the abruptness of distance fall-offs. We determined the three free parameters by fixing ρ_0 using the average density of the first five sample points and then inverted for r_c and γ using a maximum likelihood approach that assumes Poissonian uncertainty in each distance bin (*24*). The distribution of the resulting decay exponents exhibits a natural break between γ = 3.1 and 4.3, separating sequences into steady decay with γ = 1.5 to 3.1 and abrupt decay with γ = 4.3 to 5.9 (fig. S5). The difference between these two populations exceeds the maximum expected uncertainty of ±0.34 based on the 95% confidence intervals of the maximum likelihood fits. In addition to having larger γ values than steady-decay sites, sites with abrupt decay are characterized by corner distances that are substantially closer to the maximum extent of the sequence (fig. S6).

Several observations indicate that the result is not an artifact of specific instrumental, statistical, or measurement practices: (i) The maximum distance of the earthquakes is smaller than array aperture in all cases, so that the spatial decay is not a product of limited array extent, which may truncate the data. We further tested the influence of array geometry by computing M_c as a function of distance from wells and found no systematic bias. (ii) Where a single site allows for comparison between instrumentation (i.e., surface versus borehole arrays), the results are identical. Specifically, the spatial decay at the Soultz site shows abrupt fall-offs in both the borehole and the wide-aperture surface array catalogs when corrected for M_c. (iii) In addition to fitting Eq. 1, the two types of decay can be identified by using pure power laws. A power law is well fit over a large distance range for sites with steady decay, whereas abrupt-decay sites show power law–like behavior only over a limited range of distances that are near the maximum distance of the datasets (fig. S6). (iv) Lastly, stacking seismicity density estimates on the basis of 2D or 3D distance leads to a visible separation of steady and abrupt decay (Fig. 1B and fig. S11).

To connect the spatial decay to perturbing stresses, we considered induced pore-pressure changes and poroelastic coupling. We expressed the observed seismicity fall-off within a Coulomb failure framework as the product of stress perturbation and the number of faults sufficiently close to failure (*25*)

$$\rho(r) \propto \Delta\sigma \times \frac{N_{fl}}{2\pi r \Delta r} \quad (2)$$

where ρ is seismicity density, Δσ is the induced

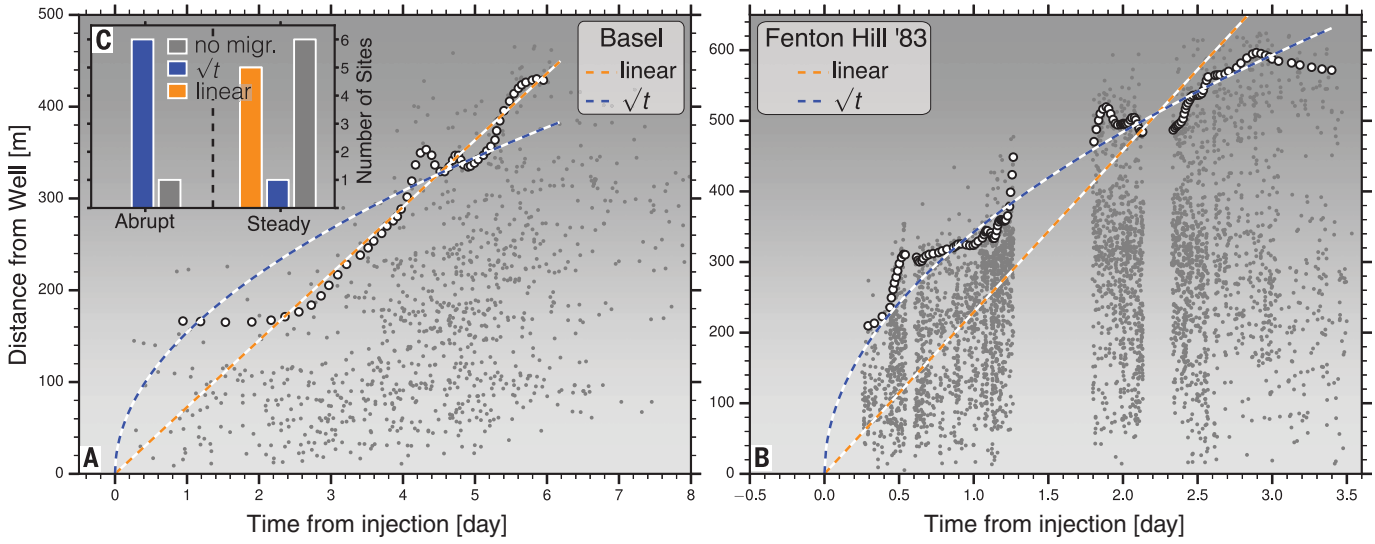


Fig. 2. Sequences with abrupt spatial decay are dominated by square-root migration, an indication of pressure diffusion. Examples of (**A**) linear and (**B**) square-root migration. Gray markers show event time and distance from injection; white markers show the 95th percentile of distances in specific time bins. (**C**) Histogram of the number of abrupt- and steady-decay sites with square-root (blue), linear (orange), or no (gray) migration.

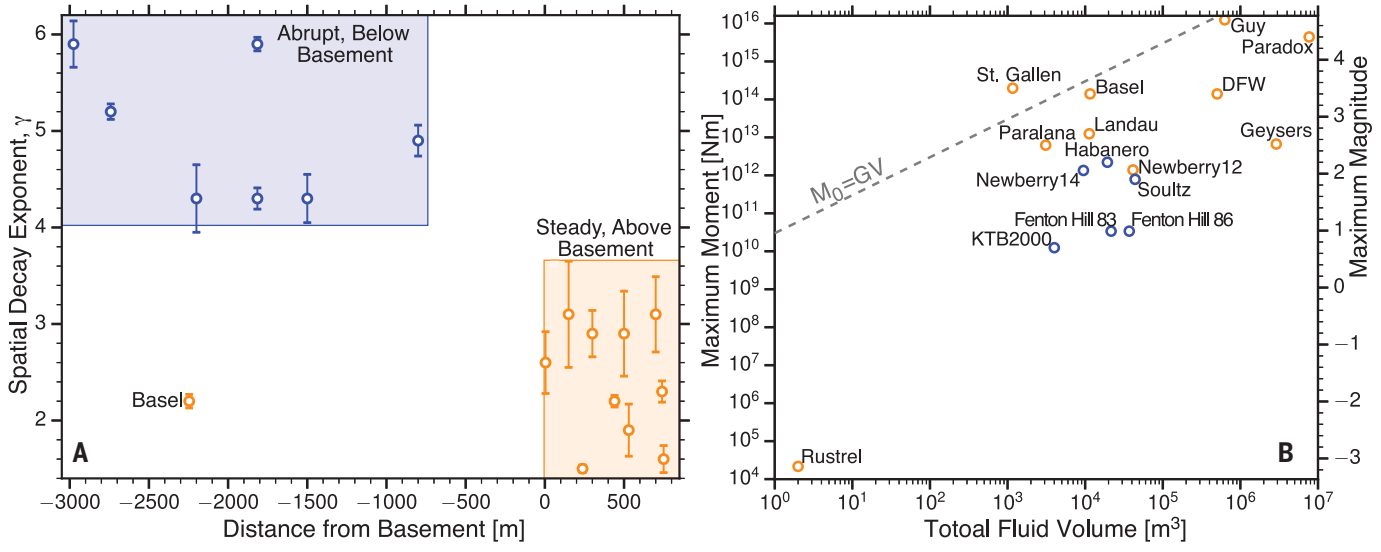


Fig. 3. Above-basement injection commonly results in larger spatial footprints and a higher probability of inducing larger-magnitude earthquakes. (**A**) The spatial decay, separated into abrupt (blue) and steady (orange), is controlled by the distance between injection and crystalline basement. (**B**) Maximum magnitude of each sequence as a function of total injected volume for steady (orange) and abrupt (blue) decays. The dashed line is the theoretically expected maximum moment based on (*35*), for which G is shear modulus and V is total injected volume. (For Fenton Hill, we report the largest recorded magnitude during all stimulations. No stimulation-specific fluid volume is available for Raft River.)

stress perturbation, and $N_{fl}/(2\pi r\Delta r)$ is the density of faults per area. To build intuition for the influence of changes in $\Delta\sigma$, we first assumed that the fault density is constant and investigated the distance fall-off expected for the direct pressure effect and poroelastic coupling. For the direct pressure case, we can write the distance fall-off of the first term in Eq. 2 in a vertically confined reservoir as $\Delta\sigma = \Delta P_p \sim W(u)$, for which ΔP_p is pore pressure, $W(u)$ is the exponential integral, and the argument $u = r^2/(4Dt)$; r is distance, t is time from injection, and D is hydraulic diffusivity (*26*). The corresponding distance fall-off is substantially faster than poroelastic stress decay in the far field of injection, which decays as $1/r^2$ (*26*). The expected distance decay based on pure pressure and poroelastic models approximately matches the shapes of observed density fall-offs for abrupt and steady decays (Fig. 1, B and C).

Although the similarity between the simple models and the observed spatial decay is compelling, additional processes are expected to contribute. Such processes include the coupling between pressure and permeability, as well as heterogeneity in the permeability structure in the presence of faults (*8*, *14*, *27*). In addition, elastic stress transfer and event-event interaction of both seismic and aseismic ruptures can increase the extent of induced seismicity sequences (*12*).

We further analyzed sequences with extended, power law–like decay by fitting the joined densities above corner distances (r_c) of individual sequences (Fig. 1D). The corresponding linear least-square fit of the log-transformed data has a value of $\gamma = 1.8$. This value is substantially smaller (i.e., more gradual) than the spatial decay of aftershocks from mainshock epicenters in California, for which $\gamma = 2.4$ (*25*) (Fig. 1D). The latter value was determined by identifying mainshock-aftershock clusters by using a nearest-neighbor

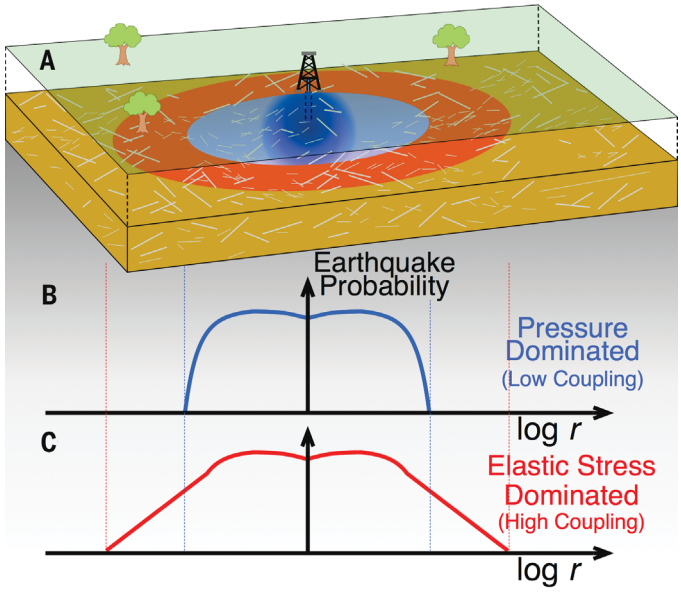


Fig. 4. The probability of inducing an earthquake at distance *r* is controlled by fault availability and amplitude of stress perturbation. (**A**) Schematic representation of injection operation, footprint of poroelastic response (blue and red ellipses), and fault network (gray lines). (**B**) Earthquake probability (in events per area) as a function of distance from the injection well for pressure-dominated triggering. (**C**) Same as (B) in a coupled system with elastic stress dominance in the far field. [Both the *x* and *y* axes in (B) and (C) are logarithmic.]

distance in a space-time-magnitude domain (*28*). This difference in γ exponents indicates that spatial seismicity decay contains information on the distinctive forcing stresses specific to injection-induced sequences. The stresses from a mainshock that produce aftershocks cannot explain the data, so we must invoke additional processes such as fluid migration and poroelasticity. The smaller γ value for induced sequences points to a more gradual spatial relaxation of the underlying stress field.

To test the influence of location uncertainty on the shape and extent of spatial decay, we performed Monte Carlo simulations of random power law–distributed data convolved with random normal distributions with standard deviations that correspond to the observed location uncertainty (section S4). Our simulations reveal that decay exponents can in principle be inflated owing to location uncertainty. However, our location uncertainties are small enough that the effect is negligible for this study (fig. S9). We can robustly identify sites with abrupt decay in the 2D data because both relative and absolute uncertainties are smaller than the determined r_c values (fig. S10).

We next investigated potentially systematic migratory behavior of seismicity in each sequence. Migration is particularly interesting because the most commonly invoked mechanism for induced seismicity—direct pore-pressure diffusion from a nearly instantaneous injection—has a well-understood square-root dependence of distance on time (*9*). Other cases worth considering are no migration, as might be expected for a rapidly applied elastic stress perturbation, and linear migration. These three models were fitted to the seismicity envelope using a maximum-likelihood method, and the preferred model was chosen on the basis of a Bayesian information criterion (Fig. 2 and section S3).

Square-root migration is most common at abruptly decaying sites, consistent with direct pore-pressure effects (Fig. 2). Cases with steady decay, on the other hand, show linear migration (five sites; Fig. 2C) or a lack of migration (six sites). The latter may be characteristic for poroelastic stress–dominated sequences during which earthquakes are triggered beyond the pressure-dominated region even shortly after injection starts, leading to a breakdown of square-root migration (*1*). Processes that may contribute to linear migration are related to elastic stress transfer, including event-event triggering, slow slip, and aseismic slip (*12*, *18*, *29*). Additional mechanisms that likely affect seismicity migration include fracture creation, changes in permeability structure, and thermally induced stresses. These mechanisms are expected to be most pronounced during geothermal injection and close to injection wells (*16*).

Distance decay and space-time migration highlight two distinct groups of induced seismicity sequences, consistent with direct pressure effects on the one hand and far-field elastic stresses on the other hand. To better understand the physical controls on this distinct behavior, we investigated operational and geologic parameters at all injection sites. We evaluated the ability of injection volume, average flow rate, peak well head pressure, injectivity (which is approximated by flow rate divided by well head pressure), injection depth, distance to basement (i.e., vertical distance to the crystalline basement rock), reservoir temperature, stress state, and injection duration to discriminate the two populations defined on the basis of spatial decay (fig. S12) (*24*). The most important governing parameter for the separation of abrupt and steady decay is distance to basement: All sites with abrupt decay were injecting below the upper basement surface, and sites with steady decay (except for Basel) were located in sedimentary rocks above the basement (Fig. 3).

This difference in behavior based on lithology at injection depth can potentially be explained by a difference in the ability of the rocks to transmit fluid stresses into the solid material. The Biot-Willis α coefficient (*26*) captures the poroelastic coupling so that larger α values correspond to more effective coupling

$$\alpha = 1 - \frac{K}{K_s} \qquad (3)$$

where K is the bulk modulus and K_s is the modulus of the solid, so that α is expected to be close to 0 for stiff, low-porosity rocks in the basement and close to 1 for soft sediments (*30*). Experimental results confirm that bulk strains in low-porosity rocks are less sensitive to fluid pressure changes (*31*). In other words, pressure perturbations within the larger, interconnected pore space in sedimentary rocks are much more efficient in changing bulk stresses than pressure changes in smaller, isolated pores in crystalline rock. Sedimentary units may also have higher permeability, which influences the spatial extent of injection but is less consequential for the observed separation into steady and abrupt decay.

A notable exception to the lithology-controlled seismic response is the Basel injection site. At this site, steady spatial decay of earthquakes up to local magnitude (M_L) 3.4 was triggered by injection into the crystalline basement. A potential explanation for this observation is injection directly into a highly fractured fault damage zone, as indicated by seismicity and focal mechanism analysis (*13*). Such damage zones likely exhibit higher α values, even within the crystalline basement, but will rarely be encountered unless specifically targeted during injection because of their hydrogeologic properties. The 1967 Rocky Mountain Arsenal sequence is arguably in a similar category given the reported extent of the seismicity, but it is not included in this analysis because of the data quality issues discussed earlier.

In addition to differences in local geology, we also observed variations in maximum event magnitudes between the two populations of induced earthquakes. Previous studies showed maximum magnitudes above *M*4 for earthquakes induced by hydraulic fracturing and above *M*5 for earthquakes induced by geothermal operations or fluid disposal (*1*, *2*, *15*, *32*). Peak magnitudes for a given injection volume are generally larger at sites with steady decay, which also produced the largest event in our dataset, with a magnitude of M_L 4.7 (Fig. 3B). The difference in maximum magnitudes is expected if more

extensive spatial footprints at sites with steady decay increase the probability of encountering larger faults (*33*). Even for injection into a single well, poroelastic stresses may promote the activation of such distant, large faults that are close to failure without requiring a direct hydraulic connection. This effect is further enhanced by closely spaced, high-rate injection wells, which act as a finite source with stress decay out to ~2 to 3 source dimensions (*11*).

The spatial decay of induced sequences can be expressed by the product of stress perturbation and available prestressed faults that fail because of this perturbation (Fig. 4). At sites with high poroelastic coupling, we expect elastic stresses to dominate earthquake triggering in the far field of injection, showing a scale-invariant spatial decay of $\sim r^{-2}$, whereas direct pressure effects dominate close to injection wells. If we assume that the term for fault availability in Eq. 2 is expressed by $N_{\mathrm{fl}}/(2\pi r\Delta r) \propto r^{d_f-D}$, where d_f is the fractal dimension of the fault network and D is the geometric dimension of the density measurement (here, $D = 2$), we find that $d_f = 2.2$. Thus, fault density increases gradually as $r^{0.2}$ with distance, so that for pressure-dominated sequences, the expected point of peak seismicity occurs at some distance from the well (Fig. 4B). The model proposed here captures large-scale, average trends in seismicity decay with distance from injection wells. Nevertheless, additional processes such as coupling between changes in pore pressure and permeability, as well as variations in regional fault network geometry, may also contribute to the functional form of the decay.

Our global compilation of fluid injection–induced seismicity allows for a better understanding of the maximum earthquake-triggering distance from an injection site. We can describe the shape of spatial seismicity decay as either abrupt or steady. Steady decay shows power law–like behavior to distances of more than 10 km and is more gradual than spatial aftershock decay. Sites with abrupt decay are dominated by square-root space-time migration, which is consistent with pressure diffusion. Abrupt decay is limited to sites where injection is within the crystalline basement, whereas steady decay primarily occurs above the basement. The maximum magnitude is larger for sites with steady decay owing to the greater probability of activating bigger faults within the extended spatial footprint of the injection wells.

Previous strategies to mitigate induced seismicity encouraged injecting in sedimentary units instead of directly into the basement (*34*). However, our results suggest that injection into sedimentary rocks leads to more distant and larger earthquakes for a given volume of injection, perhaps owing to more efficient pressure and stress transmission. The larger spatial footprints of above-basement injection may be responsible for the extensive seismogenic response in some areas, such as Alberta and Oklahoma (*1*, *4*). Nevertheless, injection into the basement also poses a source of seismic hazard. Large earthquakes, such as the 1967 Rocky Mountain Arsenal M_L 5.3 (*8*) or the recent South Korean M_w 5.4 events (*32*), can occur if fluid is directly injected into a basement fault that is either specifically targeted or encountered by chance. The key result from this analysis is that injection in sedimentary units immediately overlying basement rocks is more likely to encounter a large fault by chance because of the larger spatial footprint. Far-reaching poroelastic effects complicate the assessment of distance-based induced seismic hazard and should be included in mitigation strategies.

REFERENCES AND NOTES

1. X. Bao, D. W. Eaton, *Science* **354**, 1406–1409 (2016).
2. W. L. Yeck *et al*., *Geophys. Res. Lett.* **44**, 711–717 (2017).
3. W. L. Ellsworth, *Science* **341**, 1225942 (2013).
4. K. M. Keranen, M. Weingarten, G. A. Abers, B. A. Bekins, S. Ge, *Science* **345**, 448–451 (2014).
5. R. J. Skoumal, M. R. Brudzinski, B. S. Currie, J. Levy, *Earth Planet. Sci. Lett.* **405**, 274–280 (2014).
6. T. H. W. Goebel *et al*., *Geophys. Res. Lett.* **43**, 1092–1099 (2016).
7. C. B. Raleigh, J. H. Healy, J. D. Bredehoeft, *Science* **191**, 1230–1237 (1976).
8. P. A. Hsieh, J. D. Bredehoeft, *J. Geophys. Res.* **86**, 903–920 (1981).
9. P. Segall, S. Lu, *J. Geophys. Res. Solid Earth* **120**, 5082–5103 (2015).
10. D. F. Sumy, E. S. Cochran, K. M. Keranen, M. Wei, G. A. Abers, *J. Geophys. Res. Solid Earth* **119**, 1904–1923 (2014).
11. T. H. W. Goebel, M. Weingarten, X. Chen, J. Haffener, E. E. Brodsky, *Earth Planet. Sci. Lett.* **472**, 50–61 (2017).
12. Y. Guglielmi, F. Cappa, J.-P. Avouac, P. Henry, D. Elsworth, *Science* **348**, 1224–1226 (2015).
13. T. Kraft, N. Deichmann, *Geothermics* **52**, 59–73 (2014).
14. P. O. Ogwari, H. R. DeShon, M. J. Hornbach, *J. Geophys. Res. Solid Earth* **123**, 553–563 (2018).
15. L. V. Block, C. K. Wood, W. L. Yeck, V. M. King, *Seismol. Res. Lett.* **85**, 609–624 (2014).
16. P. Martínez-Garzón *et al*., *J. Geophys. Res.* **119**, 8378–8396 (2014).
17. S. Horton, *Seismol. Res. Lett.* **83**, 250–260 (2012).
18. L. Duboeuf *et al*., *J. Geophys. Res. Solid Earth* **122**, 8285–8304 (2017).
19. S. Baisch, R. Weidler, R. Vörös, D. Wyborn, L. De Graaf, *Bull. Seismol. Soc. Am.* **96**, 2242–2256 (2006).
20. J. Albaric *et al*., *Geothermics* **52**, 120–131 (2014).
21. S. Baisch, M. Bohnhoff, L. Ceranna, Y. Tu, H.-P. Harjes, *Bull. Seismol. Soc. Am.* **92**, 2369–2380 (2002).
22. T. Diehl, T. Kraft, E. Kissling, S. Wiemer, *J. Geophys. Res. Solid Earth* **122**, 7272–7290 (2017).
23. K. F. Evans, A. Zappone, T. Kraft, N. Deichmann, F. Moia, *Geothermics* **41**, 30–54 (2012).
24. Materials and methods are available as supplementary materials.
25. K. R. Felzer, E. E. Brodsky, *Nature* **441**, 735–738 (2006).
26. H. Wang, *Theory of Linear Poroelasticity with Applications to Geomechanics and Hydrogeology* (Princeton Univ. Press, 2000).
27. Y. Mukuhira, H. Moriya, T. Ito, H. Asanuma, M. Häring, *Geophys. Res. Lett.* **44**, 3109–3118 (2017).
28. I. Zaliapin, Y. Ben-Zion, *J. Geophys. Res. Solid Earth* **118**, 2847–2864 (2013).
29. H. Houston, B. G. Delbridge, A. G. Wech, K. C. Creager, *Nat. Geosci.* **4**, 404–409 (2011).
30. E. Detournay, A. H. D. Cheng, "Fundamentals of poroelasticity," in *Comprehensive Rock Engineering: Principles, Practice and Projects, Vol. II: Analysis and Design Method*, C. Fairhurst, Ed. (Pergamon Press, 1993), pp. 113–171.
31. X. Ma, M. D. Zoback, *J. Geophys. Res. Solid Earth* **122**, 2478–2503 (2017).
32. K.-H. Kim *et al*., *Science* **360**, 1007–1009 (2018).
33. N. J. VanDerElst, M. T. Page, D. A. Weiser, T. H. W. Goebel, S. M. Hosseini, *Geophys. Res. Lett.* **121**, 4575–4590 (2015).
34. M. R. M. Brown, S. Ge, A. F. Sheehan, J. S. Nakai, *J. Geophys. Res. Solid Earth* **122**, 6569–6582 (2017).
35. A. McGarr, *J. Geophys. Res. Solid Earth* **119**, 1008–1019 (2014).

ACKNOWLEDGMENTS

We thank P. Ogwari, H. De Shon, P. Martinez-Garcon, J. Ritter, M. Bohnhoff, L. De Barros, S. Wiemer, T. Kraft, J. Albaric, V. Oye, S. White, J. Riffault, and M. Fehler for their help with accessing the data and H. Shaddox, T. Lay, and S. Siman-Tov for discussions. **Funding:** We gratefully acknowledge funding from the U.S. Department of Energy, grant DE-SC0015539. **Author contributions:** T.H.W.G. and E.E.B. contributed equally to this work. **Competing interests:** The authors are not aware of any competing interests. **Data and materials availability:** Seismicity and injection data are available in the public domain (the supplementary materials contain references for each site).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/899/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S12
Table S1
References (*36*–*84*)
Data S1

9 March 2018; accepted 2 July 2018
10.1126/science.aat5449

SOLAR CELLS

High-performance perovskite/Cu(In,Ga)Se_2 monolithic tandem solar cells

Qifeng Han[1]*†, Yao-Tsung Hsieh[1]†, Lei Meng[1]†, Jyh-Lih Wu[2], Pengyu Sun[1], En-Ping Yao[1], Sheng-Yung Chang[1], Sang-Hoon Bae[1], Takuya Kato[2], Veronica Bermudez[2], Yang Yang[1]*

The combination of hybrid perovskite and Cu(In,Ga)Se_2 (CIGS) has the potential for realizing high-efficiency thin-film tandem solar cells because of the complementary tunable bandgaps and excellent photovoltaic properties of these materials. In tandem solar device architectures, the interconnecting layer plays a critical role in determining the overall cell performance, requiring both an effective electrical connection and high optical transparency. We used nanoscale interface engineering of the CIGS surface and a heavily doped poly[bis(4-phenyl)(2,4,6-trimethylphenyl)amine] (PTAA) hole transport layer between the subcells that preserves open-circuit voltage and enhances both the fill factor and short-circuit current. A monolithic perovskite/CIGS tandem solar cell achieved a 22.43% efficiency, and unencapsulated devices under ambient conditions maintained 88% of their initial efficiency after 500 hours of aging under continuous 1-sun illumination.

Constructing a tandem solar cell with minimal thermalization losses has proved to be a successful approach for overcoming the Shockley-Queisser limit of a single-junction cell. This device can realize the superposition of the open-circuit voltage (V_{OC}) of both subcells while simultaneously preserving high short-circuit current (J_{SC}) by using photoactive materials with complementary absorption characteristics to harvest a broader solar spectrum (*1*–*9*). Thin-film photovoltaic (PV) technologies applying various inorganic and organic photoactive materials have attracted considerable attention (*10*–*13*). Perovskite compounds debuted as an efficient light harvester for photoelectrochemical cells (*14*) and later evolved from liquid to solid-state junctions that enabled a large boost in performance (*15*, *16*), reaching power conversion efficiencies (PCEs) >22% in just 5 years (*17*–*21*). However, the well-established Cu(In,Ga)Se_2 (CIGS) solar cells have also yielded a maximum PCE >22% (*22*, *23*). Both PV materials have widely tunable bandgaps, from 1.0 to 1.7 eV for CIGS and 1.2 to 2.3 eV for perovskite (*24*–*29*). These characteristics provide the capability to achieve the highest efficiency of double-junction tandem solar cells, where the ideal rear and front cells should have bandgaps of 1.1 and 1.7 eV, respectively (*30*, *31*).

Several studies on four-terminal mechanically stacked perovskite/CIGS tandem solar cells have been reported (*32*–*34*). The highest efficiency for this type of perovskite/CIGS architecture is 22.1%, obtained with a 19% PCE CIGS rear cell and a 16% PCE perovskite front cell (*35*). By contrast, few studies on two-terminal perovskite/CIGS tandem solar cells have been reported, even though a two-terminal monolithic tandem architecture is potentially preferable for industrial applications because of the reduced number of electrodes and transparent conducting layers necessary. In 2015, Todorov *et al.* reported a two-terminal perovskite/CIGS tandem solar cell with an efficiency of 10.9%, which was much lower than the performance of the individual CIGS or perovskite subcells (*36*). There are three main reasons for this inferior efficiency. First, optical losses can be caused by top opaque metal electrodes. Second, the intrinsic ZnO (i-ZnO) and aluminum-doped ZnO (AZO) layers of typical CIGS cells were removed as zinc oxides can cause deterioration of the perovskite layer. However, by doing so, the original CIGS device architecture was compromised and elimination of the ZnO layer would inevitably degrade the CIGS device performance. Third, the fill factor (FF) was reduced to 60% because of a high series resistance (R_s) caused by poor contact between the two subcells.

The smoothness of the interconnecting layer (ICL) is equally crucial to create a reliable contact between the two subcells, because the planar perovskite solar cell is composed of several functional layers, with thicknesses from a few tens to hundreds of nanometers, that are sensitive to substrate roughness. Thus, in order to take advantage of these two technologies for two-terminal tandem solar cells, the challenge is to ensure the integrality of the two subcells, which relies heavily on the transparent top electrode of the perovskite front cell, maintaining the integrity of the original CIGS device structure to preserve its superior efficiency, and a well-designed ICL with a smooth surface.

We developed a transport top electrode, suitable ICL, and hole-transporting layer (HTL) for our tandem device and present a high-performance monolithic perovskite/CIGS tandem solar cell without modification of the CIGS device structure, i.e., preserving its TCO layers (i-ZnO and boron-doped ZnO (BZO) layers). For the two subcells, we applied a semitransparent perovskite with a bandgap of 1.59 eV as the front cell, and CIGS with a bandgap of 1.00 eV as the rear cell. The certified tandem device achieves a PCE of 22.43%.

To design a functional ICL, the CIGS device surface must be taken carefully into consideration. In this study, BZO is used as the top layer of the CIGS device, which has a surface roughness of about 60 nm, and the maximum vertical distance (VD) of the natural BZO layer texture can reach more than 250 nm (Fig. 1A). We speculate that such considerable roughness and VD may originate from the difference between peaks and valleys of the CIGS absorber layer. In addition, inhomogeneous nucleation of the bottom CdS buffer layer can also enhance BZO roughness, as shown in fig. S1.

The maximum VD is comparable to the length of the perovskite absorber layer, which is usually between 300 and 600 nm, and even larger than the thickness of the perovskite charge-transporting layers. With these large VDs, it becomes challenging to stack the perovskite solar cell on top of the CIGS with a homogeneous layer-by-layer structure. The rough BZO surface would cause perovskite subcell failure as the BZO peaks and rods can easily entangle the functional layers in the perovskite device to induce electrical shorting pathways between the top contact of perovskite subcell and the BZO layer. We confirmed that the nature of the CIGS device surface is problematic for building a smooth ICL on top of it, and hence the ICL roughness is pivotal in realizing high-performance perovskite/CIGS tandem solar cells.

To address this issue, we first deposited an indium tin oxide (ITO) layer, followed by chemical mechanical polishing (CMP) to smooth out the ITO surface. The addition of a sufficiently thick ITO layer can serve as a buffer layer for the CMP process to level out the huge VD of the BZO layer. The ITO layer was polished with a commercialized SiO_2 slurry. The detailed parameters for CMP processing are provided in the supplementary materials (*37*, *38*). After polishing, the maximum VD of the ITO layers was reduced to 40 nm (Fig. 1B), which rendered the ITO surface smooth enough for subsequent fabrication of the functional perovskite front cell. Notably, the CMP process did not polish the BZO layer such that we retained the original CIGS solar cell structure (Fig. 1C). Furthermore, the BZO work function (−4.0 eV, fig. S2) was lower than that of poly[bis(4-phenyl)(2,4,6-trimethylphenyl)amine] (PTAA) (−5.1 eV, the HTL of perovskite subcell),

[1]Department of Materials Science and Engineering, University of California–Los Angeles, Los Angeles, CA 90095, USA. [2]Atsugi Research Center, Solar Frontier K.K., Atsugi, Kanagawa 243-0206, Japan.
*Corresponding author. Email: yangy@ucla.edu (Y.Y.); qfhan@ucla.edu (Q.H.) †These authors contributed equally to this work.

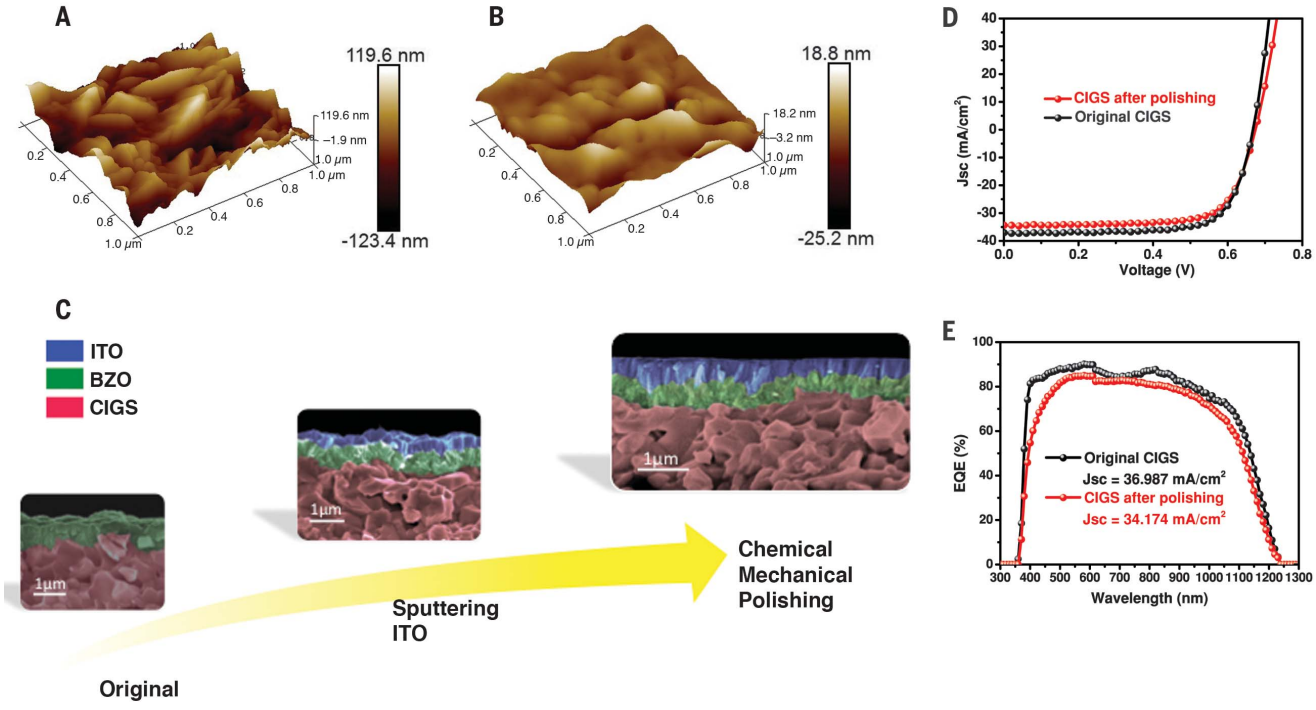


Fig. 1. Effects of CMP on CIGS surface and resulting performance of CIGS solar cells. (**A**) Atomic force microscopy (AFM) image of the CIGS surface before CMP polishing. (**B**) AFM image of the CIGS surface after CMP polishing. (**C**) Cross-sectional SEM images of the CMP processing on the CIGS surface. (**D**) *J-V* curves of original CIGS solar cells and after CMP polishing with a step size of 0.02 V and a scan velocity of 0.1 V/s, measured under AM1.5G illumination. (**E**) EQE of original CIGS solar cells and after CMP polishing.

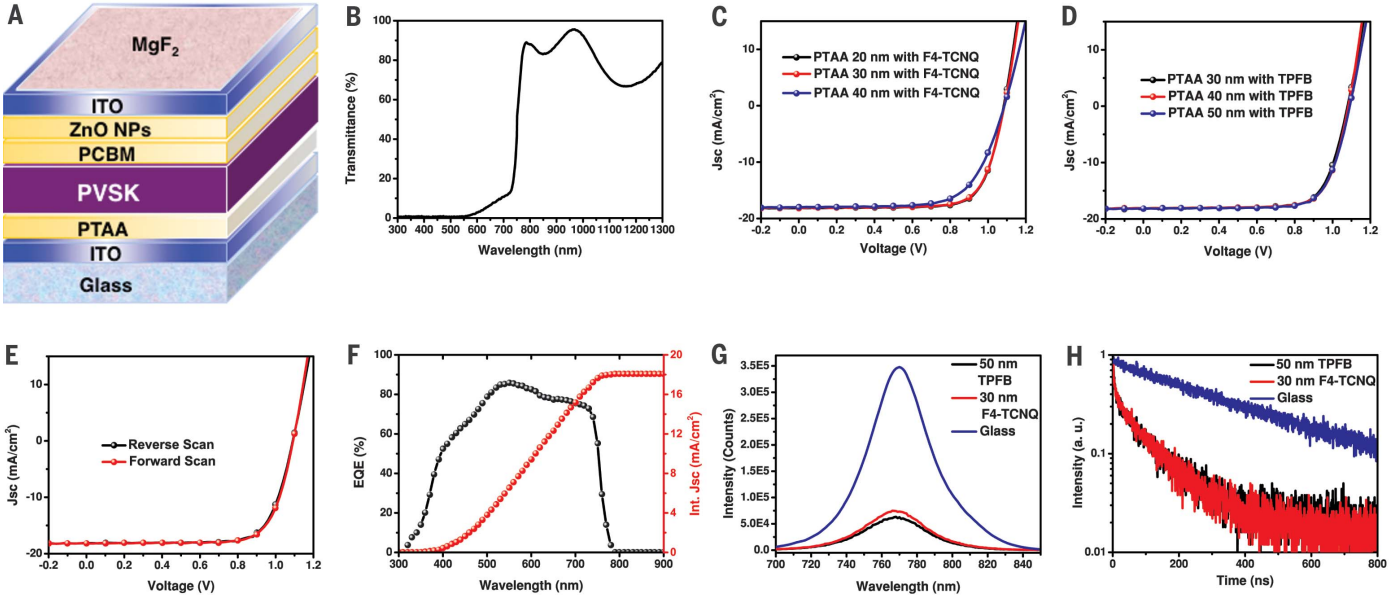


Fig. 2. Performance of the semitransparent $Cs_{0.09}FA_{0.77}MA_{0.14}Pb(I_{0.86}Br_{0.14})_3$ perovskite solar cell. (**A**) Schematic of the semitransparent single junction perovskite solar cell. (**B**) Transmittance spectrum through the entire device stack. (**C**) *J-V* curves of the perovskite solar cell using different thicknesses of 1 wt % F4-TCNQ–doped PTAA with illumination through the MgF_2 side. (**D**) *J-V* curves of the perovskite solar cell using different thicknesses of 10 wt % TPFB-doped PTAA with illumination through the MgF_2 side. (**E**) *J-V* curves in the forward (−0.1 to 1.2 V) and reverse (1.2 to −0.1 V) scan of the perovskite solar cell using 10 wt % TPFB-doped PTAA with illumination through the MgF_2 side. (**F**) EQE spectrum of the perovskite solar cell using 10 wt % TPFB-doped PTAA. The J_{SC} calculated from the EQE curve is 18.062 mA/cm^2. (**G**) Photoluminescence of the perovskite layer on top of glass and PTAA doped with 1 wt % F4-TCNQ or 10 wt % TPFB. (**H**) TRPL data for the perovskite layer in contact with glass and PTAA doped with 1 wt % F4-TCNQ or 10 wt % TPFB.

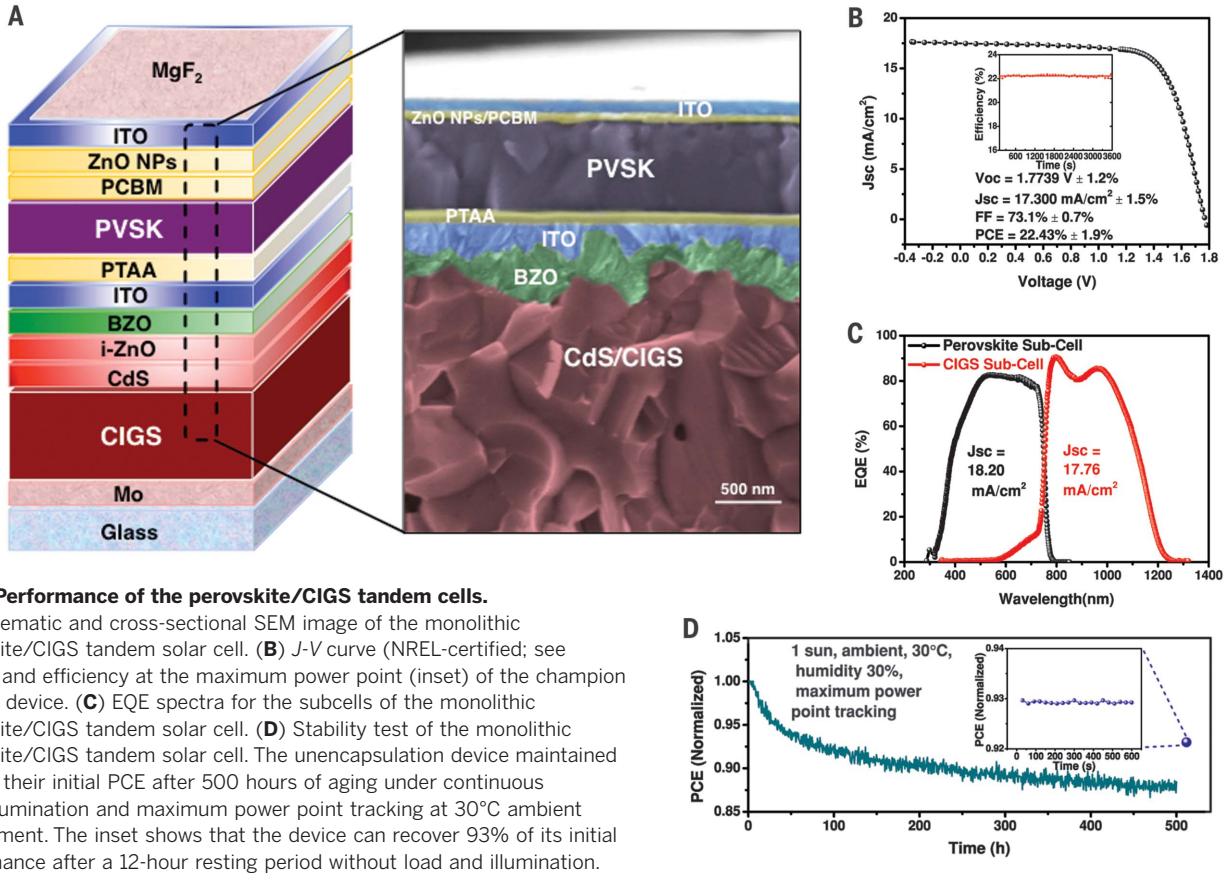


Fig. 3. Performance of the perovskite/CIGS tandem cells. (**A**) Schematic and cross-sectional SEM image of the monolithic perovskite/CIGS tandem solar cell. (**B**) *J-V* curve (NREL-certified; see fig. S8) and efficiency at the maximum power point (inset) of the champion tandem device. (**C**) EQE spectra for the subcells of the monolithic perovskite/CIGS tandem solar cell. (**D**) Stability test of the monolithic perovskite/CIGS tandem solar cell. The unencapsulation device maintained 88% of their initial PCE after 500 hours of aging under continuous 1-sun illumination and maximum power point tracking at 30°C ambient environment. The inset shows that the device can recover 93% of its initial performance after a 12-hour resting period without load and illumination.

which causes a large contact potential barrier. This ITO layer can efficiently modify the surface work function to create a better ohmic contact for hole transportation. On the basis of our experiments, a 300-nm ITO layer proved sufficient to carry out the CMP process and fully cap the BZO peaks and rods.

The current density–voltage (*J-V*) curves of the stand-alone original and polished CIGS solar cells are compared in Fig. 1D, and the data are shown in Table 1. The V_{OC} remained constant, which further implies that CMP processing did not damage the CIGS device structure. The J_{SC} decreased from 37.10 to 34.34 mA/cm^2 for the ITO-polished device. We attributed the J_{SC} drop to the additional light absorption by the ITO layer, as we observed with the smaller external quantum efficiency (EQE) intensity across the entire response region (Fig. 1E). The EQE of the ITO-polished device was lower than that of the original CIGS device at wavelengths from 400 to 500 nm. This response region corresponds to the smaller bandgap of ITO compared to BZO and provides evidence that the ITO layer absorbed a fraction of the incident light. However, the ITO absorption in the short-wavelength region was negligible as the CIGS subcell is designed as a rear cell in the tandem device structure. After polishing, the FF decreased from 74.7 to 72.4%. The FF reduction was mainly induced by the R_s increase induced by mediocre ITO lateral conductivity rather than the shunt resistance decrease (Fig. 1D). However, because this ITO layer is used as the ICL for the tandem solar cell, the lateral conductivity will not affect the charge carrier transportation between the front and rear subcells. After polishing the ITO layer of the CIGS device, the PCE was reduced from 18.73 to 16.76%. If we exclude the FF deficit, the current loss was only 1.386 mA/cm^2 in the wavelength region from 750 to 1250 nm.

We fabricated a semitransparent perovskite solar cell with an inverted structure (i.e., p-i-n) (Fig. 2A). Instead of using a metal electrode, we used a 100-nm ITO layer as the top contact in this structure to allow for sufficient light transmission (*39*). Various bandgaps of perovskites were tried in order to achieve current matching between the two subcells in the monolithic cell, and the best performance was achieved with a composition of $Cs_{0.09}FA_{0.77}MA_{0.14}Pb(I_{0.86}Br_{0.14})_3$, which had a bandgap of 1.59 eV from ultraviolet-visible (UV-vis) measurements (fig. S4). According to the optical simulation results (fig. S6), a 600-nm perovskite layer is needed to provide adequate current density to match the CIGS rear-cell current density. The average transmittance of the semitransparent perovskite cell in the wavelength region between 770 and 1300 nm is >80% (Fig. 2B), allowing most of the long-wavelength light to be absorbed by the CIGS rear cell. The transmittance gradually decreased from 770 to 550 nm, and the light was fully absorbed by the perovskite cell below 550 nm.

The thickness and coverage of the first layer on top of the ICL play a critical role for tandem device performance because the planar perovskite device structure has a limited tolerance to the substrate roughness. Given that the polished ICL still preserved about 40 nm for its VD, we studied the solar cell performance versus the HTL thickness by using PTAA as the HTL and two different molecules, 2,3,5,6-tetrafluoro-7,7,8,8-tetracyanoquinodimethane (F4-TCNQ) and 4-isopropyl-4'-methyldiphenyliodonium tetrakis (pentafluorophenyl)borate (TPFB), as dopants to enhance the HTL conductivity. We deposited PTAA with a low-temperature annealing process (110°C) to avoid damage to the rear CIGS solar cell (*40*–*42*). The *J-V* curves of semitransparent devices using F4-TCNQ and TPFB with different thicknesses measured under 100-mW/cm^2 illuminations are shown in Fig. 2, C and D, respectively, and the corresponding device parameters are summarized in Table 2.

With the same concentration of F4-TCNQ, the device performance obtained by applying 30 nm PTAA is similar to that achieved with 20 nm PTAA; however, the R_s increased when the PTAA reached 40 nm, which led to a FF and PCE reduction. The best device performance was achieved by using 20 nm F4-TCNQ–doped PTAA that gave a V_{OC} of 1.084 V, J_{SC} of 18.10 mA/cm^2, and FF of 75.6%, leading to an overall device efficiency of 14.83%. The desired F4-TCNQ/PTAA ratio was 1 weight % (wt %), and increasing the F4-TCNQ concentration did not

Table 1. Performance of CIGS solar cells before and after CMP polishing.

Device configuration	V_{OC} (V)	J_{SC} (mA/cm^2)	FF (%)	PCE (%)
Original CIGS	0.676	37.10	74.7	18.73
CIGS solar cell after ITO polishing	0.674	34.34	72.4	16.76

Table 2. Performance of perovskite solar cells using different dopants.

Device configuration	HTL thickness (nm)	V_{OC} (V)	J_{SC} (mA/cm^2)	FF (%)	PCE (%)
Perovskite solar cell using F4-TCNQ–doped PTAA as HTL	20	1.084	18.10	75.6	14.83
	30	1.086	18.12	74.5	14.66
	40	1.086	18.01	67.6	13.22
Perovskite solar cell using TPFB-doped PTAA as HTL	30	1.083	18.14	74.0	14.54
	40	1.084	18.13	74.9	14.72
	50	1.091	18.15	75.5	14.95
	50 (forward scan)	1.092	18.14	75.3	14.92

decrease the R_s because F4-TCNQ aggregates within the PTAA film. The *J-V* curves of the devices with different F4-TCNQ doping levels in the 30-nm PTAA layers are presented in fig. S7, and their device data are summarized in table S1.

The *J-V* curves of the perovskite device using TPFB-doped PTAA (TPFB/PTAA = 10 wt %) were less sensitive to HTL thickness. The best device had 50-nm PTAA, with a V_{OC} of 1.091 V, J_{SC} of 18.15 mA/cm^2, FF of 75.5%, and PCE of 14.95%. The high FF of this device resulted from the R_s reduction from the HTL, in which the TPFB-doped PTAA has a higher conductivity than the F4-TCNQ-doped PTAA.

The semitransparent perovskite device using TPFB-doped PTAA was scanned from positive to negative (reverse scan) and negative to positive (forward scan) voltages with a step size of 20 mV and a delay time of 0.2 s for each data point in the *J-V* measurement (Fig. 2E). The photocurrent hysteresis is negligible as the perovskite grain boundaries are well passivated by [6,6]-phenyl-C61-butyric acid methyl ester (PCBM), which agrees with other reported results (*43*, *44*). EQE data (Fig. 2F) for the semitransparent perovskite cell show an offset position at 780 nm, which was consistent with the UV-vis results. The integrated J_{SC} from the EQE using the AM 1.5 reference spectra reached 18.062 mA/cm^2.

Device performance is closely related to charge carrier dynamics in perovskite solar cells. We analyzed the charge collection and transportation by using steady-state photoluminescence (PL) and time-resolved PL (TRPL). Figure 2G shows the steady-state PL spectroscopy of perovskite films on three different substrates (glass, ITO/F4-TCNQ-doped PTAA, and TPFB-doped PTAA). Quenching was observed on both types of PTAA compared to the perovskite layers on glass, indicative of efficient charge transfer from the photoactive layer to the transport layer on contact with these two types of PTAA. From the TRPL responses, a decrease in the PL lifetime from 335 ns to 84 and 78 ns in the presence of 30-nm F4-TCNQ and 50-nm TPFB-doped PTAA, respectively, indicated that charge carriers within the perovskite layer were extracted effectively by these two types of PTAA.

Figure 3A illustrates the schematics and cross-sectional scanning electron micrograph (SEM) images of the tandem devices studied here. The polished ITO layer was used as the ICL to bridge two subcells together without the need for a tunneling junction. The *J-V* curve for the perovskite/CIGS champion tandem solar cell with 0.042 cm^2 is shown in Fig. 3B, certified by the National Renewable Energy Laboratory (NREL). The tandem cell exhibited a V_{OC} of 1.774 V (equal to the sum of the stand-alone V_{OC} of subcells), a FF of up to 73.1%, and a J_{SC} of 17.3 mA/cm^2, leading to an overall device PCE of 22.43%. We observed negligible hysteresis (fig. S9). In Fig. 3C, the integrated J_{SC} from the EQE curves for the top and rear cells is 18.20 and 17.76 mA/cm^2, respectively, showing well the current-matched subcells, although the rear solar cell slightly limited the overall tandem solar cell current. The EQE of the ITO-polished CIGS is lower than 80% from 800 to 1100 nm in Fig. 1E, but it can be restored to ~85% by applying a MgF_2 layer (Fig. 3C). This improvement can minimize the efficiency loss of the polished CIGS. Tandem devices with larger area (0.52 cm^2) were also made, and the best one with negligible hysteresis had a 20.8% PCE measured in-house (fig. S11). The device parameters are summarized in table S2. To realize the full potential of perovskite/CIGS tandem devices, we suggest three key improvements. First, reducing the V_{OC} loss in the perovskite subcell is required. An effective defect passivation can help to provide a higher V_{OC} (*45*, *46*). Second, using vapor-based processes to deposit the electron transport layer is more desirable than conventional solution process to prevent shunt pathways and improve interface contacts (*39*). Lastly, use of a higher PCE for the CIGS subcell is needed, especially considering that the limiting current could be increased by tailoring the CIGS bandgap in our case.

In addition to a high PCE, long-term stability is another crucial benchmark for industrialization of perovskite solar cells. Several aging routines have been suggested to estimate a conclusive stability in which the ion migration effects are excluded (*47*, *48*). We monitored the unencapsulated tandem device performance by aging for 500 hours under continuous 1-sun illumination and maximum power point tracking at 30°C ambient environment. The device started with 22.0% PCE and retained >88% of its initial efficiency after aging, and it recovered 93% of its initial PCE after being kept in the dark for 12 hours without load (shown in Fig. 3D). We believe that the top transparent metal oxide layers (composed of ZnO nanoparticles and sputtered ITO) can effectively resist moisture ingress (*39*, *49*), so that this structure can help the perovskite compounds remain stable without severe degradation.

Our approach can also help to alleviate the environmental impact of cadmium and extend the working period of the perovskite/CIGS tandem solar cell. The CIGS solar cells can be reused after washing out the degraded perovskite front cell. Because here we apply PTAA as the HTL material, the whole perovskite front cell can be removed from the CIGS rear cell by dissolving in chlorobenzene and *N*,*N*-dimethylformamide. Details of the washing process are described in the supplementary materials. The CIGS rear cell maintains the same performance when the front perovskite cell is removed, demonstrating that the fabrication and dissolving processes of the front subcell do not damage the CIGS device. Similar PCEs are obtained for the reused tandem devices (fig. S12).

REFERENCES AND NOTES

1. C. H. Henry, *J. Appl. Phys.* **51**, 4494–4500 (1980).
2. A. Brown, M. A. Green, *Prog. Photovolt. Res. Appl.* **10**, 299–307 (2002).
3. J. Y. Kim *et al.*, *Science* **317**, 222–225 (2007).
4. S. Albrecht *et al.*, *Energy Environ. Sci.* **9**, 81–88 (2016).
5. G. E. Eperon *et al.*, *Science* **354**, 861–865 (2016).
6. K. A. Bush *et al.*, *Nat. Energy* **2**, 17009 (2017).
7. J. Werner *et al.*, *J. Phys. Chem. Lett.* **7**, 161–166 (2016).
8. C. D. Bailie, M. D. McGehee, *MRS Bull.* **40**, 681–686 (2015).
9. J. Kim *et al.*, *Nat. Commun.* **6**, 6391 (2015).
10. S. Mathew *et al.*, *Nat. Chem.* **6**, 242–247 (2014).
11. L. Dou *et al.*, *Adv. Mater.* **25**, 6642–6671 (2013).
12. M. Graetzel, R. A. J. Janssen, D. B. Mitzi, E. H. Sargent, *Nature* **488**, 304–312 (2012).
13. A. Chirilă *et al.*, *Nat. Mater.* **12**, 1107–1111 (2013).
14. A. Kojima, K. Teshima, Y. Shirai, T. Miyasaka, *J. Am. Chem. Soc.* **131**, 6050–6051 (2009).
15. H. S. Kim *et al.*, *Sci. Rep.* **2**, 591 (2012).
16. M. M. Lee, J. Teuscher, T. Miyasaka, T. N. Murakami, H. J. Snaith, *Science* **338**, 643–647 (2012).
17. NREL, Best Research Cell Efficiencies; https://www.nrel.gov/pv/assets/images/efficiency-chart.png (2017).
18. W. S. Yang *et al.*, *Science* **348**, 1234–1237 (2015).
19. S. D. Stranks, P. K. Nayak, W. Zhang, T. Stergiopoulos, H. J. Snaith, *Angew. Chem. Int. Ed.* **54**, 3240–3248 (2015).
20. H. S. Jung, N.-G. Park, *Small* **11**, 10–25 (2015).
21. M. D. McGehee, *Nature* **501**, 323–325 (2013).
22. P. Jackson *et al.*, *Phys. Status Solidi Rapid Res. Lett.* **10**, 583–586 (2016).
23. A. Chirilă *et al.*, *Nat. Mater.* **10**, 857–861 (2011).

24. T. K. Todorov, O. Gunawan, T. Gokmen, D. B. Mitzi, *Prog. Photovolt. Res. Appl.* **21**, 82–87 (2013).
25. S. Jung *et al.*, *Curr. Appl. Phys.* **10**, 990–996 (2010).
26. D. P. McMeekin *et al.*, *Science* **351**, 151–155 (2016).
27. N. J. Jeon *et al.*, *Nature* **517**, 476–480 (2015).
28. J. W. Lee *et al.*, *Adv. Energy Mater.* **5**, 1501310 (2015).
29. R. E. Beal *et al.*, *J. Phys. Chem. Lett.* **7**, 746–751 (2016).
30. T. J. Coutts *et al.*, *Prog. Photovolt. Res. Appl.* **11**, 359–375 (2003).
31. F. Meillaud, A. Shah, C. Droz, E. Vallat-Sauvain, C. Miazza, *Sol. Energy Mater. Sol. Cells* **90**, 2952–2959 (2006).
32. L. Kranz *et al.*, *J. Phys. Chem. Lett.* **6**, 2676–2681 (2015).
33. C. Zuo *et al.*, *Adv. Sci.* **3**, 1500324 (2016).
34. A. Guchhait *et al.*, *ACS Energy Lett.* **2**, 807–812 (2017).
35. F. Fu *et al.*, *Nat. Energy* **2**, 16190 (2016).
36. T. Todorov *et al.*, *Adv. Energy Mater.* **5**, 1500799 (2015).
37. G. W. Choi *et al.*, *Microelectron. Eng.* **83**, 2213–2217 (2006).
38. K. Y. Lee, G. W. Choi, Y. J. Kim, Y. O. Choi, N. O. Kim, *J. Korean Phys. Soc.* **60**, 388–392 (2012).
39. K. A. Bush *et al.*, *Adv. Mater.* **28**, 3937–3943 (2016).
40. Q. Wang, C. Bi, J. Huang, *Nano Energy* **15**, 275–280 (2015).
41. C. Bi *et al.*, *Nat. Commun.* **6**, 7747 (2015).
42. Y. Shao, Y. Yuan, J. Huang, *Nat. Energy* **1**, 15001 (2016).
43. J. Xu *et al.*, *Nat. Commun.* **6**, 7081 (2015).
44. Y. Li *et al.*, *J. Am. Chem. Soc.* **137**, 15540–15547 (2015).
45. X. Zheng *et al.*, *Nat. Energy* **2**, 17102 (2017).
46. M. Abdi-Jalebi *et al.*, *Nature* **555**, 497–501 (2018).
47. M. Saliba, *Science* **359**, 388–389 (2018).
48. K. Domanski, E. A. Alharbi, A. Hagfeldt, M. Grätzel, W. Tress, *Nat. Energy* **3**, 61–67 (2018).
49. J. You *et al.*, *Nat. Nanotechnol.* **11**, 75–81 (2016).

ACKNOWLEDGMENTS

We thank T. Moriarty (National Renewable Energy Laboratory) for conducting device measurements and providing valuable discussions on *I-V* and EQE results; G. Li, J.-W. Lee, and O. Lin for constructive technical discussion; and N. De Marco for editing. We thank UCLA Nanoelectronics Research Facility for assistance with CMP equipment and supplies. **Funding:** This work is financially supported by grants from the National Science Foundation (grant no. ECCS-1509955; program director: N. El-Masry), and the Air Force Office of Scientific Research (grant no. FA9550-15-1-0333; program manager: C. Lee) **Author contributions:** Q.H. and Y.-T.H. conceived the experiments, performed data analysis, and wrote the manuscript. L.M. developed the hole transport layer and assisted in perovskite fabrication. J.-L.W., T.K., and V.B. led the fabrication of the CIGS solar cell. P.S. performed the optical modeling. E.-P.Y. assisted with the measurement data analysis. S.-Y.C. assisted in CIGS surface analysis. S.-H.B. assisted in polishing processes. All authors discussed the results and commented on the manuscript. Y.Y. directed and supervised the entire research. **Competing interests:** All authors declare no competing interests. **Data and materials availability:** All data are available in the manuscript and the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/904/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 and S2
References (*50*)

5 March 2018; accepted 27 June 2018
10.1126/science.aat5055

MOLECULAR MOTORS

Directional control of a processive molecular hopper

Yujia Qing, Sandra A. Ionescu, Gökçe Su Pulcu, Hagan Bayley*

Intrigued by the potential of nanoscale machines, scientists have long attempted to control molecular motion. We monitored the individual 0.7-nanometer steps of a single molecular hopper as it moved in an electric field along a track in a nanopore controlled by a chemical ratchet. The hopper demonstrated characteristics desired in a moving molecule: defined start and end points, processivity, no chemical fuel requirement, directional motion, and external control. The hopper was readily functionalized to carry cargos. For example, a DNA molecule could be ratcheted along the track in either direction, a prerequisite for nanopore sequencing.

Processivity lies at the heart of biological machines. A replicative DNA polymerase can incorporate thousands of nucleotides before dissociating from its template (*1*). Molecular motors, such as kinesin and dynein, travel directionally along microtubules over hundreds of steps without detaching from the track (*2*–*4*). For years, scientists have been trying to build moving molecules that resemble their biomolecular counterparts but use simpler components (*5*). The ultimate goals are to achieve true processivity, which can be defined as directional motion without leaving a track and the performance of useful work such as the transport of a cargo. Ideally, a synthetic system should exhibit the reversibility of stepping seen in various biological systems (*6*, *7*) to enable the direction of motion to be switched through external control.

We report the design of a one-legged molecular hopper that is ratcheted by dynamic covalent chemistry along a protein track (Fig. 1A) with robust processivity (table S1). Further, the direction in which the hopper moves is subject to external control by an electrical potential. The track is built inside a protein nanopore, α-hemolysin (αHL), and consists of a series of cysteine footholds facing the lumen of the transmembrane β barrel (Fig. 1B). The cysteines are evenly spaced along a β strand with an average interfoothold distance of 6.8 Å (Cα-Cα) and an average vertical spacing of 5.6 Å. The hopper uses consecutive thiol-disulfide interchange reactions to move in the direction in which the DNA cargo has been oriented by an applied potential. To execute the S_N2 reaction, the three participating sulfur atoms must align in a near-linear configuration (*8*–*10*). Under the applied potential, the DNA inside the barrel is pulled in the electric field with a force of ~10 pN (see supplementary materials). The force sets the overall direction of motion by flipping the DNA (see below) and helps to orient the disulfide for cleavage by the neighboring downstream cysteine thiolate, which moves the hopper one step forward, although other forces can contribute to the forward motion (Fig. 1C). Backstepping is disfavored and overstepping is impossible. Release of the hopper from the linear track was not observed, presumably because the track is too rigid to accommodate the resulting disulfide bridge between adjacent footholds on the same β strand. In short, each step is chemically directional because the hopper's "foot" is positioned to favor the forward reaction. Further, the motion is autonomous, requiring no chemical fuel.

Under +150 mV, the hopper was delivered to the track from the cis compartment as a hopper-carrier conjugate (Fig. 1D) capped with a single traptavidin, which was arrested at the pore entrance (Fig. 2A). The disulfide in the construct reacted strictly regioselectively with Cys^{115}, releasing the carrier and placing the hopper-DNA cargo on the starting foothold (Fig. 2A and fig. S1). The location of the hopper was ascertained from the residual current passing through the nanopore, which reflected the length of the DNA located within the β barrel when the hopper was at a particular foothold (fig. S1). By monitoring current changes, we followed the stepwise hopping motion at the single-molecule level in real time.

The voltage-controlled hopping motion was directional and processive. On a track containing five cysteine footholds (at positions 113, 115, 117, 119, 121), a hopper carrying an oligoadenosine 40-mer (A40, hopper 1) moved cis

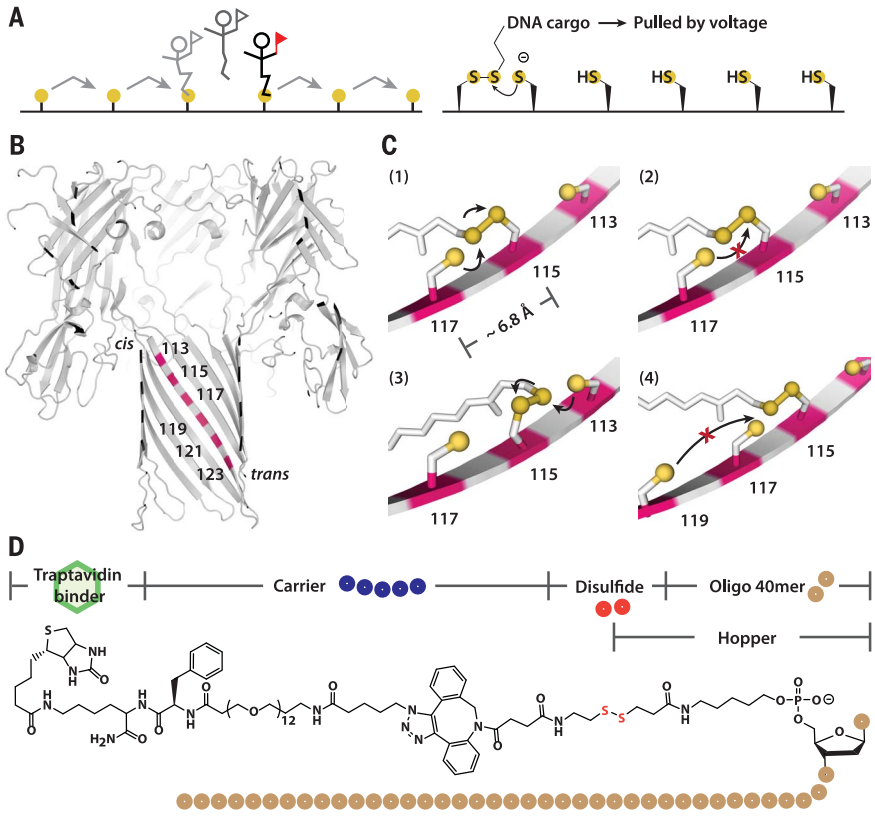


Fig. 1. A molecular hopper on a protein track. (**A**) A hopper carrying a cargo (red flag) moves along a track by means of consecutive thiol-disulfide interchange reactions. The overall direction is set by the applied potential. (**B**) A six-foothold track comprising odd-numbered cysteine residues on a β strand inside the αHL protein nanopore. (**C**) The applied potential exerts a force on the DNA cargo, which helps to align the three sulfur atoms (yellow) participating in the interchange. The collinear geometry promotes hopping (1) (movie S1) but not the formation of an intrastrand disulfide, which would release the hopper from the track (2). Occasionally, backstepping is observed (3). Overstepping (4) does not occur. (**D**) The hopper enters the nanopore as a carrier-hopper disulfide conjugate.

Department of Chemistry, University of Oxford, Oxford OX1 3TA, UK.
*Corresponding author. Email: hagan.bayley@chem.ox.ac.uk

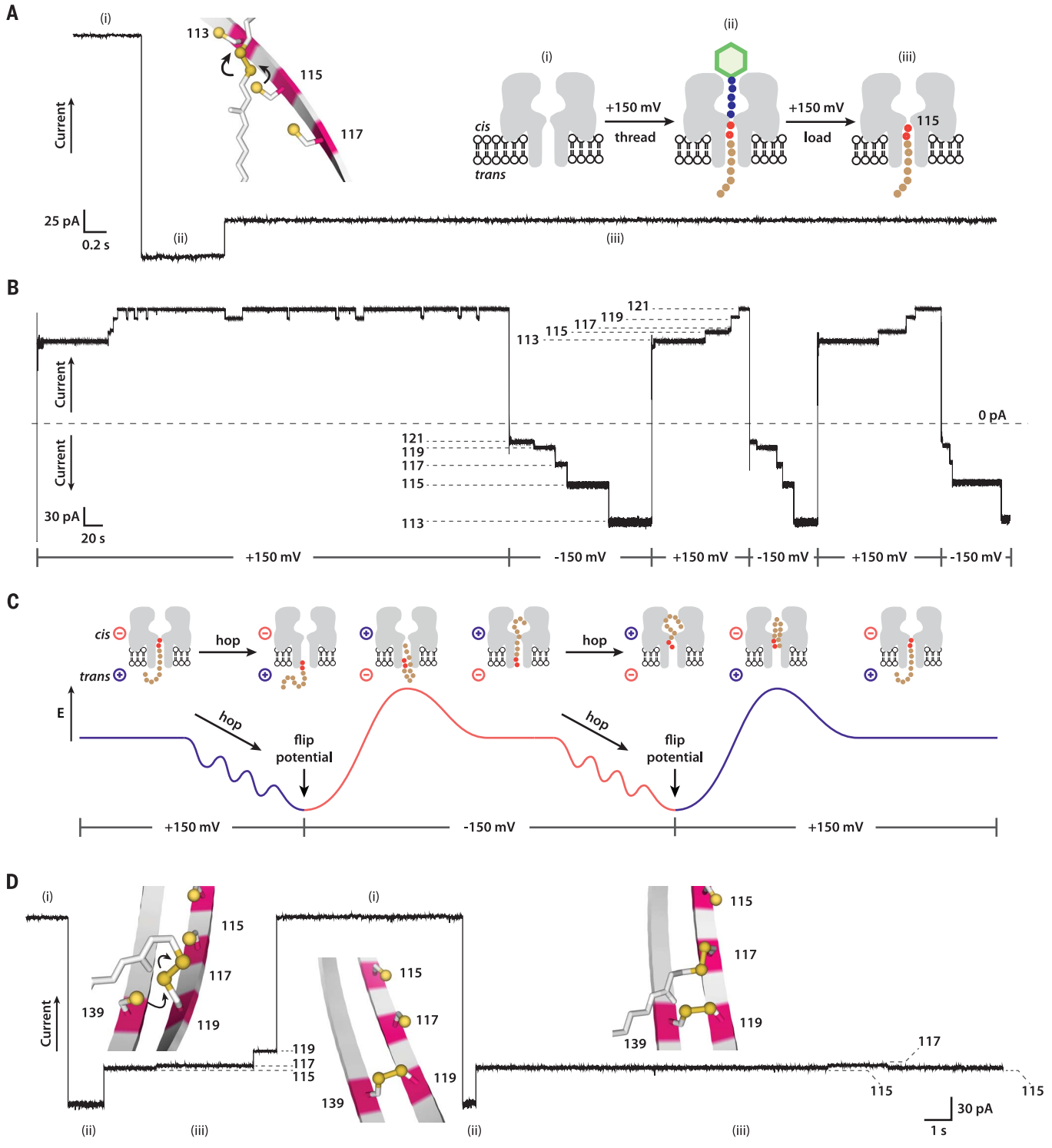


Fig. 2. Monitoring individual hopper steps. (**A**) Under +150 mV, a hopper-carrier conjugate capped with traptavidin was pulled from the cis compartment into an αHL nanopore containing cysteines at positions 113, 115, 117, 119, and 121 in one of the seven subunits. The resultant blockade reduced the ionic current from (i) to (ii). Reaction of the disulfide in the hopper-carrier with Cys^{115} covalently attached the hopper to the track, and the ionic current increased to (iii). (**B**) With a five-cysteine track, four hopping steps were observed at ±150 mV. Every forward step moved part of the DNA cargo outside the β barrel, producing an increase in conductance (movie S2). Alternation of the applied potential drove the hopper repeatedly up and down the track. (**C**) A hypothetical free energy diagram (not to scale) of the controlled hopping motion. (**D**) On an L-shaped track consisting of cysteines at positions 115, 117, 119, and 139, the hopper moved along the track from Cys^{115} to Cys^{119}, where it was released by the side chain of Cys^{139}. Subsequently, a second hopper became loaded at Cys^{115}, but its motion was arrested at Cys^{117} because Cys^{119} was now engaged in an interstrand disulfide bond [(i), (ii), (iii) as in (A)]. Backstepping to Cys^{115} was also seen. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20° ± 1°C.

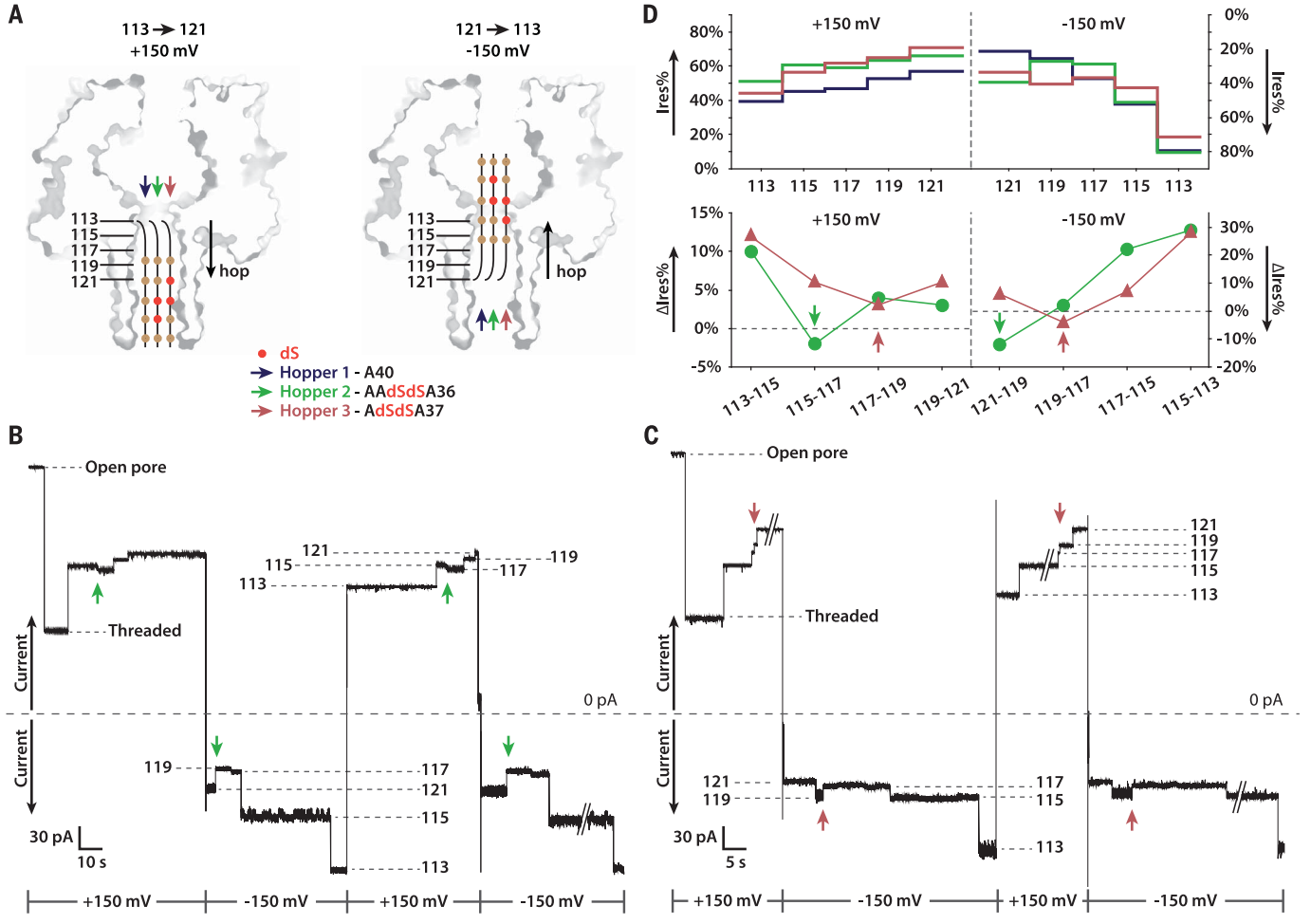


Fig. 3. Discrimination of different DNA cargos. (**A**) Two sequential abasic nucleotides (dSdS) were substituted at positions 3 and 4 (hopper 2) or at positions 2 and 3 (hopper 3). The numbers of nucleotides (brown circles, dA; red circles, dS) placed inside the β barrel are based on PyMOL modeling. (**B**) With a five-cysteine track, four-step hopping was observed with hopper 2 at ±150 mV. The current decreases for hops from 115 to 117 and from 121 to 119 are marked (green arrows). (**C**) Four-step hopping with hopper 3 at ±150 mV. The current transitions for hops from 117 to 119 and from 119 to 117 are marked (red arrows). (**D**) Top: Overlaid current traces of hoppers 1, 2, and 3 [colors as in (A)] with step durations normalized. The current levels are given as the residual current with respect to the open pore level (Ires%). Bottom: Step sizes of hoppers 2 and 3 plotted as ΔIres%. Minima in the plots showing the single nucleotide offset are marked. Conditions: 2 M KCl, 20 mM HEPBS, 20 μM EDTA, pH 8.5, 20° ± 1°C.

to trans under +150 mV, and trans to cis under –150 mV (Fig. 2B and fig. S2). When the hopper reached a terminal foothold, the sign of the applied potential was reversed in order to re-orient (flip) the DNA cargo, and hence the hopper. Alternation between positive and negative potentials repeatedly drove the hopper toward the trans or the cis end of the track. For the DNA to experience a force, at least one negatively charged phosphodiester bond must lie within the electric field, which drops along the length of the pore's β barrel (*11*). Therefore, in the present nanopore construct, the length of the track was limited and voltage-controlled hopping could only be demonstrated with up to six footholds (fig. S3). The hopping direction could be changed by reversing the applied potential at any foothold; thus, our system offers complete control over directionality and the ability to move a hopper back to the initial foothold after an outing. Moreover, the applied potential provides an external energy source to produce directional motion (Fig. 2C; see supplementary materials). Limited by bilayer stability, the longest records of processive hopping were documented with hopper 1, which completed 249 forward steps in 93 min on a five-cysteine track (113 to 121) with a mean dwell time of ~22 s per foothold (fig. S2). Dissociation of the hopper from the track was never seen (n > 30 outings on different tracks), which implies that substantial improvements on step numbers would be achieved if the stability of the bilayer were improved. In comparison, previous synthetic small-molecule walkers moved directionally for less than 10 steps (*5*, *12*). Wild-type kinesins typically exhibit a mean step number of 75 to 175 before dissociation (*2*, *3*).

The hopping rates for each of the four steps on the five-cysteine track were derived for both the cis-to-trans and trans-to-cis directions at pH 8.5 and displayed differences of less than a factor of 40 (0.0081 to 0.30 s^{-1}; tables S2 to S4). Because a thiolate is the reactive nucleophile in disulfide interchanges, the rate differences might arise from variations in the pK_a values of the foothold thiols, which will be affected by neighboring residues. Previously, an arsenic(III) walker showed a factor of ≤50 difference in attachment rates with the footholds on the same five-cysteine track (113 to 121) at pH 8.0 (*12*). In the future, tracks of thiols might be engineered with optimized interfoothold distances and enhanced chemical reactivity to speed up the hopping process. Alternatively, the properties of the reactive sulfur atom in the hopper might be manipulated by flanking functional groups. With both the two-cysteine and three-cysteine tracks, the influence of voltage on hopping was examined at ±100 mV, ±150 mV, and ±180 mV (tables S5 and S6). The rates showed weak nonexponential voltage

dependences, suggesting that the applied potential might not be the only source of propulsion (see supplementary materials). However, an electrical potential is essential (i) to flip the DNA over a large barrier to set the direction of motion, and (ii) to aid in the orientation of the three participating sulfur atoms to favor "forward" reactions over backsteps. With respect to the latter, the estimated "effective concentrations" of participating downstream thiols are not especially high (see supplementary materials) (*13*) and indeed need not be high to produce overall forward motion (see below).

Although disfavored, backstepping was occasionally detected, which we attributed to conformational lability of the hopper within the nanopore even under an applied potential. During a recording with hopper 1 on a five-cysteine track (113 to 121), there were 33 backward steps on the nonterminating footholds out of 282 steps in total (12%). Of the 33 backward steps, 29 occurred from 115 to 117 at −150 mV (table S3). Despite the large forward equilibrium constant for 117-to-115 stepping ($K = k_{117-115}/k_{115-117} = 22$), 115-to-117 backstepping was observed because of the comparatively slow forward movement to the next foothold, 113 (at −150 mV, $k_{115-117} = 0.0094\ s^{-1}$, $k_{115-113} = 0.0081\ s^{-1}$; tables S2 and S3). Backstepping was observed when the hopper was left on a terminal foothold, as no forward footholds remained (Fig. 2B). These backsteps were quickly reversed by the hopper, which preferentially resided at the final station ($K = k_{119-121}/k_{121-119} = 5.2$; table S4). The overall motion of the hopper is governed by the product of the K values for each step. A modest value ($K > 1$) at each step produces a considerable overall tendency toward forward movement.

Each subunit of the αHL pore offers two antiparallel β strands to the transmembrane β barrel with an interstrand distance of ~5 Å (Cα-Cα). Given that the formation of cross-strand disulfides has been reported (*14*), we reasoned that the addition of a cysteine on an adjacent strand would compel hopper release from the track at a designated foothold. Indeed, with an L-shaped track consisting of cysteines at positions 115, 117, 119, and 139, the hopper attached to the track at foothold 115 by regioselective disulfide formation and dissociated from the track when it reached foothold 119. The release was initiated by Cys^{139} through thiol-disulfide interchange to form a cross-strand disulfide bridge, which blocked the access of subsequent hoppers to foothold 119 (Fig. 2D). The preference for hopper release versus hopper transfer to the adjacent strand is attributed to an unfavorable collinear alignment of the three participating sulfur atoms necessary for transfer. In the future, the engineering of footholds on a surface will allow the construction of more complex hopping pathways where hoppers are transferred to new tracks at designed junctions and cargos are released at predesignated depots.

The ability to translocate a stretched DNA cargo while maintaining a covalent bond with the nanopore suggests a method for the chemical ratcheting of a nucleic acid during nanopore sequencing (*15*), which was explored in a proof-of-concept experiment. To provide a marker, we incorporated two adjacent abasic residues (1′,2′-dideoxyribose, dS) (*16*) into the cargo oligos carried by hoppers 2 and 3 and recorded current patterns during four-step hopping between Cys^{113} and Cys^{121} (Fig. 3A). By comparison with hopper 1, hoppers 2 and 3 showed different patterns of current modulation (Fig. 3, B and C, fig. S4, and table S7). The conductance patterns generated by the four-step hopping motion could be repeated with different molecules of hoppers 2 and 3 ($n = 3$ for each hopper), establishing the patterns as clear identifiers of each cargo sequence. The residual currents (Ires%, the remaining current as a percentage of the open pore current) for the three hoppers residing at each foothold were plotted for comparison (Fig. 3D). Hoppers 1 and 2 gave almost identical current blockades at each of footholds 115 and 113 under −150 mV, implying that the dSdS sequence had been transported well out of the sensing region by hopper 2. Moreover, hoppers 2 and 3 have a single nucleotide offset in the dSdS positions, and we observed a one-step offset between hoppers 2 and 3 in ΔIres%, the difference in Ires% between two successive steps [Fig. 3D; the vertical step size, 5.6 Å, is similar to the internucleotide distance in stretched single-stranded DNA, 6.9 Å (*17*)].

These observations demonstrate that the hopper system reported here has the potential to discriminate bases for sequencing purposes (*16*). An advantage of a processive hopper, which might improve sequencing accuracy, is the ability to reverse the chemical ratcheting process and thereby obtain many-fold coverage of an individual DNA strand. Of course, the present system is limited by its short track. Although longer β-barrel pores exist (*18*, *19*), a viable sequencing process will require protracted ratcheting of numerous DNA strands in parallel, perhaps by using footholds on an extended crystalline surface or internal thiophosphate feet to transport long replica strands over relatively short tracks.

REFERENCES AND NOTES

1. W. A. Breyer, B. W. Matthews, *Protein Sci.* **10**, 1699–1711 (2001).
2. R. D. Vale *et al*., *Nature* **380**, 451–453 (1996).
3. R. B. Case, D. W. Pierce, N. Hom-Booher, C. L. Hart, R. D. Vale, *Cell* **90**, 959–966 (1997).
4. A. Gennerich, R. D. Vale, *Curr. Opin. Cell Biol.* **21**, 59–67 (2009).
5. S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, A. L. Nussbaumer, *Chem. Rev.* **115**, 10081–10206 (2015).
6. W. J. Walter, M. P. Koonce, B. Brenner, W. Steffen, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 5289–5293 (2012).
7. R. A. Cross, A. McAinsh, *Nat. Rev. Mol. Cell Biol.* **15**, 257–271 (2014).
8. R. D. Bach, O. Dmitrenko, C. Thorpe, *J. Org. Chem.* **73**, 12–21 (2008).
9. P. A. Fernandes, M. J. Ramos, *Chem. Eur. J.* **10**, 257–266 (2004).
10. A. P. Wiita *et al*., *Nature* **450**, 124–127 (2007).
11. S. Howorka, H. Bayley, *Biophys. J.* **83**, 3202–3210 (2002).
12. G. S. Pulcu, E. Mikhailova, L.-S. Choi, H. Bayley, *Nat. Nanotechnol.* **10**, 76–83 (2015).
13. C. A. Hunter, H. L. Anderson, *Angew. Chem. Int. Ed.* **48**, 7488–7499 (2009).
14. J. Venkatraman, G. A. Nagana Gowda, P. Balaram, *J. Am. Chem. Soc.* **124**, 4987–4994 (2002).
15. H. Bayley, *Clin. Chem.* **61**, 25–31 (2015).
16. B. Gyarfas *et al*., *ACS Nano* **3**, 1457–1466 (2009).
17. A. Bosco, J. Camunas-Soler, F. Ritort, *Nucleic Acids Res.* **42**, 2064–2074 (2014).
18. I. Iacovache *et al*., *Nat. Commun.* **7**, 12062 (2016).
19. J. Jiang, B. L. Pentelute, R. J. Collier, Z. H. Zhou, *Nature* **521**, 545–549 (2015).

ACKNOWLEDGMENTS

We thank H. Tamagaki and J. Lee for their help with peptide design and synthesis; S. Lee for her help with data analysis using QuB; L. Taemaitree for her help with LC-MS characterization; and M. Howarth for providing a sample of traptavidin for preliminary tests. **Funding:** Supported by a European Research Council Advanced Grant; a China Scholarship Council–University of Oxford Scholarship (Y.Q.); and an Amelia Jackson Studentship, Exeter College, Oxford (S.A.I.). **Author contributions:** Y.Q. and H.B. designed the study; Y.Q. performed the experiments; S.A.I. repeated the single-channel recordings; G.S.P. contributed to the protein preparation; and Y.Q., S.A.I., and H.B. wrote the manuscript. **Competing interests:** H.B. is the founder of, a consultant for, and a shareholder of Oxford Nanopore Technologies, a company engaged in the development of nanopore sensing and sequencing technologies. Y.Q. and H.B. have filed patents describing the molecular hopper and applications thereof. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/908/suppl/DC1
Materials and Methods
Figs. S1 to S12
Tables S1 to S7
Movies S1 and S2
References (*20*–*51*)

25 February 2018; accepted 21 June 2018
10.1126/science.aat3872

ORGANOMETALLICS

Observation of alkaline earth complexes $M(CO)_8$ (M = Ca, Sr, or Ba) that mimic transition metals

Xuan Wu[1]*, Lili Zhao[2]*, Jiaye Jin[1], Sudip Pan[2], Wei Li[1], Xiaoyang Jin[1], Guanjun Wang[1], Mingfei Zhou[1]†, Gernot Frenking[2,3]†

The alkaline earth metals calcium (Ca), strontium (Sr), and barium (Ba) typically engage in chemical bonding as classical main-group elements through their *n*s and *n*p valence orbitals, where *n* is the principal quantum number. Here we report the isolation and spectroscopic characterization of eight-coordinate carbonyl complexes $M(CO)_8$ (where M = Ca, Sr, or Ba) in a low-temperature neon matrix. Analysis of the electronic structure of these cubic O_h-symmetric complexes reveals that the metal–carbon monoxide (CO) bonds arise mainly from $[M(d_\pi)] \rightarrow (CO)_8$ π backdonation, which explains the strong observed red shift of the C-O stretching frequencies. The corresponding radical cation complexes were also prepared in gas phase and characterized by mass-selected infrared photodissociation spectroscopy, confirming adherence to the 18-electron rule more conventionally associated with transition metal chemistry.

The periodic table of the elements is conventionally divided according to the valence atomic orbitals (AOs) into main-group s and p blocks, a transition metal d block, and a lanthanide and actinide f block. A useful set of guidelines for understanding the structures and stabilities of molecules encompasses the associated 8-, 18-, and 32-electron rules introduced by Langmuir (*1*, *2*) before the advent of quantum theory. These rules were later explained by attributing particular stability to filled sp, spd, or spdf valence shells, respectively (*3*).

The alkaline earth elements beryllium, magnesium, calcium, strontium, and barium have a ns^2 valence-shell configuration, where *n* is the principal quantum number, and, as such, typically engage in chemical bonding as ionic salt compounds or in polar bonds via their two *n*s valence electrons in divalent M(II) species (*4*), where M is an alkaline earth metal. Earlier studies suggested that the heaviest atom barium may use its 5d AOs to some extent in chemical bonds (*5*), which led to the suggestion that barium be designated an "honorary transition metal" (*6*). Previously, we reported the experimental observation of barium carbonyl ions $Ba(CO)^q$ (where charge q = +1 and −1) (*7*). The analysis of the electronic structure showed that the cation binds the ligand mainly through $Ba^+(5d_\pi^1) \rightarrow CO(\pi^*$ LUMO) backdonation (LUMO, lowest unoccupied molecular orbital), with Ba^+ in the excited ^2D ($5d^1$) electronic reference state. In that respect, the $Ba(CO)^+$ complex behaves similarly to a transition metal carbonyl. The Ba-CO interactions in the radical anion $Ba(CO)^-$ were consistent with dominant contributions of Ba ($5d_\pi$) ← $CO^-(\pi^*$ SOMO) π donation (SOMO, singly occupied molecular orbital) and Ba($5d_\sigma$/6s) ← CO^-(σ HOMO) σ donation (HOMO, highest occupied molecular orbital). The most important valence functions of barium in $Ba(CO)^+$ cation and $Ba(CO)^-$ anion thus appeared to be the 5d orbitals (*7*).

These findings inspired us to search for the 18-electron octacarbonyl complex $Ba(CO)_8$. Surprisingly, we found that not only barium but also the lighter homologs strontium and calcium formed octacarbonyl complexes $M(CO)_8$ (M = Ca, Sr, or Ba) that can be stabilized in a low-temperature neon matrix.

The neutral alkaline earth–carbonyl complexes were prepared by the reactions of pulsed laser-evaporated metal atoms and carbon monoxide (CO) in solid neon and were investigated using Fourier transform infrared absorption spectroscopy. The experiments were carried out with a wide range of CO concentrations (from 0.02 to 2% relative to Ne on the basis of volume). In the experiments with relatively low CO concentrations, terminally bonded mononuclear low-coordinate carbonyl complexes with C≡O stretching frequencies in the 2050 to 1800 cm^{-1} region were observed. Experiments with isotopically substituted CO samples allowed the unambiguous identification of some low-coordinate complexes through isotopic shifts and splittings. The barium di-, tri-, and tetracarbonyls can clearly be identified on the basis of spectra in the experiments with 0.03% $^{12}C^{16}O$ (fig. S1); 0.05% ^{12}CO and 0.05% ^{13}CO (fig. S2); and 0.05% $C^{16}O$ and 0.05% $C^{18}O$ (fig. S3). The monocarbonyls were theoretically predicted to be unstable (*8*) and were not observed in the experimental vibrational spectra. Intense absorption bands centered at 1987 cm^{-1} for Ca, 1995 cm^{-1} for Sr, and 2014 cm^{-1} for Ba were observed upon progressive annealing of the samples to temperatures of 10 to 13 K under relatively high CO concentrations (Table 1). These absorptions become the dominant features in the spectra with high CO concentrations (see Fig. 1A for Ca and figs. S4 and S5 for Sr and Ba), suggesting that the absorber is the coordinatively saturated 18-electron octacarbonyl complex. The observation of only one carbonyl stretching band suggests that these neutral octacarbonyls have the highest cubic O_h symmetry. Experiments with mixtures of $^{12}C^{16}O$ and $^{13}C^{16}O$ and also $^{12}C^{16}O$ and $^{12}C^{18}O$ provided conclusive identification of these cubic octacarbonyl complexes. Although the bands of the Sr and Ba complexes are too broad to resolve isotopic splittings, the band of the Ca complex is sharp and intense in the spectra with relatively low CO concentrations (Fig. 1A); well-resolved mixed isotopic spectra could therefore be compared with calculations. The experimentally observed spectra are in good agreement with the simulated isotopic spectral features shown in figs. S6 and S7.

The radical cations of the alkaline earth–carbonyl complexes were prepared in the gas phase by using a pulsed laser vaporization–supersonic-expansion ion source and studied by mass-selected infrared photodissociation spectroscopy in the carbonyl stretching-frequency region. Typical mass spectra are shown in figs. S8 to S10. Mononuclear metal–carbonyl cation complexes $[M(CO)_n]^+$ (M = Ca, Sr, or Ba) with *n* as high as 10 to 15 were observed. These larger complexes contain both strongly bound CO ligands directly coordinated to the central metal ion and weakly bound, or tagged, CO ligands (*9*, *10*). All of these complexes dissociated by elimination of a CO ligand after photoexcitation at the C≡O stretching vibrational frequencies. The infrared spectra of $[Sr(CO)_n]^+$ (n = 6 to 9) are shown in Fig. 1B and the corresponding spectra of the Ca and Ba complexes in figs. S11 and S12. All the spectra for the $[M(CO)_n]^+$ complexes with n = 6 to 8 feature a broad single band [full width at half maximum (FWHM) of 49 cm^{-1} for Ca, 52 cm^{-1} for Sr, and 29 cm^{-1} for Ba for the n = 8 complexes] that is slightly red shifted relative to the free-CO absorption at 2143 cm^{-1}. The dissociation efficiency increases substantially (about 70% for Ca, 220% for Sr, and 107% for Ba) from n = 8 to 9, and the n = 9 complexes have a much narrower bandwidth [FWHM of 36 cm^{-1} for Ca, 10 cm^{-1} for Sr, and 16 cm^{-1} for Ba]. Along with the intense band around 2100 cm^{-1}, an additional weak band in the 2160 to 2180 cm^{-1} region is also observed for the n = 9 complexes. The bands in this latter frequency region are assigned to the vibrations of weakly tagged external CO ligands (*9*, *10*). The appearance of the tagged CO band at n = 9 indicates that the n = 8 complexes are coordinatively saturated.

We carried out quantum chemical calculations using density functional theory and ab initio

[1]Department of Chemistry, Collaborative Innovation Center of Chemistry for Energy Materials, Shanghai Key Laboratory of Molecular Catalysts and Innovative Materials, Fudan University, Shanghai 200433, China. [2]Institute of Advanced Synthesis, School of Chemistry and Molecular Engineering, Jiangsu National Synergetic Innovation Center for Advanced Materials, Nanjing Tech University, Nanjing 211816, China. [3]Fachbereich Chemie, Philipps-Universität Marburg, Hans-Meerwein-Strasse, D-35043 Marburg, Germany.
*These authors contributed equally to this work.
†Corresponding author. Email: mfzhou@fudan.edu.cn (M.Z.); frenking@chemie.uni-marburg.de (G.F.)

methods to support the assignments of the vibrational spectra to the observed species and to examine the electronic structure of the carbonyl complexes. Figure 2A shows the optimized geometries of the neutral octacarbonyls calculated at the M06-2X-D3/def2-TZVPP level. The molecules have cubic (O_h) symmetry and a triplet ($^3A_{1g}$) electronic ground state with the valence electron configuration $a_{1g}^2t_{1u}^6t_{2g}^6a_{2u}^2e_g^2$. Calculations of the singlet state (fig. S13) gave structures with D_{4d} (Ca and Sr) or D_{4h} symmetry (Ba), which were between 6.5 and 7.5 kcal/mol higher in energy than the corresponding triplet state species. Figure 2, B and C, shows the equilibrium geometries of the cations $[M(CO)_8]^+$, which resemble the neutral species in the singlet state. Thus, $[Ca(CO)_8]^+$ and $[Sr(CO)_8]^+$ possess a D_{4d} structure and a $^2A_1$ electronic ground state, whereas $[Ba(CO)_8]^+$ has D_{4h} symmetry and a $^2B_{2g}$ electronic state. Figure 2 also shows the calculated zero-point energy (ZPE)–corrected bond dissociation energies (D_0) of the octacarbonyls for loss of one and eight CO ligand(s). At the M06-2X-D3/def2-TZVPP level, the D_0 values for the dissociation of one CO lie between 9.1 kcal/mol for Ca and 11.5 kcal/mol for Sr for the $M(CO)_8$ neutral complexes and between 8.4 kcal/mol for Ba and 9.6 kcal/mol for Sr for the $[M(CO)_8]^+$ cation complexes. The theoretical D_0 values for loss of eight CO ligands yielding M/M^+ in the electronic ground state were between 58.8 kcal/mol for Sr and 63.3 kcal/mol for Ca for the neutrals and between 73.5 kcal/mol for Ba and 87.6 kcal/mol for Ca for the cations. The calculated values at the CCSD(T)/def2-TZVPP level using the M06-2X-D3/def2-TZVPP optimized geometries were slightly smaller. The basis set superposition error is quite small (0.3 to 0.4 kcal/mol for single CO dissociation and 0.6 to 1.0 kcal/mol for loss of eight COs).

The theoretical wave numbers for the C≡O stretching modes of $M(CO)_8$ and $[M(CO)_8]^+$ are shown in Table 1 along with the experimental values. The calculated values refer to the harmonic antisymmetric stretching frequencies scaled by a factor of 0.941, which comes from the ratio of the calculated stretching mode of free CO (2278 cm^{-1}) to the experimental value (2143 cm^{-1}) (*11*). The calculated harmonic wave numbers are slightly higher than the experimental anharmonic values, but the trends for the different metals and for the isotopes of the neutral systems are in excellent agreement with the recorded values. The calculations suggest only one infrared (IR) active mode for the neutral complexes $M(CO)_8$ and two closely lying IR active modes for the cations $[M(CO)_8]^+$. The latter splitting is too small to be experimentally observed. Theory and experiment indicate a considerable red shift of the C≡O frequency for the neutral systems and a much smaller red shift for the cations. The calculated red shifts and the isotopic variations of the calculated and experimental wave numbers match each other.

Figure 3 shows the splitting of the valence orbitals of atoms with a spd valence shell in an octacoordinate cubic field with O_h symmetry (*12*). The breakdown of the AOs into irreducible representations of the O_h point group in an eight-coordinate ligand field is similar to the splitting in a six-coordinate octahedral field (*13*, *14*), but there are two important differences. One concerns the splitting of the $(n-1)$d AOs of the metal. The degenerate e_g AOs in the octacoordinate cubic field are the $(n-1)d_\pi$ AOs, and the triply degenerate t_{2g} AOs are the $(n-1)d_\sigma$ AOs. By contrast, in the six-coordinate octahedral field, the degenerate e_g AOs are the $(n-1)d_\sigma$ AOs, and the triply degenerate t_{2g} AOs have $(n-1)d_\pi$ character. The second difference concerns the appearance of the a_{2u} molecular orbital (MO) in the cubic field (Fig. 3), which is absent in the octahedral field. The a_{2u} MO is a ligand-only orbital, because there is no valence AO of the spd shell that possesses this symmetry. This has consequences for the number of electrons that are required to fulfill the 18-electron rule. Because two valence electrons of the ligands in a cubic field are not available for donation to the central metal atom M in ML_8, where L is a ligand, the complex must provide a total of 20 electrons to fully occupy the metal's valence shell. This scenario has recently been explored in the transition

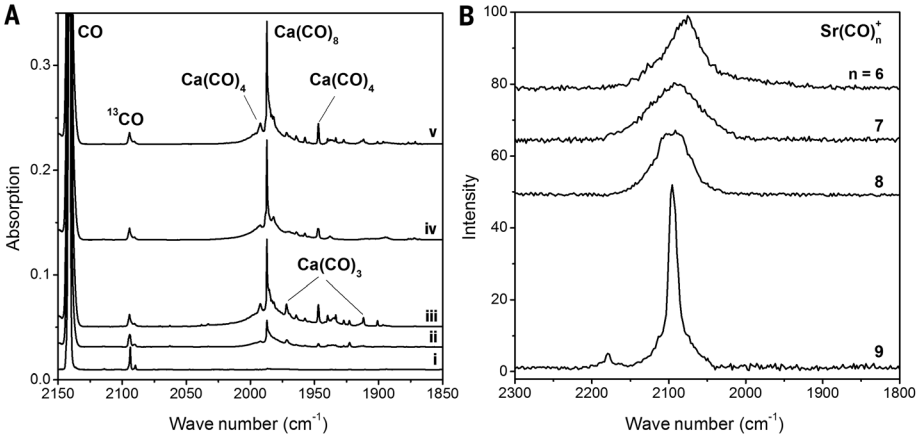
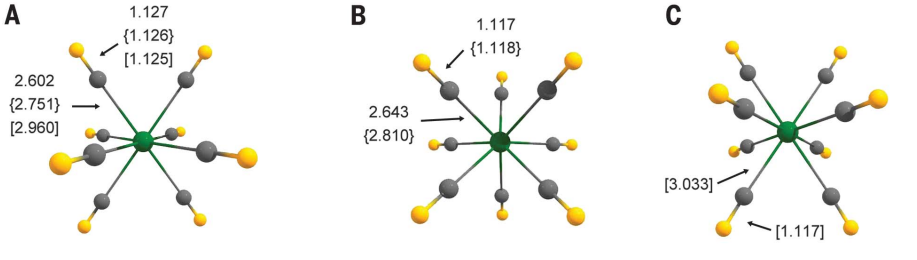


Fig. 1. IR spectra of alkaline earth carbonyl complexes. (**A**) IR absorption spectra of calcium-carbonyl complexes in the 2150 to 1850 cm^{-1} region from codeposition of laser-evaporated calcium atoms with 0.1% CO in neon. Spectral lines: (i) after 30 min of sample deposition at 4 K, (ii) after a 12-K annealing period, (iii) after a 13-K annealing period, (iv) after 15 min of visible light irradiation, and (v) after another 12-K annealing period. (**B**) IR photodissociation spectra of the $Sr(CO)_n^+$ (n = 6 to 9) complexes in the 2300 to 1800 cm^{-1} region.



D_0 = 9.1 (9.0) *63.3 (42.6)* kcal/mol, M = Ca
D_0 = 11.5 (11.1) *58.8 (41.5)* kcal/mol, M = {Sr}
D_0 = 10.4 (8.7) *62.8 (49.2)* kcal/mol, M = [Ba]

D_0 = 9.5 (9.2) *87.6 (76.0)* kcal/mol, M = Ca
D_0 = 9.6 (9.7) *77.8 (68.4)* kcal/mol, M = {Sr}

D_0 = 8.4 (8.5) *73.5 (66.2)* kcal/mol

Fig. 2. Calculated equilibrium geometries of alkaline earth octacarbonyls. (**A**) $M(CO)_8$ (M = Ca, Sr, or Ba), (**B**) $[M(CO)_8]^+$ (M = Ca or Sr), and (**C**) $[Ba(CO)_8]^+$. Bond lengths are in angstroms. The D_0 values in roman type are the ZPE-corrected bond dissociation energies for loss of one CO ligand; the italicized values are the corresponding energies for the loss of eight CO ligands and M/M^+ in the ground state. The values without parentheses are from M06-2X-D3/def2-TZVPP calculations; the values in parentheses are from CCSD(T)/def2-TZVPP using M06-2X-D3/def2-TZVPP optimized geometries.

metal octacarbonyl anions $[TM(CO)_8]^-$ (TM = Sc, Y, or La) (*15*). Only 18 electrons are available in $M(CO)_8$ (M = Ca, Sr, or Ba), so the degenerate e_g MO is occupied by two electrons with the same spin, giving a triplet ($^3A_{1g}$) electronic ground state. Because the e_g correlates with the $(n-1)d_\pi$ AO, it becomes clear that the electronic reference state of the alkaline earth atom in $M(CO)_8$ is a triplet state with $ns^0(n-1)d^2$ electron configuration. The metal center has a zero formal oxidation state.

We analyzed the nature of the metal-CO bonds with the energy decomposition analysis–natural orbitals for chemical valence (EDA-NOCV) (*16*) method, a powerful tool that provides detailed insight into chemical bonding (*17*). A description is given in the methods section. Table 2 shows the calculated results for the interactions between the metal atom M in the triplet electronic reference state with a $ns^0(n-1)d^2$ electron configuration and the $(CO)_8$ cage at the frozen geometry of $M(CO)_8$. The interaction energies ΔE_{int} suggest that the intrinsic attraction M-$(CO)_8$ is strong and varies in the order Ca > Sr >> Ba. The dominant contribution to the total interaction ΔE_{int} comes from the orbital term ΔE_{orb}. The preparation energies (ΔE_{prep}) involved in the formation of the $(CO)_8$ cage from free CO molecules are low, ranging from 3.3 kcal/mol for Ba to 13.9 kcal/mol for Ca, whereas the electronic excitation energies for the atoms into the spherically symmetric $ns^2 \rightarrow ns^0(n-1)d^2$ triplet state of M are quite high, lying between 68.2 kcal/mol for Ba and 159.5 kcal/mol for Ca. The strongly stabilizing interaction energies ΔE_{int} caused by the attraction between the CO ligands and electronically excited M overcompensate the ΔE_{prep} values. The adiabatic interaction energy ΔE (=$-D_e$, where D_e is the dissociation energy without zero-point vibrational energy correction) with respect to the electronic ground state of M and eight CO is between −65.5 kcal/mol for Sr and −73.7 kcal/mol for Ba. The D_e values in Table 2 exhibit a similar trend as the D_0 data in Fig. 2. The former values do not consider ZPE contributions, and they are obtained from calculations with different basis sets using Slater-type orbital basis functions.

The most important insight from the EDA-NOCV calculations is the breakdown of ΔE_{orb} into pairwise orbital interactions. Table 2 shows that the metal-CO bonding comes mainly from the $[M(d_\pi)] \rightarrow (CO)_8$ π backdonation of the degenerate (e_g) set of singly occupied $(n-1)$d AOs of the metal into the antibonding π* MOs of CO. This explains the large red shift of the C≡O stretching mode of the octacarbonyls. The contribution of the $[M(x)] \leftarrow (CO)_8$ σ donation into x, where x denotes the valence acceptor AO of M, is much smaller than the $[M(d_\pi)] \rightarrow (CO)_8$ π backdonation. The order of the acceptor AOs of the metal atoms for the interaction energy is $(n-1)d_\sigma > ns > np$. There is also a small stabilizing contribution of the a_{2u} MO, which is due to the polarization of the $(CO)_8$ ligand orbitals.

The dominating orbital interactions by the valence d electrons of M can be explained with the energetically very-high-lying occupied orbitals, which make M atoms excellent donor species. In our previous study of $Ba(CO)^+$, we found that the cation Ba^+ ($5d_\pi^1$) is a donor to neutral CO; the atomic partial charge of Ba in $Ba(CO)^+$ was calculated as +1.39 *e*. The experimental values for the energetically lowest-lying ^3F state with the electron configuration $ns^0(n-1)d^2$ are not far from the ionization limit (*18*). The valence d electrons of the M atoms are only weakly bonded to the atoms.

Figure 3 also shows the occupied valence MOs of $Ca(CO)_8$ that are relevant for the metal-CO bonding. The shape of the five (*19*) MOs reveals that the contributions of the metal AOs are very small, except in the SOMO, which clearly exhibits the shape of the $(n-1)d_\pi$ AOs. The valence MOs of

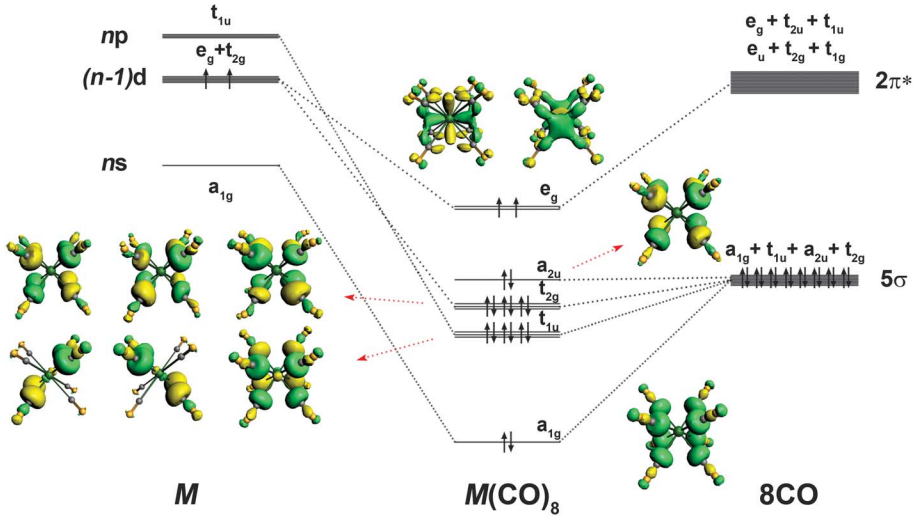


Fig. 3. Bonding scheme and shape of the occupied valence orbitals of $M(CO)_8$ (M = Ca, Sr, or Ba). Splitting of the spd valence orbitals of an atom M with the configuration $(n-1)d^2ns^0np^0$ in the octacoordinate cubic (O_h) field of eight CO ligands is also given. Only the occupied valence orbitals that are relevant for the M-CO interactions are shown. Up and down arrows indicate electrons with opposite spin.

Table 1. Calculated and experimental wave numbers. Calculated (M06-2X-D3/def2-TZVPP) and experimental wave numbers (cm^{-1}) of the C-O stretching mode and frequency shifts of $M(CO)_8$ and $[M(CO)_8]^+$ (M = Ca, Sr, or Ba). The calculated values are scaled by 0.941. Blank spaces indicate that isotope values are not available.

Complex	Calculated wave numbers						Experimental wave numbers					
	CO	Δ*	^{13}CO	Δ†	$C^{18}O$	Δ†	CO	Δ*	^{13}CO	Δ†	$C^{18}O$	Δ†
$Ca(CO)_8$	2018	−125	1974	−44	1969	−49	1987	−156	1944	−43	1939	−48
$Sr(CO)_8$	2027	−116	1982	−45	1977	−50	1995	−148	1952	−43	1946	−49
$Ba(CO)_8$	2044	−99	1999	−45	1994	−50	2014	−129	1970	−44	1965	−49
$[Ca(CO)_8]^+$	2116 2117	−27 −26					2094	−49				
$[Sr(CO)_8]^+$	2113 2119	−30 −24					2096	−47				
$[Ba(CO)_8]^+$	2127 2129	−16 −14					2113	−30				

*Frequency shift relative to free CO. †Isotopic frequency shift.

Table 2. EDA-NOCV results. EDA-NOCV results for triplet state $M(CO)_8$ (M = Ca, Sr, or Ba) complexes at the M06-2X/TZ2P-ZORA level using M06-2X-D3/def2-TZVPP optimized geometries, taking $(CO)_8$ in singlet ground state and M in triplet excited state with a $ns^0(n-1)d^2$ valence electronic configuration as interacting fragments. Energy values are given in kcal/mol.

Energy term	Assignment	Interacting fragments		
		Ca + $(CO)_8$	Sr + $(CO)_8$	Ba + $(CO)_8$
ΔE_{int}		−243.1	−224.1	−145.2
ΔE_{hybrid}*		41.8	46.9	37.4
ΔE_{Pauli}		19.5	30.5	30.8
ΔE_{elstat}†		−65.3 (21.5%)	−61.7 (20.5%)	−78.5 (36.8%)
ΔE_{orb}†		−239.1 (78.5%)	−239.7 (79.5%)	−135.0 (63.2%)
$\Delta E_{orb(1)}$‡§ (e_g)	[M(d)] → $(CO)_8$ π backdonation	−206.2 (86.2%)	−206.4 (86.2%)	−95.0 (69.8%)
$\Delta E_{orb(2)}$‡§ (t_{2g})	[M(d)] ← $(CO)_8$ σ donation	−21.3 (9.0%)	−20.7 (8.7%)	−22.8 (16.8%)
$\Delta E_{orb(3)}$‡ (a_{1g})	[M(s)] ← $(CO)_8$ σ donation	−2.4 (1.0%)	−2.9 (1.2%)	−3.2 (2.4%)
$\Delta E_{orb(4)}$‡§ (t_{1u})	[M(p)] ← $(CO)_8$ σ donation	−0.9 (0.3%)	−0.9 (0.3%)	−2.7 (2.1%)
$\Delta E_{orb(5)}$‡ (a_{2u})	$(CO)_8$ polarization	−0.6 (0.3%)	−1.0 (0.4%)	−2.3 (1.7%)
$\Delta E_{orb(rest)}$‡		−7.7 (3.2%)	−7.8 (3.3%)	−9.0 (6.7%)
$\Delta E_{prep(a)}$	8 CO → $(CO)_8$	13.9	7.7	3.3
$\Delta E_{prep(b)}$‖	M, $ns^2 \rightarrow ns^0(n-1)d^2$ (T)	159.5	150.9	68.2
ΔE ($\Delta E_{int} + \Delta E_{prep}$) = $-D_e$	$M(CO)_8$ → M (^{1}S) + 8 CO	−69.7	−65.5	−73.7

*Contribution of the metahybrid term in M06-2X. †The values in parentheses show the contribution to the total attractive interactions ΔE_{elstat} plus ΔE_{orb}, where ΔE_{elstat} is the electrostatic interaction energy. ‡The values in parentheses show the contribution to the total orbital interaction, ΔE_{orb}. §The sum of the two (e_g) or three (t_{2g} or t_{1u}) components is given. ‖The EDA calculations give a triplet state with spherically symmetrical distribution of the d electrons. The experimental values for excitation into the energetically lowest-lying ^{3}F state with $ns^0(n-1)d^2$ configuration are 124.2 and 59.8 kcal/mol for Ca and Ba, respectively. There is no experimental value for the relevant ^{3}F state of Sr. The data are taken from (*18*).

the heavier homologs $Sr(CO)_8$ and $Ba(CO)_8$ look very similar (see figs. S14 and S15). The effect of the orbital interactions on the charge distribution is evident from the shape of the deformation densities Δρ, which are associated with the orbital interactions. Figure S16 shows the contour line plots of the deformation densities $\Delta\rho_{(1)}$ to $\Delta\rho_{(5)}$, which are connected to the pairwise interactions $\Delta E_{orb(1)}$ to $\Delta E_{orb(5)}$ in $Ca(CO)_8$ (Table 2). Note that only one component of the orbital terms is shown, and the color code of the charge flow is from red to blue. Figure S16A displays a large charge flow in the direction $[Ca(d_\pi)] \rightarrow (CO)_8$, which comes from the π backdonation. Figure S16, B to D, exhibits the charge flow in the opposite direction $[Ca(x)] \leftarrow (CO)_8$ into the valence AOs of calcium $(n-1)d_\sigma$, ns, and np. Figure S16E shows the charge polarization of the $(CO)_8$ a_{2u} ligand orbital, where electronic charge is shifted from oxygen toward carbon. The deformation densities $\Delta\rho_{(1)}$ to $\Delta\rho_{(5)}$ of the orbital interactions in $Sr(CO)_8$ and $Ba(CO)_8$ are shown in figs. S17 and S18, respectively.

The analysis of the electronic structure of the octacarbonyls $M(CO)_8$ provides a straightforward explanation for why the molecules have a cubic (O_h) equilibrium geometry and a triplet ($^3A_{1g}$) electronic ground state with a triplet reference state and $ns^0(n-1)d^2$ electron configuration. The coordination by eight CO ligands in a cubic field fills the spd valence AOs of the alkaline earth atoms and the a_{2u} ligand-only MO. The energetically highest-lying bonding MO is the degenerate e_g orbital. Because 16 electrons from the ligands and 2 electrons from metal are available, the e_g MO has SOMO components following Hund's rule. Simple electron counting indicates that the octacarbonyls $M(CO)_8$ fulfill the 18-electron rule. The EDA-NOCV calculations of the cations $[M(CO)_8]^+$ agree with the bonding analysis of the neutral complexes. The $[M(d_\pi)]^+ \rightarrow (CO)_8$ π backdonation is weaker than in the neutral species, because there is only one $(n-1)d_\pi$ electron available (tables S1 and S2).

The bonding situation in the alkaline earth–octacarbonyl complexes shows that not only barium but also strontium and calcium may effectively use their $(n-1)$d AOs in chemical bonding. It is conceivable that the chemical reactivity of heavier alkaline earth elements is more diverse than hitherto thought. Recent reports about unusual structures and reactivities of calcium and strontium compounds (*20*, *21*) could be a hint that d-orbital participation of the alkaline earth metals plays an important role. The design of future experiments should consider the capacity of the heavier alkaline earth elements to behave like transition metals.

REFERENCES AND NOTES

1. I. Langmuir, *Science* **54**, 59–67 (1921).
2. I. Langmuir, *J. Am. Chem. Soc.* **41**, 868–934 (1919).
3. P. Pyykkö, *J. Organomet. Chem.* **691**, 4336–4340 (2006).
4. N. Wiberg, *Holleman-Wiberg: Inorganic Chemistry* (Academic Press, 2001).
5. M. Kaupp, *Angew. Chem. Int. Ed.* **40**, 3534–3565 (2001).
6. L. Gagliardi, P. Pyykkö, *Theor. Chem. Acc.* **110**, 205–210 (2003).
7. X. Wu *et al.*, *Angew. Chem. Int. Ed.* **57**, 3974–3980 (2018).
8. Geometry optimizations of M(CO) using density functional theory and ab initio methods gave only weakly bonded van der Waals complexes with long M-CO distances.
9. A. M. Ricks, Z. E. Reed, M. A. Duncan, *J. Mol. Spectrosc.* **266**, 63–74 (2011).
10. G. J. Wang, M. F. Zhou, *Chin. J. Chem. Phys.* **31**, 1–11 (2018).
11. K. P. Huber, G. Herzberg, *Molecular Spectra and Molecular Structure IV. Constants of Diatomic Molecules* (Van Nostrand-Reinhold, 1979).
12. J. K. Burdett, R. Hoffmann, R. C. Fay, *Inorg. Chem.* **17**, 2553–2568 (1978).
13. A. Diefenbach, F. M. Bickelhaupt, G. Frenking, *J. Am. Chem. Soc.* **122**, 6449–6458 (2000).
14. T. A. Albright, J. K. Burdett, M. H. Whangbo, *Orbital Interactions in Chemistry* (Wiley, ed. 2, 2013).
15. J. Jin *et al.*, *Angew. Chem. Int. Ed.* **57**, 6236–6241 (2018).
16. M. P. Mitoraj, A. Michalak, T. Ziegler, *J. Chem. Theory Comput.* **5**, 962–975 (2009).
17. L. Zhao, M. von Hopffgarten, D. M. Andrada, G. Frenking, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1345 (2018).
18. National Institute of Standards and Technology, NIST atomic spectra database levels form; https://physics.nist.gov/PhysRefData/ASD/levels_form.html.
19. The doubly or triply degenerate orbitals SOMO (e_g), HOMO-1 (t_{2g}), and HOMO-2 (t_{1u}) are single MOs, which have two and three components, respectively.
20. A. S. S. Wilson, M. S. Hill, M. F. Mahon, C. Dinoi, L. Maron, *Science* **358**, 1168–1171 (2017).
21. B. Maitland *et al.*, *Angew. Chem. Int. Ed.* **56**, 11880–11884 (2017).

ACKNOWLEDGMENTS

Funding: The authors gratefully acknowledge financial support from the National Natural Science Foundation of China (grant nos. 21688102, 21433005, and 21703099), Nanjing Tech University and SICAM Fellowship from Jiangsu National Synergetic Innovation Center for Advanced Materials (L.Z. and G.F., grant nos. 39837132 and 39837123), and the Natural Science Foundation of Jiangsu Province for Youth (grant no. BK20170964). S.P. thanks Nanjing Tech University for a postdoctoral fellowship and the High Performance Computing Center of Nanjing Tech University for supporting the computational resources. **Author contributions:** X.W. did the matrix isolation experiments. J.J., W.L., X.J., and G.W. did the gas-phase infrared photodissociation experiments. M.Z. contributed to conception and design of experiments;

data analysis and interpretation; and drafting and critical revision of the manuscript. S.P. carried out the calculations and contributed to data collection and analysis of the results. L.Z. and G.F. contributed to conception and analysis of the theoretical results and drafting and critical revision of the manuscript. **Competing interests:** The authors have no competing interests. **Data and materials availability:** All data for this work are provided in the manuscript and in the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/912/suppl/DC1
Materials and Methods
Figs. S1 to S19
Tables S1 to S3
References (*22*–*40*)
Data S1

6 May 2018; accepted 13 July 2018
10.1126/science.aau0839

CLIMATE CHANGE

Increase in crop losses to insect pests in a warming climate

Curtis A. Deutsch[1,2]*†, Joshua J. Tewksbury[3,4,5]†, Michelle Tigchelaar[6], David S. Battisti[6], Scott C. Merrill[7], Raymond B. Huey[2], Rosamond L. Naylor[8]

Insect pests substantially reduce yields of three staple grains—rice, maize, and wheat—but models assessing the agricultural impacts of global warming rarely consider crop losses to insects. We use established relationships between temperature and the population growth and metabolic rates of insects to estimate how and where climate warming will augment losses of rice, maize, and wheat to insects. Global yield losses of these grains are projected to increase by 10 to 25% per degree of global mean surface warming. Crop losses will be most acute in areas where warming increases both population growth and metabolic rates of insects. These conditions are centered primarily in temperate regions, where most grain is produced.

By 2050, growing-season temperatures will likely exceed those recorded during the past century and may substantially reduce crop yields (*1*–*4*). However, models assessing the effects of climate warming on crop yields rarely consider impacts on insect pests, despite the damages that result directly from pest infestations and indirectly from pesticides applied to reduce pest damage (*5*, *6*). In the future, pest species are likely to differ in their responses to warming, changing the relative impacts of pests geographically and among crops (*7*, *8*). Here we use well-established relationships between temperature and the physiology and demography of insects to project the future impact of insects on crop production globally and regionally. We estimate pest-related changes in yields of the major grain crops maize, rice, and wheat, which together account for 42% of direct calories consumed by humans worldwide (*9*).

A warmer climate will alter at least two agriculturally relevant characteristics of insect pests. First, an individual insect's metabolic rate (M) accelerates with temperature, and an insect's rate of food consumption must rise accordingly (*10*–*12*). Second, the number of insects (n) will change, because population growth rates of insects also vary with temperature. These growth rates are expected to decline as a result of warming in tropical regions while rising elsewhere (*8*) (fig. S1). The total energy consumption of a pest population (the "population metabolism") is proportional to the product of these two factors and directly relates to the crop yield loss (L) caused by insect herbivory. Fractional changes in pest-induced crop losses ($\Delta L/L$) can thus be partitioned into a metabolic component ($\Delta M/M$) and a demographic component ($\Delta n/n$) (*13*). The sum of these fractional changes approximates the total fractional change in yield loss

$$\frac{\Delta L}{L} = \frac{\Delta M}{M} + \frac{\Delta n}{n} \quad (1)$$

To evaluate how warming changes the population metabolism of insect pests, we integrated established physiological responses of insects to temperature into a spatially explicit demographic model (*13*). The metabolic and population growth rates were derived from laboratory experiments across a wide range of temperatures and for diverse insect taxa including pest species. Relationships between temperature and insect population growth rates drive logistic population increases of insects during each crop's growing season, and they also scale the fractional survival rate of insects over the rest of the year (*14*), termed the diapause survival, ϕ_o. We calibrated key demographic model parameters—population size and carrying capacity—using contemporary crop yields (*15*) and their insect-related losses, measured for our three focal crops at sites around the world (*5*). To predict future changes in population growth and metabolic rates, we added projected monthly surface temperature anomalies from climate model simulations under a "business-as-usual" emissions scenario (RCP8.5) (*16*) to the observed daily and seasonally varying temperatures from the 20th century (1950 to 2000). Results are presented for several climate models that span a range of climate sensitivities and for a range of uncertainties in biological traits and assumptions (*13*). We report yield losses as a function of global mean surface temperature change, making the results comparable across emissions scenarios, time periods, and climate sensitivities.

Crop production losses to pests increase globally with rising temperatures in all climate models and across all biological parameters (Fig. 1). When average global surface temperatures increase by 2°C, the median increase in yield losses owing to pest pressure is 46, 19, and 31% for wheat, rice, and maize, respectively, bringing total estimated losses to 59, 92, and 62 metric megatons per year. These projected losses are similar across all climate models and are thus robust to uncertainties in both global and regional warming patterns, although the time at which such damage levels are reached depends on the emissions scenario and on each model's sensitivity to increasing atmospheric CO_2 (Fig. 1D)

The differences in global grain losses between crops and across model parameters (Fig. 1) reflect the distinct spatial patterns of demographic and metabolic impacts of warming on insect pests in the climates where these crops are grown. In temperate regions, warming increases both the size of insect populations and their per capita metabolic rate (Fig. 2, right). As a result, the increase in pest-related crop loss is consistently larger than in tropical regions, where the increasing metabolic rate is offset by declining population growth rates, resulting in a smaller overall rise in crop damages. This broad geographic pattern holds across all crops, climate models, and life history parameters considered (Fig. 2 and figs. S3 and S4).

The contribution of per capita metabolic rates to the total pest-induced crop losses is projected to increase consistently across regions and over time. For each of the three crops examined here, increases in temperature vary only modestly across growing regions and seasons, causing a nearly uniform fractional rise in the metabolic rates of the insect pests (Fig. 1). The magnitude of the metabolic component (Eq. 1) is proportional to the temperature sensitivity of metabolic rates, E_{met}, which varies by <50% across insect species (E_{met} = 0.65 ± 0.15; mean ± standard deviation) (*12*). As a result, the metabolic component of insect pest population metabolism can be estimated relatively robustly at both regional and global scales.

In contrast, the demographic component of future crop loss to insect pests is spatially variable and can either exacerbate or ameliorate the impact of rising metabolic rates (Fig. 1 and figs. S3 and S4). In the lowland tropics, pest populations are predicted to decline because current temperatures there are already near optimal, so warming should reduce population growth rates (*8*) (fig. S2). On the other hand, extratropical pest populations are generally projected to grow as temperatures become closer to optimal, with a small contribution from increasing diapause survival as winters warm (*14*) (fig. S6). Because temperate populations often reach carrying capacity only late in the growing season, if at all, they have the most potential for increases in population size as temperature rises (fig. S2). How much they increase depends on baseline survival rates

[1]School of Oceanography, University of Washington, Seattle, WA 98195, USA. [2]Department of Biology, University of Washington, Seattle, WA 98195, USA. [3]Future Earth, University of Colorado, Boulder, CO 80303, USA. [4]Department of Environmental Studies, University of Colorado, Boulder, CO 80303, USA. [5]School of Global Environmental Studies, Colorado State University, Fort Collins, CO 80523, USA. [6]Department of Atmospheric Sciences, University of Washington, Seattle, WA 98195, USA. [7]Department of Plant and Soil Science, University of Vermont, Burlington, VT 05405, USA. [8]Department of Earth System Science and the Center on Food Security and the Environment, Stanford University, Stanford, CA 94305, USA.
***Corresponding author. Email: cdeutsch@uw.edu**
†These authors contributed equally to this work.

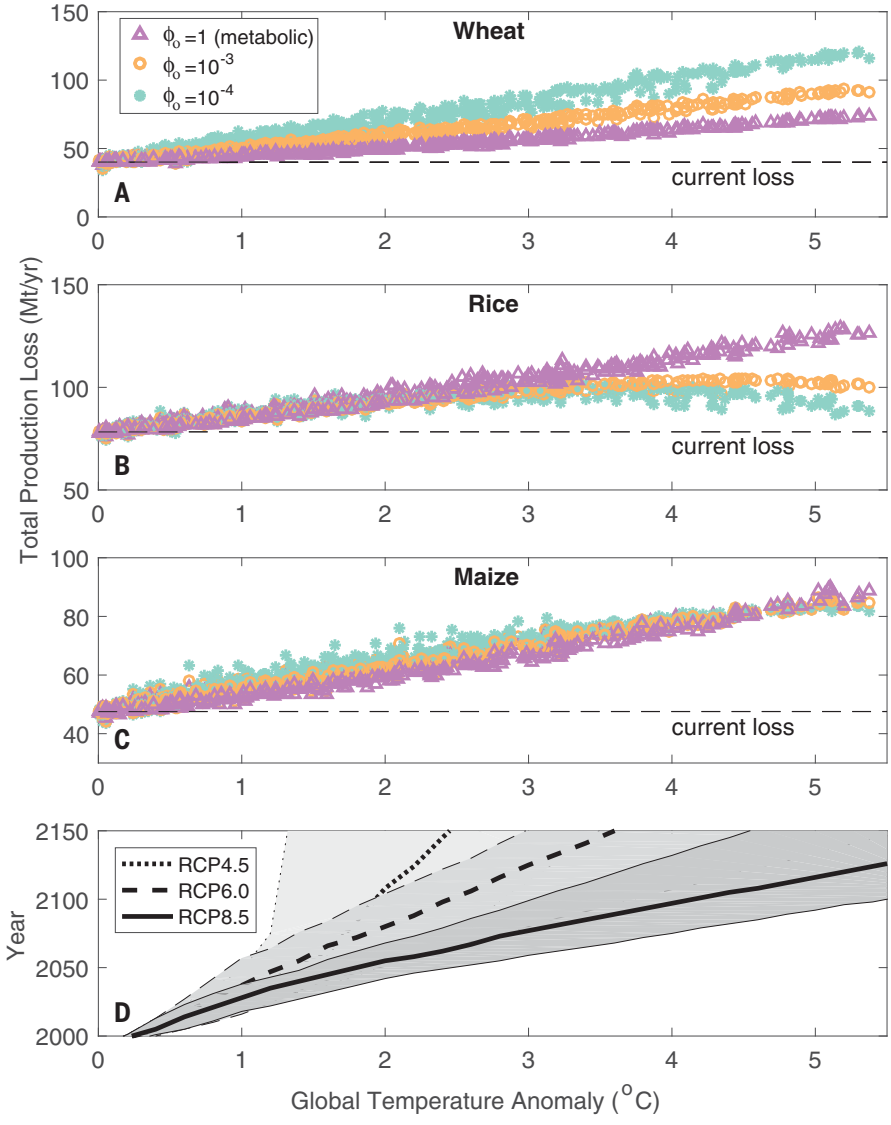


Fig. 1. Global loss of crop production owing to the impact of climate warming on insect pests. Crop production losses for (**A**) wheat, (**B**) rice, and (**C**) maize are computed by multiplying the fractional change in population metabolism by the estimated current yield loss owing to insect pests, summed over worldwide crop locations. Results are plotted versus mean global surface temperature change, for four climate models (*13*), for two different values of the demographic parameter governing survival during diapause (ϕ_o = 0.0001, asterisks; ϕ_o = 0.001, circles), and for the metabolic effect alone (triangles). Mt/yr, metric megatons per year. The year in which a given global mean temperature anomaly is reached (**D**) depends on the greenhouse gas emissions scenario (RCP, representative concentration pathway) and varies across models (shading) owing to uncertainty in climate sensitivity to those emissions (*13*).

during the nongrowing season (ϕ_o), which can be highly variable. However, the pattern of weak demographic impacts in tropical regions and strong deleterious impacts in northern temperate regions is consistent across a wide range of plausible ϕ_o values, from 0.0001 to 0.01 (fig. S3).

Because our three focal crops are grown in different climates, where warming can induce opposite changes in insect population growth rates, diapause survival differentially affects losses of these three crops. For wheat, which is typically grown in relatively cool climates, warming will increase pest population growth and overwinter survival rates, leading to large population increases in the growing season (Fig. 1A). In rice, which is grown in relatively warm tropical environments, the same population dynamic has the opposite impact; warming there should reduce insect population growth rates and thus partly counteract the rising crop losses due to increased insect metabolism, allowing global rice production lost to insects to stabilize for warming exceeding ~3°C (Fig. 1B). For maize, the demographic effect has only a small net impact on global production losses, because this crop is grown in some regions where population rates will increase and in other regions where population rates will decline, in nearly equal measure (Fig. 1C).

The spatial patterns of modeled changes in insect population metabolism also predict differential impacts across major geopolitical boundaries (Fig. 3). The most substantial yield declines will occur in many of the world's most productive agricultural regions, thus reducing global grain availability (Fig. 3 and table S5). France, the United States, and China—countries that produce most of the world's maize—are also among the countries projected to experience the largest increases in pest-related crop losses (Figs. 1C and 3C). These countries have among the highest yields per hectare today (Fig. 3). In addition, France and China are responsible for a considerable fraction of global wheat and rice production, respectively, and are projected to suffer large increases in yield loss of these grains owing to climate impacts on pests (Figs. 1C and 3C and table S5).

Our analysis focuses on the changing impacts of insect pests on crop yields with an increase in global temperature, accounting for the most robust general responses of insect pests to temperature. The full scope of physiological and ecological impacts is likely to be complex and sensitive to particular crop-pest interactions for which more physiological data will be needed, especially among tropical pest species (fig. S1). These interactions will occur in conjunction with direct plant responses to warming and rising CO_2 levels, which, for the three major crops that we considered, are predominantly negative (*17*). However, scenarios with added or alternative biological dynamics, such as thermoregulation by insects (*18*) or increased diapause mortality with warming (*19*), suggest that the dominant patterns described here are robust (figs. S5 and S6), and species-specific predictions for pests that affect these three crops generally agree with our predictions (*13*).

Agricultural practices will shift as the climate warms. Changes in planting dates, cultivar use, and planting locations are already under way (*20*) and will become more pronounced as the rate of climate warming increases (*21*). Our results suggest that farmers will need to make additional changes, such as introducing new crop rotations, to maintain yields in the face of rising insect pest pressure. In intensive agricultural environments, adaptation measures may involve greater pesticide use, at the cost of associated health and environmental damage and the elevated threat of pesticide resistance. Without wider attention to how climate warming will affect crop breeding and sustainable pest management strategies, insect-driven yield losses will result in reduced global grain supplies and higher staple food prices. Poor grain consumers and farming households, who account for a large share of the world's 800 million people living in chronic hunger (*9*), will suffer most.

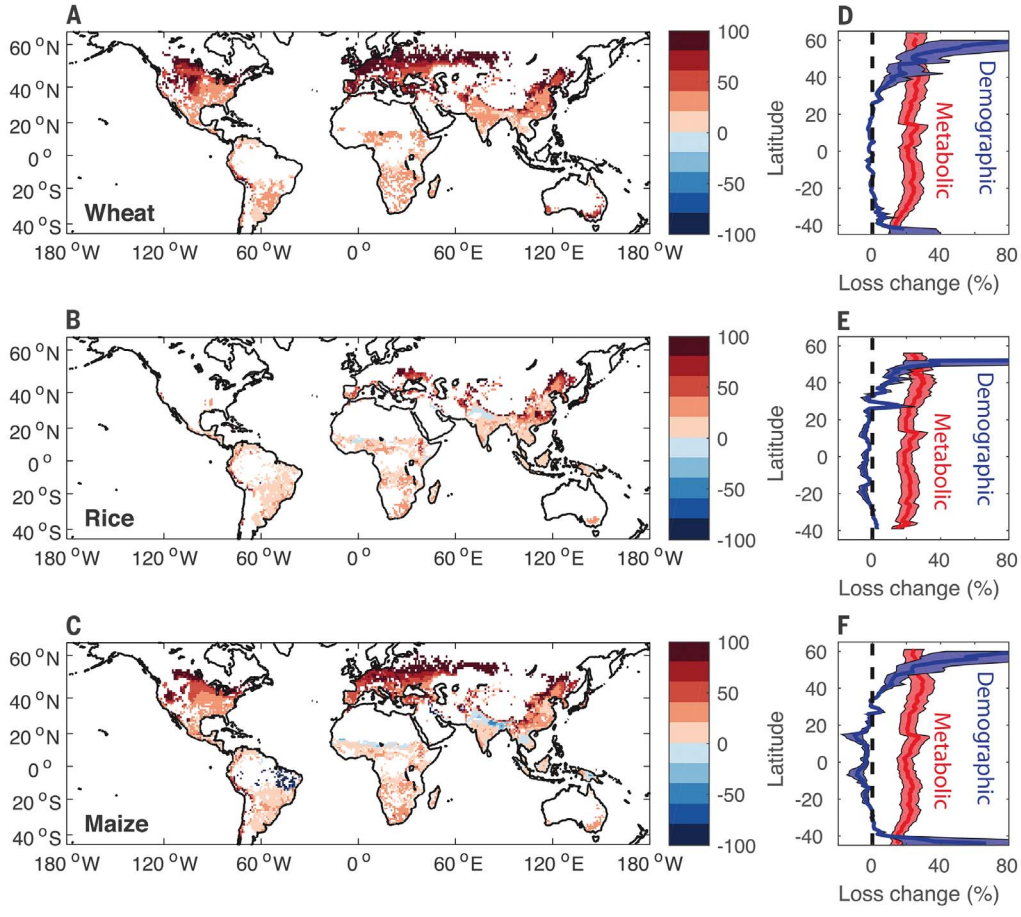


Fig. 2. Projected geographic pattern of change in crop yield losses to insect pests in a 2°C-warmer climate. Results are mapped for the fractional (percent) increase in crop yield loss owing to pests from both metabolic and demographic effects ($\Delta L/L$) for (**A**) wheat, (**B**) rice, and (**C**) maize. The zonal median change is plotted for the separate contribution of demographic effects ($\Delta n/n$, blue) and metabolic effects ($\Delta M/M$, red) for (**D**) wheat, (**E**) rice, and (**F**) maize. Results are shown for a range of life history traits in the longitudinal average curves (right panels). The metabolic effect uses activation energies (E_o) with a mean (0.65 eV) and standard deviation (±0.15 eV) for insects (*12*). The demographic effect assumes a range of ϕ_o values from 0.0001 to 0.01. All results are averaged over multiple climate models (*13*), in all years when the global mean surface temperature is 2 ± 0.1°C greater than in the late 20th century.

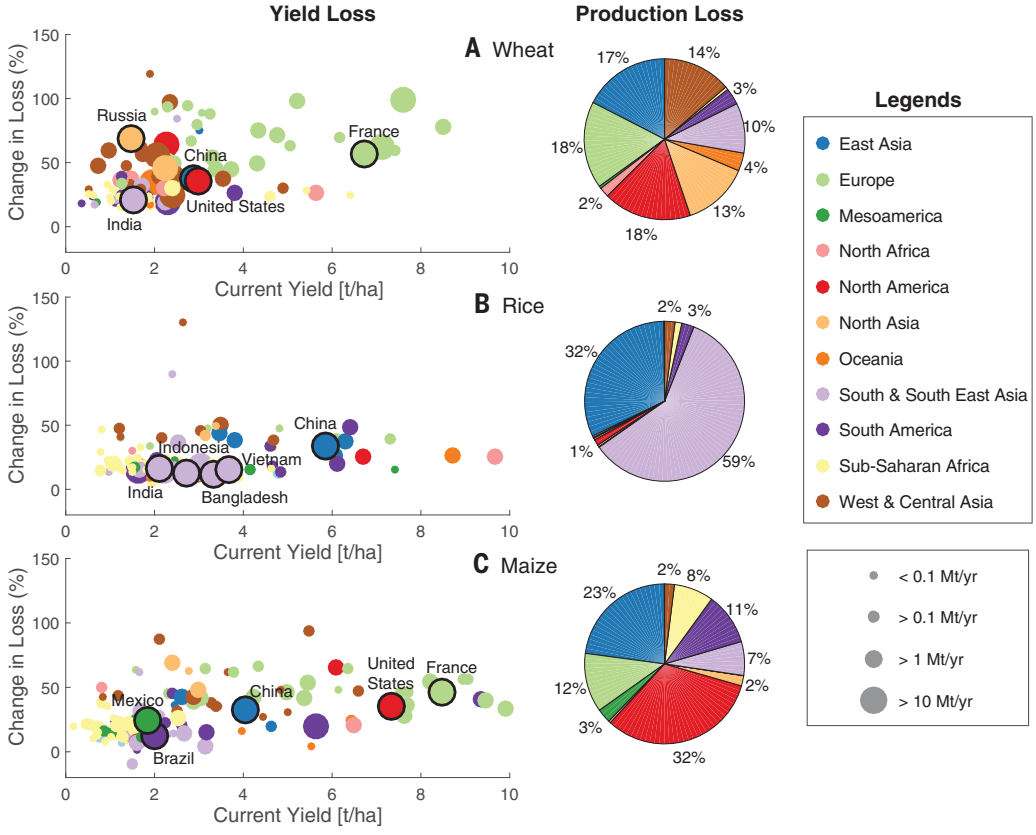


Fig. 3. Predicted regional increases in crop losses to insect pests in a 2°C-warmer climate. The change in future yield loss for each country is shown for the median grid cell within each country and plotted as a function of its median present-day crop yield per unit of planted area for (**A**) wheat, (**B**) rice, and (**C**) maize. The symbol size is scaled to total current production for each country, and color indicates the United Nations region. For each crop, the five countries with the highest current production are labeled and circled. The geographic burden of additional future production losses is shown in the pie charts. A full list of effects by region and country can be found in tables S1 to S5.

REFERENCES AND NOTES

1. J. R. Porter, L. Xie, A. J. Challinor, K. Cochrane, S. M. Howden, M. M. Iqbal, D. B. Lobell, M. I. Travasso, in *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field *et al*., Eds. (Cambridge Univ. Press, 2014), pp. 485–533.
2. D. S. Battisti, R. L. Naylor, *Science* **323**, 240–244 (2009).
3. J. Liu *et al*., *PLOS ONE* **8**, e57750 (2013).
4. C. Rosenzweig *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **111**, 3268–3273 (2014).
5. E.-C. Oerke, *J. Agric. Sci.* **144**, 31–43 (2006).
6. P. J. Gregory, S. N. Johnson, A. C. Newton, J. S. I. Ingram, *J. Exp. Bot.* **60**, 2827–2838 (2009).
7. E. L. Zvereva, M. V. Kozlov, *Glob. Change Biol.* **12**, 27–41 (2006).
8. C. A. Deutsch *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 6668–6672 (2008).
9. Food and Agriculture Organization of the United Nations, FAOSTAT database collections; www.fao.org/faostat/en/.
10. M. E. Dillon, G. Wang, R. B. Huey, *Nature* **467**, 704–706 (2010).
11. C. Petersen, H. A. Woods, J. G. Kingsolver, *Physiol. Entomol.* **25**, 35–40 (2000).
12. U. M. Irlich, J. S. Terblanche, T. M. Blackburn, S. L. Chown, *Am. Nat.* **174**, 819–835 (2009).
13. See the supplementary materials.
14. C. Lesk, E. Coffel, A. W. D'Amato, K. Dodds, R. Horton, *Nat. Clim. Chang.* **7**, 713–717 (2017).
15. C. Monfreda, N. Ramankutty, J. A. Foley, *Global Biogeochem. Cycles* **22**, GB1022 (2008).
16. M. R. Collins, R. Knutti, J. Arblaster, J.-L. Dufresne, T. Fichefet, P. Friedlingstein, X. Gao, W. J. Gutowski, T. Johns, G. Krinner, M. Shongwe, C. Tebaldi, A. J. Weaver, M. Wehner, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al*., Eds. (Cambridge Univ. Press, 2013), pp. 1029–1136.
17. C. Zhao *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9326–9331 (2017).
18. K. A. Potter, H. Arthur Woods, S. Pincebourde, *Glob. Change Biol.* **19**, 2932–2939 (2013).
19. C. M. Williams, H. A. L. Henry, B. J. Sinclair, *Biol. Rev. Camb. Philos. Soc.* **90**, 214–235 (2015).
20. Z. Liu, K. G. Hubbard, X. Lin, X. Yang, *Glob. Change Biol.* **19**, 3481–3492 (2013).
21. D. Deryng, W. J. Sacks, C. C. Barford, N. Ramankutty, *Global Biogeochem. Cycles* **25**, GB2006 (2011).
22. C. Deutsch, J. J. Tewksbury, M. Tigchelaar, D. S. Battisti, S. Merrill, R. B. Huey, R. L. Naylor, Dataset for "Increase in crop losses to insect pests in a warming climate." Dryad (2018); https://dx.doi.org/10.5061/dryad.b7q3g2q.

ACKNOWLEDGMENTS

We gratefully acknowledge H. Frenzel for technical support, the World Climate Research Programme for producing and making available the CMIP5 (Coupled Model Intercomparison Project Phase 5) model output, and the editorial suggestions of three anonymous reviewers. **Funding:** This work was made possible by grants to C.A.D. from the Gordon and Betty Moore Foundation (GBMF#3775) and the National Science Foundation (OCE-1419323, OCE-1458967, and OCE-1542240). **Author contributions:** C.A.D. and J.J.T. conceived of the study. C.A.D. constructed the model with input from J.J.T. and S.C.M. R.B.H. contributed data. C.A.D. and J.J.T. analyzed model results with input from all authors. All authors contributed to writing the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data used in this study are described in the supplementary materials, and gridded model output is available in the Dryad repository (*22*).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/916/suppl/DC1
Materials and Methods
Figs. S1 to S7
Tables S1 to S5
References (*23*–*36*)

16 February 2018; accepted 18 July 2018
10.1126/science.aat3466

CLIMATE CHANGE

Past and future global transformation of terrestrial ecosystems under climate change

Connor Nolan[1], Jonathan T. Overpeck[2,1], Judy R. M. Allen[3], Patricia M. Anderson[4], Julio L. Betancourt[5], Heather A. Binney[6], Simon Brewer[7], Mark B. Bush[8], Brian M. Chase[9], Rachid Cheddadi[9], Morteza Djamali[10], John Dodson[11,12], Mary E. Edwards[6,13], William D. Gosling[14,15], Simon Haberle[16], Sara C. Hotchkiss[17], Brian Huntley[3], Sarah J. Ivory[18], A. Peter Kershaw[19], Soo-Hyun Kim[17], Claudio Latorre[20], Michelle Leydet[10], Anne-Marie Lézine[21], Kam-Biu Liu[22], Yao Liu[23], A. V. Lozhkin[24], Matt S. McGlone[25], Robert A. Marchant[26], Arata Momohara[27], Patricio I. Moreno[28], Stefanie Müller[29], Bette L. Otto-Bliesner[30], Caiming Shen[31], Janelle Stevenson[32], Hikaru Takahara[33], Pavel E. Tarasov[29], John Tipton[34], Annie Vincens[35], Chengyu Weng[36], Qinghai Xu[37], Zhuo Zheng[38], Stephen T. Jackson[39,1]*

Impacts of global climate change on terrestrial ecosystems are imperfectly constrained by ecosystem models and direct observations. Pervasive ecosystem transformations occurred in response to warming and associated climatic changes during the last glacial-to-interglacial transition, which was comparable in magnitude to warming projected for the next century under high-emission scenarios. We reviewed 594 published paleoecological records to examine compositional and structural changes in terrestrial vegetation since the last glacial period and to project the magnitudes of ecosystem transformations under alternative future emission scenarios. Our results indicate that terrestrial ecosystems are highly sensitive to temperature change and suggest that, without major reductions in greenhouse gas emissions to the atmosphere, terrestrial ecosystems worldwide are at risk of major transformation, with accompanying disruption of ecosystem services and impacts on biodiversity.

Terrestrial ecosystem function is governed largely by the composition and physical structure of vegetation (*1*–*3*), and climate change impacts on vegetation can potentially cause disruption of ecosystem services and loss of biodiversity (*4*, *5*). It is critical to assess the likely extent of ecosystem transformation as global greenhouse gas (GHG) emissions increase (*6*) and to understand the full potential magnitude of impacts should current GHG emission rates continue unabated.

Ecosystem transformation generally involves the replacement of dominant plant species or functional types by others, whether recruited locally or migrating from afar. Observations from around the globe indicate that current climate change may already be driving substantial changes in vegetation composition and structure (*3*). Ecosystem change is accelerated by mass mortality of incumbent dominants (*7*, *8*), and widespread dieback events and other large disturbances are already under way in many forests and woodlands (*9*–*11*), with further mortality events predicted under increasing temperatures and drought (*3*, *9*, *10*, *12*). Replacement of predisturbance dominants by other species and growth forms has been widely documented (*8*, *13*, *14*). In addition, evidence is accumulating for geographic range shifts in individual species, and climate change is interacting with invasive species, fire regimes, land use, and CO_2 increase to drive vegetation changes in many regions (*15*, *16*).

Beyond observations of recent and ongoing change, models indicate ecosystem transformation under climate projections for the 21st century. These include dynamic global vegetation models (*3*, *17*), species distribution models (*18*), and comparison of the multivariate climate distance between biomes with that between modern and future climates (*19*). However, the capacity for assessing the magnitudes of ecosystem transformation under future climate scenarios is limited by the difficulty of evaluating model performance against empirical records, particularly when projected climate states are novel (*19*, *20*).

Paleoecological records of past ecological responses to climate change provide an independent means for gauging the sensitivity of ecosystems to climate change. High-precision time-series studies indicate that local and regional ecosystems can shift rapidly, within years to decades, under abrupt climate change (*21*–*23*), but sites with such detailed chronologies are scarce. In this study, we used published reports to compile a global network of radiocarbon-dated paleoecological records of terrestrial vegetation composition and structure since the Last Glacial Maximum (LGM), ~21,000 years before the present (yr B.P.) (*24*). Most postglacial warming happened 16,000 to 10,000 yr B.P., although it commenced earlier in parts of the Southern Hemisphere (*25*, *26*). Global warming between the LGM and the early Holocene (10,000 yr B.P.) was on the order of 4 to 7°C, with more warming over land than oceans (*26*, *27*). These estimates are roughly comparable to the magnitude of warming that Earth is projected to undergo in the next 100 to 150 years if GHG emissions are not reduced substantially (*28*). The magnitudes of changes in vegetation composition and structure since the last glacial period (LGP) provide an index of the magnitude of ecosystem change that may be expected under warming of similar magnitude in the coming century (*29*). Although the rate of projected future global warming is at least an order of magnitude greater than that of the last glacial-to-interglacial transition (*26*), a glacial-to-modern comparison provides a conservative estimate of the extent of ecological transformation to which the planet will be committed under future climate scenarios.

We reviewed and evaluated paleoecological (pollen and macrofossil) records from 594 sites

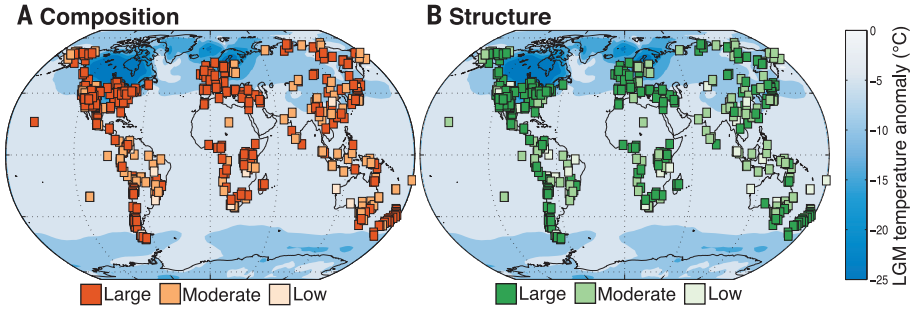


Fig. 1. Vegetation differences between the LGP and the present. Each square represents an individual paleoecological site. The color density indicates the magnitude of estimated vegetation change since the LGP (21,000 to 14,000 yr B.P.). Background shading denotes the estimated temperature anomaly between the LGM 21,000 years ago and today on the basis of assimilated proxy-data and model estimates (*27*). (**A**) Composition. (**B**) Structure.

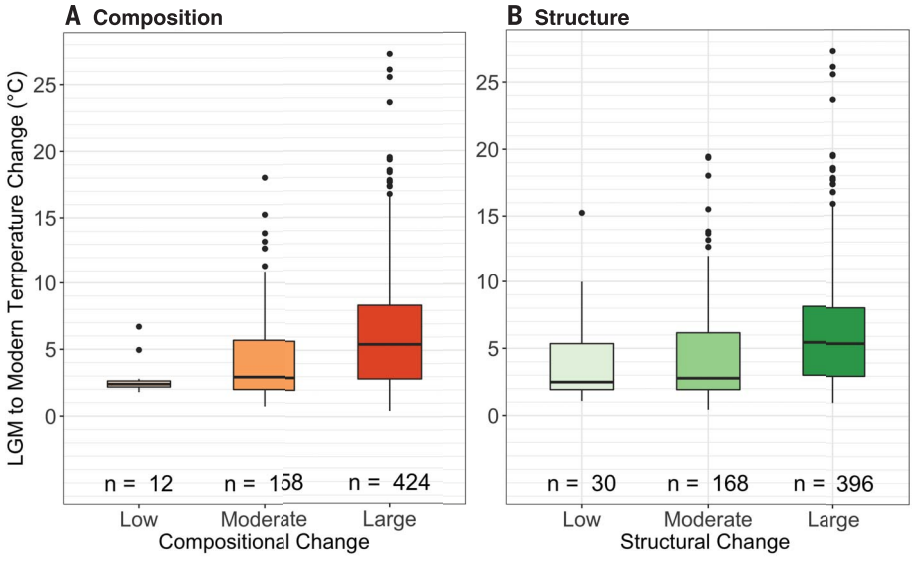


Fig. 2. Estimated temperature differences for different categories of vegetation response. Box plots of the estimated mean annual temperature differences between the LGM and today in each of the three vegetation change categories (low, moderate, and large) for (**A**) composition and (**B**) structure. Low vegetational changes are associated largely with relatively small temperature anomalies, whereas moderate and large changes are associated with larger post-LGM temperature differences, indicating that the magnitude of temperature change plays an important role in the magnitude of vegetation change. The glacial temperature anomalies are from data in (*27*). Analyses using the TraCE-21ka simulation show similar patterns (fig. S4).

worldwide (fig. S1), all drawn from peer-reviewed published literature, to determine the magnitude of postglacial vegetation change. We adopted an expert-judgment approach in which paleoecologists with relevant regional experience compiled published records (table S1); reviewed the data, diagrams, and accompanying papers; and inferred the composition and structure of the glacial-age and Holocene vegetation at each site (*24*). For the purposes of our analyses, we defined the LGP as the interval between 21,000 and 14,000 yr B.P. Although postglacial warming was under way in many regions by 16,000 yr B.P. (*25*), continental ice sheets were still extensive 14,000 yr B.P., and some climate regimes remained essentially "glacial" in nature, particularly in the Northern Hemisphere (*30*). Extending the LGP window to 14,000 yr B.P. provides a larger array of records for the assessment, both in glaciated and unglaciated terrains, and renders our analysis more conservative (climatic and vegetation contrasts with the Holocene are likely to decrease between 21,000 and 14,000 yr B.P.).

For each record, experts were asked to classify the magnitudes of compositional change and structural change since the LGP as large, moderate, or low and to provide detailed justification for their judgments (*24*) (table S2). This placed all the diverse records into a common framework for comparison. For sites that experienced moderate to large ecological change, experts were also asked to assess the role of climate change (large, moderate, or none) in driving the observed vegetation change. For each of these four judgments, experts were asked to state their level of confidence as high, medium, or low. In assessing the role of climate change, experts were asked to focus specifically on whether climate change since the LGP was sufficient to drive the observed changes, acknowledging that other factors (e.g., human activity, postglacial CO_2 increase, and megafaunal dynamics) may have also played important roles. For sites with a long history of human land use, experts used Holocene records predating widespread land clearance as a benchmark for comparison with the LGP records.

Our results indicate that the magnitude of past glacial-to-interglacial warming was sufficient at most locations across the globe to drive changes in vegetation composition that were moderate (27% of sites) to large (71%), as well as moderate (28%) to large (67%) structural changes (Fig. 1 and table S3). These changes were particularly evident at mid- to high latitudes in the Northern Hemisphere, as well as in southern South America, tropical and temperate southern Africa, the Indo-Pacific region, Australia, Oceania, and New Zealand (Fig. 1A). Compositional change at most sites in the Neotropics was moderate to large, but three sites showed little or no compositional change, all

[1]Department of Geosciences, University of Arizona, Tucson, AZ 85721, USA. [2]School for Environment and Sustainability, University of Michigan, Ann Arbor, MI 48109, USA. [3]Department of Biosciences, University of Durham, Durham DH1 3LE, UK. [4]Department of Earth and Space Sciences, University of Washington, Seattle, WA 98195, USA. [5]National Research Program, U.S. Geological Survey, Reston, VA 20192, USA. [6]Geography and Environment, University of Southampton, Southampton SO17 1BJ, UK. [7]Department of Geography, University of Utah, Salt Lake City, UT 84112, USA. [8]Department of Biological Sciences, Florida Institute of Technology, Melbourne, FL 32901, USA. [9]Centre National de la Recherche Scientifique, UMR 5554, Institut des Sciences de l'Evolution de Montpellier, Université Montpellier, Bat. 22, CC061, Place Eugène Bataillon, 34095 Montpellier, France. [10]Aix Marseille Université, Avignon Université, CNRS, IRD, Institut Méditerranéen de Biodiversité et d'Ecologie, 13545 Aix-en Provence, France. [11]Palaeontology, Geobiology and Earth Archives Research Centre (PANGEA), University of New South Wales, Sydney, NSW 2052, Australia. [12]State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, Chinese Academy of Sciences, Xi'an 71002, Shaanxi, China. [13]College of Natural Sciences and Mathematics, University of Alaska–Fairbanks, Fairbanks, AK 99775, USA. [14]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, 1090 GE Amsterdam, Netherlands. [15]School of Environment, Earth and Ecosystem Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK. [16]Department of Archaeology and Natural History, Australian National University, Canberra, Australia. [17]Department of Botany, University of Wisconsin, Madison, WI 53706, USA. [18]Department of Geosciences, Pennsylvania State University, State College, PA 16802, USA. [19]School of Earth, Atmosphere, and Environment, Monash University, Melbourne, VIC 3800, Australia. [20]Departamento de Ecología, Institute of Ecology and Biodiversity (IEB), Pontificia Universidad Católica de Chile, Santiago, Chile. [21]Sorbonne Université, CNRS-IRD-MNHN, LOCEAN/IPSL Laboratory, 4 Place Jussieu, 75005 Paris, France. [22]Department of Oceanography and Coastal Sciences, Louisiana State University, Baton Rouge, LA 70803, USA. [23]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ 86011, USA. [24]North-East Interdisciplinary Scientific Research Institute, Far East Branch Russian Academy of Sciences, Magadan 685000, Russia. [25]Landcare Research, Lincoln 7640, New Zealand. [26]Department of Environment, York Institute for Tropical Ecosystems, University of York, York YO10 5NG, UK. [27]Graduate School of Horticulture, Chiba University, Matsudo-shi, Chiba 271-8510, Japan. [28]Departamento de Ciencias Ecológicas, IEB and (CR)2, Universidad de Chile, Santiago, Chile. [29]Institute of Geological Sciences, Freie Universität Berlin, D-12249 Berlin, Germany. [30]National Center for Atmospheric Research, Climate and Global Dynamics Laboratory, Boulder, CO 80307, USA. [31]Yunnan Normal University, Key Laboratory of Plateau Lake Ecology and Global Change, Kunming, Yunnan 650092, China. [32]School of Culture, History, and Language, Australian National University, Canberra, Australia. [33]Graduate School of Life and Environmental Sciences, Kyoto Prefectural University, Kyoto, 606-8522, Japan. [34]Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR 72701, USA. [35]Centre Européen de Recherche et d'Enseignement des Géosciences de l'Environnement (CEREGE), 13545 Aix-en-Provence, France. [36]School of Ocean and Earth Science, Tongji University, Shanghai, China. [37]Institute of Nihewan Archaeology and College of Resource and Environmental Sciences, Hebei Normal University, Shijiazhuang 050024, China. [38]School of Earth Science and Engineering, Guangdong Provincial Key Lab of Geodynamics and Geohazards, Sun Yat-Sen University, Guangzhou 510275, China. [39]Southwest Climate Adaptation Science Center, U.S. Geological Survey, Tucson, AZ 85721, USA.
***Corresponding author. Email: stjackson@usgs.gov**

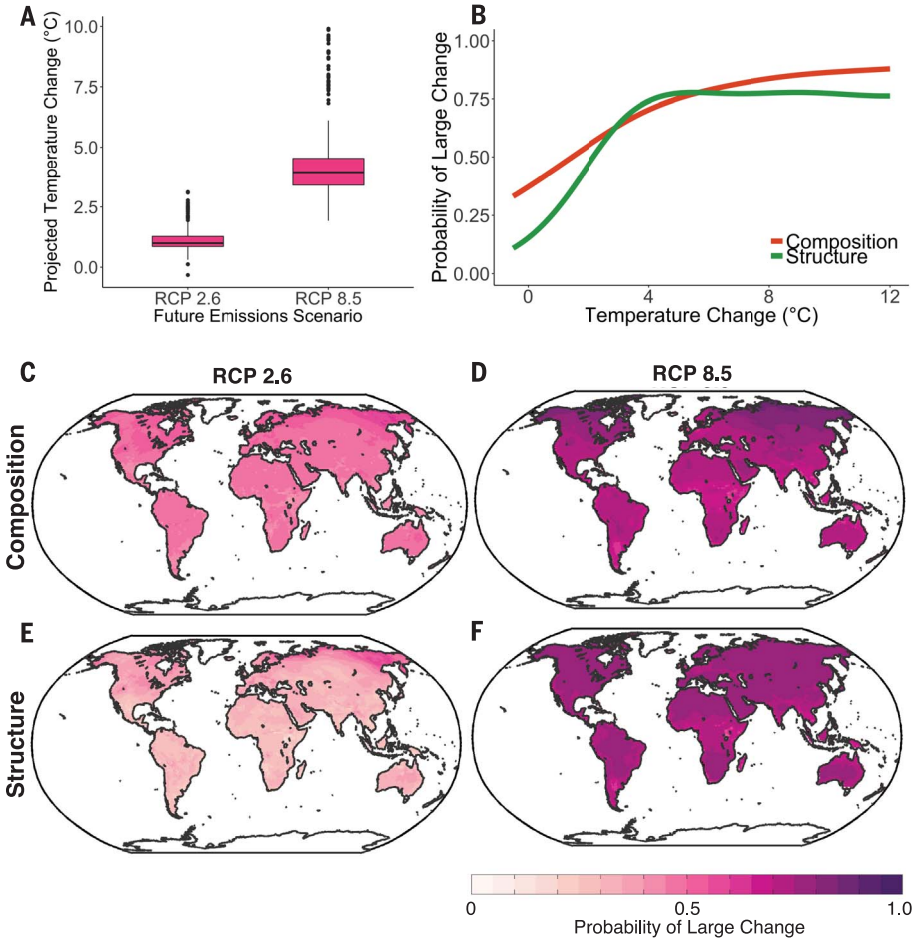


Fig. 3. Estimated vegetation change under future climate scenarios. (**A**) Box plots of the estimated mean annual temperature differences between today and future climate simulations for individual sites (as determined by using the nearest grid point). Most sites show relatively small temperature change under the low-emission scenario (RCP 2.6), with substantially greater change under the high-emission scenario (RCP 8.5). (**B**) Probabilities of large changes in vegetation composition and structure as a function of temperature change. (**C** to **F**) Estimated probabilities of large compositional and structural changes by the end of the 21st century (the average of the period from 2081 to 2100) under RCP 2.6 (C and E) and RCP 8.5 (D and F). Probabilities (B to F) are estimated from a logistic spline regression model fit by using LGM-to-modern temperature change as a predictor variable and observed LGP-to-modern vegetation changes (large versus not large) as the response variable. Future temperature increases are calculated as an average for 2081 to 2100 under the model scenarios, minus an average for 1985 to 2005 from the CCSM4 historical simulation. Analyses using the TraCE-21ka simulation show similar patterns (fig. S7).

with medium to high confidence (fig. S2). Shifts in vegetation structure were also moderate to large at mid- to high-latitude sites, although a few sites showed low change (Fig. 1B). The Neotropics had nine sites with little or no structural change (Fig. 1B), all with high-confidence assessments (fig. S2). These sites have been occupied by tropical forest ecosystems since the LGM, although most have undergone moderate to large compositional change (*31*, *32*). For nearly all sites that experienced moderate or large ecological change, climate change since the LGP was judged to be sufficient to explain the observed changes with high confidence (table S4). Atmospheric CO_2 concentrations also increased from 190 to 280 parts per million during the deglaciation, interacting with and in some cases modulating ecological responses to climate change. However, CO_2 changes alone cannot account for postglacial vegetation changes (supplementary text).

Independently of the expert-judgment process, we used the estimated anomaly in mean annual temperatures between the LGM and the present (preindustrial) as a proxy for the overall magnitude of climate change since the LGP (*24*). LGM temperature estimates were derived using an assimilated data–model integration (*27*). Low-change sites were largely concentrated in regions where the estimated temperature anomaly was relatively small (Fig. 1). To explore this relationship further, we plotted the frequency distribution of the difference between estimated LGM and present-day mean annual temperatures for individual sites in each of the three ecological-response categories. Nearly all sites with low compositional change between the LGP and today are associated with small estimated temperature anomalies (median, 2.4°C), whereas sites with moderate to high compositional change have larger temperature anomalies (Fig. 2A). Results for structural changes are similar, although a greater number of sites with low structural change include larger temperature anomalies (Fig. 2B). This difference is not surprising, because compositional change in vegetation can occur without an accompanying change in vegetation structure (Fig. 1). Europe and eastern North America experienced unusually large temperature changes since the LGM, owing to depressed temperatures near the large ice sheets, and these regions show substantial compositional and structural changes since the LGP. However, results from other parts of the globe indicate that widespread ecosystem changes were driven by much smaller temperature changes (fig. S3). We repeated our analysis using the TraCE-21ka model simulations (*33*, *34*), which yield a lower magnitude of LGP-to-Holocene climate change (*35*); despite the potential conservative bias, results for compositional and structural changes (fig. S4) were similar to those in Fig. 2. Temperature differences between the LGP and the present were substantially greater for sites with large ecological change than for those with low to moderate change, by both paleoclimate estimates (*27*, *33*) (table S2).

We also used our database of ecological change since the LGM to assess the global distribution of the probabilities of large compositional and structural changes given GHG emission scenarios [representative concentration pathways (RCPs) 2.6, 4.5, 6.0, and 8.5, each as simulated by the Community Climate System Model version 4 (CCSM4)] (*24*, *36*). The range of LGM-to-present temperature changes (Fig. 2) overlaps with the range of temperature changes projected for the coming century under these scenarios (Fig. 3A and fig. S5). We quantified the relationship between temperature and ecological change by using a logistic spline regression with ordered categories (*37*). We fit models for compositional and structural change by using the temperature change since the LGM as the independent predictor variable. In both models, LGM-to-modern temperature change is a significant predictor of ecosystem change ($P < 0.001$). We then used these models to predict the risk of large change for the future range of projected global temperature changes (Fig. 3B) and to map the probability of large change under RCP 2.6 and RCP 8.5 (Fig. 3, C to F) at the end of the 21st

century (see fig. S6 for RCP 4.5 and RCP 6.0). Under RCP 2.6, the probability of large compositional change is less than 45% over most of the globe (Fig. 3C) and the probability of large structural change is generally less than 30% (Fig. 3E). By contrast, under the business-as-usual emissions scenario, RCP 8.5, the probabilities of large compositional change and large structural change are both greater than 60% (Fig. 3, D and F). Analyses using the TraCE-21ka model yielded similar patterns (fig. S7).

Our study uses a single variable, mean annual temperature, as a metric for the broader array of climatic changes that can drive vegetation change, and it compares vegetation and climate states separated by 10,000 to 20,000 years. Future climate change, like that in the past, will be multivariate, involving shifts in seasonal temperatures, seasonal precipitation, climate extremes, and variability regimes. As mean annual temperature increases, other ecologically important variables will change, often in complex or counter-intuitive ways (*20*, *38*, *39*), and ecological responses will often be episodic or nonlinear (*8*, *13*–*15*). Although the temperature increases since the LGP provide crude analogs for ongoing and future climate changes—for example, boundary conditions and forcings are different now (*26*, *40*, *41*)—our results nevertheless provide concrete evidence that vegetation composition and structure are sensitive to changes in mean annual temperature of the magnitudes forecast for the coming century and that vegetation transformations will become increasingly extensive as temperatures increase. Under the RCP 8.5 scenario, the rate of warming will be on the order of 65 times as high as the average warming during the last deglaciation (*26*). Furthermore, the warming between the LGP and the Holocene occurred within the range of previous glacial and interglacial temperatures, whereas projected future changes will exceed those experienced over the past 2 million years (*26*). Although many ecological responses (e.g., species migration, colonization, and succession) will likely lag behind climate changes, ecosystem transformations will often be accelerated by disturbance and mortality events, land use, and invasive species (*7*–*15*).

We therefore conclude that terrestrial vegetation over the entire planet is at substantial risk of major compositional and structural changes in the absence of markedly reduced GHG emissions. Much of this change could occur during the 21st century, especially where vegetation disturbance is accelerated or amplified by human impacts (*7*). Many emerging ecosystems will be novel in composition, structure, and function (*42*), and many will be ephemeral under sustained climate change; equilibrium states may not be attained until the 22nd century or beyond. Compositional transformation will affect biodiversity via disintegration and reorganization of communities, replacement of dominant or keystone species, pass-through effects on higher trophic levels, and ripple effects on species interactions (*16*, *43*). Structural transformation will have particularly large consequences for ecosystem services (*4*), including the achievement of nature-based development solutions under the United Nations' Sustainable Development Goals (*44*). Structural changes will also influence biodiversity, driving alterations in habitats and resources for species at higher trophic levels. Compositional and structural changes may also induce potentially large changes to carbon sources and sinks, as well as to atmospheric moisture recycling and other climate feedbacks. Our results suggest that impacts on planetary-scale biodiversity, ecological functioning, and ecosystem services will increase substantially with increasing GHG emissions, particularly if warming exceeds that projected by the RCP 2.6 emission scenario (1.5°C).

REFERENCES AND NOTES

1. F. S. Chapin III, P. A. Matson, H. A. Mooney, *Principles of Terrestrial Ecosystem Ecology* (Springer, 2002).
2. S. Díaz *et al.*, *J. Veg. Sci.* **15**, 295–304 (2004).
3. J. Settele, *et al.*, in *Climate Change 2014—Impacts, Adaptation, and Vulnerability: Part A: Global and Sectoral Aspects: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, C. B. Field *et al.*, Eds. (Cambridge Univ. Press, 2014), pp. 271–360.
4. Millennium Ecosystem Assessment, *Ecosystems and Human Well-Being: Synthesis* (Island Press, 2005).
5. B. J. Cardinale *et al.*, *Nature* **486**, 59–67 (2012).
6. United Nations, *United Nations Framework Convention on Climate Change* (United Nations, 1992).
7. J. T. Overpeck, D. Rind, R. Goldberg, *Nature* **343**, 51–53 (1990).
8. C. D. Allen, D. D. Breshears, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 14839–14842 (1998).
9. C. D. Allen, D. D. Breshears, N. G. McDowell, *Ecosphere* **6**, 129 (2015).
10. W. R. L. Anderegg, J. M. Kane, L. D. L. Anderegg, *Nat. Clim. Chang.* **3**, 30–36 (2013).
11. G. F. Midgley, W. J. Bond, *Nat. Clim. Chang.* **5**, 823–829 (2015).
12. G. P. Asner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E249–E255 (2016).
13. J. F. Johnstone *et al.*, *Front. Ecol. Environ.* **14**, 369–378 (2016).
14. C. H. Guiterman, E. Q. Margolis, C. D. Allen, D. A. Falk, T. W. Swetnam, *Ecosystems* (2017).
15. C. I. Millar, N. L. Stephenson, *Science* **349**, 823–826 (2015).
16. C. Parmesan, M. E. Hanley, *Ann. Bot.* **116**, 849–864 (2015).
17. I. C. Prentice, A. Bondeau, W. Cramer, S. P. Harrison, T. Hickler, W. Lucht, S. Sitch, B. Smith, M. T. Sykes, "Dynamic global vegetation modeling: Quantifying terrestrial ecosystem responses to large-scale environmental change," in *Terrestrial Ecosystems in a Changing World*, J. G. Canadell, D. E. Pataki, L. F. Pitelka, Eds. (Springer, 2007), pp. 175–192.
18. G. E. Rehfeldt, N. L. Crookston, M. V. Warwell, J. S. Evans, *Int. J. Plant Sci.* **167**, 1123–1150 (2006).
19. J. W. Williams, S. T. Jackson, J. E. Kutzbach, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 5738–5742 (2007).
20. S. T. Jackson, J. W. Williams, *Annu. Rev. Earth Planet. Sci.* **32**, 495–537 (2004).
21. J. R. M. Allen, W. A. Watts, B. Huntley, *Quat. Int.* **73–74**, 91–110 (2000).
22. J. W. Williams, D. M. Post, L. C. Cwynar, A. F. Lotter, A. J. Levesque, *Geology* **30**, 971–974 (2002).
23. A. Correa-Metrio *et al.*, *Quat. Sci. Rev.* **38**, 63–75 (2012).
24. Materials and methods are available as supplementary materials.
25. J. D. Shakun, A. E. Carlson, *Quat. Sci. Rev.* **29**, 1801–1816 (2010).
26. V. Masson-Delmotte *et al.*, in *Climate Change 2013—The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al.*, Eds. (Cambridge Univ. Press, 2013), pp. 383–464.
27. J. D. Annan, J. C. Hargreaves, *Clim. Past* **9**, 367–376 (2013).
28. M. Collins *et al.*, in *Climate Change 2013—The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al.*, Eds. (Cambridge Univ. Press, 2013), pp. 1029–1136.
29. J. Guiot, W. Cramer, *Science* **354**, 465–468 (2016).
30. P. U. Clark *et al.*, *Science* **325**, 710–714 (2009).
31. P. A. Colinvaux, P. E. De Oliveira, J. E. Moreno, M. C. Miller, M. B. Bush, *Science* **274**, 85–88 (1996).
32. M. B. Bush, P. E. De Oliveira, P. A. Colinvaux, M. C. Miller, E. Moreno, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **214**, 359–393 (2004).
33. Z. Liu *et al.*, *Science* **325**, 310–314 (2009).
34. F. He, "Simulating transient climate evolution of the last deglaciation with CSM3," dissertation, University of Wisconsin, Madison, WI (2011).
35. Z. Liu *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3501–E3505 (2014).
36. G. A. Meehl *et al.*, *J. Clim.* **25**, 3661–3683 (2012).
37. S. N. Wood, N. Pya, B. Säfken, *J. Am. Stat. Assoc.* **111**, 1548–1563 (2016).
38. S. T. Jackson, J. T. Overpeck, *Paleobiology* **26** (Supplement), 194–220 (2000).
39. S. T. Jackson, J. L. Betancourt, R. K. Booth, S. T. Gray, *Proc. Natl. Acad. Sci. U.S.A.* **106** (suppl. 2), 19685–19692 (2009).
40. D. L. Hartmann *et al.*, in *Climate Change 2013—The Physical Science Basis: Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker *et al.*, Eds. (Cambridge Univ. Press, 2013), pp. 159–254.
41. P. U. Clark *et al.*, *Nat. Clim. Chang.* **6**, 360–369 (2016).
42. R. J. Hobbs, E. S. Higgs, C. Hall, Eds., *Novel Ecosystems: When and How Do We Intervene in the New Ecological World Order?* (Wiley-Blackwell, 2013).
43. C. Bellard, C. Bertelsmeier, P. Leadley, W. Thuiller, F. Courchamp, *Ecol. Lett.* **15**, 365–377 (2012).
44. United Nations, Sustainable Development Goals; https://sustainabledevelopment.un.org/sdgs.

ACKNOWLEDGMENTS

The paper benefitted from the thoughtful comments of S. T. Gray and three anonymous reviewers. **Funding:** This research was supported by the NSF (DEB-1241851, AGS-1243125, and EAR-1304083) and by the Department of the Interior's Southwest Climate Adaptation Science Center. Research in northeast Siberia was funded by the Russian Academy of Sciences, FEB (15-I-2-067), and the Russian Foundation for Fundamental Research (15-05-06420). **Author contributions:** S.T.J., C.N., and J.T.O. designed the project; all authors collected data; C.N. analyzed data with advice from J.T.O., S.T.J., S.B., and J.T.; and S.T.J., C.N., and J.T.O. wrote the paper with text contributions in the supplementary materials from B.M.C., M.B.B., M.E.E., J.L.B., B.H., Y.L., and S.J.I. and further contributions from all authors. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/920/suppl/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S8
Tables S1 to S4
References (*45*–*95*)
Data S1

27 April 2017; resubmitted 24 April 2018
Accepted 30 July 2018
10.1126/science.aan5360

GENOME STRUCTURE

Three-dimensional genome structures of single diploid human cells

Longzhi Tan[1,2*], Dong Xing[1*], Chi-Han Chang[1], Heng Li[3], X. Sunney Xie[1,4,5†]

Three-dimensional genome structures play a key role in gene regulation and cell functions. Characterization of genome structures necessitates single-cell measurements. This has been achieved for haploid cells but has remained a challenge for diploid cells. We developed a single-cell chromatin conformation capture method, termed Dip-C, that combines a transposon-based whole-genome amplification method to detect many chromatin contacts, called META (multiplex end-tagging amplification), and an algorithm to impute the two chromosome haplotypes linked by each contact. We reconstructed the genome structures of single diploid human cells from a lymphoblastoid cell line and from primary blood cells with high spatial resolution, locating specific single-nucleotide and copy number variations in the nucleus. The two alleles of imprinted loci and the two X chromosomes were structurally different. Cells of different types displayed statistically distinct genome structures. Such structural cell typing is crucial for understanding cell functions.

The nucleus of a human diploid cell contains 46 chromosomes, 23 maternal and 23 paternal, together carrying 6 Gb of genomic DNA. The three-dimensional (3D) genome structure is thought to be crucial for the regulation of gene expression and other cellular functions (*1*). For example, the nuclei of sensory neurons assume unusual architectures in the mouse visual (*2*) and olfactory systems (*3*). Chromatin conformation capture assays, such as 3C (*4*) and Hi-C (*5*), allow for studies of 3D genome structures in bulk samples through proximity ligation of DNA (*6*). However, the difference between cells can only be observed by single-cell measurements. Single-cell chromatin conformation capture methods avoid ensemble averaging (*7*–*12*) and have yielded 3D genome structures of haploid mouse cells (*10*, *11*). However, characterizing the 3D genome structures of diploid mammalian cells remains challenging (*13*). Here, we used an improved chromatin conformation capture method and phased (haplotype-resolved) single-nucleotide polymorphisms (SNPs) to distinguish between the two haplotypes of each chromosome. This allowed us to examine the cell type dependence of 3D genome structures of diploid cells.

Obtaining high-resolution 3D genome structures of single diploid cells requires resolving a large number of chromatin "contacts"—pairs of genomic loci that are joined by proximity ligation. We developed a chromatin conformation capture method, termed Dip-C (Fig. 1A), that can detect more contacts than existing methods with minimal false positives. In particular, we omitted biotin pulldown (*8*, *9*) and conducted high-coverage whole-genome amplification with multiplex end-tagging amplification (META), which introduced few artifactual chimeras (*14*, *15*). We detected a median of 1.04 million contacts per single cell (n = 17, minimum = 0.71 million, maximum = 1.48 million) from GM12878, a female human lymphoblastoid cell line, and a median of 0.84 million contacts (n = 18, minimum = 0.67 million, maximum = 1.08 million) from peripheral blood mononuclear cells (PBMCs) of a male human donor (*16*). This exceeds the medians achieved with existing methods by a factor of ~5 (fig. S4 and table S1). Most cells were in the G_1 or G_0 phase of the cell cycle. In addition, we simultaneously detected copy number variations (CNVs), losses of heterozygosity (LOHs), DNA replication, and V(D)J recombination with a 10-kb bin size (figs. S2 and S3).

Another challenge in reconstructing diploid genomes is to determine which haplotypes are involved in each chromatin contact (*17*–*20*) (table S1). To assign haplotypes, we developed an imputation algorithm (Fig. 1B). We reasoned that unknown haplotypes can be imputed from "neighboring" (in terms of genomic distances) contacts by assuming that the two homologs would typically contact different partners. Using a statistical property of interchromosomal and long-range intrachromosomal contacts (*15*), we defined a contact neighborhood as a superellipse with an exponent of 0.5 and a radius of 10 Mb, where haplotypes of nearby contacts were weighted in imputing the haplotypes of each contact (fig. S7). In the Dip-C algorithm, after removing 3C/Hi-C artifacts [contacts with few neighbors (*11*)] and initial imputation, haplotypes can be optionally refined through a series of draft 3D models (*15*) (fig. S5). Imputation accuracy was estimated to be ~96% for each haplotype by cross-validation

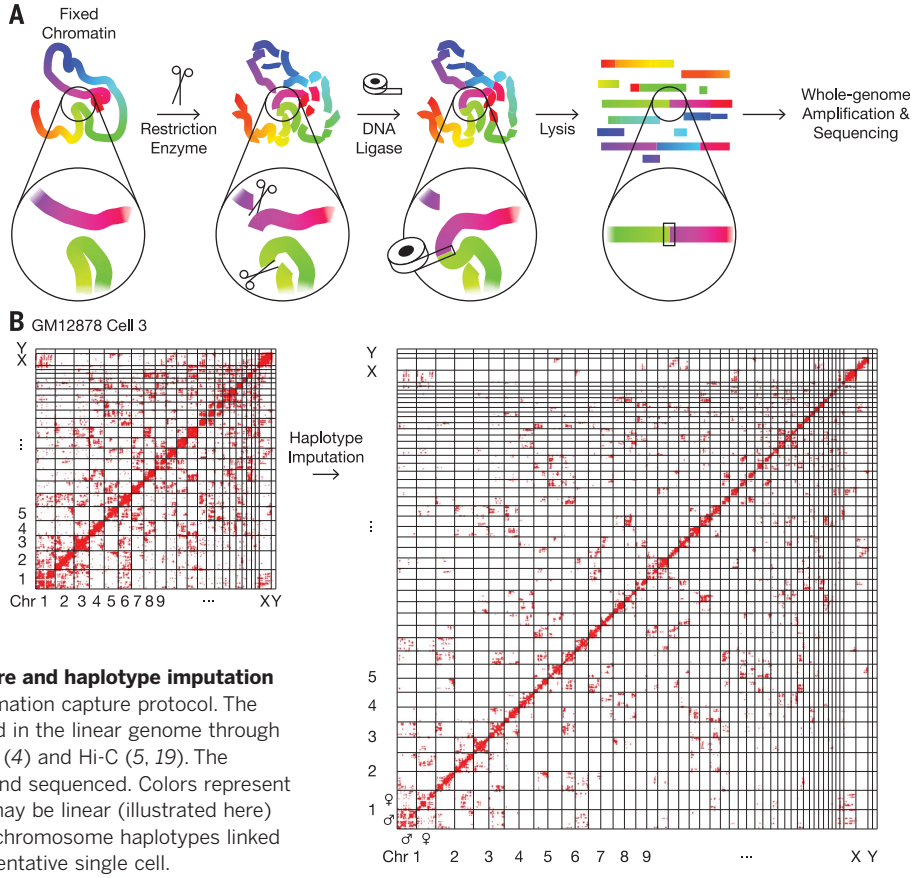


Fig. 1. Single-cell chromatin conformation capture and haplotype imputation by Dip-C. (**A**) Schematics of the chromatin conformation capture protocol. The 3D information of chromatin structure was encoded in the linear genome through proximity ligation of chromatin fragments, as in 3C (*4*) and Hi-C (*5*, *19*). The ligation product was then amplified by META (*15*) and sequenced. Colors represent genomic coordinates. Note that ligation products may be linear (illustrated here) or circular (not shown). (**B**) Imputation of the two chromosome haplotypes linked by each chromatin "contact" (red dots) in a representative single cell.

(*15*) (table S1). Regions harboring CNVs or LOHs, as well as an apparently damaged GM12878 cell, were excluded from reconstruction (table S1).

We reconstructed the 3D diploid human genomes at 20-kb resolution. Reconstruction was successful without supervision for 94% (15 of 16) of the GM12878 cells and 67% (12 of 18) of the PBMCs, and after removal of small problematic regions for 6% (1 of 16) of the GM12878 cells and 22% (4 of 18) of the PBMCs (table S1 and fig. S8) (*15*). Note that because chromatin conformation capture—the process of converting 3D coordinates to chromatin contacts—is intrinsically lossy and noisy, our 3D structures harbored additional uncertainties including perturbations of chromatin structures during the experiments, inaccuracies in the energy function used by 3D modeling, and nuclear volumes inaccessible to DNA sequencing (e.g., centromeres, nucleoli, and nuclear speckles). These uncertainties are common to all 3C/Hi-C studies and are difficult to estimate, and imputation may be less successful when two homologs are nearby or adopt similar shapes. Therefore, other problematic regions might persist even after manual removal.

Figure 2A shows a representative cell. Each particle, displayed as a colored point, represents 20 kb of chromatin, or a radius of ~100 nm. A lower bound for reconstruction uncertainty was estimated from the median deviation of ~0.4 particle radii (~40 nm) across all 20-kb particles between three replicates (fig. S9 and table S1). Well-known nuclear morphologies were observed in an M/G_1-phase GM12878 cell, where chromosomes retained their characteristic V shapes after recent mitosis, and in several PBMCs, where multiple nuclear lobes were reminiscent of the partially segmented nuclei of low-density neutrophils and other blood cell types (Fig. 2B).

We also used published data on mouse embryonic stem cells (mESCs) (*10*) to reconstruct 3D diploid mouse genomes despite fewer contacts (~0.3 million per cell, or ~0.2 million under our definition) (table S1), because the mouse line harbored more SNPs than humans (*15*).

Similar to previously described haploid mouse genomes (*10*, *11*), the diploid human genomes exhibited chromosome territories (Fig. 2A) and chromatin compartments [visualized by CpG frequency as a proxy (*21*)], with the heterochromatic compartment B (*5*) concentrated at the nuclear periphery and around foci in the nuclear center (Fig. 2C). Spatial clustering of DNA sequences with similar CpG frequencies suggests a correlation between primary sequence features and 3D genome folding (*1*).

Our 3D structures revealed different radial preferences across the human genome (black dots

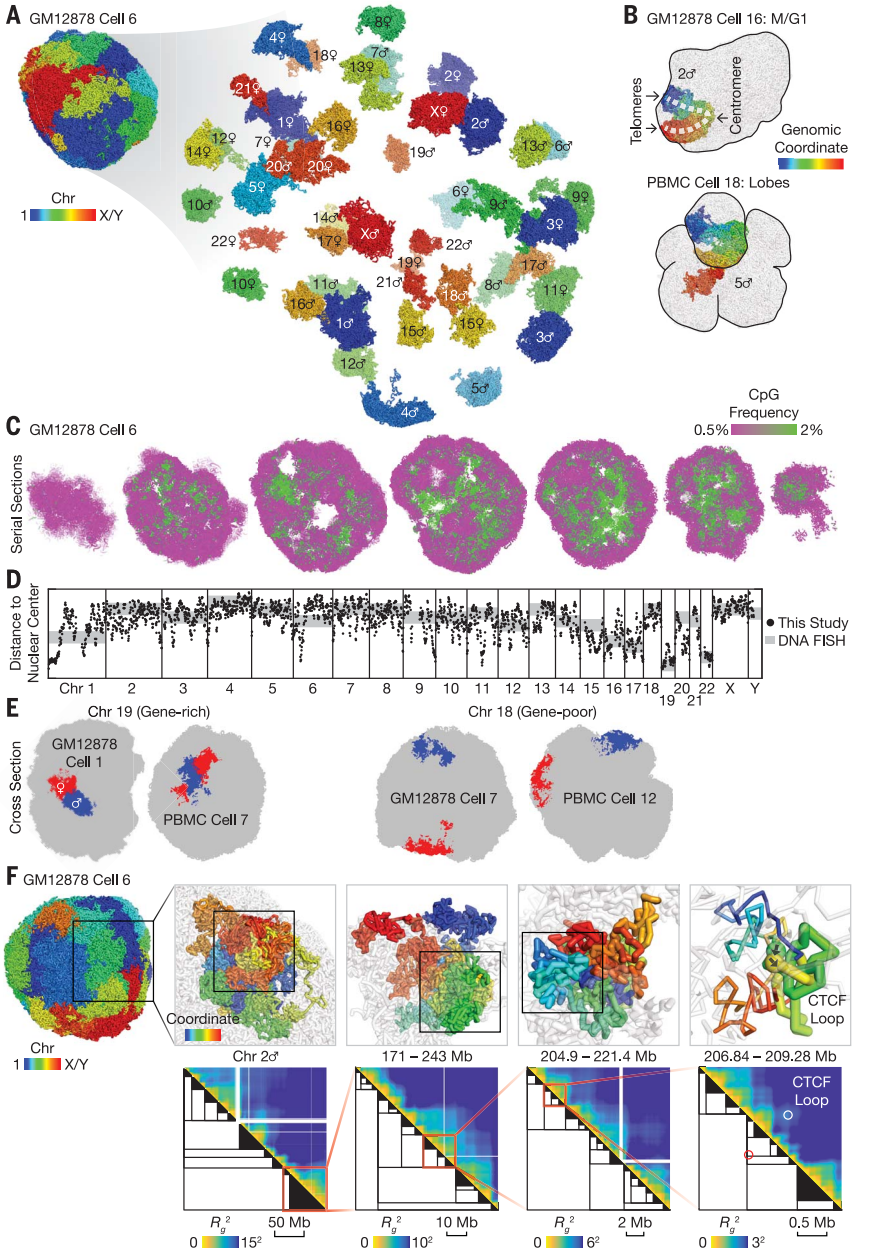


Fig. 2. 3D genome structures of single diploid human cells. (**A**) 3D genome structure of a representative GM12878 cell. Each particle represents 20 kb of chromatin, or a radius of ~100 nm. (**B**) Peculiar nuclear morphology in a cell that recently exited mitosis (top) and in a cell with multiple nuclear lobes (bottom). (**C**) Serial cross sections of a single cell showing compartmentalization of euchromatin (green) and heterochromatin (magenta), visualized by CpG frequency as a proxy (*21*). (**D**) Radial preferences across the human genome, as measured by average distances to the nuclear center of mass. Our results (black dots, smoothed by 1-Mb windows) agree well with published DNA FISH data (gray lines) on whole chromosomes (*22*) (shifted and rescaled) and provide fine-scale information. Lower and upper axis limits were 20 and 50 particle radii, respectively, for the black dots. GM12878 cell 4 (extensive chromosomal aberrations) and cell 16 (M/G_1 phase) were excluded. (**E**) Example radial preferences of two chromosomes. The gene-rich chromosome 19 preferred the nuclear interior (left), whereas the gene-poor chromosome 18 almost always resided on the nuclear surface (right). (**F**) Stochastic fractal organization of chromatin was quantified by a matrix of radii of gyration of all possible subchains of each chromosome (heat maps). We identified a hierarchy of single-cell domains across genomic scales (black trees). A subtree was simplified as a black triangle if either of its two subtrees was below a certain size (from left to right: 10 Mb, 2 Mb, 500 kb, 100 kb). In each panel, the region from the previous panel is shown in transparent gray. In the rightmost panel, thick sticks (top) and circles (bottom) highlight the formation of a known CTCF loop (*19*). Spheres with arrows (top) indicate the positions and orientations of the two converging CTCF sites. Genomic coordinates are for the human genome assembly hg19.

[1]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA. [2]Systems Biology Ph.D. Program, Harvard Medical School, Boston, MA 02115, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [4]Innovation Center for Genomics, Peking University, Beijing 100871, China. [5]Biodynamic and Optical Imaging Center, Peking University, Beijing 100871, China.
*These authors contributed equally to this work.
†Corresponding author. Email: sunneyxie@pku.edu.cn

in Fig. 2D). Our results agree well with whole-chromosome painting data by DNA fluorescence in situ hybridization (FISH) (*22*) (gray lines in Fig. 2D). Both methods show that the gene-rich chromosome 19 prefers the nuclear interior, while the gene-poor chromosome 18 prefers the nuclear periphery (Fig. 2E). Within each chromosome, different segments could have distinctly different radial preferences, which were correlated with chromatin compartments (fig. S11A). For example, the CpG-rich euchromatic end (left) of chromosome 1 was heavily biased toward the nuclear center, whereas some other regions on the same chromosome were biased toward the nuclear periphery (Fig. 2D). Such fine-scale information cannot be obtained from whole-chromosome painting (*22*, *23*) experiments.

Our Dip-C results provide a holistic view of the stochastic, fractal organization of chromatin across different genomic scales. Bulk Hi-C suggests that chromatin forms a "fractal globule" with compartments (*5*, *19*) and domains such as topologically associating domains (TADs) (*24*) and CCCTC-binding factor (CTCF) loop domains (*19*). However, such fractal organization has not been visualized in single human cells in a genome-wide manner. We observed spatial clustering (globules) and segregation (insulation) of consecutive chromatin particles along each chromosome (Fig. 2F, upper panels). Such organization could be quantified by a matrix of radii of gyration of all possible subchains in each chromosome (Fig. 2F, lower panels). Single-cell domains could then be identified as squares that had relatively small radii [partly similar to (*8*)] (*15*). We found single-cell domains across all genomic scales and therefore identified them through hierarchical merging, yielding a tree of domains [partly similar to (*25*, *26*) in bulk Hi-C] (Fig. 2F). On the smallest scale, some domains coincided with CTCF loop domains from bulk Hi-C (*19*) (rightmost panels in Fig. 2F). Single-cell domains were highly heterogeneous between cells, frequently breaking and merging bulk domains (fig. S19), consistent with a recent study on tetraploid mouse cells (*8*).

Traditional methods such as bulk Hi-C and two-color DNA FISH are pairwise measurements and thus cannot study multichromosome intermingling. In our 3D models, we quantified multichromosome intermingling by the diversity of chromosomes (Shannon index) near each 20-kb particle (fig. S20A), revealing genomic regions that frequently contacted multiple chromosomes (fig. S20B). These regions were similar between the human cell types despite their different average extents of intermingling (fig. S10), and they were mostly euchromatic (CpG-rich) (fig. S11B) for two reasons: (i) Euchromatin more frequently resided on the surface of chromosomes than did heterochromatin [consistent with (*7*)] (fig. S11D), and (ii) even when heterochromatin resided on the surface, it tended to face the nuclear periphery (*11*) (fig. S11A) and thus had no partners to intermingle with. The intermingling regions partially overlapped with "hubs" identified by a recent report (*27*).

We examined the structural relationship between the maternal and paternal alleles, which can only be studied in diploid cells. Our data captured the structural difference between the two alleles caused by genomic imprinting. At imprinted loci, the two alleles can differ drastically in transcriptional activity (*28*). Near the maternally transcribed *H19* gene and the paternally transcribed *IGF2* gene, bulk Hi-C identified different contact profiles and different use of CTCF loops between the two homologs (*19*). We directly visualized this ~0.6-Mb region in single cells (Fig. 3A). Despite cell-to-cell heterogeneity, the maternal allele more frequently separated *IGF2* from both *H19* and the nearby *HIDAD* site and disrupted the *IGF2-HIDAD* CTCF loop, whereas the paternal allele more frequently stayed fully intermingled.

X chromosome inactivation (XCI) presents a striking example of the difference between two homologs (*28*). As expected, we found in the female GM12878 cell line that the active X chromosome [the maternal allele based on RNA expression (*15*)] tended to exhibit an extended morphology, and the inactive X a compact one (Fig. 3B), although in some cells this morphological difference was not obvious. More consistently, the two X chromosomes in each cell were characterized by their distinct patterns of chromatin compartments. The active X featured clear compartmentalization of euchromatin and heterochromatin, resembling that of the male X (in PBMCs); in contrast, compartments along the inactive X were more uniform (fig. S12E). Individual X chromosomes could be clearly separated into active and inactive clusters by principal components analysis (PCA) of single-cell compartments (Fig. 3C). Our conclusion held if single-cell compartments were defined on the basis of contacts [partly similar to (*10*)] rather than 3D structures (fig. S15, A and B). We also visualized the simultaneous formation

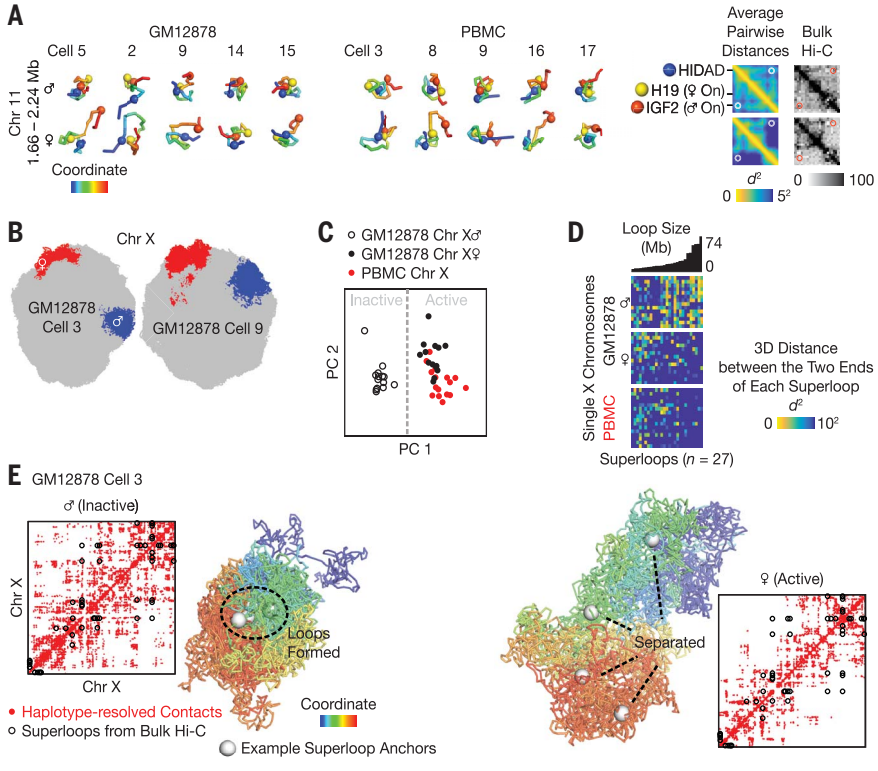


Fig. 3. Distinct 3D structures of the maternal and paternal alleles. (**A**) Structural difference between the two alleles of the imprinted *H19/IGF2* locus. Despite cell-to-cell heterogeneity, the maternal allele more frequently separated *IGF2* from both *H19* and the nearby *HIDAD* site and disrupted the *IGF2-HIDAD* CTCF loop (white and red circles). Spheres highlight three CTCF sites from bulk Hi-C. Heat maps show the root-mean-square average pairwise distances between all 20-kb particles. Haplotype-resolved bulk Hi-C (black heat map with 25-kb bins) is adapted from figure 7C of (*19*). (**B**) Active (red) and inactive (blue) X chromosomes prefer extended and compact morphologies, respectively, as shown by cross sections of two representative cells. (**C**) Individual active and inactive X chromosomes can be distinguished by PCA of single-cell chromatin compartments, defined for each 20-kb particle as the average CpG frequency of nearby (within 3 particle radii) particles. (**D**) The inactive X chromosome tends to form the previously reported "superloops," 27 very-long-range (5 to 74 Mb) chromatin loops identified by bulk Hi-C (*19*, *20*, *29*). Superloops are sorted by size. (**E**) Haplotype-resolved contact maps (red dots) and 3D structures of the two X chromosomes in an example cell. Black circles denote all superloops (*19*). White spheres denote four example superloop anchors (DXZ4, x75, ICCE, and FIRRE). GM12878 cells 4 and 16 are excluded from (C) and (D).

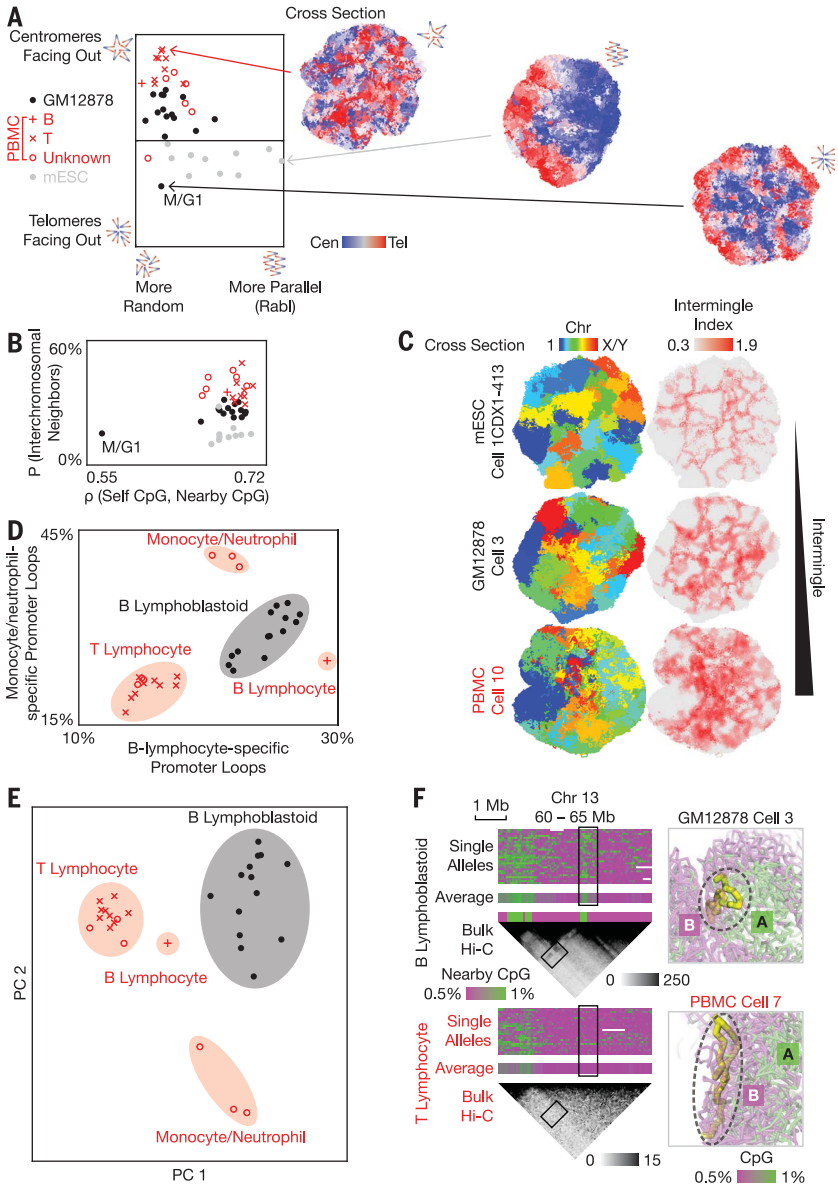


Fig. 4. Cell type–specific chromatin structures. (**A**) Quantification of the organization of centromeres and telomeres. The mESCs exhibit stronger Rabl configuration (horizontal axis; the length of summed centromere-to-telomere vectors normalized by the total particle number, which differs between human and mouse; axis limit = 0.005 particle radii), whereas the PBMCs tend to point centromeres outward relative to telomeres (vertical axis; the summed centromere-to-telomere difference in distance from the nuclear center of mass normalized by the total particle number; axis limit = 0.007 particle radii). Each marker represents a single cell and was inferred by V(D)J recombination in PBMCs (table S1 and fig. S3B). (**B**) Quantification of chromosome intermingling (vertical axis; the average fraction of nearby particles that are not from the same chromosome) and chromatin compartmentalization (horizontal axis; Spearman correlation between each particle's own CpG frequency and the average of nearby particles). (**C**) Example cross sections of three cell types, colored according to chromosome (left) or by the multichromosome intermingle index (right). (**D**) Among the human cells, four cell type clusters (shaded)—B lymphoblastoid cells, presumable T lymphocytes, B lymphocytes, and presumable monocytes/neutrophils (PBMC cells 9, 14, and 18)—could be distinguished from the differential formation (defined as end-to-end distance ≤ 3 particle radii) of known cell type–specific promoter-enhancer loops from published bulk promoter capture Hi-C (*35*). (**E**) The same four clusters could also be distinguished by unsupervised clustering via PCA of single-cell chromatin compartments, without the need for bulk data. The two alleles of each locus were treated as two different loci. GM12878 cell 16 was excluded from (D) and (E). (**F**) An example region that was differentially compartmentalized between two cell types (black, B lymphoblastoid cells; red, presumable T lymphocytes). Right panels visualize the configuration of the ~0.5-Mb region (chr 13: 62.5 to 63 Mb, thick yellow sticks) with respect to the rest of the genome (transparent, colored by CpG frequencies) in two representative cells. Only the paternal alleles are shown. Bulk Hi-C (black heat map with 50-kb bins) is from (*19*, *41*). GM12878 cell 4 was excluded.

of multiple "superloops" (*19*, *20*, *29*) in the inactive X chromosome (Fig. 3, D and E). Averaged contact matrices of the inactive and active X chromosomes agreed well with bulk Hi-C (*19*) (fig. S15, C and D).

In contrast to XCI, it is unknown whether single-cell compartments of two autosomal alleles may vary in a coordinated manner. By decomposing the variability of single-cell compartments into between-cell and within-cell differences (fig. S12A), we found that autosomal alleles fluctuate (with respect to their median compartments) almost independently from each other, exhibiting on average near-zero Spearman correlation (fig. S12D). Our conclusion held if compartments were defined on the basis of contacts (fig. S16).

We can pinpoint genomic changes, such as SNPs and CNVs, to their precise spatial locations in the cell nucleus. The donor of the GM12878 cell line carried a heterozygous G-to-A mutation (rs4244285) in the cytochrome P450 gene *CYP2C19*, leading to a truncated, nonfunctional protein variant CYP2C19*2 and affecting metabolism of hormones and drugs (*30*). Figure S18A shows the 3D localization of this drug-response SNP on the paternally inherited chromosome 10 of a GM12878 cell. In addition to inherited mutations, single cells also harbor somatic changes. In lymphocytes, somatic V(D)J recombination generates diversity of immunoglobulins and T cell receptors by DNA deletions and inversions. Figure S18B shows the 3D localization of two V(D)J recombinations at a T cell receptor locus, leading to two different DNA deletions on the two alleles of chromosome 14 of a T lymphocyte. The capability to spatially localize genomic changes is important for studying cancers and inherited diseases, where mutations can have severe consequences and may disrupt the chromatin structure of nearby regions.

We also examined the cell type dependence of 3D genome structures. Similar to haploid mESCs (*11*), chromosomes in diploid mESCs preferred the Rabl configuration (centromeres pointing toward one side of the nucleus and telomeres toward the other), albeit to a different extent in each cell (Fig. 4A). In contrast, we found the Rabl configuration to be weak in most GM12878 cells and PBMCs. Most PBMCs pointed their centromeres toward the nuclear periphery and telomeres toward the nuclear center, consistent with previously reported arrangements in human lymphocytes (*31*). By contrast, the M/G_1-phase GM12878 cell pointed centromeres toward the outer rim of a characteristic mitotic rosette.

The overall extent of chromosome intermingling also differed among the cell types. Chromosomes tended to intermingle less in mESCs and more in PBMCs, with GM12878 intermediate between them (Fig. 4, B and C), consistent with

previous reports that chromosomes intermingle less in the pluripotent mESCs than in terminally differentiated fibroblasts (*32*) and that chromosomes intermingle more in resting human lymphocytes than in activated ones (which resembled GM12878) (*33*). As expected (*10*, *34*), the M/G_1-phase cell exhibited a low level of chromosome intermingling and the lowest level of chromatin compartmentalization.

Cell type–dependent promoter-enhancer looping has been suggested to underlie differential gene expression (*35*). Among the human cells, differential formation of known cell type–specific promoter-enhancer loops [based on cell type–purified bulk Hi-C (*15*, *35*)] clearly separated the single cells into four cell type clusters: B lymphoblastoid cells (GM12878), presumable T lymphocytes, B lymphocytes, and presumable monocytes/neutrophils (Fig. 4D). Defining loop formation on the basis of contacts rather than 3D structures yielded similar results (fig. S17A).

Cell type clusters could be equally well separated in an unsupervised manner, without prior knowledge of the cell types. Unlike ensemble-averaged structures such as protein crystal structures, single-cell 3D genomes are intrinsically stochastic and dynamic. Statistical characterization such as PCA is necessary to distinguish different cell types, in which clusters of single cells correspond to valleys in a Waddington landscape (*36*) of certain cellular phenotypes. This kind of cell typing has been carried out using phenotype variables such as single-cell transcriptomes (*37*) and open chromatin regions (*38*, *39*), each of which must have underlying structural differences in the 3D genome.

With Dip-C, we are in a position to carry out cell typing with genome structure as the sole variable. Given the high information content of 3D structures, many possible features might be used in cluster analysis. Here, we chose single-cell chromatin compartments as the input variable of PCA. The four cell type clusters were clearly separated (Fig. 4E), with one of the most differentially compartmentalized regions shown in Fig. 4F. Our conclusion held if compartments were defined on the basis of contacts (fig. S17, B and C). Previous reports (*7*, *8*, *10*–*12*) had focused on defining the width (or spread) of a single Waddington valley, studying, for example, cell cycle dynamics within a cell type and domain stochasticity within a cell cycle phase. Our PCA result, in contrast, highlighted the consistent difference among cell types, signifying the separation between Waddington valleys.

Our initial examination of only a handful of cell types has clearly shown the tissue dependence of 3D genome structures. A systematic survey of more cell types under various conditions will likely lead to new discoveries in cell differentiation, carcinogenesis, learning and memory, and aging.

REFERENCES AND NOTES

1. T. Cremer, C. Cremer, *Nat. Rev. Genet.* **2**, 292–301 (2001).
2. I. Solovei *et al.*, *Cell* **137**, 356–368 (2009).
3. E. J. Clowney *et al.*, *Cell* **151**, 724–737 (2012).
4. J. Dekker, K. Rippe, M. Dekker, N. Kleckner, *Science* **295**, 1306–1311 (2002).
5. E. Lieberman-Aiden *et al.*, *Science* **326**, 289–293 (2009).
6. K. E. Cullen, M. P. Kladde, M. A. Seyfred, *Science* **261**, 203–206 (1993).
7. T. Nagano *et al.*, *Nature* **502**, 59–64 (2013).
8. I. M. Flyamer *et al.*, *Nature* **544**, 110–114 (2017).
9. X. Li, J. Zhang, H. Zhao, Z. Pei, Z. Xuan, patent application (2017); https://patents.google.com/patent/WO2017066908A1/en.
10. T. Nagano *et al.*, *Nature* **547**, 61–67 (2017).
11. T. J. Stevens *et al.*, *Nature* **544**, 59–64 (2017).
12. V. Ramani *et al.*, *Nat. Methods* **14**, 263–266 (2017).
13. S. Carstens, M. Nilges, M. Habeck, *PLOS Comput. Biol.* **12**, e1005292 (2016).
14. C. Chen *et al.*, *Science* **356**, 189–194 (2017).
15. See supplementary materials.
16. S. Lu *et al.*, *Science* **338**, 1627–1630 (2012).
17. J. R. Dixon *et al.*, *Nature* **518**, 331–336 (2015).
18. N. Servant *et al.*, *Genome Biol.* **16**, 259 (2015).
19. S. S. Rao *et al.*, *Cell* **159**, 1665–1680 (2014).
20. L. Giorgetti *et al.*, *Nature* **535**, 575–579 (2016).
21. W. J. Xie *et al.*, *Sci. Rep.* **7**, 2818 (2017).
22. S. Boyle *et al.*, *Hum. Mol. Genet.* **10**, 211–219 (2001).
23. Y. Zhou *et al.*, *Cell Res.* **27**, 298–301 (2017).
24. J. R. Dixon *et al.*, *Nature* **485**, 376–380 (2012).
25. J. Fraser *et al.*, *Mol. Syst. Biol.* **11**, 852 (2015).
26. C. Weinreb, B. J. Raphael, *Bioinformatics* **32**, 1601–1609 (2016).
27. S. A. Quinodoz *et al.*, *Cell* **174**, 744–757.e24 (2018).
28. J. T. Lee, M. S. Bartolomei, *Cell* **152**, 1308–1323 (2013).
29. E. M. Darrow *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E4504–E4512 (2016).
30. Coriell Institute for Medical Research, "Remarks" in record of female subject 12878 (2017); www.coriell.org/0/Sections/Search/Sample_Detail.aspx?Ref=GM12878.
31. C. Weierich *et al.*, *Chromosome Res.* **11**, 485–502 (2003).
32. S. Maharana *et al.*, *Nucleic Acids Res.* **44**, 5148–5160 (2016).
33. M. R. Branco, T. Branco, F. Ramirez, A. Pombo, *Chromosome Res.* **16**, 413–426 (2008).
34. N. Naumova *et al.*, *Science* **342**, 948–953 (2013).
35. B. M. Javierre *et al.*, *Cell* **167**, 1369–1384.e19 (2016).
36. C. H. Waddington, *The Strategy of the Genes: A Discussion of Some Aspects of Theoretical Biology* (Allen & Unwin, 1957).
37. F. Tang *et al.*, *Nat. Methods* **6**, 377–382 (2009).
38. J. D. Buenrostro *et al.*, *Nature* **523**, 486–490 (2015).
39. D. A. Cusanovich *et al.*, *Science* **348**, 910–914 (2015).
40. J. T. Robinson *et al.*, *Cell Syst.* **6**, 256–258.e1 (2018).
41. M. Jeong *et al.*, BioRxiv 212928 [Preprint]. 9 November 2017. https://doi.org/10.1101/212928.

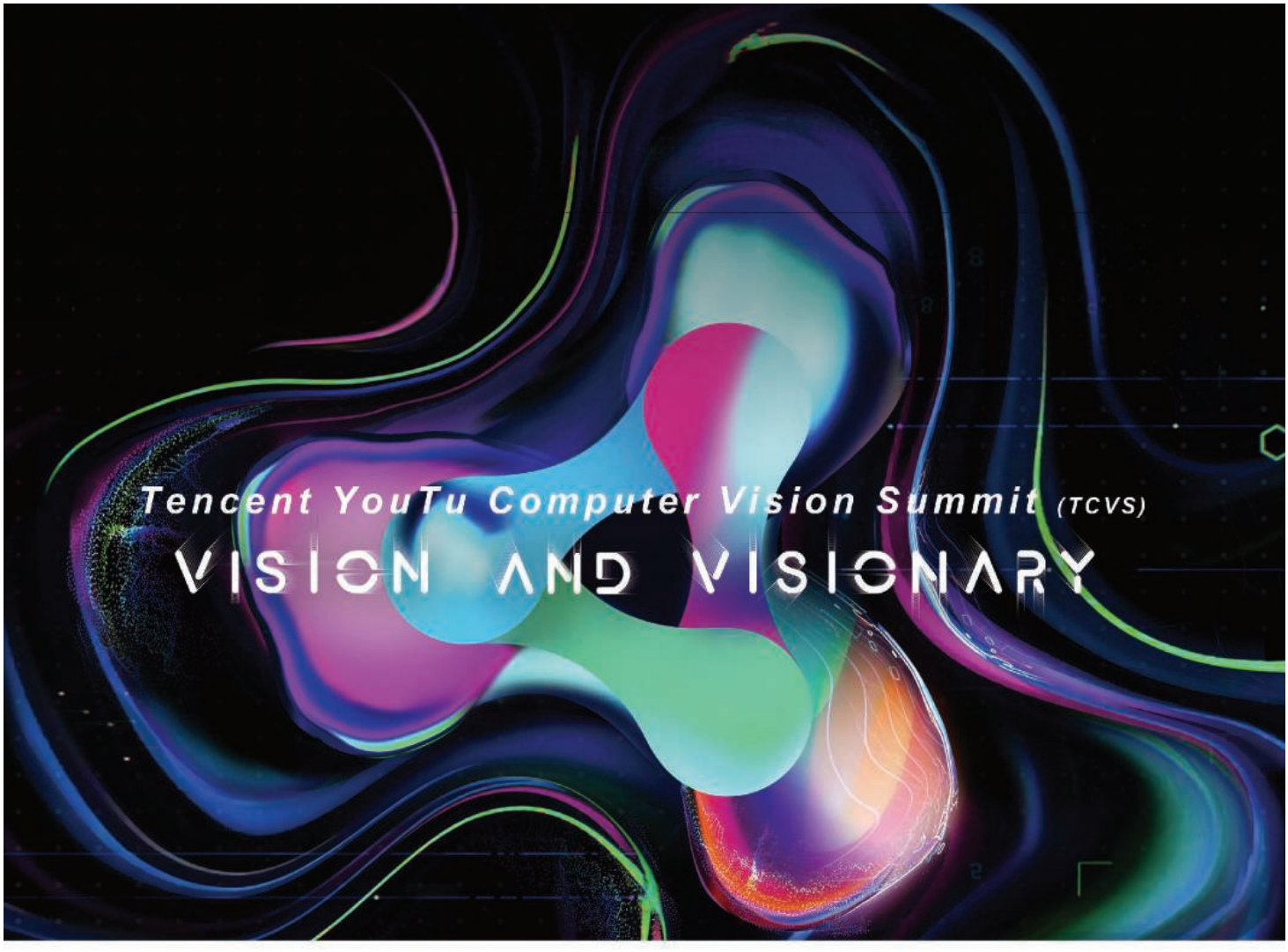
ACKNOWLEDGMENTS

We thank C. Zong (Harvard University, currently Baylor College of Medicine) for his involvement at the early stage; E. Lieberman Aiden (Baylor College of Medicine) and E. Stamenova (Broad Institute) for advice about their in situ Hi-C protocol; T. Stevens (University of Cambridge) for help on his simulated annealing software ("nuc_dynamics"); M. Yang (Peking University) for published phased genotypes of the blood donor; T. Nagano and S. Wingett (Babraham Institute) for answering questions about published raw data; and Y. Gao, L. Meng, and S. Liu (Peking University) for helpful discussion. **Funding:** Supported by the Beijing Advanced Innovation Center for Genomics at Peking University, an NIH Director's Pioneer Award (DP1 CA186693), a Harvard Brain Initiative (HBI) Collaborative Seed Grant, and two grants from the National Science Foundation of China (21390412 and 21327808) (X.S.X.); an HHMI International Student Research Fellowship (L.T.); and NHGRI grant R01 HG010040 (H.L.). **Author contributions:** L.T., D.X., C.-H.C., and X.S.X. designed the experiments; L.T. and D.X. performed the experiments; L.T. and H.L. analyzed the data. L.T., D.X., and X.S.X. wrote the manuscript. **Competing interests:** L.T., D.X., C.-H.C., and X.S.X. are inventors on a provisional patent application US 62/509,981 filed by Harvard University that covers META and Dip-C. **Data and materials availability:** Raw sequencing data were deposited at the National Center for Biotechnology Information with accession number SRP149125 at www.ncbi.nlm.nih.gov/sra/SRP149125. Processed data were deposited with GEO Series accession number GSE117876 at www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE117876. Code is available at GitHub (https://github.com/tanlongzhi/dip-c and https://github.com/lh3/hickit). Chromatin contacts can be viewed interactively in the Juicebox web browser (*40*); please see the manual page of our codes.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/361/6405/924/suppl/DC1
Materials and Methods
Figs. S1 to S20
Tables S1 and S2
References (*42*–*44*)

12 March 2018; accepted 6 August 2018
10.1126/science.aat5641



About Tencent YouTu Computer Vision Summit:

On September 6, 2018, Tencent YouTu Lab, an artificial intelligence team under Tencent, will cooperate with the world's top scientific journal *Science* to organize an advanced computer vision summit - Tencent YouTu Computer Vision Summit (TCVS). In this summit, world-renowned scholars, experts and partners of Tencent YouTu Lab from all over the world will gather together to share their perspectives on the latest technology development, scientific innovation, application and the future trend of computer vision. Moreover, Tencent YouTu Lab will also announce its recent research achievement and development strategy.

Time

2018.9.6 9:30-17:30 (GMT+8)

Webcast

Visit http://daxue.qq.com for subscription and live stream.

Tencent YouTu Lab | Science AAAS

Department of Defense

Integral in the Democratic Republic of the Congo Ebola Outbreak Response

By Dr. Clay Holloway & Hannah Feldman, Medical Countermeasure Systems

A Navy laboratory technician pipettes drug product for use in a filovirus vaccine. Photograph: Navy Visual News Service/MC3 Jake Berenguer

In the past decade, we have seen an increase in infectious disease outbreaks, most notably with the West Africa Ebola outbreak in 2014, which left more than 11,000 dead and 17,000 survivors, many of whom still require post-recovery medical care. The threat continues today. Public health officials and business leaders like Bill Gates have long warned that the world is not ready for the next pandemic.

Global health officials say it's more important than ever for disease outbreaks to be stopped at their source before they come full-blown epidemics. And that is exactly what the Department of Defense—in collaboration with the World Health Organization (WHO), interagency, and other international partners—did in response to the latest Ebola outbreak in the Democratic Republic of the Congo (DRC).

On May 8, the Ministry of Health of the DRC notified the WHO of an Ebola virus outbreak in Bikoro Health Zone in the northwest Equateur Province. A total of 54 cases and 33 deaths were reported with over 900 identified contacts. The WHO worked closely with the DRC government to rapidly scale up its operations and mobile health partners.

Ebola is endemic to the DRC. The latest outbreak occurred in 2017, and was quickly contained. The WHO attributed this success to the quick testing of blood samples, announcing the outbreak early, and a rapid response from health authorities.

Department of Defense Response

The Joint Program Executive Office for Chemical, Biological, Radiological, and Nuclear Defense (JPEO-CBRND) and the Joint Program Management Office for Medical Countermeasure Systems (JPM-MCS) played a significant role in the U.S. Government's response to the recent outbreak. The JPEO-CBRND is the Department of Defense's single focal point for the developing and fielding of chemical, biological, radiological, and nuclear defense equipment and medical countermeasures.

The JPEO recognized the importance of contributing to the Ebola effort based on its previous and continued development of medical countermeasures to protect, identify, and treat the Ebola virus. "We are utilizing a whole-of-government approach to this outbreak. While our mission space is ultimately to support the warfighter, it is our imperative to support our domestic and international public health partners in this response," said Mr. Douglas Bryce, Joint Program Executive Officer for JPEO-CBRND. "We are supporting the global community in whatever way we can."

Prevention

The MCS product offices responsible for vaccine and platform technology development engaged in multiple efforts to accelerate development of Ebola vaccine candidates. Of note, the WHO and Médecins Sans Frontières (MSF) utilized the Department of Defense-funded r-VSVΔG-ZEBOV vaccine in a ring vaccination campaign. MCS, in collaboration with the Defense Threat Reduction Agency, provided early testing and development activities for the vaccine. During the exercise, more than 3,000 people were vaccinated. MCS currently maintains a 23,000 multi-dose vial stockpile of the investigational vaccines in case additional doses are required.

Diagnosis

The WHO developed a National Laboratory Strategy in response to the outbreak. As part of this effort, the Defense Biological Product Assurance Office deployed two co-funded Nebraska Strategic Research Institute personnel to support sequencing efforts. The personnel took sequencing equipment, consumables, and reagents to establish a sequencing capability at the Institut National de Recherche Biomédicale within the DRC. This capability sequenced Ebola samples from the outbreak so the data can be shared with the biodefense community on critical reagents for the Ebola virus.

Treatment

MCS's biological therapeutics office worked closely with the WHO and MSF on their Ebola drug treatment in development. The drug Remdesivir, developed by Gilead Sciences and funded by the Department of Defense, was deployed to the DRC for emergency use during the outbreak.

In addition, the MCS platform technologies office partnered with the National Allergy and Infectious Disease Vaccine Research Center to rapidly manufacturing their VRC-114 anti-Ebola monoclonal antibody in support of conducting clinical trials for a vaccine and therapeutic in the DRC. By leveraging the DOD's existing monoclonal antibody platform and Advanced Development and Manufacturing facility, MCS can rapidly produce additional doses of the VRC-114 to be used in treating Ebola infected individuals, thereby accelerating FDA approval of an Ebola virus therapy to U.S. Forces and the WHO.

Results

The current Ebola outbreak is contained, thanks in part to the JPEO-CBRND's support to U.S. Government and international partners. The DRC Ministry of Health and WHO announced the official end of the outbreak on July 24. According to the WHO, there still remains a risk for resurgence or flare-ups posed by potentially undetected transmission chains, such as the latest outbreak in eastern DRC that began at the beginning of August. However, strengthened surveillance mechanisms and survivor monitoring programs are in place to rapidly mitigate, detect, and respond to such events.

COL David Hammer, Joint Project Manager for MCS praised the Department's response to the outbreak. "Our team contributed greatly to containing the outbreak. The WHO's medical countermeasure options would've been constrained if the products developed by JPEO-CBRND and the DOD were not utilized. We couldn't have provided support without collaborating with our inter- and intra-agency partners, local teams, and the international community. The DOD's response saved lives," he said.

About Medical Countermeasure Systems

JPM-MCS, a component of the U.S. Department of Defense's Joint Program Executive Office for Chemical, Biological, Radiological, and Nuclear Defense, aims to provide U.S. military forces and the nation with safe, effective, and innovative medical solutions to counter chemical, biological, radiological, and nuclear threats.

To contact MCS: usarmy.detrick.dod-jpeo-cbrnd.list.communications@mail.mil

A novel target for accelerating drug development: Biomedical science training

Cherie Butts[1] and Avery August[2]

Demand for a qualified biomedical science workforce to tackle the challenges of making better medicines remains high; however, few scientists and clinicians learn about drug development during their training. To assist trainees with appreciating differences between basic science (understanding disease mechanisms) and applied science (drug development), Biogen and the Cornell Broadening Experiences in Scientific Training (BEST) program convened a conference in June 2018 at the Biogen headquarters in Cambridge, Massachusetts (#Biogen BESTDDConf2018).

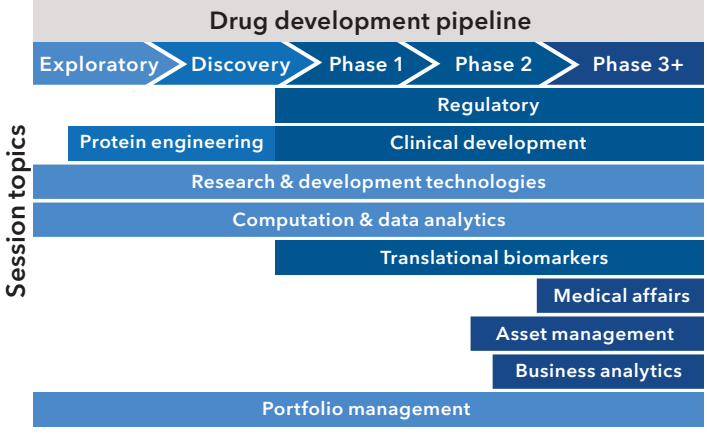


FIGURE 1.

Selection process

Participants were identified primarily from academic institutions with U.S. National Institutes of Health (NIH) BEST programs (www.nihbest.org; https://commonfund.nih.gov/workforce), as they are familiar with biopharma career pathways. Trainees* were exposed to key drug development questions, different roles in and out of the laboratory or clinic, and skills needed to be successful in biopharma. To ensure that information from the conference extended beyond those who attended, a requirement was that trainees share key concepts with others at their home institutions.

Unique approach

The average time for developing a new drug is approximately 12 years and costs over USD 1 billion, predominantly due to failures at each stage of drug development (*1*). An appropriately trained workforce is one mechanism for accelerating timelines and reducing the risk of failure. As many biomedical sciences training programs do not offer activities related to drug development, trainees must opt for additional specialized fellowships (ranging in duration from several weeks to a few years) or transition to industry with little knowledge of the skills necessary to be successful in this sector. As an initiative of Biogen's Portfolio Transformation, a short-term, intensive conference was developed to demystify drug development for academic trainees. The goal was to create a model for similar events across the country.

High-performing project teams are a hallmark of biopharma, but are less common in academia (*2*). Conference activities, therefore, focused on providing participants with a project-team experience that highlighted key drug development questions; stage-appropriate composition of project teams; the importance of team dynamics and of maximizing the strengths of each member; and how the biopharma ecosystem supports project teams.

Purposeful outcomes

Participants were introduced to the drug development process (from concept to approval), and sessions were led by individuals from across Biogen, who offered insight on their roles—including how they support project teams. The topics included asset management, biomarker development, business and data analytics, clinical development, medical affairs, portfolio management, protein engineering, and regulatory affairs and policy (Figure 1). In addition, participants served on teams that generated a business case and recommendations for progression of a mock project to the next drug development stage (Figure 2).

Refining and reframing

A new training model is needed to strengthen and refine the necessary skills for those who wish to translate new biomedical discoveries into beneficial drugs. More trainees with the right experience will increase the pace of drug development, reducing the burden of debilitating medical conditions on society. Such a reframing of the training experience will positively change the conduct of science and expand the ways that meaningful contributions to biomedical science are defined. This conference emphasizes the importance of experiential learning and serves as a model for such training.

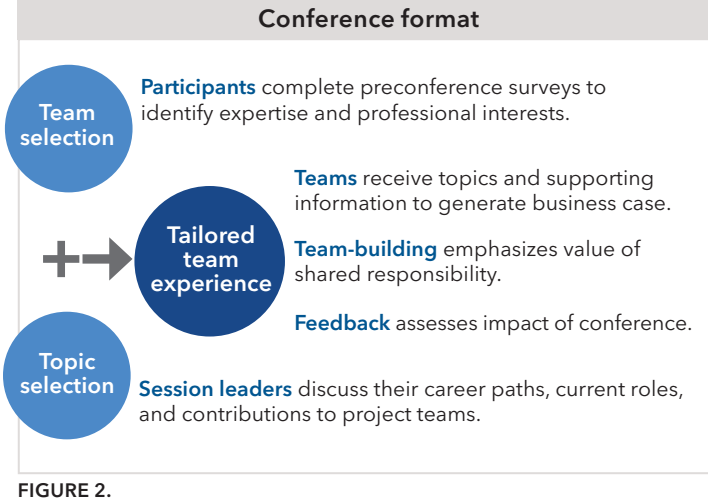


FIGURE 2.

References

1. G. A. Van Norman, *JACC: Basic to Translational Science* **1,** 170–179 (2016).
2. J. R. Katzenbach, D. K. Smith, *Harvard Business Review* **71**, 111–120 (1993).

For additional information and to explore future opportunities with this drug development training model, please contact the authors: cherie.butts@biogen.com and averyaugust@cornell.edu.

*40 Ph.D. students, postdocs, and medical students from Boston University; Cornell University; Emory University; Georgia Institute of Technology; Johns Hopkins University; Meharry Medical College; Michigan State University; New York University; Rutgers University; Universities of California at Davis, Irvine, and San Francisco; Universities of Colorado at Anschutz and Denver; University of North Carolina at Chapel Hill; University of Massachusetts Medical School; University of Rochester; Vanderbilt University; Virginia Polytechnic Institute and State University; and Wayne State University.

[1]Portfolio Transformation and Late-Stage Clinical Development, Biogen, Cambridge, MA; cherie.butts@biogen.com [2]Cornell BEST Program and Department of Microbiology and Immunology, Cornell University, Ithaca, NY; averyaugust@cornell.edu

Biogen | CORNELL BEST PROGRAM

2019 MRS® SPRING MEETING & EXHIBIT

April 22–26, 2019 | Phoenix, Arizona

CALL FOR PAPERS

Abstract Submission Opens
September 28, 2018

Abstract Submission Closes
October 31, 2018

Spring Meeting registrations include MRS Membership July 1, 2019 – June 30, 2020

BROADER IMPACT

BI01 High Impact Practice—Increasing Ethnic and Gender Diversification in Engineering Education

CHARACTERIZATION, PROCESSING AND THEORY

CP01 Advances in *In Situ* Experimentation Techniques Enabling Novel and Extreme Materials/Nanocomposite Design
CP02 Design and *In Situ* TEM Characterization of Self-Assembling Colloidal Nanosystems
CP03 Advances in *In Situ* Techniques for Diagnostics and Synthetic Design of Energy Materials
CP04 Interfacial Science and Engineering—Mechanics, Thermodynamics, Kinetics and Chemistry
CP05 Materials Evolution in Dry Friction—Microstructural, Chemical and Environmental Effects
CP06 Smart Materials for Multifunctional Devices and Interfaces
CP07 From Mechanical Metamaterials to Programmable Materials
CP08 Additive Manufacturing of Metals
CP09 Mathematical Aspects of Materials Science—Modeling, Analysis and Computations

ELECTRONICS AND PHOTONICS

Soft Organic and Bimolecular Electronics

EP01 Liquid Crystalline Properties, Self-Assembly and Molecular Order in Organic Semiconductors
EP02 Photonic Materials and Devices for Biointerfaces
EP03 Materials Strategies and Device Fabrication for Biofriendly Electronics
EP04 Soft and Stretchable Electronics—From Fundamentals to Applications
EP05 Engineered Functional Multicellular Circuits, Devices and Systems
EP06 Organic Electronics—Materials and Devices

Semiconductor Devices, Interconnects, Plasmonic and Thermoelectric Materials

EP07 Next-Generation Interconnects—Materials, Processes and Integration
EP08 Phase-Change Materials for Memories, Photonics, Neuromorphic and Emerging Application
EP09 Devices and Materials to Extend the CMOS Roadmap for Logic and Memory Applications
EP10 Heterovalent Integration of Semiconductors and Applications to Optical Devices
EP11 Hybrid Materials and Devices for Enhanced Light-Matter Interactions
EP12 Emerging Materials for Plasmonics, Metamaterials and Metasurfaces
EP13 Thermoelectrics—Materials, Methods and Devices

ENERGY AND SUSTAINABILITY

Energy Storage

ES01 Organic Materials in Electrochemical Energy Storage
ES02 Next-Generation Intercalation Batteries
ES03 Electrochemical Energy Materials Under Extreme Conditions
ES04 Solid-State Electrochemical Energy Storage

Catalysis, Alternative Energy and Fuels

ES05 Cooperative Catalysis for Energy and Environmental Applications
ES06 Atomic-Level Understanding of Materials in Fuel Cells and Electrolyzers
ES07 New Carbon for Energy—Materials, Chemistry and Applications
ES08 Materials Challenges in Surfaces and Coatings for Solar Thermal Technologies
ES10 Rational Designed Hierarchical Nanostructures for Photocatalytic System
ES11 Advanced Low Temperature Water-Splitting for Renewable Hydrogen Production via Electrochemical and Photoelectrochemical Processes
ES12 Redox-Active Oxides for Creating Renewable and Sustainable Energy Carriers

Water-Energy Materials and Sustainability

ES09 Advanced Materials for the Water-Energy Nexus
ES13 Materials Selection and Design—A Tool to Enable Sustainable Materials Development and a Reduced Materials Footprint
ES14 Materials Circular Economy for Urban Sustainability

Photovoltaics and Energy Harvesting

ES15 Fundamental Understanding of the Multifaceted Optoelectronic Properties of Halide Perovskites
ES16 Perovskite Photovoltaics and Optoelectronics
ES17 Perovskite-Based Light-Emission and Frontier Phenomena—Single Crystals, Thin Films and Nanocrystals
ES18 Frontiers in Organic Photovoltaics
ES19 Excitonic Materials and Quantum Dots for Energy Conversion
ES20 Thin-Film Chalcogenide Semiconductor Photovoltaics
ES21 Nanogenerators and Piezotronics

QUANTUM AND NANOMATERIALS

QN01 2D Layered Materials Beyond Graphene—Theory, Discovery and Design
QN02 Defects, Electronic and Magnetic Properties in Advanced 2D Materials Beyond Graphene
QN03 2D Materials—Tunable Physical Properties, Heterostructures and Device Applications
QN04 Nanoscale Heat Transport—Fundamentals
QN05 Emerging Thermal Materials—From Nanoscale to Multiscale Thermal Transport, Energy Conversion, Storage and Thermal Management
QN06 Emerging Materials for Quantum Information
QN07 Emergent Phenomena in Oxide Quantum Materials
QN08 Colloidal Nanoparticles—From Synthesis to Applications

SOFT MATERIALS AND BIOMATERIALS

SM01 Materials for Biological and Medical Applications
SM02 Progress in Supramolecular Nanotheranostics
SM03 Growing Next-Generation Materials with Synthetic Biology
SM04 Translational Materials in Medicine—Prosthetics, Sensors and Smart Scaffolds
SM05 Supramolecular Biomaterials for Regenerative Medicine and Drug Delivery
SM06 Nano- and Microgels
SM07 Bioinspired Materials—From Basic Discovery to Biomimicry

www.mrs.org/spring2019

Meeting Chairs

Yuping Bao The University of Alabama
Bruce Dunn University of California, Los Angeles
Subodh Mhaisalkar Nanyang Technological University
Ruth Schwaiger Karlsruhe Institute of Technology—Institute for Applied Materials
Subhash L. Shinde University of Notre Dame

Don't Miss These Future MRS Meetings!

2019 MRS Fall Meeting & Exhibit
December 1–6, 2019, Boston, Massachusetts

2020 MRS Spring Meeting & Exhibit
April 13–17, 2020, Phoenix, Arizona

MRS MATERIALS RESEARCH SOCIETY®
Advancing materials. Improving the quality of life.
506 Keystone Drive • Warrendale, PA 15086-7573
Tel 724.779.3003 • Fax 724.779.8313 • info@mrs.org • www.mrs.org

Glass Media Bottles

PUREGRIP borosilicate glass media bottles use a cap with a protruding ridge that improves a user's ability to safely handle the bottle with gloved hands, especially when the bottles are wet. The patented cap is easy to open or close, and provides a wide, flat surface for writing and labeling. Constructed of durable 3.3 low-expansion ASTM E438 Type I Class A borosilicate glass, PUREGRIP bottles come in sizes of 100 mL, 250 mL, 500 mL, 1 L, 2 L, 3 L, 5 L, 10 L, and 20 L. Each bottle in the product line includes volumetric gradations that have been individually certified for accuracy.

Foxx Life Sciences
For info: 603-890-3699
www.foxxlifesciences.com

Contained Filtration System

The SupaClean system from Amazon Filters is designed for filtration applications where quick cleandowns are required in a multipurpose production site, or where hazardous materials are being processed and risk of operator exposure needs to be minimized. The system consists of Amazon Filter's standard range of high-performance filter cartridges enclosed inside two sealed plastic bags, all of which are contained inside a stainless-steel housing that provides the pressure vessel required for the filtration process. SupaClean is ideally suited to the production of coatings. Extensive validation testing ensures that the plastic bags enclosing the filters and product always remain attached even when full of liquid. Filter changeout is quick and simple—the filter assembly can be simply lifted out of the housing and replaced by a new unit, with minimal risk of the operator or surrounding area encountering the product.

Amazon Filters
For info: +44-(0)-1276-670600
www.amazonfilters.com

Secondary Antibodies

Brilliant Violet 421 and Brilliant Violet 480 conjugated secondary antibodies allow you to add more colors to your multiple labeling assays in the violet–blue region of the spectrum. When combined with Alexa Fluor 488, Rhodamine Red-X, and Alexa Fluor 647 conjugates, effective five-color fluorescent labeling is possible. If nuclear counterstaining is desired, four-color antibody staining is possible using Brilliant Violet 421, Brilliant Violet 480, Alexa Fluor 488, and Rhodamine Red-X. Switching the nuclear stain from DAPI (emission in the blue region) to DRAQ5 (which has red emission) frees the violet–blue region of the spectrum to accommodate the two Brilliant Violet dyes. DRAQ5's excitation and emission profiles overlap those of Alexa Fluor 647. We offer a range of secondary antibodies that are recommended for multiple labeling due to their minimal cross-reactivity.

Jackson ImmunoResearch
For info: 800-367-5296
www.jacksonimmuno.com

Chemically Competent and Electrocompetent Cells

Vmax Express is a novel, fast-growing bacterial strain designed and optimized for high-level recombinant protein expression. This rationally engineered, next-generation prokaryotic protein expression system can serve as a replacement for slow-growing *Escherichia coli* systems that are prone to low yields and the expression of proteins as insoluble inclusion bodies. Vmax cells are derived from the marine microorganism *Vibrio natriegens*. This gram-negative, nonpathogenic bacterium has a doubling time twice that of *E. coli* and produces three to four times its biomass. Utilizing a robust transcription and translation system to support this rapid growth rate, Vmax Express generates more biomass and more recombinant protein per liter faster. It is compatible with plasmids and antibiotics commonly used with bacterial expression systems such as *E. coli* BL21(DE3). Designed with a tightly controlled, IPTG-inducible T7 promoter system, Vmax Express cells can be cultured using routine growth medium as well as commercial autoinduction media, or our Vmax Enriched Growth Media.

SGI-DNA
For info: 855-474-4362
www.sgidna.com/vmax-express.html

Lentivector Plasmid and Prepackaged Virus

System Biosciences' family of Lenti-Labeler constructs facilitate a wide range of studies—including cell tracking, high-throughput assays, and more—by enabling efficient, reliable labeling of your cells. The pLL-CMV-GFP-T2A-Puro Lenti-Labeler construct expresses copepod green fluorescent protein (GFP) from the cytomegalovirus (CMV) promoter, which delivers strong expression in most commonly used cell lines (HeLa, HEK293, HT1080, etc.), and coexpresses the puromycin resistance gene for selection in vitro prior to in vivo use. Available as either fully propagatable, sequence-verified plasmid DNA or ready-to-transduce prepackaged lentivirus, pLL-CMV-GFP-T2A-Puro Lenti-Labeler is designed for reliability, so you can get valuable insights faster.

System Biosciences
For info: 888-266-5066
www.systembio.com

F-Actin Recombinant Protein

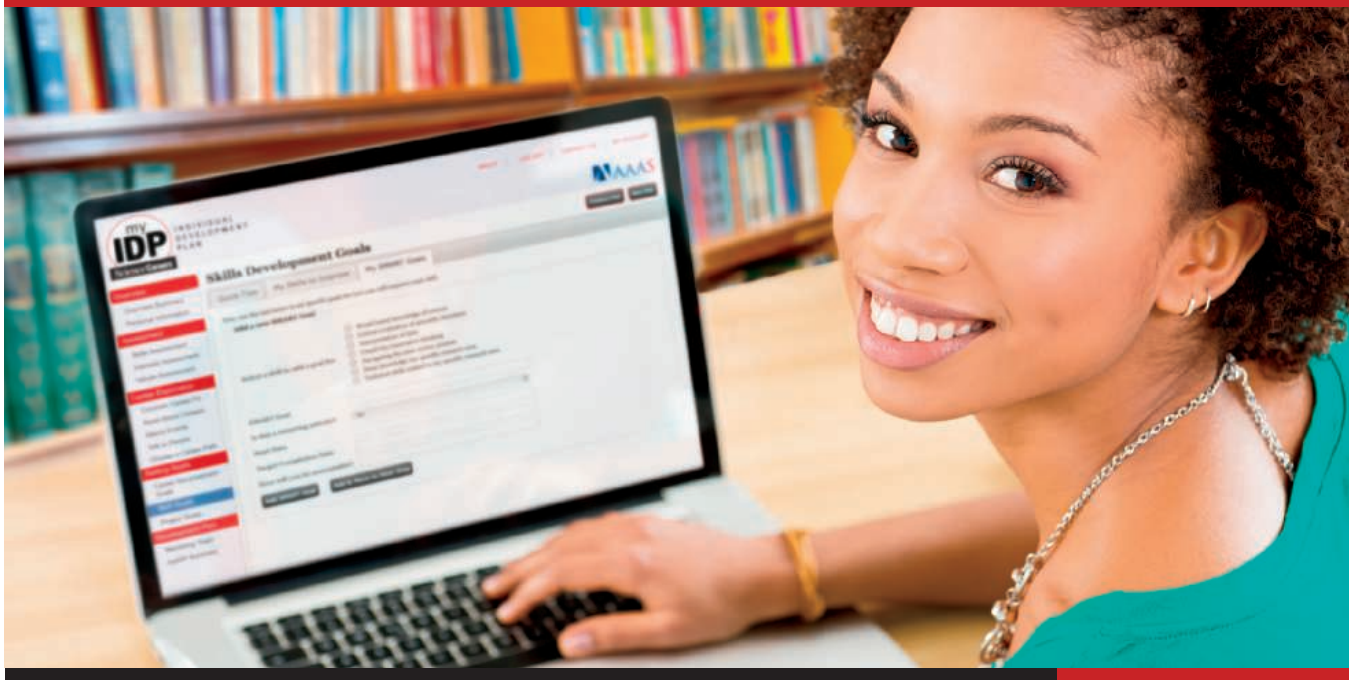
Achieve fast staining and immediate functional analysis of filamentous actin (F-actin) in living and fixed cells with ibidi's new reagent, LifeAct-TagGFP2 Protein. The addition of this recombinant protein extends a nontoxic, noninvasive alternative to phalloidin. F-actin is a key component of the cytoskeleton and is involved in many important cellular processes, such as cell division, cell migration, and endocytosis. LifeAct—a 17-amino-acid peptide derived from a protein found in yeast—specifically binds F-actin structures in living or fixed eukaryotic cells, while uniquely retaining full actin dynamics. Conjugated with green fluorescent protein (GFP), it can be conveniently introduced into cells, making it the ideal tool for F-actin labeling. With LifeAct-TagGFP2 Protein, scientists can easily transfect living cells, then get brilliant fluorescence microscopy images of cellular F-actin structures for analysis within minutes.

ibidi
For info: 844-276-6363
ibidi.com

Electronically submit your new product description or product literature information! Go to www.sciencemag.org/about/new-products-section 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.

myIDP: A career plan customized for you, by you.



For your career in science, there's only one Science

my IDP ScienceCareers

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

ScienceCareers In partnership with:

advancing the life sciences 1912-2012 FASEB Federation of American Societies for Experimental Biology

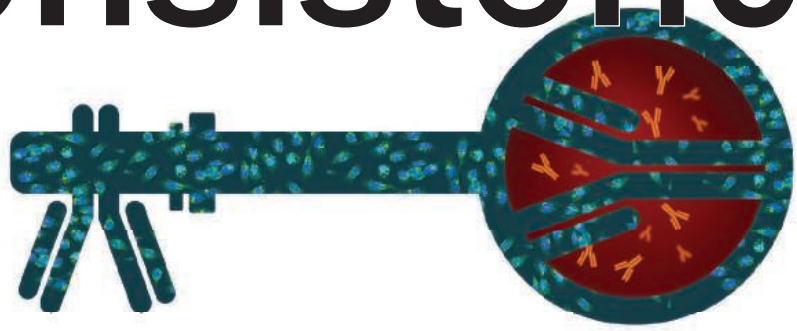
UCSF University of California San Francisco

Medical College of Wisconsin

Burroughs Wellcome Fund

R&D SYSTEMS
a bio-techne® brand

Consistency Is



Get Results The First Time And Every Time

With R&D Systems® Antibodies

Learn more | rndsystems.com/antibodies

bio-techne®

Global bio-techne.com info@bio-techne.com TEL +1 612 379 2956 **North America** TEL 800 343 7475
Europe | Middle East | Africa TEL +44 (0)1235 529449 **China** info.cn@bio-techne.com TEL +86 (21) 52380373

Peking University Third Hospital——

Healthcare, Education and Innovation for 60 years

北京大学第三医院建院60周年
60th Anniversary of Peking University Third Hospital
1958-2018

Jie Qiao, **MD.**
President of The Peking University Third Hospital

2018 marks the 60th anniversary of the founding of Peking University Third Hospital (PUTH), which was founded in October 1958 under the supervision of the Ministry of Health. As a hospital affiliated to China' s first national comprehensive university in China, PUTH is young and energetic, enjoying a rapid pace of development. Consistently ranked as a top tertiary hospital nationwide, PUTH integrates world-class medical services with cutting-edge medical research and first-rate medical education. Currently, it has a total of 7 campuses and over 6000 faculty and staff members. It has produced many notable scientists with national and international reputations, including one academician of the Chinese Academy of Sciences (CAS) and one academician of the Chinese Academy of Engineering (CAE).

Over the past 60 years, PUTH has always ranked among the capital' s top hospitals by providing high quality medical care. In 2017, the hospital received more than 4 million ambulatory and emergency visits. The annual discharge number was over 100,000, the annual operation number over 60,000, and the average length of stay was 5.78 days. Among the inpatients with intractable diseases, about 1/3 are from the other parts of China.

In accordance with the principle of academic excellence, PUTH endeavors to establish world-class disciplines and research platforms, cultivate high-level experts with global vision, and provide excellent medical education. PUTH boasts one National Clinical Research Center, three Ministry of Education Key Laboratories, one Ministry of Health Key Laboratory, and over 5,000 square meters for Public Laboratory Service. Its faculty members have held over 40 positions as chief editors of top journals which are included in the China Scientific and Technical Papers and Citations Database (CSTPCD), over 110 positions in academic associations/societies, and won many international awards.

PUTH has a strong sense of social responsibility and is actively engaged in China' s public hospital reform. In recent years, it has undertaken many key national projects and programs, with research findings widely adopted by relevant government sectors. With improved management skills, it has played an exemplary role in clinical pathway, care service, counterpart support, and the Aid Program for Tibet and Xinjiang. Meanwhile, PUTH is also a care provider for unexpected emergencies on important occasions, such as the Beijing 2008 Summer Olympics, the 60th anniversary of the People's Republic of China, and many earthquake and disaster relief projects.

With 60 years of history, PUTH has created its distinct culture. All PUTH people will follow the motto of "Unity, Dedication, Practicality and Creativity" , and through high medical skills and a strong sense of duty, strive to become a medical center of national prominence, a diagnosis and treatment center for rare and grave diseases, a research center for clinical medicine, and a training base for medical experts.

The Department of Cardiology

The Department of Cardiology, as a key clinical discipline, boasts a key Cardiovascular Molecular Biology and Regulatory Peptides Laboratory. Committed to addressing the clinical problems of major cardiovascular diseases, our department conducts basic research on the risk factors and pathogenesis of coronary heart disease and heart failure, precautionary function of biomarkers, and the protective mechanisms of cardiac rehabilitation, from the aspects of susceptibility genes, biomarkers, receptors in cardiovascular diseases, drug therapy and exercise rehabilitation. Taking full advantage of its asset as part of a comprehensive hospital, our department has made progress in many fields, such as the diagnosis and treatment of critical cardiovascular disease, the evaluation of coronary artery disease, hybrid coronary intervention (HCR) and minimally invasive surgical bypass grafting, cardiac rehabilitation and cardiovascular drug safety. Over the past decade, the department has built and improved its many platforms, i.e. for cardiovascular function analysis, omics and medical bioinformatics, clinical trial on internationally standardized drugs, and cardiac rehabilitation. It has established a coronary heart disease cohort, a biorepository containing over 13, 000 samples, as well as a big data platform for Cardiovascular Disease Resource Repository, which collects research data on coronary heart disease, heart failure, hypertension, and arrhythmia with a follow-up system for quality assurance, laying a solid foundation for clinical research.

The Department of Gynecology and Obstetrics

Founded in 1958, the Obstetrics and Gynecology (OB-GYN) Department at PUTH is famous as the birthplace of the first IVF-ET baby in mainland China. As a national clinical and applied research unit for reproductive health and related diseases, it has been designated as the national Clinical Research Center for Obstetrics and Gynecology as well as the base for medical cooperation on Reproductive Health and Population between WTO and Peking University (PKU). In 2016 and 2017, our department ranked No.1 in terms of its scientific and technologic influence.

Funded by the Ministry of Science and Technology (MOST), the Ministry of Education (MOE), the Education and the Natural Science


Medical research


Operating Theatre


Outpatient Clinic

Foundation of China (NSFC), our department has embarked on research into the molecular mechanism of human gamete development control and the pathogenesis of common reproductive diseases, and built comprehensive platforms for reproductive endocrine epidemiology research, preimplantation genetic screening (PGS) development , single-cell multi-omics sequencing and bioinformatics analysis. Among its many other achievements is the genetic, epigenetic and transcriptomic map construction of human gamete in its embryonic development, the establishment of gene regulatory networks (GRN), together with several clinical RCT studies. By now 315 SCI papers have been published by such journals as Cell, Nature, Lancet, and JAMA.

In the future, the OB-GYN Department will carry out more clinical and transformational research and establish more platforms for international cooperation for the improvement of reproductive health in China.

The Department of Ophthalmology

Established in 1958, the Department of Ophthalmology has been committed to clinic service, professional training and medical research. Our highly skilled ophthalmologists and staff provide a complete range of services on cataract, glaucoma, cornea and external disease, retinal diseases, pediatric ocular disorders, refractive and plastic surgery. Our clinical expertise and sophisticated diagnostic and treatment procedures make our department a nationwide referral center.

Currently, we have more than 170 faculty members, including 13 full professors and 21 associate professors, with an annual average of about 260,000 outpatients and 16,000 surgeries. In 2017, our department was ranked 6th nationwide in terms of its scientific and technological influence.

In the last five years, our physician-scientists have received more than 20 research grants worth about RMB 15 million from the government and 41 grants approximately RMB 11.62 million from NGOs. We published more than 90 peer-reviewed papers and boasted 20 patent claims. As the base of Beijing Key Laboratory of Ophthalmology, we focus our research on the following fields:

1.Stem cell study in glaucoma and Age-related Macular Degeneration (AMD)

2. Neuroprotection of retina ganglion cell

3. Clinical trial on ocular trauma, secondary glaucoma and corneal diseases

4. Visual stimulation and vision-related neuroplasticity

5. Ocular Surface Bioengineering

We will keep updating knowledge and technology to provide greater customer service.

The Department of Orthopedics

Since its establishment in 1958, the Department of Orthopedics at PUTH has been the pioneer in the surgical treatment of lumbar disc herniation and cervical spondylosis, and won great fame ever after, particularly for its spine surgery, at home and abroad. After 60 years of unremitting efforts, our department has become an integrative medical center specialized in spine, joints and trauma related diseases. Committed to medical treatment, training, research and recuperation, our department is one of the most important orthopedic centers in China and ranks top three nationally in its discipline recognition and scientific & technologic influence.

Guided by the motto, “Healing with compassion, leading with excellence” , our department continues to gather and cultivate talents as the main strategy for improving its overall strength, to develop and adopt innovative treatment technologies to facilitate clinical diagnosis, and to conduct research on both common and intractable orthopedic diseases.

In recent years, our department has intensified the establishment of research platforms. Now we boast a complete set of platforms for clinical research, basic research and translational research, having made breakthroughs in the development of orthopedic implant devices, particularly by adopting microporous titanium alloy in spine surgery with the aid of 3D printing technology.

The Department of Sports Medicine

Peking University Institute of Sports Medicine, established in January 1959 by Professor Qu Mianyu, is the oldest and most prestigious of its kind in China. It provides a full range of medical services, teaching and research in four subspecialties: sports traumatology, sports rehabilitation, sports nutrition, and medical supervision.

As the only Prevention and Treatment Center for Sports Injuries and Diseases designated by China General Administration of Sport (SGAS), our institute takes the lead in every aspect, providing medical coverage for 36 national and Beijing municipal sports teams, attending to about 15,000 athletes (over 60 Olympic champions) every year. It also provides top-level healthcare for the general community, with over 120,000 outpatient visits and 7,200 surgical operations per year. For more than a half century, it has made exceptional contributions to the success of China’ s competitive sports in the world.

As the base of the Beijing Key Laboratory of Sports Injuries, the institute has formed a multidisciplinary research platform including molecular biology, cell biology, histopathology, biomechanics, molecular imaging, materials science and tissue engineering, etc. Its research interests mainly include:

1. understanding the pathogenesis of common sports injuries and degenerative joint diseases, and identifying novel therapeutic approaches;

2. tissue engineering technology for treatment of joint injuries;

3. the effects of exercise on human physiology and biochemistry and the underlying mechanism;

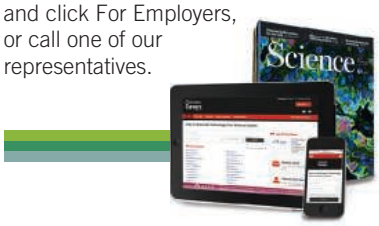
4. functional assessment and human performance.

PUTH cordially welcomes job applicants and visiting scholars with expertise in related areas. Feel free to contact us:
Website: http://www.puh3.net.cn
Email: puthdyb@bjmu.edu.cn
Tel: +86-10-82266699
Fax: +86-10-62017700
Address: Peking University Third Hospital, 49 North Garden Road, Beijing, China. 100191

Science Careers

SCIENCE CAREERS ADVERTISING

For full advertising details, go to ScienceCareers.org and click For Employers, or call one of our representatives.



AMERICAS

+1 202 326-6577
+1 202 326-6578
advertise@sciencecareers.org

EUROPE, INDIA, AUSTRALIA, NEW ZEALAND, REST OF WORLD

+44 (0) 1223 326527
advertise@sciencecareers.org

CHINA, KOREA, SINGAPORE, TAIWAN, THAILAND

+86 131 4114 0012
advertise@sciencecareers.org

JAPAN

+81 3-6459-4174
advertise@sciencecareers.org

CUSTOMER SERVICE

AMERICAS

+1 202 326-6577

REST OF WORLD

+44 (0) 1223 326528

advertise@sciencecareers.org

ScienceCareers
FROM THE JOURNAL SCIENCE AAAS

ScienceCareers.org

POSTDOCTORAL OPPORTUNITIES

CDC NIOSH

Research Fellow Position with CDC/NIOSH
Morgantown, WV

The National Institute for Occupational Safety and Health (NIOSH) at CDC is pleased to announce a research fellowship opening in the Receptor Biology Laboratory, located in Morgantown, WV. The research aims at understanding molecular and mechanistic events involved in the development of occupational diseases, with potential emphasis on pulmonary inflammation, fibrosis, and cancer that result from exposure to particulates, nanomaterials, and small chemicals.

QUALIFICATIONS include: (1) A Ph.D. or Master's degree in molecular biology, biochemistry, immunology, toxicology, pharmacology, or cancer research; (2) Experience conducting in vivo and in vitro studies evaluating molecular and mechanistic changes as a result of exposure; (3) Experience in developing and characterizing rodent models of pathologic phenotypes is desirable; (4) Excellent communication skills and the ability to work effectively and collegially. A strong background in molecular and transgenic approaches relating to disease and pathologic effects, and a good understanding of and experience in biochemistry, immunology, toxicology, and lung disease are highly desirable.

Salary is dependent upon academic degree and experience. *NIOSH is an Affirmative Action/Equal Opportunity Employer.* Send a letter with research experience and interests, Curriculum Vitae, and contacts of three referees via email to: **Dr. Qiang Ma** at **qam1@cdc.gov**.

VANDERBILT UNIVERSITY MEDICAL CENTER

An immediate opening for a postdoctoral position exists at Vanderbilt University Medical Center. The project is to examine the role of tumor-initiating cells in mouse models of colonic neoplasia using unique reporter mice generated in the lab. Individuals with experience in mouse genetics and confocal imaging/high resolution microscopy are encouraged to apply. Familiarity with modern methods for precise lineage-tracing, as well as examining high-dimensional data with multiplex immunofluorescence, smFISH, and single-cell RNA-seq, are desirable. Funding is secure, along with a highly competitive salary and fringe benefits. Those interested should apply online via Science Careers and may contact **Bob Coffey** at e-mail: **robert.coffey@vumc.org**.

The **National PKU Alliance (NPKUA)** works to improve the lives of individual with phenylketonuria and pursue a cure. NPKUA is pleased to release its 2019 Call for Research Proposals and Fellowships to continue this mission. Since 2010, NPKUA has invested $3 million in research that has led to new knowledge, acceleration of new therapies, and supported pilot studies that have enabled larger federal funding opportunities. More information can be found on the NPKUA.org website under Scientific Grant Request. For more information email **lex.cowsert@npkua.org**.

POSTDOCTORAL OPPORTUNITIES

STANFORD SCHOOL OF MEDICINE

POSTDOCTORAL FELLOW

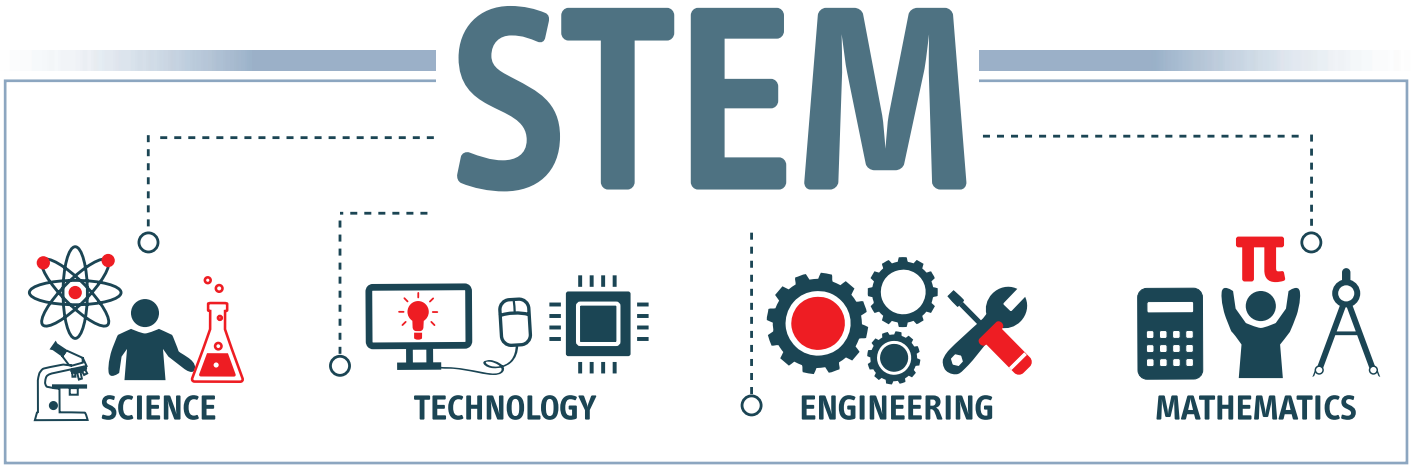
The **Lab of Dr. Idoyaga in the Microbiology and Immunology Department at Stanford University** (**http://idoyagalab.stanford.edu/**) seeks a **Postdoctoral Fellow** whose position will focus on the role of dendritic cells and microbiota during skin cancer progression. The successful candidate will use, among other techniques, CyTOF and RNAseq. The applicant must have, or be nearing completion of, a Ph.D. or M.D./Ph.D., and be familiar with FACS, immune cell isolation and mouse handling. Please send Curriculum Vitae and a list of 3 references to e-mail: **jidoyaga@stanford.edu**. *Stanford University is an Equal Opportunity Employer.*

MEDICAL SCHOOL
UNIVERSITY OF MICHIGAN

POSTDOC PREVIEW
March 7-8, 2019
University of Michigan, Ann Arbor

Fully funded travel for senior graduate students to interview with prospective mentors and tour campus. Applications open **November 20, 2018**.

Learn more at website: **http://bit.ly/UMPostdocPreview** or contact us at e-mail: **postdocoffice@med.umich.edu**



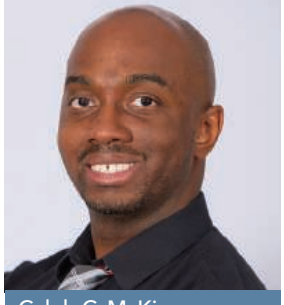
A unique calling: Careers in career development for STEM doctorates

Some people find they are more geared to careers helping others in their field than pursuing the field itself. In that vein, some Science, Technology, Engineering, and Mathematics (STEM) Ph.D.s are inspired to pursue careers in counseling STEM postdocs about their careers—sometimes by STEM and other Ph.D.s who are career counselors themselves. **By Alaina G. Levine**

Caleb C. McKinney, assistant dean of graduate and postdoctoral training and development at Georgetown University Medical Center in Washington, D.C., laughs when he thinks about how he maneuvered his Ph.D. in virology into a career in career development. As a postdoc at the U.S. National Institutes of Health (NIH), he wa s training students in his group on the practicalities of conducting research and found the experience to be personally transformative.

"I was writing a letter of recommendation for a student I had helped when I realized that I wanted to have these 'realized' moments on a bigger scale," McKinney says. He approached the Office of Training and Diversity at the National Institute of Allergy and Infectious Diseases (NIAID) about assisting them with their efforts, and subsequently volunteer-coordinated activities that fostered the professional development of NIAID fellows. His career, and his bliss, were on their way.

Some scientists and engineers who have navigated the knotty question of "what should I do with my life?" choose to remain on "Rue de Research" and pursue traditional academic professions. Then there are a few, like McKinney, who respond by deciding they want to help other Ph.D.s find impactful careers. And that's why we are seeing the emergence of the still fledgling field of career development for doctorates in STEM now buoyed by STEM doctorates themselves.


Caleb C. McKinney

One career, many paths

Along these lines, there is now a growth in formal institutional administrative divisions such as postdoc affairs offices, which help STEM grad students and postdocs think about their career opportunities. There are also organizations with missions to advance the careers of those in the career development profession, such as the Graduate Career Consortium (GCC) and the National Postdoctoral Association (NPA). Additionally, "train-the-trainers" programs, such as those organized by NIH that support the knowledge growth of investigators and other mentors who are working with protégés, are making an impact in expanding this profession.

"It's growing by leaps and bounds," says **Patrick Brandt,** director of career development and training in the Office of Graduate Education at the University of North Carolina at Chapel Hill. "10 years ago, there weren't many institutions hiring Ph.D.-trained professionals in this area."

Natalie Lundsteen, assistant dean for career and professional development at the University of Texas Southwestern Medical Center in Dallas (UT Southwestern), says her career was launched when she noticed "a need for someone with career skills and a Ph.D. to work with grad students, cont.>

IMAGES: (FROM TOP) © TRUEFFELPIX/SHUTTERSTOCK.COM; COURTESY OF GEORGETOWN UNIVERSITY MEDICAL CENTER

Upcoming features

Faculty: Alternative Funding Resources–September 14 ■ Top Employers Survey–October 26

which aligned with a big explosion in the world of career development in the late 2000s." Her dissertation research followed students pursuing internships at London banks, through which she discovered that the skills students need to succeed in a workplace are not necessarily linked to skills gained in academia. This finding inspired her to take on a career development role at the Massachusetts Institute of Technology, and she was subsequently recruited to UT Southwestern to build its career development division from scratch. Today, she assists biosciences Ph.D. students with their career strategies and is actively involved with GCC. Many of her advisees have gone on to careers in career development themselves.

To paraphrase an old Paul Simon tune, there are "50,000 ways" to leave your research and arrive at career development as a career. Immunologist **Lia Paola Zambetti** used communications to do so. She was a research fellow at the Singapore Immunology Network (SIgN) of the Agency for Science, Technology and Research (A*STAR) in Singapore, but wanted "to get out of the academic grind and find a job that was not related to the bench." She had already been engaged in science communications and had been freelancing as a popular science writer for several years. Through networking, she found a position in a new communications office at A*STAR. After three years there, she secured her current position at the University of Sydney, where she manages a fellowship program and organizes trainings in soft skills, such as public speaking, networking, and leadership for early-career researchers.

Tracy Costello's path was crystalized as a postdoc in genetics and biostatistics at the University of Texas MD Anderson Cancer Center in Houston. While still a fellow, she volunteered with and later served on the board of directors of NPA. "It solidified for me that I had the ability to impact people beyond my particular time, in that some of the things I was working on might not benefit me or my peers, but will benefit postdocs in the future," she says. "It was very freeing when I realized this was the direction I wanted to go."

After finishing her appointment, Costello did a brief foray in industry, but was quickly recruited back to MD Anderson to shape and launch its postdoc affairs office. An invitation to do the same at the Moffitt Cancer Center in Tampa, Florida, came four years later. Today, in addition to her job, she is serving as the chair of NPA.

"It solidified for me that I had the ability to impact people beyond my particular circle, in that some of the things I was working on might not benefit me or my peers, but will benefit postdocs in the future."
– Tracy Costello

What it takes

Each career development job is slightly different from the next, depending on many diverse factors including the institution's culture, funding stream, and history of career advising. Some STEM career development specialists report to the dean of the graduate school or vice president for student affairs and serve on a team of career advisors. Alternatively, they could be organized under the vice provost for research or graduate education and enjoy the title of dean or assistant dean themselves. Other positions are more solitary, such as in postdoc affairs offices, which could consist of a single person—the director—who serves as both program manager and support staff.

"You will wear many hats," explains Lundsteen, "including event planner, conflict resolver, manager, public relations manager, public speaker, salesperson, counselor, resource gatherer, and writer." Other responsibilities could include strategic planning, stakeholder relations, communications, project and program management, logistics, negotiation, policy, and fundraising. Luckily, many of the skills required to be successful in the career development arena mirror those that Ph.D. scientists and engineers have acquired through their research training, such as proposal development and data collection.

Being comfortable with event planning is generally a must, as the career development professional will be organizing all kinds of trainings, workshops, networking mixers, and speaker series for their constituents—often several events at once. Brandt adds that having a "service mindset" will help with event-planning. "You have to be OK with doing mundane tasks, such as reserving rooms and caterers and inviting keynote speakers and explaining to them what you want them to speak about," he says.

Naturally, strong communication skills are critical. "It's all about communicating backwards and forwards with the students, employers, alumni, and entities in the community," says Lundsteen, who regularly stays apprised of industry trends for her charges by interfacing with the regional chamber of commerce and reading the *Dallas Business Journal*. "I am a clearinghouse for opportunities and information, and my job is to be objective and to present alternatives to students."

Costello adds that the most important skills for success in this sector are being able to listen, to ask insightful questions, and to provide clients "a safe space to explore what they want to do."

Making the transition

For many career development professionals, the seeds for their career advancement were planted in the institutions in which they conducted their Ph.D. or postdoc research. They sought out opportunities to volunteer, assisted in career counseling efforts, did informational interviews, and demonstrated to the community that this was their passion and aspiration.

"The number one way to position yourself for a role like this is to get involved, whether it is at your local university or at the national level," urges Costello. "It's critical that people have a microexperience of a career, whether it is an internship or volunteer experience. It's a huge plus, because when you're actually applying for the job, you can say you are already engaged in it." cont.>

PHOTO: NICHOLAS GOULD

Faculty Cluster Hire in Artificial Intelligence

The University of Texas at San Antonio (UTSA) is seeking candidates to fill eight (8) faculty positions to foster collaborative research, education and outreach and to create interdisciplinary areas of knowledge that will advance the field of Artificial Intelligence (AI). All positions are either Tenure-Track Assistant, Associate or Full Professor level.

College of Architecture, Construction & Planning

Construction Science (1) — https://jobs.utsa.edu/postings/10151
Specializing in decision-making, decision support, unsupervised deep learning, current and future data for Internet of Things (IoT) and smart development of urban environments.

College of Business

Management Science & Statistics (1) — https://jobs.utsa.edu/postings/10148
Specializing in applied statistics and experience with AI, machine learning, and operational research within a multidisciplinary environment.

Information Systems and Cyber Security (2) — https://jobs.utsa.edu/postings/10100
Specializing in conducting research and developing tangible solutions to security challenges, particularly with gathering intelligence from unstructured data sets in real-time.

College of Engineering

Electrical and Computer Engineering (1) — https://jobs.utsa.edu/postings/10159
Specializing in AI, as it relates to smart health systems, electronic health records, digital health science, cloud computing, public health, or diagnostic radiology/computational imaging/image processing.

College of Liberal & Fine Arts

Psychology (1) — https://jobs.utsa.edu/postings/10099
Specializing in learning in complex data environments, resources-constrained AI processing, generalizable and predictable AI, deep learning, natural language processing, machine intelligence, super-intelligence, logics for intelligent interaction, logic for multi-agent systems in AI human factors, cyber psychology, privacy issues and healthcare applications.

College of Public Policy

Demography (1) — https://jobs.utsa.edu/postings/10101
Specializing in predictive modeling and data visualization for healthcare demand/burden, urban environments and planning, food environments and systems applications.

College of Sciences

Computer Science (1) — https://jobs.utsa.edu/postings/10158
Specializing in cyber adversarial learning, resource constrained AI, or AI as it relates to cloud computing, bioinformatics and other health-related applications.

Details/To apply: **http://research.utsa.edu/ai**

UTSA®
The University of Texas at San Antonio™

Featured participants

Georgetown University Biomedical Graduate Education, Postdoctoral Development
biomedicalprograms.georgetown.edu/postdoc

Moffitt Cancer Center
www.moffitt.org

Peters Lab, Maastricht University
www.maastrichtuniversity.nl/peter.peters

UNC School of Medicine, Office of Graduate Education, Science Training and Diversity (STaD)
www.med.unc.edu/oge/stad/about-us

University of Sydney
sydney.edu.au

UT Southwestern Medical Center, Graduate Career Development Office
www.utsouthwestern.edu/education/graduate-school/about-us/career-services

Additional resources

Graduate Career Consortium
gradcareerconsortium.org

National Postdoctoral Association
www.nationalpostdoc.org

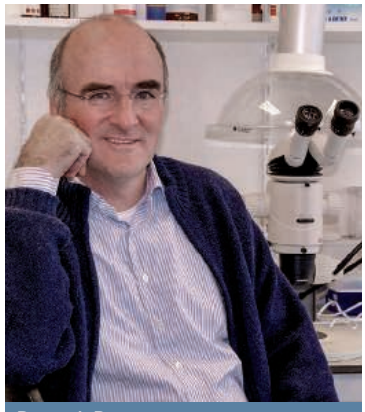

Peter J. Peters

PHOTO: COURTESY OF NICO ONG

Seek out these opportunities early, says Lundsteen, by getting involved in your university's grad student or postdoc association. It's even better to join these organizations' career committees so you get a greater "understanding the mechanics of the job," she advises. And take heed—if your institution doesn't have one of these organizations or committees, why not be an innovator and start one yourself?

One important aspect of making the transition is to ensure proper communications with your mentor. Regarding principal investigators (PIs) who have already demonstrated that they are open to you pursuing nonbench careers, it would be prudent to start a discussion with them early on to safeguard that smooth transition. With mentors on the other end of the spectrum, who may be less than enthused if you suggest you want to do anything outside the ivory tower, you should be careful about when and how you broach the subject, and try to do so in a safe manner that doesn't damage your relationship or endanger your employment arrangement.

McKinney took extra care to involve his PI in his plans early on. "I started developing project management platforms and working with my PI every week to make sure I was on target with my experimental deliverables, so I could get that extra time in the volunteer experience," he says. "I kept her informed as a key stakeholder."

But while you are engaging your PI and looking to do a side gig to gain experience, it's important not to sacrifice your research. "Be good in the lab because you want to have high credibility. You have to be taken seriously," warns **Peter J. Peters,** university professor and Limburg Chair at Maastricht University in The Netherlands.

While serving as the dean of postdoctoral affairs at the Netherlands Cancer Institute, Peters built the Postdoc Career Development Initiative (PCDI) to mentor and inspire young researchers at early stages of their scientific careers; it was later formally funded by the Dutch Ministry of Economic Affairs and became an independent organization. "People need to recognize you as someone who is good at science and a good citizen in the institution. Then, the director will give you money for your ideas," he says. "If you are sloppy in your work as a postdoc, you won't get momentum for your work at the institution."

Network to work

Some career development jobs for STEM postdocs are advertised, but Costello notes that relatively few open up each year—and as is the case for most of the sources interviewed here, they usually get their positions because their reputation is known among decision-makers and institutions. "It's gratifying to see the number of positions increase and the growth in this area, although it is still difficult to get into it," says Brandt. "You have to be willing to move where the jobs are."

You also have to network. "It doesn't matter if you are shy or introverted, nobody will do the work for you if you don't take ownership of your path," says Zambetti. "The only way out is networking—it's painful and tough and excruciating at the start, but it does get easier with practice."

Fortunately, most counselors are happy to help others who want to explore this profession. "Anyone in this profession would be willing to have someone shadow them or sit in on appointments," says Lundsteen. "We are the best people to ask for an informational interview because what we do for a living is tell people to do informational interviews."

The payoff for this path

One of the features of this career path that is especially gratifying for those with STEM degrees is that they get to remain a part of the scientific enterprise while they influence the next generation of scientists. Brandt loves the fact that he is still an active participant in higher education. He still publishes, although now it is in education research.

"I like the flow of the academic year, being on campus, and hearing the bell tower chime—I feel like I'm an academician," he says. "I also love the science and still feel like I get to vicariously enjoy it."

And then there is the definitive payoff. "I love my job every day because I get to help people figure things out. Hopefully they don't feel the pain and struggle I felt at not knowing where to go if it's not going to be faculty," says Costello. "Most mentors don't know how to guide you in any of the other career opportunities."

Lundsteen agrees: "My life's mission is showing people their capabilities and possibilities, and that brings me the greatest joy. I see I'm making a difference. I have the luxury of helping people and having them write me and say, 'You played a part.'"

Alaina G. Levine is a science writer, science careers consultant, professional speaker, and author of *Networking for Nerds* (Wiley, 2015).

WARREN ALPERT FOUNDATION

NEW POSTDOCTORAL FELLOWSHIP

The Warren Alpert Foundation announces the creation of the **Warren Alpert Distinguished Scholars program** that supports individual scientists of exceptional creativity who have an MD or PhD degree and who are post-doctoral fellows in the neurosciences in a medical school.

These awards are given as transitional awards before recipients become a member of a faculty at the Assistant Professor level or higher. Deans of medical schools are invited to submit one nomination.

Please see web page **www.warrenalpertfoundation.org** for details.

Applications are due on January 15, 2019.

Stony Brook University

Multiple Postdoctoral Positions Fall/Winter 2018

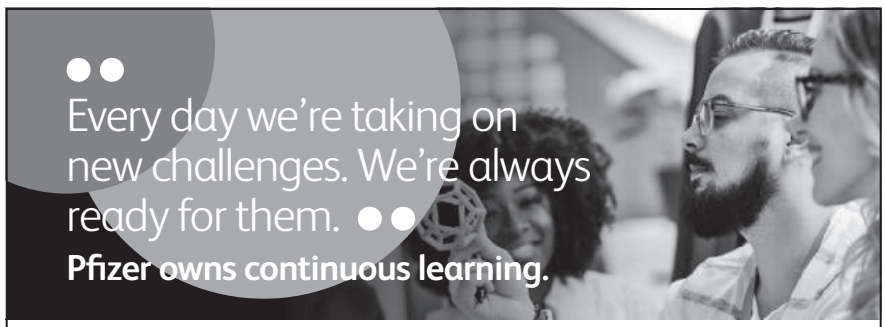
Stony Brook University is recruiting for multiple postdoctoral positions in various sub-specialties, for the upcoming fall and winter months.

Stony Brook has been characterized by innovation, energy and progress, and making ground-breaking discoveries since its beginning half a century ago.

Any interested candidates are invited to visit our JOBS page.

www.stonybrook.edu/postdocjobs

Stony Brook University/SUNY is an equal opportunity, affirmative action employer.



Pfizer Worldwide Research and Development Postdoctoral Program

At Pfizer, postdocs are trained to become successful, independent investigators, capable of formulating and addressing important scientific hypotheses. In addition, trainees receive broad exposure to the process of drug discovery, from idea to clinical trials. Areas of scientific focus include cardiovascular and metabolic diseases, comparative medicine, drug safety, biotherapeutics/protein engineering, inflammation and immunology, medicinal chemistry, oncology, pharmacology, vaccines, and clinical, computational, and genomic sciences.

We recruit highly motivated Ph.D. recipients with an outstanding track record of scientific productivity and a passion for ground-breaking, fast-paced research that facilitates the development of innovative therapies for human diseases. Our program promotes dissemination of research results through publications and participation in scientific meetings, provides opportunities for collaboration with leading academic labs and industry consortia, and offers exceptional professional development training and networking opportunities.

To explore our program and research, visit us online at:
www.pfizer.com/careers/en/postdoctoral-program

pfizercareers.com

Pfizer OWNIT!

POSTDOCTORAL OPPORTUNITIES

Jefferson Science Fellowship

USAID
FROM THE AMERICAN PEOPLE

The National Academies of Sciences, Engineering, and Medicine is pleased to announce a call for applications for the 2019 Jefferson Science Fellows (JSF) program. Initiated by the Secretary of State in 2003, this fellowship program engages the American academic science, technology, engineering and medical communities in the design and implementation of U.S. foreign policy and international development.

Jefferson Science Fellows spend one year on assignment at the U.S. Department of State or the U.S. Agency for International Development (USAID) as science advisors on foreign policy/international development issues. Assignments are tailored to the needs of the hosting office, while taking into account the Fellows' interests and areas of expertise.

The fellowship is open to tenured, or similarly ranked, academic scientists, engineers, and physicians from U.S. institutions of higher learning. Applicants must hold U.S. citizenship and will be required to obtain a security clearance prior to beginning the fellowship.

The deadline for applications for the 2019-2020 program year is **October 31, 2018**. To learn more about the Jefferson Science Fellows program and to apply, visit **www.nas.edu/jsf**

The Jefferson Science Fellows program is administered by the National Academies of Sciences, Engineering, and Medicine and supported by the U.S. Department of State and the United States Agency for International Development.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

NRC Research Associateship Programs

The National Academy of Sciences, Engineering, and Medicine offers postdoctoral and senior research awards on behalf more than 20 U.S. federal research agencies and affiliated institutions with facilities at over 100 locations throughout the U.S. and abroad.

We are actively seeking highly qualified candidates including recent doctoral recipients and senior researchers. Applications are accepted during four annual review cycles (with deadlines of November 1, February 1, May 1, and August 1,).

Awardees have the opportunity to:

- conduct independent research in an area compatible with the interests of the sponsoring laboratory
- devote full-time effort to research and publication
- access the excellent and often unique facilities of the federal research enterprise
- collaborate with leading scientists and engineers at the sponsoring laboratories

Benefits of an NRC Research Associateship award include:

- 1 year award, renewable for up to 3 years
- Stipend ranging from $45,000 to $80,000, higher for senior researchers
- Health insurance, relocation benefits, and professional travel allowance

Applicants should hold, or anticipate receiving, an earned doctorate in science or engineering. Degrees from universities abroad should be equivalent in training and research experience to a degree from a U.S. institution. Some awards are open to foreign nationals as well as to U.S. citizens and permanent residents.

The National Academies of Sciences, Engineering, and Medicine's Fellowships Office has conducted the NRC Research Associateship Programs in cooperation with sponsoring federal laboratories and other research organizations approved for participation since 1954. Through national competitions, the Fellowships Office recommends and makes NRC Research Associateship awards to outstanding postdoctoral and senior scientists and engineers for tenure as guest researchers at participating laboratories. A limited number of opportunities are available for support of graduate students in select fields.

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

POSTDOC OPPORTUNITIES

The University of Georgia

Multiple Postdoctoral positions available in Genetics at the University of Georgia in diverse areas of molecular and population genetics and genomics. **http://www.genetics.uga.edu/**; **https://postdocs.uga.edu/**.

The Ye lab studies genetic adaptation to diet during human evolution and the genetic basis of complex metabolic diseases. Experience in population genomics or bioinformatics preferred Contact: Kaixiong (Calvin) Ye, **Kaixiong.Ye@uga.edu**

The White lab studies mechanisms underlying the evolution of young sex chromosomes in the threespine stickleback fish. Contact: Dr. Mike White, **whitem@uga.edu**

The Goll lab studies the molecular bases of heterochromatin establishment during embryonic development. Experience with RNA-seq and ChIP preferred. Contact: Dr. Mary Goll, **Mary.Goll@uga.edu**

The Sweigart and Parrott labs are developing new transgenic approaches in the wildflower genus Mimulus (monkeyflower). Contact: Dr. Andrea Sweigart, **sweigart@uga.edu**

The Terns lab studies the basic biology and applications of CRISPR-based prokaryotic anti-viral immune systems. Contact: Dr. Mike Terns, **mterns@uga.edu**

PHYSICAL CHEMISTRY FACULTY POSITIONS
Boston College
Chemistry Department

The Chemistry Department of Boston College invites applications for two ***tenure-track positions*** to be effective in the fall of 2019. Applicants will be evaluated based on their potential to establish a prominent and well-funded research program and to excel in teaching at the graduate and undergraduate levels. Successful applicants will join a department of approximately 120 doctoral students, 30 postdoctoral fellows, 200 undergraduate majors, and an internationally recognized faculty.

Assistant Professor in the area of Physical Chemistry with a focus on computational/theoretical chemistry requires a Ph.D. in Chemistry or related areas; postdoctoral experience is desirable but not required. *Exceptional candidates with expertise in other areas are encouraged to apply, as well.* The candidates are expected to have published in top-refereed journals and demonstrated the ability to perform outstanding independent research.

Interested applicants must submit a cover letter (which includes the names of three references), a graphical executive summary of research plans (one page), curriculum vitae, a summary of research plans (eight pages maximum), a statement of teaching philosophy and arrange to have three letters of reference submitted via the online faculty application at: **apply.interfolio.com/53188**. All application materials must be submitted electronically on or prior to **October 15, 2018**.

Boston College, a university of eight schools and colleges, is an Equal Opportunity Employer and supports Affirmative Action.

TENURE-TRACK ASSISTANT PROFESSOR PHYSICAL CHEMISTRY

Harvard University Faculty of Arts and Sciences
Department of Chemistry and Chemical Biology

Position Description: Candidates are invited to apply for a tenure-track assistant professorship in physical chemistry, broadly defined, including experimental and theoretical research in areas such as but not limited to atomic and molecular physics, biophysical chemistry, condensed matter, quantum science and ultrafast spectroscopy. The appointment is expected to begin on July 1, 2019. The tenure-track professor will be responsible for teaching at the undergraduate and graduate levels. We are seeking candidates who have an outstanding research record and a strong commitment to undergraduate and graduate teaching.

Basic Qualifications: Doctorate or terminal degree in chemistry or related discipline required by the time the appointment begins.

Additional Qualifications: Demonstrated experience in teaching is desired.

Special Instructions: Please submit the following materials through the ARIeS portal (http://academicpositions.harvard.edu/postings/8371). Applications must be submitted no later than October 15, 2018.

1. Cover letter
2. *Curriculum Vitae* with publications list
3. Teaching statement (describing teaching approach and philosophy)
4. Outline of future research plans
5. Names and contact information of 3-5 references. Three letters of recommendation are required, and the application is complete only when all three letters have been received.
6. Selected publications

Contact Information: Susan M. Kinsella, Search Administrator, Department of Chemistry and Chemical Biology, Faculty of Arts and Sciences, Harvard University, 12 Oxford St., Cambridge, MA 02138. Phone: 617-496-4088. kinsella@chemistry.harvard.edu

Harvard is an Equal Opportunity Employer and all qualified applicants will receive consideration for employment without regard to race, color, religion, sex, national origin, disability status, protected veteran status, gender identity, sexual orientation, pregnancy and pregnancy-related conditions, or any other characteristic protected by law.

SCHOOL OF MEDICINE & DENTISTRY
UNIVERSITY of ROCHESTER MEDICAL CENTER

Tenure-Track/Tenured Faculty Position

The Center for Oral Biology in the Eastman Institute for Oral Health invites applications for a faculty position at the early or mid-career level. Successful applicants should have a PhD, MD, DDS, or combined degrees, and demonstrated ability to conduct an innovative research program to investigate an area of science relevant to human disease/oral biology; including: tooth and craniofacial development; salivary gland biology; orofacial pain; or, oral bacteriology/immunology. Preference will be given to applications that complement ongoing programs or bring novel expertise and research perspectives. Individuals seeking an appointment must have demonstrated the ability to conduct independent research. The Center of Oral Biology is located in the state-of-the-art Arthur Kornberg Medical Research Building at the University of Rochester School of Medicine and Dentistry. Faculty members in the Center carry joint appointments in appropriate academic departments and participate in graduate student training in several graduate programs at the University of Rochester. More information about the Center and available positions can be found on the internet (http://www.urmc.rochester.edu/center-oral-biology/).

For further details and to apply online, please go to: **http://www.rochester.edu/working/hr/jobs/** (Job ID **#203687**). Please provide your curriculum vitae, statement of current and future research interests, and names and addresses of at least three references.

The University of Rochester is an Equal Opportunity Employer. Women and minorities are encouraged to apply.

Yale University
School of Medicine

FACULTY POSITION AT THE ASSISTANT PROFESSOR LEVEL

DEPARTMENT OF CELLULAR AND MOLECULAR PHYSIOLOGY

The Department of Cellular and Molecular Physiology is conducting a search for new faculty members at the assistant professor level.

The search seeks candidates whose research connects the properties of molecules to the properties of physiological systems.

Excellent opportunities are available for collaborative research, as well as for graduate and medical student teaching. Candidates must hold a Ph.D., M.D., or equivalent degree. Applicants should include a curriculum vitae, a statement of research interests and goals, and should arrange to have three letters of reference sent. Applicants should apply at the following website:
http://apply.interfolio.com/53471

Application Deadline: **October 19, 2018**

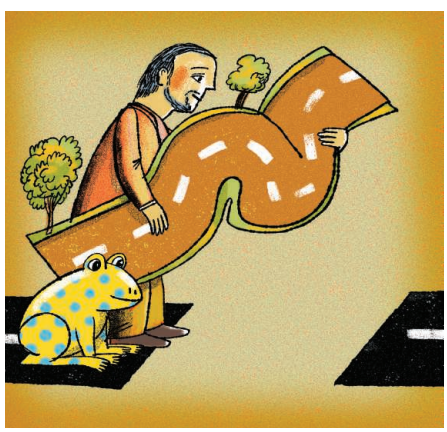
Yale University is an Affirmative Action/Equal Opportunity Employer and welcomes applications from women, persons with disabilities, covered veterans, and members of minority groups.

By **Sidney F. Gouveia**

The detour that became a shortcut

Like many science students, I had always envisioned a pretty straightforward career path: a graduate degree, postdoctoral research, and—if all went well—a faculty position. But I was thrown off this track before I even completed my bachelor's degree in biology. A university strike delayed my graduation, and as a result I missed the graduate school application deadline. Suddenly I had no idea what my future might hold, and I needed to make a living. I was relieved to be offered a job managing a newly established conservation area in my home state of Sergipe in Brazil, and I was excited about working to support biodiversity. But in the back of my mind, I worried that the job would take me in the wrong direction, away from the academic career I still desired.

The idea of managing a protected area was appealing, but my everyday workload was far from inspiring. I handled some interesting challenges, such as reaching a compromise with the ranchers whose cattle needed to cross the reserve for water. But I spent more time on paperwork and meetings than on ecosystems and biodiversity. And the only opportunities for career advancement were administrative positions, one step away from becoming a career bureaucrat. That was not how I wanted to spend my life.



"I worried that the job would take me in the wrong direction."

So, 3 years in, I decided that I needed to make a change. I had managed to complete a master's degree in ecology and conservation on the side while working at my day job, and in my spare time, I studied the reserve's frogs. But it was time to get back on the academic ladder full time. Applying to Ph.D. programs was the obvious next step.

When I was accepted into a program in ecology and evolution, I couldn't wait to trade government paperwork for the intellectual stimulation of being fully immersed in research. Yet I was a bit unsure how well my transition back to academia would go. Would the skills I developed during my years at the reserve be of any use in my new endeavor, or would I be hopelessly rusty and lost?

At first, as I had feared, I felt a little behind my fellow students. Despite the supportive environment, I couldn't escape the fact that I lacked skills vital to my new research field, such as programming and advanced statistics. I doubted that I would ever make any progress in my research or produce a decent thesis.

But I soon realized that, during my time at the reserve, I had developed my own valuable skills. Managing the conservation area, which relied on community participation and compromise, had taught me to work collaboratively. Through juggling reserve management, community meetings, and endless paperwork, I had learned to work creatively and, above all, to get things done. I soon realized that doing multivariate analyses was no harder than dealing with the multidimensional problems of reserve management, and that writing scientific papers was no more demanding than compiling environmental policy reports. And my collaborative approach served me well as I worked closely with my new peers. In time, I gained the confidence I needed to succeed.

Three years after starting my Ph.D., I found what I hoped would be my next career step: a permanent faculty position at my alma mater. As I went into overdrive to finish my thesis and put together a compelling application, I drew on abilities honed during my time managing the conservation area—including meeting deadlines and multitasking effectively—to wrap up my degree and land the job.

Looking back, I appreciate how my precocious experience as a reserve administrator has contributed to my progress in academia. I had been thrown into the deep end, alone at a completely new reserve, where I was expected to mediate conflicts and solve problems with next to no resources. In turn, I developed creativity, persuasiveness, and patience. My initial detour from my academic goals ended up being a shortcut to the career I have always wanted. ■

Sidney F. Gouveia is an adjunct professor at the Federal University of Sergipe in São Cristóvão, Brazil. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER